

ECOLE DES HAUTES ETUDES EN SCIENCES SOCIALES

Thèse  
pour l'obtention du grade de  
Docteur de l'EHESS  
Discipline : Sciences Cognitives

**LE ROLE DE LA PROSODIE  
DANS LE TRAITEMENT SYNTAXIQUE ADULTE  
ET L'ACQUISITION DE LA SYNTAXE**

par Séverine Millotte

10 Octobre 2005

Thèse dirigée par Emmanuel DUPOUX  
et co-dirigée par Anne CHRISTOPHE

Membres du Jury :

Emmanuel BIGAND (rapporteur)  
Emmanuel DUPOUX (directeur de thèse)  
Michel FAYOL  
Uli FRAUENFELDER (rapporteur)  
Thierry NAZZI

## REMERCIEMENTS

Me voilà arrivée en fin de thèse, la fin d'une étape mais aussi le début d'une nouvelle aventure, et je voudrais profiter de ce moment pour remercier les personnes sans qui je n'en serais pas là aujourd'hui...

Tout d'abord, un grand merci à la Direction Générale de l'Armement qui a financé cette thèse pendant 3 ans et qui a permis que ma recherche se passe dans de bonnes conditions.

C'est donc le fruit de trois années de thèse que je présente ici et que je soutiendrai bientôt : merci à Michel Fayol et Thierry Nazzi d'avoir accepté de faire partie de mon jury de thèse, ainsi qu'à Emmanuel Bigand et Uli Frauenfelder d'avoir accepté d'être mes rapporteurs de thèse, en plus de faire partie de mon jury. J'espère que vous apprécierez cette thèse, autant que j'ai apprécié de la réaliser.

Ma thèse s'est déroulée dans un endroit merveilleux, propice à un épanouissement tant professionnel que personnel : le Laboratoire de Sciences Cognitives et Psycholinguistique. Et je voudrais remercier tous ceux que j'ai pu y rencontrer : tous, qu'ils soient chercheurs ou étudiants, permanents ou de passage, m'ont apporté soutien et encouragements, m'ont permis d'approfondir les idées que j'avais et d'en découvrir de nouvelles grâce à de nombreuses discussions toujours constructives...

Je voudrais remercier tout particulièrement Emmanuel Dupoux : c'est lui qui m'a fait confiance, qui a cru en moi et qui m'a ouvert les portes de son laboratoire. A notre première rencontre, il m'a conseillé d'entreprendre le DEA de Sciences Cognitives pour compléter tout ce que j'avais acquis pendant ma formation universitaire en psychologie. Même si j'étais déçue de ne pas commencer ma thèse directement, avec le recul je sais que je lui dois beaucoup car cette année de DEA m'a ouvert l'esprit sur d'autres domaines de recherche et m'a permis d'approfondir et de compléter tout ce que j'avais pu apprendre jusqu'à présent. J'étais donc dans les meilleures conditions pour commencer ma thèse qu'il a accepté d'encadrer. Merci pour tout.

Travailler au LSCP m'a permis de rencontrer et de nouer des liens avec de nombreuses personnes : merci à Sharon Peperkamp, Franck Ramus et Sid Kouider pour leur aide et leur soutien, mais aussi pour leur écoute et leur bonne humeur. Des remerciements tout particuliers à Susana Franck, Carmen Diop et Michel Dutat, qui m'ont permis de surmonter et de ne pas sombrer dans les rouages administratifs et problèmes informatiques. Savoir qu'ils étaient là pour m'aider et me conseiller était très important. Merci aussi à Sylvie Margules, notre reine du recrutement et des manip bébés : pendant ces mois que nous avons passés à tester ensemble, nous avons connu des moments de joie et d'exaltation, mais aussi des moments de doute et de découragement... nous les avons vécus ensemble, partageant nos problèmes, nos craintes, nos espoirs, mais aussi nos moments de délire : vive les MApou et les maPI !!! Quant aux expériences menées auprès des adultes, elles ne se seraient pas déroulées aussi bien sans l'aide de Anne-Marie Argenti et de Anne-Caroline Fievet, alors merci.

Au LSCP, de nombreux étudiants peuplent les bureaux (parfois trop petits pour tous nous contenir !!!). Je voudrais leur dire à tous qu'ils ont toujours été un soutien pour moi et que les moments d'échanges, de discussions et de rigolade que j'ai vécus avec eux étaient très importants. Des merci tout particuliers à Savita Bernal, Inga Vendelin, Gayaneh Szenkovits, Rozenn Le Calvez, Emmanuel Chemla... Je voudrais aussi dire à Isabelle Darcy et Charlotte Jacquemot que leur départ a laissé un grand vide : quand je suis arrivée au labo, c'est vers elles que je me tournais pour trouver conseils et soutien, et si je me suis si vite sentie intégrée, c'est en partie grâce à elle, alors merci beaucoup.

Et enfin (last but not least), il manque une personne du LSCP dont je n'ai pas encore parlé, mais elle méritait à elle toute seule un paragraphe de remerciements : c'est bien entendu Anne Christophe. Je ne sais pas comment je vais faire pour exprimer en quelques mots tout ce que je lui dois... Anne est une personne extraordinaire, que cela soit sur un plan professionnel ou personnel : elle m'a appris énormément pendant ces années, sur mon travail (me guidant sur

le chemin de la connaissance des différentes compétences et théories nécessaires à ma recherche) mais aussi sur moi-même. Elle a toujours su me rassurer et m'écouter quand je doutais de moi ou de mon travail, elle a toujours été là pour me soutenir dans les moments difficiles. Elle a su me guider et m'accompagner sur le chemin de mon épanouissement personnel et professionnel. Elle a aussi répondu présente pour m'accueillir chez elle et me faire partager des instants de sa vie privée (merci aussi à Pierre, Julien, Hélène, Charles et Alexandre, avec qui j'ai partagé de nombreuses soirées de détente et qui m'ont accueillie à bras ouverts chez eux pendant de nombreux mois). Je dois énormément à Anne Christophe : des conversations lumineuses et productives, aux échanges plus personnels et éclats de rire, elle a su créer le climat idéal pour s'épanouir et travailler dans les meilleures conditions possibles. Elle sera à jamais ma référence et mon mentor, et je voudrais donc la remercier très sincèrement pour ce qu'elle est et ce qu'elle m'a donné.

Ces années de thèse m'ont aussi permis de rencontrer de nombreux chercheurs, qui tous ont pris le temps de m'écouter, de me conseiller et de me faire progresser : je voudrais remercier en particulier Francois Dell, Jeff Lidz, Janet Pierrehumbert, Roger Wales, Sandy Waxman et Pauline Welby.

J'ai aussi eu la chance d'aller passer un mois aux Etats-Unis, à Providence, dans le laboratoire de Jim Morgan qui m'a appris à tester les jeunes enfants. Merci donc à Jim (et à Gina) qui m'ont accueillie pendant un mois. Merci aussi à Katherine Demuth, Katherine White, Melanie Soderstrom, Lorie Pinault et Karen Rathbun pour leur accueil chaleureux, leur soutien et leur conseils. J'espère vous revoir tous bientôt.

Sur un plan plus privé, je tiens à remercier mes amis et ma famille qui ont toujours été là pour m'écouter, me soutenir, m'aider, créer les moments de détente et de bonne humeur qui me permettaient de me ressourcer et de repartir plus forte. Je pense plus particulièrement à Caroline Floccia, qui m'a donné envie de faire de la recherche en psycholinguistique et sans qui je ne serais pas là. Et je tiens aussi à remercier particulièrement mon Papa et Roselyne : c'est grâce à eux que j'ai pu venir finir mes études sur Paris. Ils ont toujours eu confiance en moi et ont tout fait pour que ma thèse se déroule bien : nos soirées passées à discuter sérieusement ou à rire aux larmes étaient des moments privilégiés, que nous n'avions pas pu vivre pleinement jusqu'à cette période, et qui resteront un très bon souvenir pour moi.

Et enfin (et surtout...), je voudrais remercier du fond du cœur celui sans qui rien n'aurait été possible et sans qui rien ne serait aussi beau : mon Nico. J'ai la chance de partager ma vie avec l'homme le plus merveilleux qui existe sur cette Terre : il a toujours cru en moi, et il a toujours été là pour moi, présent pour m'écouter (et je parle beaucoup...), me conseiller, me soutenir, prendre soin de moi. Il a accepté que je parte sur Paris pour faire ma thèse, ce qui voulait dire de nombreux moments de séparation. Ca n'a pas toujours été facile mais nous nous en sommes sortis et nous voilà encore plus forts qu'avant. Savoir que j'allais le retrouver les week-ends était ce qui me donnait du courage durant les semaines de travail : il est mon oasis, il est tout ce qui compte le plus, et je suis fière de lui et heureuse de partager sa vie. J'espère que lui aussi sera fier de moi et du travail que j'ai accompli pendant ces années. C'est donc tout naturellement que je lui dédie cette thèse.

*A Nicolas,*

# TABLE DES MATIERES

<b>INDEX DES FIGURES .....</b>	<b>V</b>
<b>INDEX DES TABLEAUX.....</b>	<b>VII</b>
<b>INTRODUCTION .....</b>	<b>1</b>
<b>CHAPITRE 1</b>	
<b>LE TRAITEMENT SYNTAXIQUE ADULTE .....</b>	<b>8</b>
1.1 LES DIFFERENTS MODELES DE TRAITEMENT SYNTAXIQUE .....	8
1.1.1 <i>Les modèles en série (modulaires)</i> .....	9
1.1.2 <i>Les modèles en parallèle (non modulaires)</i> .....	13
Rôle des informations lexicales .....	14
Rôle du contexte.....	17
Rôle des informations statistiques.....	19
1.2 LE ROLE DE LA PROSODIE DANS LE TRAITEMENT SYNTAXIQUE .....	22
1.2.1 <i>Qu'est-ce que la prosodie ?</i> .....	22
Définition .....	22
Liens entre prosodie et syntaxe .....	23
1.2.2 <i>Le rôle des indices prosodiques dans la résolution d'ambiguïtés syntaxiques</i> .....	34
Résolution d'ambiguïtés syntaxiques temporaires : stimuli synthétisés .....	36
Résolution d'ambiguïtés syntaxiques temporaires : stimuli naturels .....	40
<b>CHAPITRE 2</b>	
<b>UTILISATION DES FRONTIERES DE GROUPES PHONOLOGIQUES DANS</b>	
<b>LA RESOLUTION D'AMBIGUITES SYNTAXIQUES (ETUDE 1).....</b>	<b>45</b>
2.1 EXPERIENCE 1 : COMPLETION DE PHRASES.....	47
2.1.1 <i>Méthode</i> .....	47
Matériel .....	47
Procédure .....	49
Participants.....	50
2.1.2 <i>Résultats</i> .....	51
Analyses acoustiques des stimuli .....	51
Résultats de la tâche de complétion de phrases .....	59
2.1.3 <i>Discussion</i> .....	61
2.2 EXPERIENCE 2 : DETECTION DE MOT .....	63

2.2.1 Méthode.....	64
Matériel.....	64
Procédure.....	67
Participants.....	68
2.2.2 Résultats.....	68
Analyses des temps de réaction et des oublis.....	68
Analyses des fausses alarmes.....	72
Analyses des réponses rapides.....	75
2.2.3 Discussion.....	78
2.3 DISCUSSION GENERALE.....	79
2.3.1 Résultat général.....	79
2.3.2 Modèles de traitement syntaxique : quels indices sont utilisés ?.....	79
2.3.3 Locuteurs experts versus locuteurs naïfs.....	83

### **CHAPITRE 3 UTILISATION D'INDICES PROSODIQUES PRODUITS PAR DES LOCUTEURS NAÏFS (ETUDE 2)..... 85**

3.1 LE DEBAT ACTUEL SUR LA PRODUCTION D'INDICES PROSODIQUES INFORMATIFS PAR DES LOCUTEURS NAÏFS ET NON-EXPERTS.....	85
3.2 NOTRE ETUDE : PRODUCTION D'INDICES PROSODIQUES PAR DES LOCUTEURS NAÏFS ET UTILISATION DE CES INDICES POUR RESOUDRE DES AMBIGUÏTES SYNTAXIQUES.....	91
3.2.1 Production des phrases ambiguës.....	91
Choix des stimuli à enregistrer.....	92
Enregistrement des stimuli.....	94
3.2.2 Complétion de phrases produites par les locuteurs naïfs.....	97
Méthode.....	97
Résultats.....	99
Discussion.....	103

### **CHAPITRE 4 ACQUISITION DE LA SYNTAXE DE SA LANGUE MATERNELLE..... 108**

4.1 INITIALISATION DISTRIBUTIONNELLE POUR L'ACQUISITION DES CATEGORIES GRAMMATICALES.....	109
4.1.1 Approche théorique.....	109
4.1.2 Analyses de corpus.....	111
Rôle des mots adjacents.....	111
Comparaison de paires minimales de phrases.....	112

Rôle des cadres lexicaux .....	113
4.2 INITIALISATION PHONOLOGIQUE POUR L'ACQUISITION DES CATEGORIES SYNTAXIQUES.....	114
4.2.1 <i>Différence entre les mots de contenu et les mots grammaticaux</i> .....	114
Analyses de corpus.....	115
Données expérimentales sur l'utilisation de ces informations .....	117
4.2.2 <i>Différence entre les noms et les verbes</i> .....	122
Analyses de corpus.....	122
Données expérimentales sur l'utilisation de ces informations .....	125
4.3 INDICES PROSODIQUES DE LA STRUCTURE SYNTAXIQUE DES PHRASES.....	127
4.3.1 <i>Approche théorique</i> .....	127
4.3.2 <i>Sensibilité des jeunes enfants aux indices prosodiques</i> .....	128
4.4 LIENS ENTRE PROSODIE ET MOTS GRAMMATICaux.....	133

## CHAPITRE 5

### PERCEPTION ET UTILISATION DES FRONTIERES DE GROUPES PHONOLOGIQUES PAR DES ENFANTS DE 16 MOIS (ETUDE 3)..... 137

5.1 METHODE.....	140
5.1.1 <i>Matériel</i> .....	140
5.1.2 <i>Equipement</i> .....	141
5.1.3 <i>Procédure</i> .....	142
5.1.4 <i>Participants</i> .....	145
5.2 RESULTATS .....	146
5.2.1 <i>Analyses acoustiques des stimuli</i> .....	146
Analyse de la durée des segments.....	146
Analyse de la fréquence fondamentale .....	148
Analyse de l'énergie des voyelles .....	150
5.2.2 <i>Résultats comportementaux</i> .....	151
5.3 DISCUSSION.....	157

## CHAPITRE 6

### ROLE DES MOTS GRAMMATICaux ET DES FRONTIERES PROSODIQUES DANS L'ACQUISITION DE LA SYNTAXE (ETUDE 4)..... 161

6.1 METHODE.....	165
6.1.1 <i>Matériel</i> .....	165
6.1.2 <i>Procédure</i> .....	167
6.1.3 <i>Participants</i> .....	169
6.2 RESULTATS .....	169

6.2.1 Analyses des réponses noms et verbe .....	169
6.2.2 Analyses des temps de réaction .....	173
6.2.3 Analyses détaillées par sous-conditions .....	174
Comparaison des sous-conditions « avec clitique » .....	175
Comparaison des sous-conditions « sans clitique » .....	176
6.3 DISCUSSION .....	178
6.3.1 Comparaison du rôle des indices utilisés .....	178
6.3.2 Qu'est-ce qui peut être considéré comme un mot grammatical ? .....	183
<b>CONCLUSION GENERALE .....</b>	<b>187</b>
<b>ANNEXES .....</b>	<b>209</b>
<b>REFERENCES BIBLIOGRAPHIQUES .....</b>	<b>253</b>

# INDEX DES FIGURES

FIGURE 1 : ANALYSES ACOUSTIQUES DES STIMULI – ETUDE 1 .....	52
FIGURE 2 : ANALYSES ACOUSTIQUES DES STIMULI – ETUDE 1 .....	54
FIGURE 3 : RESULTATS DE L'EXPERIENCE 1 – ETUDE 1 .....	60
FIGURE 4 : RESULTATS DE L'EXPERIENCE 2 – ETUDE 1 (MOYENNE DES TEMPS DE REACTION ET POURCENTAGE D'OUBLIS).....	69
FIGURE 5 : RESULTATS DE L'EXPERIENCE 2 – ETUDE 1 (POURCENTAGE MOYEN DE FAUSSES ALARMES).....	74
FIGURE 6 : RESULTATS DE L'EXPERIENCE 2 – ETUDE 1 (NOMBRE MOYEN DE REPONSES ADJECTIF ET VERBE DONNEES) .....	76
FIGURE 7 : RESULTATS DE L'ETUDE 2 (NOMBRE MOYEN DE REPONSES ADJECTIF ET VERBE, TOUS LOCUTEURS CONFONDUS.....	101
FIGURE 8 : RESULTATS DE L'ETUDE 2 (NOMBRE MOYEN DE REPONSES ADJECTIF ET VERBE, POUR LES DIFFERENTS LOCUTEURS .....	103
FIGURE 9 : ANALYSES ACOUSTIQUES DES STIMULI – ETUDE 3).....	148
FIGURE 10 : RESULTATS DE L'ETUDE 3 (POURCENTAGE MOYEN D'ORIENTATION DE LA TETE DU GROUPE BISYLLABIQUE ET DU GROUPE MONOSYLLABIQUE .....	152
FIGURE 11 : RESULTATS DE L'ETUDE 4 (POURCENTAGE MOYEN DE REPONSES NOM ET VERBE DANS CHAQUE CONDITION .....	170
FIGURE 12 : RESULTATS DE L'ETUDE 4 (TEMPS DE REACTION MOYEN POUR REPENDRE AUX PHRASES NOM ET VERBE DANS LES DEUX CONDITIONS EXPERIMENTALES).....	173
FIGURE 13 : RESULTATS DE L'ETUDE 4 (POURCENTAGE MOYEN DE REPONSES NOM ET VERBE DANS LES TROIS SOUS-CONDITIONS DE LA CONDITION « AVEC CLITIQUE »..	175
FIGURE 14 : RESULTATS DE L'ETUDE 4 (POURCENTAGE MOYEN DE REPONSES NOM ET VERBE DANS LES DEUX SOUS-CONDITIONS DE LA CONDITION « SANS CLITIQUE » ..	177
FIGURE 15 : MODELE D'ACQUISITION DE LA SYNTAXE DERIVANT DE NOTRE HYPOTHESE DE RECHERCHE.....	197
FIGURE 16 : RESULTATS DE L'ETUDE 3 DETAILLES PAR ITEMS (phrases bisyllabiques)..	231
FIGURE 17 : RESULTATS DE L'ETUDE 3 DETAILLES PAR ITEMS (monosyllabiques).....	231
FIGURE 18 : RESULTATS DES ENFANTS AMERICAINS DE 10 ET 13 MOIS .....	244
FIGURE 19 : RESULTATS 13 MOIS FRANÇAIS – EXPERIENCE PILOTE.....	245

FIGURE 20 : RESULTATS 16 MOIS FRANÇAIS – EXPERIENCE PILOTE.....	246
FIGURE 21 : COMPARAISON DES RESULTATS AMERICAINS ET FRANÇAIS.....	249

# INDEX DES TABLEAUX

TABLEAU 1 : VALEUR MOYENNE DE LA FREQUENCE FONDAMENTALE DES DEUX VOYELLES PLACEES AVANT LA PREMIERE POSITION DE FRONTIERE DE GROUPE PHONOLOGIQUE (ETUDE 1).....	56
TABLEAU 2 : VALEUR MOYENNE DE LA FREQUENCE FONDAMENTALE DES DEUX VOYELLES PLACEES AVANT LA DEUXIEME POSITION DE FRONTIERE DE GROUPE PHONOLOGIQUE (ETUDE 1).....	57
TABLEAU 3 : RECAPITULATIF DES QUATRE POSSIBILITES QUE RENCONTRAIENT LES PARTICIPANTS DE L'EXPERIENCE 2, SELON LA CIBLE QU'ILS DEVAIENT DETECTER ET LA PHRASE QU'ILS ENTENDAIENT .....	64
TABLEAU 4 : COEFFICIENTS DE CORRELATION ENTRE LES ESTIMATIONS DE FREQUENCE ET DE PLAUSIBILITE DONNEES POUR CHAQUE ITEM AMBIGU, ET LES RESULTATS PAR ITEM OBTENUS DANS L'ETUDE 1 .....	82
TABLEAU 5 : ANALYSES ACOUSTIQUES – ETUDE 3 .....	147
TABLEAU 6 : RESULTATS INDIVIDUELS PAR SUJETS (EXPERIENCE 1 - ETUDE 1).....	212
TABLEAU 7 : RESULTATS INDIVIDUELS PAR ITEMS (EXPERIENCE 1 - ETUDE 1).....	213
TABLEAU 8 : RESULTATS INDIVIDUELS PAR SUJETS (EXPERIENCE 2 - ETUDE 1).....	215
TABLEAU 9 : RESULTATS INDIVIDUELS PAR SUJETS (EXPERIENCE 2 - ETUDE 1 - ANALYSES DES REPONSES DONNEES AVANT LA FIN DU MOT AMBIGU) .....	216
TABLEAU 10 : RESULTATS INDIVIDUELS PAR ITEMS (EXPERIENCE 2 - ETUDE 1).....	217
TABLEAU 11 : RESULTATS INDIVIDUELS PAR ITEMS (EXPERIENCE 2 - ETUDE 1 - ANALYSES DES REPONSES DONNEES AVANT LA FIN DU MOT AMBIGU) .....	218
TABLEAU 12 : RESULTATS INDIVIDUELS PAR SUJETS TOUS LOCUTEURS (ETUDE 2) .....	221
TABLEAU 13 : RESULTATS INDIVIDUELS PAR SUJETS PAR LOCUTEURS (ETUDE 2).....	223
TABLEAU 14 : RESULTATS INDIVIDUELS PAR ITEMS TOUS LOCUTEURS (ETUDE 2) .....	224
TABLEAU 15 : RESULTATS INDIVIDUELS PAR ITEMS PAR LOCUTEURS (ETUDE 2).....	227
TABLEAU 16 : RESULTATS INDIVIDUELS PAR SUJETS (ETUDE 3 groupe bisyllabique) ....	230
TABLEAU 17 : RESULTATS INDIVIDUELS PAR SUJETS (ETUDE 3 gr. monosyllabique).....	230
TABLEAU 18 : RESULTATS INDIVIDUELS PAR SUJETS (ETUDE 4).....	241
TABLEAU 19 : RESULTATS INDIVIDUELS PAR ITEMS (ETUDE 4) .....	242

## INTRODUCTION

Comprendre le langage qui nous entoure nous semble être une tâche simple et automatique, ne requérant pas de véritable effort de notre part. Pourtant, imaginons un instant que vous vous réveilliez un matin, dans un pays inconnu, une terre loin de tout, peuplée d'hommes et de femmes qui parlent une langue que vous n'avez jamais entendue : vous les entendez produire des sons, vous vous doutez que ces sons forment des mots et que ces mots, mis bout à bout, veulent sûrement dire quelque chose... Mais vous, qui avez faim et qui aimeriez bien boire votre petit café du matin, êtes incapable de les comprendre ni de vous faire comprendre, alors que dans votre vie de tous les jours cela vous semble extrêmement facile. Alors vous vous escrimez à faire de grands signes, essayant de mimer l'action de boire et de manger, et dépité vous retournez vous allonger dans un coin... vous vous endormez... et vous vous réveillez bien au chaud dans votre lit... Quel cauchemar ! Une fois que votre petit café, tant désiré, est avalé, vous oubliez définitivement cet abominable cauchemar. Pourtant, si l'on y réfléchit un instant, cette mésaventure pose des questions fondamentales sur l'humain, et plus particulièrement sur ses compétences linguistiques.

Il est vrai qu'il nous semble tellement facile de communiquer et de comprendre le langage qui nous entoure, que nous ne nous posons pas, par exemple, les questions suivantes : comment faisons-nous en fait pour reconnaître les mots qui composent les phrases qui nous sont adressées ? Et comment faisons-nous pour comprendre aussi rapidement un discours même s'il comporte des mots ou des structures ambiguës ? Il est évident que nous, adultes, avons à notre disposition de nombreuses connaissances (phonologiques, lexicales, syntaxiques, pragmatiques...) que nous pouvons utiliser pour comprendre le message qui nous est adressé. Mais sans ces connaissances, comment ferions-nous ? Rappelez-vous ce désespoir qui vous étrennait quand vous ne pouviez ni comprendre, ni vous faire comprendre, dans cette île éloignée de tout ! Plus simplement, regardons attentivement ce qui nous entoure, et une question nous saute aux yeux : comment font les enfants qui arrivent au monde, pour apprendre leur langue maternelle alors qu'ils n'ont pas encore accès à toutes ces connaissances que, nous, nous possédons ?

En effet, la tâche des enfants est vraiment complexe : ils doivent apprendre la phonologie de leur langue maternelle (découvrir quels sont les sons pertinents dans cette

langue), trouver les mots dans les phrases qu'ils entendent (donc résoudre le problème de la segmentation lexicale), découvrir le sens de ces mots (pour construire leur lexique mental), et comprendre la syntaxe de cette langue (c'est-à-dire comprendre les règles qui permettent d'accéder au sens des phrases à partir du sens des mots). Malgré la difficulté de cette tâche, les enfants deviennent rapidement des locuteurs compétents de leur langue maternelle. Le but de cette thèse est de proposer et de tester une hypothèse précise concernant l'acquisition du langage par les enfants.

Un des premiers problèmes rencontrés par les enfants est le fait que la parole soit continue. Les indices que sont les espaces entre les mots, pour délimiter chaque unité lexicale à l'écrit, ne se trouvent pas dans le langage parlé. Dans un discours oral, il n'existe pas de signaux universels et systématiques permettant de trouver les frontières des mots. Mais, même une fois que les mots ont pu être isolés, il faut encore trouver leur signification. Lila Gleitman (1990) a montré que découvrir le sens des mots pouvait être facilité si l'on avait accès à certains aspects de la structure syntaxique des phrases ; mais puisque la syntaxe définit les relations entre les mots dans les phrases, il semble logique de postuler qu'il est nécessaire d'avoir accès aux mots et à leur signification pour pouvoir apprendre la syntaxe de sa langue maternelle. Ceci met en évidence un problème de circularité dans l'acquisition du langage, ou problème d'initialisation : les enfants auraient besoin du lexique pour apprendre la syntaxe, tandis que la syntaxe serait nécessaire pour construire le lexique. Plus généralement, l'apprentissage de chacune des composantes du langage pourrait être facilité si l'on suppose des connaissances préalables au niveau des autres composantes ; encore faut-il que ces connaissances aient pu être acquises indépendamment les unes des autres.

Pour résoudre ce paradoxe, l'hypothèse de l'initialisation phonologique a été proposée (Morgan, 1986 ; Morgan & Demuth, 1996a), et postule un rôle important des informations de bas niveau, directement accessibles dans le signal de parole, pour l'acquisition du langage. Ainsi, une analyse purement phonologique du signal de parole pourrait aider les enfants à acquérir le lexique et la syntaxe de leur langue maternelle.

Notre travail de recherche s'est directement intéressé à cette hypothèse d'initialisation phonologique (dont un sous-ensemble est formé par les modèles d'initialisation prosodique), en étudiant plus particulièrement la question de l'acquisition de la syntaxe.

Pour comprendre puis maîtriser la syntaxe de leur langue maternelle, les enfants doivent entre autres réussir à associer les mots de leur langue maternelle aux catégories syntaxiques existant dans cette langue, et comprendre comment ces catégories syntaxiques sont organisées dans les phrases.

Notre hypothèse de recherche repose sur l'utilisation conjointe de deux indices, qui sont directement accessibles dans le signal de parole, et qui pourraient donc être découverts et utilisés avant même que l'enfant ne connaisse beaucoup de mots : les indices prosodiques et les mots grammaticaux.

La prosodie pourrait être définie comme la mélodie du langage. Elle concerne toutes les variations de rythme et d'intonation dans les phrases. Car heureusement, nous ne parlons pas comme des robots, d'une voix monocorde et monotone. I..ma..gi..nez..si..non..comme..ce..se..rait..di..ffi..cile..de..suivre..une..con..ver..sa..tion !

Les indices prosodiques sont directement accessibles depuis le signal acoustique, et correspondent à l'ensemble des modifications de la fréquence fondamentale, de l'intensité liée à la force de la voix, et de la durée des différentes unités linguistiques (segments, syllabes). Les mots et les syllabes sont ainsi regroupés en unités prosodiques : et même si l'on ne connaît pas ce terme, tout le monde a l'intuition que, dans les phrases, on ne trouve pas des suites de mots isolés mais plutôt des groupes de mots les uns à la suite des autres.

Ces groupes de mots sont les unités prosodiques, et celles-ci peuvent être intégrées dans une structure prosodique hiérarchique (voir Shattuck-Hufnagel & Turk, 1996, pour une revue). Parmi ces unités, celles sur lesquelles nous avons centré notre attention sont les groupes phonologiques, c'est-à-dire des unités prosodiques de taille intermédiaire, qui contiennent typiquement un ou deux mots de contenu (noms, verbes, adjectifs...) ainsi que des mots grammaticaux qui leur sont associés (articles, auxiliaires...), et qui sont délimitées par des allongements en fin d'unité et des variations de fréquence fondamentale.

Une des caractéristiques des groupes phonologiques, essentielle pour notre propos, est que leurs frontières correspondent toujours à des frontières de constituants syntaxiques. La perception de ces frontières pourrait donc permettre aux enfants de segmenter l'input de parole qu'ils reçoivent en unités syntaxiques.

Mais considérer uniquement l'information apportée par les frontières prosodiques n'est pas suffisant pour pouvoir réaliser une analyse syntaxique des phrases : en effet, prenons par exemple la phrase « le chat a mangé la souris ». Percevoir les frontières prosodiques pourrait permettre à l'enfant de diviser cette phrase complète en unités plus petites, comme « [xx] [xxx] [xxx] » (où les frontières prosodiques sont représentées par des crochets, et où les x représentent les différentes syllabes). Mais comment savoir si la première unité est un groupe nominal ou un groupe verbal ou un groupe prépositionnel... ? Heureusement, un autre indice est disponible dans le signal de parole : les mots grammaticaux.

Ces mots grammaticaux représentent la catégorie des mots de la classe fermée (en opposition aux mots de la classe ouverte, ou mots de contenu), et sont formés des articles, pronoms, auxiliaires, prépositions... Ils sont considérés comme apportant les informations sur la structure des phrases (alors que les mots de contenu sont ceux qui transmettent le sens des phrases).

Il a été montré que ces mots grammaticaux avaient des caractéristiques particulières, qui pouvaient être utilisées par les enfants pour les extraire des phrases et les différencier des mots de contenu (voir par exemple, Morgan, Shi, & Allopenna, 1996; Shi, Werker, & Morgan, 1999). Par exemple, ils sont moins accentués que les mots de contenu, plus courts, beaucoup plus fréquents, et ils apparaissent généralement au début des unités prosodiques.

Ces mots grammaticaux semblent jouer un rôle important dans la syntaxe, et plus particulièrement dans la catégorisation syntaxique des items lexicaux. Par exemple, comment sait-on que le mot « table » est un nom, ou que le mot « manger » est un verbe ? On pourrait penser que « ...c'est comme ça ! une table est une table et on a appris que c'est un nom... ». Nous l'avons appris certes, mais pas obligatoirement par un apprentissage explicite. Il est vrai qu'un dictionnaire nous donnera directement cette information (table, nom féminin...), mais il serait dommage de devoir attendre de savoir lire un dictionnaire pour acquérir ces connaissances syntaxiques qui sont nécessaires à la bonne compréhension des phrases que nous entendons. En fait, la réponse éventuelle imaginée ci-dessus, comporte un élément de réponse important à notre question : effectivement, qu'y a-t-il avant le mot « table » dans cette phrase ? Un mot grammatical, et plus précisément un article !

Ainsi, les enfants pourraient utiliser le fait que les noms sont généralement précédés d'un article, alors que les verbes sont généralement précédés d'un pronom. Même nous, adultes, pouvons utiliser cette association : si l'on nous parle de « une jaurime », nous penserons naturellement que ce nouveau mot est un nom (et qu'il se rapporte probablement à un objet) ; par contre, si l'on nous dit « elle jaurime », il nous semblera naturel et évident de supposer que ce non-mot est un verbe (et qu'il se rapporte probablement à une action)

Il est bien connu que l'union fait la force, alors regardons ce qui pourrait effectivement se passer en combinant les informations apportées par ces deux types d'indices : pris ensemble, ils pourraient permettre à l'enfant de construire une première ébauche de la structure syntaxique des phrases qu'ils entendent. Reprenons l'exemple de la phrase « le chat a mangé la souris ». Dans une première étape, nous avons vu que les frontières prosodiques pourraient permettre de délimiter les constituants syntaxiques ; on obtiendrait alors la représentation suivante : « [xx] [xxx] [xxx] ». Dans une deuxième étape, les mots grammaticaux, situés au début des unités prosodiques, pourraient être extraits du signal de parole ; on aurait alors la représentation suivante : « [le x] [a xx] [la xx] ». En considérant ces deux sources d'informations, l'enfant pourrait alors avoir accès à une première représentation syntaxique de cette phrase, sous la forme suivante : « [le x]<sub>GN</sub> [a xx]<sub>GV</sub> [la xx]<sub>GN</sub> », dans laquelle les frontières syntaxiques seraient données par les frontières de groupes phonologiques, et où la nature des unités syntaxiques (groupe nominal « GN », groupe verbal « GV ») serait trouvée grâce aux mots grammaticaux, et ce avant même que l'enfant ne connaisse les mots de contenu qui composent la phrase.

C'est donc cette hypothèse de travail qui a guidé le travail de recherche présenté ici. Comme celle-ci suppose (entre autres) que les frontières prosodiques sont calculées en temps réel et exploitées pour contraindre l'analyse syntaxique, il était important de s'en assurer chez des sujets adultes. Cette thèse peut donc être divisée en deux parties : la première qui concerne la question du traitement syntaxique adulte (les trois premiers chapitres), et la deuxième qui s'intéresse à l'acquisition de la syntaxe par les enfants (les trois derniers chapitres).

Le premier chapitre commencera par dresser une revue détaillée des modèles et recherches ayant étudié la question du traitement syntaxique adulte, pour finalement se focaliser sur le lien pouvant exister entre la prosodie et la syntaxe. Nous définirons plus précisément ce qu'est la prosodie, nous aborderons la question de la congruence (ou de la non-congruence) entre structure syntaxique et structure prosodique, puis nous présenterons différentes recherches ayant étudié le rôle des indices prosodiques dans la résolution d'ambiguïtés syntaxiques. Nous verrons ainsi que de nombreuses études ont pu démontrer le rôle des frontières prosodiques majeures dans l'analyse syntaxique des phrases, alors que le rôle des frontières phonologiques a été encore peu étudié.

Il nous semblait particulièrement justifié d'étudier le rôle que pourraient jouer les frontières des groupes phonologiques, puisque ces unités prosodiques sont particulièrement liées à la structure syntaxique des phrases. De plus, ces unités comportent typiquement un ou deux mots de contenu et permettent donc de découper des phrases relativement courtes (comme celles que les enfants entendent majoritairement...) en constituants syntaxiques. Nous avons donc réalisé deux expériences pour vérifier que, si ces indices prosodiques sont présents et bien marqués dans les phrases, ils peuvent alors être utilisés par des adultes français, et leur permettre de résoudre des ambiguïtés syntaxiques temporaires. Une expérience de complétion de phrases (expérience « off-line ») et une expérience de détection de mot (expérience « on-line ») seront présentées dans le chapitre 2.

Une troisième expérience sera ensuite décrite : celle-ci a été menée pour étudier la production « naturelle » d'indices prosodiques. En effet, les premières expériences avaient utilisé une locutrice experte et informée, et nous voulions savoir si des locuteurs non-informés et non-experts produisaient eux aussi, spontanément, des indices de frontières de groupes phonologiques, suffisamment marqués et saillants pour contraindre l'analyse syntaxique des phrases faite par des auditeurs français, dans une tâche de complétion de phrases (chapitre 3).

Les résultats de cette première partie apporteront la preuve que les adultes français produisent des indices clairs de frontières de groupes phonologiques, indices qui peuvent ensuite être utilisés par des auditeurs, même en temps réel, pour guider leur analyse syntaxique des phrases et résoudre des ambiguïtés syntaxiques temporaires. Ces résultats constituent un premier pas important pour la validation de notre hypothèse de recherche sur l'acquisition de la syntaxe.

Le chapitre 4 adressera justement la question de l'acquisition de la syntaxe par les jeunes enfants, et présentera une revue détaillée de la littérature sur l'acquisition des catégories syntaxiques et l'acquisition de la structure syntaxique des phrases. Une grande importance sera donnée aux deux indices les plus pertinents pour notre propos : les mots grammaticaux et la prosodie.

La perception et l'utilisation des frontières prosodiques seront aussi étudiées expérimentalement, dans une tâche de segmentation lexicale, réalisée par des enfants français de 16 mois (chapitre 5).

Finalement, nous testerons directement la viabilité de notre hypothèse, dans une expérience de faisabilité, menée auprès d'adultes français. Rassurez-vous, nous n'allons pas les envoyer sur une île éloignée de tout pour voir comment ils peuvent analyser les phrases qui sont dites autour d'eux et qu'ils ne comprennent pas. Mais nous allons quand même leur demander de réaliser une analyse syntaxique sur des phrases n'ayant aucun sens pour eux (des phrases en jabberwocky, qui n'existent que dans le Pays des Merveilles d'Alice, et dont la caractéristique essentielle est que tous les mots de contenu sont remplacés par des non-mots). Les seules informations qu'ils auront alors à leur disposition seront la structure prosodique des phrases et la présence des mots grammaticaux. Ces seuls indices leur permettront-ils de construire une représentation syntaxique de ces phrases et d'assigner une catégorie syntaxique aux différents mots de contenu ? Une réponse positive donnerait alors encore plus de validité et de crédit à notre hypothèse de départ sur la manière dont les enfants peuvent découvrir et comprendre la syntaxe de leur langue maternelle.

# CHAPITRE 1

## LE TRAITEMENT SYNTAXIQUE ADULTE

Pour assigner un sens aux phrases qu'ils entendent, les auditeurs doivent entre autres retrouver les mots dans les phrases et les associer dans une structure syntaxique. Une des questions fondamentales pour la recherche sur le traitement du langage est de savoir comment cette structure syntaxique est construite en temps réel. Une des principales méthodologies utilisées pour répondre à cette question est d'étudier des phrases contenant des ambiguïtés syntaxiques. En effet, de telles phrases permettent aux chercheurs d'examiner comment ces ambiguïtés sont résolues et d'isoler ainsi les mécanismes impliqués plus généralement dans la construction de la structure syntaxique des phrases. Il a été démontré que différents indices pouvaient être utilisés par les auditeurs pour guider leur analyse syntaxique, ce que nous allons développer dans notre partie sur les différents modèles de traitement syntaxique. Par contre, ce n'est qu'assez récemment que les chercheurs se sont intéressés au rôle de la prosodie des phrases dans le traitement syntaxique en temps réel. Nous présenterons donc les différentes recherches qui ont étudié le rôle des informations prosodiques dans la résolution d'ambiguïtés syntaxiques, partie qui sera précédée d'un exposé sur ce qu'est la prosodie.

### 1.1 LES DIFFERENTS MODELES DE TRAITEMENT SYNTAXIQUE

L'idée générale de l'analyse syntaxique est que les adultes construisent la structure syntaxique des phrases qu'ils lisent ou entendent de manière incrémentale. Dans certaines situations, l'analyseur syntaxique rencontre des phrases ambiguës qui permettent la construction de différentes structures syntaxiques alternatives. L'analyse syntaxique des phrases est généralement vue comme composée de trois étapes : (1) la génération de la structure syntaxique ou des structures syntaxiques possibles au point d'ambiguïté ; (2) la sélection d'une structure ; et (3) la réanalyse de la phrase si la structure sélectionnée s'avère être incorrecte (voir par exemple Boland & Blodgett, 2001).

Différents modèles d'analyse syntaxique des phrases et différentes stratégies ont été proposées pour expliquer le traitement syntaxique des phrases et la résolution des ambiguïtés syntaxiques (voir Mitchell, 1994, pour une revue détaillée). Ces différents modèles peuvent être divisés en deux groupes : d'une part les modèles sériels ou modulaires, et d'autre part les modèles en parallèle ou non-modulaires.

### *1.1.1 Les modèles en série (modulaires)*

Les modèles sériels ou modulaires proposent que l'analyseur syntaxique ne poursuit initialement qu'une seule analyse syntaxique. Au début, il n'y a donc pas de différence entre la « génération syntaxique » et la « sélection syntaxique ». L'analyse initiale se déroule comme si la phrase traitée n'était pas du tout ambiguë (Frazier, 1978; Frazier, Clifton, & Randall, 1983; Frazier & Fodor, 1978; Frazier & Rayner, 1987). L'ambiguïté n'est donc pas prise en compte et une seule structure syntaxique est construite. Cependant, dès qu'un mot ne peut pas être intégré à cette structure syntaxique initialement générée, une réanalyse de la phrase est déclenchée.

Dans le système global de compréhension du langage, Frazier (1978) postule l'existence d'un mécanisme spécifique au traitement syntaxique : le processeur syntaxique. Celui-ci a pour fonction de construire la représentation syntaxique des phrases, et utilise uniquement pour cela des informations syntaxiques. Ce module syntaxique effectue son traitement de manière autonome et fournit ensuite le produit de son traitement à un second module : le processeur thématique (Rayner, Carlson, & Frazier, 1983). La construction initiale de la représentation syntaxique d'une phrase est élaborée au fur et à mesure de l'identification des mots de cette phrase, et ne repose que sur les catégories grammaticales des mots qui sont traités. Quand un choix doit être fait à un point d'ambiguïté, il est en général fondé sur des principes syntaxiques généraux, qui permettent de choisir la solution la plus économique avec la construction des structures syntaxiques les plus simples. En cas d'ambiguïté syntaxique, une seule interprétation est donc choisie. Si celle-ci s'avère ensuite incorrecte, le processeur déclenche alors une réanalyse de la phrase, en prenant en compte cette fois-ci d'autres sources d'informations comme les informations sémantiques ou contextuelles.

L'analyseur syntaxique s'appuie donc initialement sur les seules informations syntaxiques. Les deux grands principes qui ont été développés et le plus fréquemment étudiés sont d'une part l'attachement minimal<sup>1</sup> et d'autre part la clôture tardive<sup>2</sup> (Frazier, 1978). La stratégie de l'attachement minimal consiste à construire la représentation syntaxique la plus simple en termes d'indicateurs syntagmatiques, c'est-à-dire celle qui contient le moins de nœuds syntaxiques. Elle revient à construire la structure syntaxique la moins complexe possible. Quant à la stratégie de la clôture tardive, elle s'applique quand les différentes interprétations possibles contiennent le même nombre de nœuds syntaxiques (donc quand le principe d'attachement minimal ne peut être appliqué). Elle conduit alors à choisir l'attachement le plus bas dans l'indicateur syntagmatique, c'est-à-dire à attacher les items lexicaux au dernier constituant construit ou au constituant en cours de construction. Le principe est donc de laisser « ouvert » le constituant syntaxique en cours d'élaboration, et d'y incorporer le matériel lexical arrivant tant que cela est possible (donc tant que cela est grammaticalement correct).

Ces stratégies de traitement syntaxique ont été mises en évidence expérimentalement, en étudiant des phrases dites « garden-path », dont la plus célèbre est certainement la phrase (1) utilisée par Bever (1970) et reprise dans des dizaines d'articles :

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

*Le cheval qui était monté devant l'écurie est tombé.*

L'ambiguïté de cette phrase repose sur l'interprétation du verbe « raced » qui, au moment où il est traité et doit être intégré dans la structure syntaxique de la phrase, peut être considéré soit comme le verbe « to race » conjugué au passé, soit comme le participé passé du verbe « to race ». Dans le premier cas, on obtient une phrase active simple, avec l'interprétation du mot « raced » comme le verbe de la proposition principale. Dans le deuxième cas, on obtient une proposition relative réduite, avec

---

<sup>1</sup> Attachement minimal = Minimal Attachment : « Do not postulate any potentially unnecessary nodes » (Frazier, 1987, p. 562)

<sup>2</sup> Clôture tardive = Late Closure : « If grammatically permissible attach new items into the clause or phrase currently being processed » (Frazier, 1987, p. 562)

l'interprétation du mot « raced » comme le verbe débutant une proposition subordonnée relative (dont l'équivalent non ambigu serait « The horse which was raced past the barn fell »).

Il est important de noter que cette phrase est temporairement ambiguë<sup>3</sup> puisque, à la fin de la phrase, quand tous les mots ont été traités, une seule interprétation reste possible. Dans cet exemple, la présence du verbe « fell » à la fin de la phrase désambiguïse l'interprétation de la phrase, et ne laisse possible que l'interprétation de « raced » en tant que participe passé.

Selon la théorie « garden-path », l'interprétation préférentielle du segment ambigu doit être une interprétation en phrase active, le verbe étant considéré comme le verbe de la proposition principale (ce qui donne la structure syntaxique la plus simple). Or l'item désambiguïsant en fin de phrase ne permet que l'interprétation alternative : pour interpréter correctement cette phrase, il faut donc aller à l'encontre du principe d'attachement minimal.

Si l'on compare le traitement de cette phrase, contenant cette proposition relative réduite, avec le traitement d'une phrase équivalente mais contenant une proposition relative non réduite (comme « The horse which was raced past the barn fell » qui est une phrase non-ambiguë de par la présence du pronom relatif « which »), on se rend compte que les adultes ont du mal à traiter ces phrases dites « garden-path » (temps de lecture plus longs). Ceci implique que les lecteurs ont initialement choisi l'interprétation en phrase active simple, et que l'apparition du verbe « fell » en fin de phrase a été incompatible avec la structure syntaxique choisie par les adultes. Ils ont dû alors réanalyser la structure syntaxique de la phrase pour la rendre compatible avec ce dernier mot (d'où une augmentation du temps de traitement). Ces résultats ont été répliqués dans de nombreuses expériences (Britt, Perfetti, Garrod, & Rayner, 1992; Ferreira & Clifton Jr, 1986; Frazier, 1978; Frazier & Rayner, 1982; Rayner et al., 1983).

De la même manière, des évidences expérimentales ont été apportées pour le rôle de la clôture tardive dans le traitement syntaxique initial des phrases. Prenons par exemple le début de phrase suivant :

---

<sup>3</sup> En opposition, on trouve parfois des phrases complètement ambiguës pour lesquelles deux interprétations sont possibles tout au long de la phrase (les seules informations apportées par les mots de la phrase ne peuvent permettre de choisir entre les différentes alternatives).

(2a) *Since Jay always jogs a mile...*  
*Puisque Jay court toujours un kilomètre...*

L'ambiguïté de ce début de phrase repose au niveau de l'attachement du groupe nominal « a mile » : celui-ci peut en effet être considéré soit comme l'argument du verbe de la proposition principale « jogs » (comme dans la phrase 2b), soit comme le groupe nominal sujet d'une nouvelle proposition (comme dans la phrase 2c).

(2b) *Since Jay always jogs a mile this seems like a short distance to him.*  
*Puisque Jay court toujours un kilomètre cela lui paraît une courte distance.*  
(2c) *Since Jay always jogs a mile seems like a short distance to him.*  
*Puisque Jay court toujours un kilomètre lui paraît une courte distance.*

Le principe de clôture tardive fait que l'interprétation préférée de « a mile » devrait être celle de la phrase (2b), donc en tant que complément d'objet direct du verbe principal. Quant à la phrase (2c), elle viole ce principe et devrait donc être plus difficile à traiter. De nombreuses expériences ont montré en effet que le temps de lecture, pour ces phrases qui violaient la stratégie de clôture tardive, était plus long que le temps de lecture des phrases conformes à cette stratégie. En fait, les adultes passent plus de temps pour lire le matériel suivant le groupe nominal ambigu, indiquant que ce matériel n'est pas en adéquation avec la structure syntaxique qu'ils avaient initialement construite, et qu'ils doivent donc réanalyser la phrase pour la rendre compatible (Ferreira & Henderson, 1991; Frazier, 1978; Frazier & Rayner, 1982; Mitchell & Holmes, 1985).

Les premiers modèles qui ont donc été décrits pour expliquer les premières étapes de l'analyse syntaxique des phrases sont donc modulaires, avec l'existence d'un processeur syntaxique autonome, n'utilisant que des informations de type syntaxique pour prendre ses décisions. Les stratégies proposées pour guider le choix initial de la structure syntaxique sont des principes syntaxiques comme l'attachement minimal et la clôture tardive. Leur utilité a été mise de nombreuses fois en évidence, que ce soit dans des tâches de jugement de grammaticalité ou des tâches de lecture<sup>4</sup>. Le résultat principal

---

<sup>4</sup> Les tâches de lecture sont en général des « self-paced reading tasks » dans lesquelles les participants

est que les adultes rencontrent plus de difficultés pour traiter les phrases qui violent ces stratégies syntaxiques, par rapport aux phrases dont la structure est conforme aux stratégies postulées dans ces modèles en série. L'avantage de ces théories est qu'elles postulent (et permettent) une construction des structures syntaxiques relativement facile et rapide, puisque les décisions sont prises sur la base d'heuristiques qui n'utilisent qu'une seule source d'informations.

Mais nous allons maintenant décrire d'autres recherches, dont les résultats posent certains problèmes pour cette théorie, et qui ont démontré que les choix syntaxiques initiaux pouvaient être guidés par des informations autres que syntaxiques.

### *1.1.2 Les modèles en parallèle (non modulaires)*

Contrairement aux modèles sériels, les modèles en parallèle ou non modulaires (voir par exemple, Jurafsky, 1996; MacDonald, Pearlmutter, & Seidenberg, 1994; Sedivy & Spivey-Knowlton, 1994; Tabor, Juliano, & Tanenhaus, 1997; Trueswell, Tanenhaus, & Garnsey, 1994) postulent, qu'à chaque point d'ambiguïté dans le discours, les différentes alternatives syntaxiques possibles sont générées. Beaucoup de ces modèles proposent qu'il existe une activation inégale entre les différentes structures construites : en fait, l'analyseur syntaxique renforcerait l'activation d'une structure syntaxique parmi les autres structures construites en parallèle, en utilisant différents facteurs (ou différentes contraintes) auxquels seraient assignés des poids différents.

La principale caractéristique de ces modèles est que des informations de toute nature peuvent être immédiatement prises en compte pour résoudre l'ambiguïté syntaxique des phrases.

---

contrôlent la présentation des mots de la phrase en appuyant sur une touche (les mots sont présentés un à un au centre d'un écran), ou des tâches de lecture réalisées avec un « eye-tracker » qui enregistre les mouvements des yeux.

### 1.1.2.1 Rôle des informations lexicales

Selon les modèles du « guidage lexical » (« lexical guidance », Ford, Bresnan, & Kaplan, 1982; Pritchett, 1988; Tanenhaus, Carlson, & Trueswell, 1989), les représentations lexicales associées aux mots de la phrase peuvent guider le traitement syntaxique des phrases. Ces représentations lexicales contiennent toutes les informations sémantiques, syntaxiques et thématiques des mots.

Ainsi, il a été montré que la structure argumentative des verbes pouvait jouer un rôle dans le choix initial de la structure syntaxique d'une phrase, dans le sens où la structure d'arguments la plus fréquente d'un verbe guide l'analyse syntaxique initiale (Boland, 1997; Mitchell & Holmes, 1985; Shapiro, Nagel, & Levine, 1993; Trueswell & Kim, 1998; Trueswell, Tanenhaus, & Kello, 1993). Le type de constituants qui peuvent servir d'arguments au verbe, ainsi que les arguments que le verbe admet de manière préférentielle ou de manière obligatoire, sont des informations importantes pour l'analyseur syntaxique.

Par exemple, considérons les exemples suivants :

*(3a) The man accepted the prize.*

*L'homme a accepté la récompense.*

*(3b) The man accepted the prize was not going to him.*

*(3c) The man accepted that the prize was not going to him.*

*L'homme a accepté que la récompense ne lui revienne pas.*

Une ambiguïté peut résider dans l'attachement du syntagme nominal « the prize » qui peut être considéré soit comme l'objet du verbe principal (comme dans la phrase 3a), soit comme le sujet d'une proposition relative (comme dans la phrase 3b). Le principe de l'attachement minimal prédit que la structure préférentielle est celle de la phrase (3a). Les sujets devraient donc commencer à interpréter la phrase (3b) de manière incorrecte (en pensant que le syntagme nominal est l'objet direct du verbe), et

ne se rendre compte de cette erreur qu'en entendant la matériel désambiguïsant « was not going ».

Trueswell et al. (1993) ont fait lire ces phrases à des adultes en enregistrant leurs mouvements oculaires et leurs temps de fixation. Le premier résultat trouvé était compatible avec la théorie « garden-path » puisque les participants étaient effectivement ralentis pour lire la phrase (3b) comparé à la phrase (3c), qui elle n'est pas ambiguë de par la présence du pronom « that » introduisant la proposition relative (au moment de traiter la région désambiguïsante de la phrase 3b, les auteurs ont enregistré des temps de fixation plus longs et des mouvements oculaires de retour vers le début de la phrase).

Les auteurs ont testé si la structure argumentative des verbes pouvait jouer un rôle et supprimer d'une certaine manière l'interprétation minimale préférentielle. Pour ce faire, ils ont varié le verbe utilisé dans les phrases expérimentales, en prenant soit des verbes qui avaient une plus grande tendance à être utilisés avec des objets directs - comme le verbe « to accept » dans les exemples ci-dessus, soit des verbes qui avaient plutôt tendance à être utilisés avec des propositions relatives – comme le verbe « to realize » dans les exemples 4a et 4b ci-dessous (ces biais de fréquence d'utilisation ont été obtenus grâce à des analyses de corpus et des expériences de complétion de phrases).

*(4a) The man realized the prize was not going to him.*

*(4b) The man realized that the prize was not going to him.*

*L'homme a réalisé que la récompense n'allait pas lui revenir.*

Dans la théorie « garden-path », les informations autres que syntaxiques (comme les informations sémantiques ou thématiques) ne sont jamais prises en compte dans les étapes précoces du traitement syntaxique. Le traitement des phrases 4a et 4b devrait donc être similaire à celui des phrases 3b et 3c (c'est-à-dire une première interprétation incorrecte de la phrase 4a, nécessitant une réanalyse de la phrase et donc un temps de traitement plus long). Or, Trueswell et al. (1993) n'ont mis en évidence aucune difficulté pour le traitement des phrases 4a (les lecteurs se comportaient de la même manière que ce soit pour lire la phrase localement ambiguë 4a ou pour lire la phrase non-ambiguë 4b). Pour ces auteurs, la reconnaissance du verbe a entraîné l'activation parallèle des différentes structures argumentatives possibles (structure « objet direct » et structure « proposition relative »). L'activation de ces différentes structures a été ensuite

pondérée par l'information de fréquence (structure « objet direct » plus fréquente que structure « proposition relative » pour le verbe « to accept » ; structure « proposition relative » plus fréquente que structure « objet direct » pour le verbe « to realize »). C'est alors cette information qui a été utilisée pour activer plus fortement l'une des deux alternatives possibles et guider le choix initial de l'analyseur syntaxique.

D'autres recherches ont aussi montré que la structure thématique associée aux verbes (c'est-à-dire le rôle des arguments, comme par exemple la fonction d'agent, de patient, de but, d'instrument...) jouait un rôle dès les premières étapes de l'analyse syntaxique (Just & Carpenter, 1992; McRae, Spivey-Knowlton, & Tanenhaus, 1998; Tanenhaus et al., 1989; Trueswell et al., 1994).

Par exemple, Trueswell et al. (1994) ont testé des phrases pouvant être localement ambiguës, en faisant varier le sujet de ce verbe qui pouvait être soit un sujet animé (comme dans la phrase 5a), soit un sujet inanimé (comme dans la phrase 5b).

*(5a) The defendant examined by the lawyer was unreliable.*

*Le prévenu examiné par l'avocat n'était pas sérieux.*

*(5b) The evidence examined by the lawyer was unreliable.*

*La preuve examinée par l'avocat n'était pas fiable.*

Dans la phrase 5a, le sujet « the defendant » peut avoir deux rôles thématiques différents : soit il est agent du verbe (dans ce cas, le segment ambigu, souligné dans les exemples, est considéré comme le groupe verbal de la proposition principale), soit il est le patient du verbe (dans ce cas, le segment ambigu est considéré comme une proposition relative). Selon le principe d'attachement minimal, les sujets devraient interpréter « the defendant » comme le sujet agent, et « examined » comme le verbe de la proposition principale. Les premiers résultats ont en effet montré que les sujets étaient ralentis pour traiter ces phrases, et que la présence du deuxième verbe « was » entraînait une réanalyse de la phrase (qui avait été initialement interprétée en phrase avec attachement minimal alors qu'elle ne l'était pas). Les temps de lecture étaient en effet plus longs pour la phrase 5a que pour son équivalent non ambigu (« The defendant that was examined by the lawyer was unreliable » où la présence du pronom relatif « that » désambiguïse l'interprétation de la phrase).

Les auteurs ont comparé ces résultats avec le traitement des phrases impliquant un sujet inanimé (phrase 5b). En effet, dans cette situation, le groupe nominal « the evidence » ne peut être que le patient du verbe (une preuve ne peut pas examiner quelque chose mais par contre peut être examinée par quelqu'un...). Si les sujets peuvent utiliser ces connaissances dès les étapes initiales de l'analyse syntaxique de la phrase, ils ne devraient donc pas se faire piéger et devraient analyser cette phrase avec une interprétation non-minimale. Si par contre leur choix initial n'est guidé que par des stratégies syntaxiques, qui ne prennent pas en compte les informations lexicales et sémantiques, ils devraient avoir un comportement similaire pour traiter la phrase 5b et la phrase 5a (c'est-à-dire choisir initialement une interprétation erronée reposant sur l'attachement minimal). Or les auteurs n'ont mis en évidence aucune difficulté de traitement pour la phrase 5b : les participants étaient aussi rapides pour lire cette phrase potentiellement ambiguë, comparée à la lecture de son équivalent non-ambigu (« The evidence that was examined by the lawyer was unreliable). Trueswell et al. (1994) ont donc pu conclure que les informations thématiques liées aux verbes étaient utilisées dès les premières étapes de l'analyse syntaxique des phrases.

La première structure syntaxique assignée à une phrase ne reposerait donc pas (ou pas uniquement) sur des stratégies générales syntaxiques, mais serait (aussi) guidée par l'information lexicale. Il faut cependant noter que certaines expériences ont échoué à mettre en évidence le rôle précoce de l'information lexicale (Ferreira & Clifton Jr, 1986; Mitchell, 1989). Le vrai débat se situe en fait sur le moment où ces informations lexicales sont utilisées : pour les défenseurs des modèles sériels et des théories « garden-path », le fait qu'elles interviennent relativement tôt dans le traitement n'implique pas nécessairement le fait qu'elles soient utilisées lors de la phase initiale de structuration syntaxique de la phrase.

#### 1.1.2.2 Rôle du contexte

La théorie référentielle (« referential theory ») a été initialement proposée par Crain et Steedman (1985) comme un modèle de traitement du langage basé sur le discours et le contexte, en contraste avec les modèles de traitement du langage basée sur

la structure des phrases, comme les théories « garden-path ». De nombreuses études ont été réalisées par la suite et ont démontré la validité de cette hypothèse (Altmann, 1988; Altmann, Garnham, & Dennis, 1992; Altmann, Garnham, & Henstra, 1994; Altmann & Steedman, 1988; Britt, 1994; Crain & Steedman, 1985; Spivey-Knowlton & Tanenhaus, 1994; Tanenhaus, Spivey-Knowlton, Eberhard, & Sedivy, 1995).

Considérons par exemple la phrase suivante :

(6) *He told the journalist that he had recently written to to hurry up and submit his piece.*  
*Il a dit au journaliste à qui il avait récemment écrit de se dépêcher et de soumettre sa pièce.*

Généralement, et conformément aux principes des théories « garden-path », la partie ambiguë « that he had recently written to » est interprétée comme le complément du verbe « told » : l'apparition du deuxième « to » est alors considérée comme agrammaticale par des lecteurs adultes. En fait, cette séquence ambiguë devrait être interprétée comme une proposition relative se référant au mot « journalist ».

Crain et Steedman (1985) ont montré que le contexte d'apparition de cette phrase pouvait jouer un rôle sur son interprétation. En effet, si le début de la phrase « he told the journalist » est présenté dans un contexte qui introduit l'existence de plusieurs journalistes, l'utilisation de la proposition relative, spécifiant de quel journaliste le locuteur est en train de parler, semble alors très plausible, voire nécessaire pour une meilleure compréhension de la phrase de la part de l'auditeur. Les auteurs ont ainsi montré que les difficultés de traitement des phrases ambiguës de ce type pouvaient être surmontées si l'on utilisait des contextes référentiels appropriés.

D'autres expériences se sont intéressées au rôle du contexte visuel. Par exemple, Tanenhaus et al. (1995) ont enregistré les mouvements oculaires d'adultes, assis face à une table représentant différents objets, et qui entendaient différentes phrases leur indiquant une action à réaliser. Les phrases utilisées pouvaient être des phrases ambiguës, comme l'exemple suivant :

(7) *Put the apple on the towel in the box.*  
*Place la pomme qui est sur la serviette dans la boîte.*

Le syntagme nominal ambigu « on the towel » est en général interprété comme modifiant le verbe « put » et indiquant la direction de l'action (principe d'attachement minimal). L'apparition du syntagme « in the box » en fin d'énoncé indique que l'interprétation correcte de ce syntagme ambigu est en fait une modification du nom « apple ». Les auteurs ont montré que cette erreur initiale d'interprétation pouvait être supprimée si l'on utilisait un contexte référentiel visuel approprié. Ainsi, si la table face aux participants comporte deux pommes, dont une seule est posée sur une serviette, alors les auditeurs interprètent immédiatement le syntagme ambigu « on the towel » comme un complément du nom « apple » : dès qu'ils entendent « on the... », leur regard se dirige vers « the apple on the towel » (et non vers l'autre pomme possible ni vers la serviette vide).

Ainsi, les modèles postulant des systèmes de traitement encapsulés et autonomes, avec en particulier un analyseur syntaxique n'utilisant initialement que des informations syntaxiques pour assigner une structure syntaxique aux phrases, semblent une fois de plus mis en défi. Il semble en effet que d'autres informations comme les informations contextuelles puissent être rapidement intégrées et utilisées pour guider le choix initial de la structure des phrases entendues.

### 1.1.2.3 Rôle des informations statistiques

Différents auteurs ont postulé que les décisions initiales de l'analyseur syntaxique pouvaient être fondées sur l'expérience passée, c'est-à-dire sur la prévalence statistique des différentes structures dans le langage : ainsi, les structures qui sont statistiquement les plus fréquentes sont activées de manière préférentielle par rapport aux moins fréquentes (Jurafsky, 1996; MacWhinney, 1987; Mitchell & Cuetos, 1991).

C'est cette idée qui a surtout été développée pour expliquer certaines différences apparaissant entre les langues et concernant les préférences d'attachement de syntagme ambigu. Prenons par exemple, la phrase 8a ambiguë en anglais, et les phrases espagnoles 8b et 8c :

(8a) *Someone shot the servant of the actress who was on the balcony.*

(8b) *Alguien disparo contra la criada de la actriz que estaba en el balcon.*

*Quelqu'un a tiré sur la servante de l'actrice qui était au balcon.*

(8c) *Alguien disparo contra el criado de la actriz que estaba en el balcon con su marido.*

*Quelqu'un a tiré sur le serviteur de l'actrice qui était au balcon avec son mari.*

Il a été montré, dans de nombreuses études sur les préférences d'attachement des propositions relatives ambiguës, que les adultes anglais (ou américains) se comportaient conformément au principe de clôture tardive, c'est-à-dire en attachant la proposition relative au deuxième nom « actress » (par exemple, dans certaines expériences, on demandait à des sujets de répondre à la question : « Qui était au balcon ? » et la majorité des réponses était alors « l'actrice »).

Cuetos et Mitchell (1988) ont réalisé différentes études sur l'attachement ambigu de propositions relatives en espagnol. Ils ont pour cela utilisé des phrases comme les phrases 8b et 8c : si dans ces phrases, les sujets espagnols suivent une stratégie de clôture tardive, ils devraient donc attacher la proposition relative au nom « actriz » et ne rencontrer aucune difficulté de traitement. Or, quand on demandait aux sujets de dire « qui était sur le balcon » en réponse à la phrase 8b, ceux-ci répondaient « la criada ». D'autre part, les espagnols ont été ralentis pour traiter des phrases comme la phrase 8c, gênés par la présence de « con su marido » (avec son mari) en fin de phrase. Ceci indique qu'ils ont en fait attaché la proposition relative au premier nom « el criado » (le serviteur), et non au deuxième nom « la actriz » (l'actrice), contrairement à ce qui avait toujours été montré en anglais.

Pour montrer que la prévalence d'une structure sur une autre dans le langage jouait un rôle dans ces différences de stratégies, des analyses de corpus ont été réalisées, centrées sur l'attachement ambigu des propositions relatives. Il a été trouvé qu'en anglais, la majorité des propositions relatives sont attachées au second groupe nominal, alors qu'en espagnol c'est le biais opposé qui prédomine, avec un attachement préférentiel au premier groupe nominal (Mitchell, Cuetos, & Corley, 1992).

Ces différentes recherches n'établissent pas que les choix initiaux de structuration syntaxique sont complètement déterminés par des mécanismes statistiques.

Néanmoins, elles apportent des évidences quant à la relation qui existe entre les biais de traitement syntaxique des phrases ambiguës, et la prévalence de ces formes dans le langage auquel les adultes sont exposés.

Récemment, Fodor (1998; 2002) a proposé une nouvelle hypothèse pour expliquer ces différences inter-langues dans les préférences d'interprétations des sujets : l'hypothèse de la prosodie implicite (« implicit prosody hypothesis »)<sup>5</sup> qui repose donc sur les variations de rythme et d'intonation données aux phrases. Selon Fodor, le lecteur effectue un découpage prosodique des phrases qui va influencer leur analyse syntaxique, et ce même pendant une lecture silencieuse. En cas d'ambiguïté syntaxique, le lecteur choisirait donc l'analyse syntaxique la plus proche du découpage prosodique naturel de sa langue.

Les différences d'attachement, observées entre les langues pour des structures syntaxiques identiques, dépendraient donc des propriétés prosodiques propres à chaque langue. Plus précisément, la préférence d'attachement varierait en fonction de facteurs phonologiques, comme la longueur des mots et leur accentuation, qui diffèrent selon les langues.

Pour conclure cette première partie, nous avons rapporté de nombreux résultats qui suggèrent que l'analyseur syntaxique s'engage dans le choix initial d'une structure syntaxique, soit de manière unique et complète (comme dans les modèles sériels), soit de manière partielle (comme dans les modèles parallèles). Quand l'analyseur est face à une ambiguïté, il choisit alors une structure syntaxique, soit en sélectionnant une seule, soit en activant préférentiellement une des alternatives générées. Si la structure initialement construite est consistante avec l'information désambiguïsante arrivant par la suite, le traitement continue normalement. Mais au contraire, si la structure syntaxique choisie s'avère incorrecte, donc incompatible avec le matériel désambiguïsant, l'analyseur syntaxique est alors obligé de lancer une réanalyse de la phrase. Pour ce faire, de nombreuses sources d'informations peuvent être utilisées : informations syntaxiques, thématiques, contextuelles...

---

<sup>5</sup> « In silent reading, a default prosodic contour is projected onto the stimulus, and it may influence syntactic ambiguity resolution. Other things being equal, the parser favors the syntactic analysis associated with the most natural (default) prosodic contour for the construction. »

La majorité des expériences qui ont été réalisées et que nous avons abordées dans cette partie ont été menées sur du matériel écrit. Les participants devaient juger de la grammaticalité des phrases, lire des phrases alors que leurs mouvements oculaires étaient enregistrés... Un autre indice a donc longtemps été négligé : la prosodie des phrases, c'est-à-dire les variations de rythme et d'intonation de la parole. Nous allons donc maintenant aborder la notion de prosodie, ainsi que les expériences qui ont étudié le rôle des indices prosodiques dans l'analyse syntaxique des phrases, présentées cette fois-ci de manière auditive.

## 1.2 LE ROLE DE LA PROSODIE DANS LE TRAITEMENT SYNTAXIQUE

### 1.2.1 *Qu'est-ce que la prosodie ?*

#### 1.2.1.1 Définition

Selon Di Cristo (2000a),

*« La prosodie (ou la prosodologie) est une branche de la linguistique consacrée à la description (aspect phonétique) et à la représentation formelle (aspect phonologique) des éléments de l'expression orale tels que les accents, les tons, l'intonation, et la quantité, dont la manifestation concrète, dans la production de la parole, est associée aux variations de la fréquence fondamentale (F0), de la durée et de l'intensité (paramètres prosodiques physiques), ces variations étant perçues par l'auditeur comme des changements de hauteur (ou de mélodie), de longueur et de sonie (paramètres prosodiques subjectifs). Les signaux prosodiques véhiculés par ces paramètres sont polysémiques et transmettent à la fois des informations para-linguistiques et des informations linguistiques déterminantes pour la compréhension des énoncés et leur interprétation pragmatique dans le flux du discours. »*

Plusieurs fonctions de la prosodie ont été décrites (Di Cristo, 1998; Lacheret-Dujour & Beaugendre, 1999; Lhote, Barry, & Tivane, 2000), parmi lesquelles on peut

distinguer : (1) une fonction distinctive qui permet par exemple, en l'absence de marques syntaxiques, de distinguer une phrase déclarative, d'une phrase interrogative ou d'une phrase impérative, (2) une fonction démarcative, aussi appelée fonction de désambiguïsation, qui permet de segmenter les phrases et de retrouver l'organisation sémantique et syntaxique d'un énoncé, et donc de lever certaines ambiguïtés, (3) une fonction focalisatrice, ou fonction de mise en relief, qui permet au locuteur de mettre l'emphase sur un mot ou une suite de mots, et (4) une fonction expressive qui traduit les émotions, les intentions, les attitudes du locuteur.

Dans la parole orale, les mots et les syllabes ne sont donc pas prononcés de manière monotone et monocorde. Au contraire, un discours oral est caractérisé par une mélodie (une prosodie) qui s'exprime par des variations de rythme et d'intonation, par la présence de pauses plus ou moins longues dans les phrases, et dont les principaux corrélats acoustiques sont les variations de durée et les variations de fréquence fondamentale.

Les mots sont ainsi regroupés en unités prosodiques, qui peuvent être intégrées dans un modèle hiérarchique de la structure prosodique (voir Shattuck-Hufnagel & Turk, 1996, pour une revue détaillée des principales unités et des différents modèles).

#### 1.2.1.2 Liens entre prosodie et syntaxe

La première idée qui ait été proposée est que les constituants prosodiques des phrases correspondent à ceux prédits par la syntaxe. Il y aurait donc une relation isomorphe entre la structure syntaxique d'une phrase et sa prosodie. Les partisans de cette théorie proposaient que les frontières finales des constituants syntaxiques étaient le siège d'un grand nombre de phénomènes prosodiques, comme l'allongement des unités finales (segment, syllabe, mot) et la présence d'une pause (Cooper & Paccia-Cooper, 1980; Lehiste, 1973). En d'autres mots, pour eux, et en l'absence d'une véritable théorie prosodique, on pensait que la position des frontières syntaxiques déterminerait la structure prosodique des phrases.

Pour défendre cette idée, de nombreux chercheurs étudiaient les caractéristiques acoustiques du signal de parole, et recherchaient si des discontinuités prosodiques

majeures avaient effectivement tendance à apparaître aux emplacements des frontières syntaxiques majeures (Cooper & Paccia-Cooper, 1980; Goldman-Eisler, 1972). Dans certains cas, cette correspondance existait. Mais il est impossible de généraliser ce résultat à toutes les structures syntaxiques et prosodiques.

En fait, comme le soulignent par exemple Nespor et Vogel, dans leur théorie de la phonologie prosodique (Nespor & Vogel, 1986), la structure syntaxique et la structure prosodique des phrases sont étroitement liées, mais cette relation n'est pas l'identité. Par exemple, toutes les frontières syntaxiques ne sont pas obligatoirement représentées au niveau prosodique. C'est le cas assez fréquent, et très souvent cité, de la frontière syntaxique entre un pronom personnel sujet et le verbe auquel il se rapporte (voir par exemple Vaissière, 1997). En effet, considérons par exemple la phrase suivante, « il revient de la ville » ; d'un point de vue prosodique, le pronom est relié au verbe (cet item non accentué ne forme pas à lui tout seul une unité prosodique mais se rattache au verbe auquel il se rapporte). Si l'on compare la structure prosodique de cette phrase avec sa structure syntaxique, on se rend compte qu'elles ne sont pas identiques, puisque la frontière prosodique ne se situe pas entre le groupe nominal sujet (formé par le pronom) et le groupe verbal (comme dans « [Il # revient] [de la ville] », où les frontières prosodiques sont indiquées par des crochets, et où la frontière syntaxique entre le sujet et le verbe est indiquée par le symbole #. On peut noter dans ce cas que la frontière prosodique ne se situe pas au niveau de la frontière syntaxique majeure séparant le sujet du verbe ; elle est placée après le verbe, au début du groupe nominal, donnant en fait une information sur le début d'un autre constituant syntaxique de la phrase).

Un autre exemple, cité très régulièrement depuis des décennies, avait été introduit par Chomsky et Halle (1968), qui avaient noté que dans certaines phrases, la structure prosodique (et en particulier la distribution des pauses) ne correspondait pas toujours à la structure syntaxique de la phrase. Considérons l'exemple 9 : il est généralement accepté que le découpage syntaxique de cette phrase est celui indiqué en 9a (dans lequel nous avons indiqué les frontières de certains constituants syntaxiques), alors que d'un point de vue prosodique elle sera découpée de la manière indiquée en 9b (où nous avons indiqué les frontières des groupes prosodiques).

(9a) *This is [the dog that chased [the cat that caught [the rat that stole [the cheese]]]]*

(9b) *[This is the dog] [that chased the cat] [that bit the rat] [that stole the cheese]*

Cet exemple montre de manière très convaincante que les frontières syntaxiques ne sont pas toujours représentées de manière directe au niveau prosodique. Pourtant, Chomsky et Halle (1968) minimisaient l'importance de cette différence entre structure prosodique et structure syntaxique, attribuant le problème décrit ici à une question de performance.

Dans la théorie de la phonologie prosodique, que nous allons maintenant aborder de manière plus précise, de telles différences entre structure prosodique et structure syntaxique ne constituent pas un problème, mais au contraire une des motivations essentielles pour postuler une organisation phonologique interne ; car ce sont les constituants prosodiques qui déterminent le pattern prosodique d'une phrase et non ces constituants syntaxiques.

### Théorie de la phonologie prosodique :

Notre définition de la prosodie et des différentes unités prosodiques s'appuie donc sur la théorie de la phonologie prosodique (Nespor & Vogel, 1986; Selkirk, 1984; Selkirk, 1978; Shattuck-Hufnagel & Turk, 1996), qui se veut une théorie de l'interface entre la morphosyntaxe et la phonologie. Elle comporte une hiérarchisation des constituants prosodiques, qui sont dérivés des constituants morphosyntaxiques et qui forment les domaines d'application des règles phonologiques.

Plusieurs constituants peuvent être distingués avec, du plus grand au plus petit : l'énoncé, le groupe intonational, le groupe phonologique, le groupe clitique, le mot prosodique, le pied métrique et la syllabe<sup>6</sup>. Nous centrerons notre attention sur les groupes intonationnels et les groupes phonologiques, qui sont les unités les plus pertinentes pour notre propos.

---

<sup>6</sup> Ces unités ne se retrouvent pas toutes dans les différentes théories, ou apparaissent parfois sous d'autres noms : par exemple, le mot prosodique de Selkirk (1978) correspond à la fois au groupe clitique et au mot prosodique de Nespor et Vogel (1986). De même, le groupe phonologique de Nespor et Vogel (1986) se voit qualifier de syntagme majeur ou de syntagme mineur par Selkirk (1984), et de syntagme intermédiaire ou de syntagme accentuel par Beckman et Pierrehumbert (1986).

### *Les groupes intonationnels :*

Il est généralement convenu que les groupes intonationnels sont des constituants prosodiques qui sont cohérents d'un point de vue sémantique et qui possèdent un contour intonatif perceptuellement cohérent (Nicol, 1996; Shattuck-Hufnagel & Turk, 1996). Selon Selkirk, ce sont des unités de sens (Selkirk, 1984).

Les groupes intonationnels correspondent typiquement aux propositions entières, et sont généralement délimités par un allongement final et une pause (Beckman & Pierrehumbert, 1986; Bernstein Ratner, 1986; Morgan, 1986; Selkirk, 1984; Shattuck-Hufnagel & Turk, 1996; Wightman, Shattuck-Hufnagel, Ostendorf, & Price, 1992).

De plus, certaines constructions syntaxiques particulières forment généralement par elles-mêmes un groupe intonationnel (Nespor & Vogel, 1986; Selkirk, 1984), comme les énoncés parenthétiques (voir exemple 10), les propositions relatives non restrictives ou explicatives (exemple 11), et les énoncés sous forme de listes (exemple 12). Dans les exemples suivants, une frontière de groupe intonationnel sera réalisée au niveau de chaque virgule.

*(10) Les lions, comme tu le sais, sont des animaux dangereux.*

*(11) Les enfants, qui n'écoutaient pas, n'ont rien compris.*

*(12) Ils ont acheté du lait, des oeufs, du pain, et du fromage.*

### *Les groupes phonologiques :*

Les groupes intonationnels peuvent être divisés en unités plus petites : les groupes phonologiques. Ces unités sont considérées comme étant liées à la structure syntaxique des phrases (Dresher, 1996; Nespor & Vogel, 1986; Selkirk, 1982, 1984; Shattuck-Hufnagel & Turk, 1996). Elles contiennent typiquement un ou deux mots de contenu (noms, verbes, adjectifs...), regroupés avec les mots grammaticaux (articles, pronoms, auxiliaires...) qui leur sont associés.

D'un point de vue acoustique, les groupes phonologiques sont caractérisés par un allongement final et un contour intonatif unique (Wightman et al., 1992). Les segments en début d'unité sont aussi plus marqués : par exemple, le premier phonème

d'un groupe phonologique est en général plus articulé et plus long (Fougeron & Keating, 1997; Keating, Cho, Fougeron, & Hsu, 2003).

*Comment sont construits les groupes phonologiques ?*

Selkirk (1984; 1986; 1996) propose l'idée que la relation entre la structure syntaxique et la structure prosodique des phrases soit liée à des contraintes portant sur l'alignement de ces deux structures : pour chaque constituant de catégorie  $x$  dans la structure syntaxique, sa frontière gauche (ou droite, selon que la langue est réursive à droite ou réursive à gauche<sup>7</sup>) coïncide avec la frontière d'un constituant de catégorie  $y$  dans la structure prosodique.

*« Edge-based theory of the syntax-prosody interface » (Selkirk, 1986; 1996)*  
*Right / Left edge of  $\alpha \rightarrow$  edge of  $\beta$  ( $\alpha$  is a syntactic category,  $\beta$  a prosodic one)*

Nespor et Vogel (1986) suggèrent également que la formation des groupes phonologiques suit la direction de l'arbre syntaxique de la langue étudiée (récursivité à gauche ou à droite). Elles proposent l'idée que la formation des groupes phonologiques se déroule en groupant une tête lexicale (nom, verbe, adjectif) avec les mots qui lui sont rattachés sur sa gauche (pour les langues réursives à droite), jusqu'à rencontrer une frontière de constituant syntaxique, donc jusqu'à rencontrer une autre tête lexicale.

*« Relation-based mapping » (Nespor & Vogel, 1986)*  
*« A phonological phrase includes a lexical head  $X$  and all the material on its non recursive side up to the next head outside of the maximal projection of  $X$  »*

Considérons par exemple la phrase 13, « le jeune homme est arrivé avant les autres », dont la structure prosodique est représentée en 13a, et la structure prosodique en 13b.

---

<sup>7</sup> Les langues du monde se séparent en deux grandes familles : elles sont dites réursives à droite quand on peut ajouter des éléments à la fin (comme en français ou en anglais), et réursives à gauche dans le cas contraire (comme en turc ou en japonais). Cette distinction possède un corrélat acoustique : à l'intérieur des unités prosodiques, le mot le plus accentué se trouve à droite pour les langues réursives à droite, et à

(13a) *Structure syntaxique* :  $s[GN[Le\ jeune\ homme]_{GV}[est\ arrivé]_{GP}[avant\ les\ autres]]]$

(13b) *Structure prosodique* :  $[Le\ jeune\ homme] [est\ arrivé] [avant\ les\ autres]$

La structure prosodique (découpage en groupes phonologiques) est obtenue de la manière suivante : le français étant une langue réursive à droite, nous partons de la dernière tête lexicale à droite : « autres » ; un groupe phonologique est construit en intégrant les éléments situés à gauche de la tête lexicale (du côté non réursif), jusqu'à atteindre la projection maximale suivante (on a donc regroupé ensemble tous les éléments du groupe prépositionnel « avant les autres », jusqu'à atteindre la nouvelle tête lexicale, c'est-à-dire le verbe). En suivant le même principe, le verbe conjugué (attaché avec son auxiliaire) forme un groupe phonologique, de même que le groupe nominal sujet.

De même, si l'on considère la phrase turque « Mehmet cumartesinden sonra gelecek » (c'est-à-dire « Mehmet dimanche après venir-futur », donc « Mehmet viendra après Dimanche »), on obtient les structures syntaxique (14a) et prosodique (14b) suivantes :

(14a) *Structure syntaxique* :  $s[GN[Mehmet]_{GV}[GP[cumartesinden\ sonra] gelecek]]]$

(14b) *Structure prosodique* :  $[Mehmet] [cumartesinden\ sonra] [gelecek]$

Cette fois-ci, la langue utilisée est réursive à gauche : on part donc de la tête lexicale la plus à gauche « Mehmet » et l'on intègre les différents éléments du côté non réursif (à droite de la tête lexicale) jusqu'à atteindre la projection maximale suivante. Dans cet exemple, la première tête lexicale est directement suivie d'une autre tête lexicale, située en dehors du premier groupe nominal (projection maximale de « Mehmet ») : le nom « cumartesinden » (donc le premier groupe phonologique n'est constitué que du nom propre). En suivant le même principe, on obtient que le groupe prépositionnel « cumartesinden sonra » forme un groupe phonologique, de même que le verbe « gelecek ».

---

gauche pour les langues réursives à gauche. Ce mot accentué est la tête lexicale.

La raison pour laquelle les phonologues ont été amenés à définir les différentes unités de la hiérarchie prosodique, est l'application des règles phonologiques : en effet, les unités prosodiques sont les domaines d'application des règles phonologiques. En travaillant sur les règles phonologiques qui s'appliquent à travers plusieurs mots (comme la liaison en français), ils se sont rendus compte que ces règles ne s'appliquaient pas partout (elles s'appliquent entre certaines paires de mots et pas d'autres). En tentant de définir le domaine d'application de ces règles (en se basant sur les unités syntaxiques), ils se sont aperçus qu'il était impossible de formuler des domaines d'application viables. De ce fait, ils ont été amenés à construire les unités de la hiérarchie prosodique, qui servent de domaine d'application aux diverses règles phonologiques.

Concernant les groupes phonologiques, ils sont considérés comme étant, par exemple, le domaine d'application du déplacement de l'accentuation en italien (« Stress Retraction ») et en anglais (« Iambic Reversal »). Cette règle s'applique entre deux mots qui appartiennent au même groupe phonologique : si le premier mot se termine par une syllabe accentuée et que le deuxième mot commence par une syllabe accentuée, alors l'accent du premier mot est déplacé sur la gauche pour éviter une « collision » entre les deux syllabes accentuées (« stress clash »). Grâce à cet indice, on peut postuler où se trouvent les frontières de groupes phonologiques dans les deux phrases suivantes (dans lesquelles les syllabes accentuées sont indiquées en majuscules) :

*(15a) Le ciTTA MOLto nordiche non mi piacciono.*

*(15b) Le Cltta NORdiche non mi pacciono.*

Dans la première phrase, deux syllabes consécutives sont accentuées : elles n'appartiennent donc pas au même groupe phonologique. Dans le cas contraire, la règle phonologique aurait été appliquée (comme dans la phrase 15b). On obtient donc les structures prosodiques suivantes :

*(15a) [Le ciTTA] [MOLto nordiche] [non mi piacciono]*

*(15b) [Le Cltta NORdiche] [non mi pacciono]*

En fait, si l'on applique le principe de construction des groupes phonologiques que nous avons décrit précédemment (« relation-based mapping »), on s'aperçoit qu'il devrait y avoir une frontière entre « citta » et le mot suivant dans les deux phrases : les têtes lexicales prises en compte dans ces deux phrases seraient le verbe « piaccio », l'adjectif « nordiche » et le nom « citta », ce qui devraient nous donner les structures prosodiques suivantes :

(15a) [Le ciTTA] [MOLto nordiche] [non mi piaccio]

(15b) [Le ciTTA] [NORdiche] [non mi piaccio]

Dans la phrase 15b, la frontière de groupe phonologique entre le nom et l'adjectif disparaît : ceci est dû à une règle de restructuration des groupes phonologiques, proposée par Nespor et Vogel (1986) :

*Phonological phrase restructuring (optional rule) :*

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

En italien, la langue est réursive à droite : dans la phrase 15b, la règle de restructuration peut être appliquée, car l'adjectif « nordiche » est le premier complément du nom « citta », situé à sa droite, et qu'il ne « branche » pas<sup>8</sup>. L'adjectif « nordiche » peut donc être rattaché au nom « citta » : ils ne forment alors qu'un seul groupe phonologique à l'intérieur duquel la règle de déplacement de l'accent doit s'appliquer (d'où le déplacement de l'accent sur la première syllabe du nom « citta »). Par contre, la règle de restructuration ne peut pas s'appliquer entre le groupe nominal « le citta » et le complément du nom de structure plus complexe « molto nordiche », qui forment alors deux groupes phonologiques. La structure accentuelle n'est donc pas modifiée puisque la « collision » accentuelle ne traverse pas les frontières de groupes phonologiques.

La règle de déplacement de l'accent est aussi appliquée en anglais : ainsi, l'accent est reporté sur la première syllabe de « thirteen » dans la phrase « [THIRteen

---

<sup>8</sup> En termes plus intuitifs, on pourrait dire que cette règle s'applique quand une tête lexicale (nom, verbe) est suivie d'un complément unique de structure simple (un seul mot).

MEN] [WENT to SEA] », mais pas dans la phrase « [At the AGE of thirTEEN] [JULia] [WROTE a BOOK] ».

Finalement, en français, les groupes phonologiques sont le domaine d'application de la liaison. Par exemple, on trouve la consonne de liaison /z/, entre « trois » et « enfants », dans « [Cette famille] [a adopté] [trois enfants] », mais cette même consonne de liaison n'est pas prononcée entre « livres » et « assez » dans « [Jean] [possède] [des livres] [assez anciens] ». Notons que la liaison est obligatoire entre un article et un nom (comme dans le premier exemple) et interdite à la frontière des groupes phonologiques (comme dans le deuxième exemple) ; par contre, elle reste un phénomène facultatif à l'intérieur des groupes phonologiques (sa présence relève alors plutôt d'un style de langage soutenu). Par exemple, dans la phrase « Jean possède des livres anciens », la liaison entre « livres » et « anciens » est possible mais pas obligatoire.

La théorie de la phonologie prosodique propose donc que la structure prosodique d'une phrase est déterminée par la structure syntaxique, mais n'est pas isomorphe à cette dernière. Les groupes phonologiques sont construits à partir des unités syntaxiques (tête lexicale, récursivité), mais la structure prosodique résultante ne correspond pas directement à la structure syntaxique. Les groupes phonologiques définissent les domaines d'application de certaines règles phonologiques, ce qui prouve la validité de ces unités prosodiques en ce qui concerne la production de la parole (en plus, l'application de ces règles peut être utilisée pour trouver où positionner les frontières de groupes phonologiques dans le signal).

#### Autres théories sur la construction des unités prosodiques : le rôle des informations phonétiques :

D'autres informations peuvent être utilisées pour trouver où positionner les frontières prosodiques dans le signal de parole. Par exemple, Gee et Grosjean (1983) ont utilisé l'information apportée par les pauses (considérées ici comme la somme de l'allongement syllabique final et du silence) pour dériver la structure prosodique des phrases. Ils proposent un algorithme qui permet d'obtenir entre chaque mot d'une

phrase, même quand ils ne sont pas séparés par un silence, une valeur numérique ou index de durée. Plus l'index est important, plus il révèle une coupure prosodique nette entre deux unités. Cette valeur correspond à la durée du noyau vocalique de la dernière syllabe de chaque mot ajoutée à la durée de la pause qui suit (quand il y en a une).

Leur démarche est la suivante : il faut (1) calculer les index pour l'ensemble de la phrase, (2) trouver le ou les index les plus courts de la phrase, (3) grouper les unités séparées par ce ou ces index en les rattachant à un nœud commun, (4) passer aux index les plus courts suivants et effectuer le même traitement (jusqu'au regroupement de tous les mots de la phrase).

Selon les auteurs, cette procédure permet de rendre compte des structures de performance prosodiques produites par les adultes français. Celles-ci ont trois caractéristiques essentielles : elles sont symétriques (la frontière prosodique principale est située vers le milieu d'une phrase et ne correspond pas toujours à la frontière syntaxique majeure de la phrase) ; elles sont hiérarchisées (les unités de base se regroupent en unités plus grandes...) ; et enfin, elles sont équilibrées rythmiquement.

Néanmoins, même s'il semble rendre compte assez souvent des productions des sujets et du placement des frontières prosodiques dans les phrases, ce modèle peut parfois donner des résultats surprenants et peu réalistes : par exemple, concernant l'attachement des mots grammaticaux aux autres mots, les auteurs postulent que le déterminant est toujours rattaché au constituant de gauche si ce rattachement crée une unité prosodique inférieure à 4 syllabes. Ainsi, la phrase « le professeur possède un joli chat tigré » est segmentée, selon eux, en « [le professeur] [possède un] [joli chat] [tigré] ». Intuitivement, il semble que cette segmentation ne soit pas plausible (que ce soit au niveau de l'attachement de l'article « un » au verbe, ou au niveau du rythme créé par ces nombreuses unités prosodiques très courtes). Ce qui serait proposé par les modèles de la théorie de la phonologie prosodique semble plus viable : « [le professeur] [possède] [un joli chat tigré] ».

Finalement, il faut noter que les données de Gee et Grosjean (1983) concernent des phrases produites avec un débit de parole relativement lent. Or, le débit des phrases semble aussi jouer un rôle dans l'assignation de la structure prosodique des phrases : plus les locuteurs parlent lentement, plus le nombre d'unités prosodiques augmente (Ferreira, 1993; Shattuck-Hufnagel & Turk, 1996). Par exemple, à débit rapide, les suites de 7-8 syllabes tendent à ne pas être décomposées, alors qu'en débit lent, des

variations acoustiques marquent des frontières prosodiques toutes les 3-4 syllabes ; de plus, les mouvements de fréquence fondamentale et les allongements sont moins marqués en débit rapide et peuvent même disparaître (Vaissière, 1997).

Deux courants de pensée s'affrontent quant à la manière dont la structure prosodique est définie, quant aux informations qui sont utilisées pour décider de la position des frontières d'unités prosodiques. D'une part, nous avons des modèles plutôt syntaxiques (qui partent des informations de la structure syntaxique pour mettre en place la structure prosodique), et d'autre part des modèles qui partent d'une analyse plus bas niveau des informations contenues dans le signal de parole (au niveau du rythme, des pauses, des accents, de la longueur des énoncés...) pour remonter à la structure prosodique (voir Lacheret-Dujour et Beaugendre (1999), pour une présentation plus détaillée de ces différents modèles, parmi lesquels on peut citer Philippe Martin (1980) pour une approche syntaxique, et Verluoyten (1982) ou Hirst (1993) pour une approche plus bas niveau).

Ces modèles pourraient certainement être intégrés dans un modèle plus global de la construction prosodique, chacun apportant des contraintes souvent viables et plausibles quant à la position des frontières de groupes phonologiques dans les phrases : par exemple, il nous semble logique de supposer que la syntaxe donne probablement des contraintes sur les endroits où l'on peut et ne peut pas réaliser certaines frontières prosodiques (par exemple, il est impossible d'insérer une frontière de groupe phonologique entre un pronom sujet et un verbe). La réalisation effective de ces frontières dépendrait ensuite plus de facteurs phonétiques : pour des raisons liées à des contraintes de performance au niveau de la production de la parole, les indices de bas niveau seraient alors pris en compte pour réaliser certains ajustements. Par exemple, il a été bien démontré par Delais-Roussarie (1995) que le nombre de syllabes des mots influençait le découpage de parole. Selon elle, le nombre maximum de syllabes à l'intérieur d'un groupe phonologique est de 6 syllabes. Le regroupement des mots en unités prosodiques dépend donc aussi de la longueur moyenne des unités résultantes (avec une longueur moyenne de 3.8 syllabes pour les groupes phonologiques). Il semble en effet relativement intuitif que le groupe nominal « le chien gris » sera plutôt prononcé en une seule unité, alors que « des rhinocéros époustouflants » aura plutôt

tendance à être réalisé en deux unités (alors que ce sont les mêmes structures « article – nom – adjectif » dans les deux exemples).

Nous avons vu dans cette partie que même s'il existe une correspondance étroite entre structure prosodique et structure syntaxique, cette relation n'est pas l'identité. De plus, l'arbre prosodique, contrairement à l'arbre syntaxique, a une structure plate (il n'admet pas de récursivité). De ce fait, il existe de nombreux cas où la correspondance entre unités prosodiques et syntaxiques n'est pas parfaite.

On sait aussi que des organisations intonatives différentes peuvent correspondre à une seule et même structure syntaxique (et inversement). Les unités prosodiques n'entretiennent donc pas de relation bi-univoque avec les unités syntaxiques qui composent les phrases. Le fait que les locuteurs aient différentes options possibles pour le traitement prosodique d'une structure syntaxique donnée suggère bien que la syntaxe ne détermine pas entièrement la prosodie.

Même si la structure prosodique d'un énoncé peut être considérée comme une entité grammaticale à part entière, qui requiert sa propre analyse (Beckman, 1996), il ne faut pas oublier que la prosodie et la syntaxe sont très liées : par exemple, toutes les frontières de groupes phonologiques correspondent à des frontières syntaxiques (même si l'inverse n'est pas vrai). Est-il alors possible que le fait de percevoir les mots, regroupés en unités prosodiques, aide les auditeurs à réaliser une analyse syntaxique des phrases ?

### *1.2.2 Le rôle des indices prosodiques dans la résolution d'ambiguïtés syntaxiques*

Nous avons vu que toutes les frontières syntaxiques n'étaient pas obligatoirement marquées au niveau prosodique ; par contre, toutes les frontières de groupes phonologiques (ainsi que les frontières de groupes intonationnels) coïncident avec des frontières syntaxiques.

Quand et comment la prosodie pourrait-elle faciliter l'analyse syntaxique ? Pour Nicol (1996), il est probable que la résolution de certaines ambiguïtés, en particulier la résolution d'ambiguïtés syntaxiques globales, pourrait être facilitée par les informations

apportées par les frontières prosodiques. Celles-ci pourraient, par exemple, indiquer à l'auditeur le moment où se termine un constituant prosodique, donc l'endroit où se termine un constituant syntaxique. Des indices prosodiques, comme ceux marquant des unités majeures (par exemple, l'allongement final et les pauses qui indiquent les bordures des groupes intonationnels), pourraient ainsi donner des renseignements sur le moment où il est nécessaire ou obligatoire de clôturer le groupe prosodique (donc, par la même occasion, de clôturer le constituant syntaxique).

Prenons l'exemple suivant :

*(16a) When you learn gradually you worry more.*

Cette phrase est ambiguë, à cause des deux sites d'attachement possibles pour l'adverbe « gradually » (soit il modifie le verbe de la proposition subordonnée « learn », soit il modifie le verbe de la proposition principale « worry »). Cette phrase peut donc être prononcée (ou lue), puis interprétée, de deux manières différentes (comme indiqué dans les deux exemples suivants, où la présence de la virgule désambiguïse l'interprétation des phrases à l'écrit) :

*(16b) When you learn, gradually you worry more.*

*Quand on apprend, on s'inquiète de plus en plus.*

*(16c) When you learn gradually, you worry more.*

*Quand on apprend petit à petit, on s'inquiète plus.*

Considérons maintenant la première interprétation, donnée par la phrase 16b. Au point d'ambiguïté, différentes informations sont disponibles : (1) il existe une proposition à gauche qui est potentiellement complète ; (2) Cette proposition est introduite par un adverbe « when » qui signale qu'elle doit être subordonnée à une autre proposition. Supposons maintenant que les indices prosodiques sont perceptibles et disponibles, alors on a : (3) le verbe « learn » est allongé et suivi par une pause ; (4) La proposition « when you learn » est prononcée avec le contour intonatif des groupes intonationnels. Prises ensemble, ces informations donnent beaucoup de poids pour clôturer la première proposition après le verbe, et donc pour supposer que le mot ambiguë « gradually » fait partie de la proposition suivante.

En résumé, on peut dire que, intuitivement, les indices situés en fin d'unités prosodiques sont utilisés par les auditeurs pour déterminer quand et où terminer une unité syntaxique (Nicol, 1996, p.182).

Cette hypothèse a bien entendu été démontrée expérimentalement, et nous allons donc maintenant décrire différentes expériences qui ont été réalisées pour étudier le rôle des indices prosodiques dans la résolution d'ambiguïtés syntaxiques (voir aussi, Cutler, Dahan, & van Donselaar, 1997, pour une revue détaillée). Par contre, il est important de noter dès à présent, que la majorité de ces expériences ont étudié le rôle des frontières de groupes intonationnels, donc les frontières prosodiques les plus saillantes et les plus marquées d'un point de vue acoustique.

#### 1.2.2.1 Résolution d'ambiguïtés syntaxiques temporaires : stimuli synthétisés

Expériences « off-line » :

L'exemple que nous venons de décrire est un exemple de phrase complètement ambiguë. Comme le remarque Beach (1991), ces cas d'ambiguïté sont très rares dans le discours. Si les effets de la prosodie étaient restreints à ces ambiguïtés, alors la prosodie aurait un rôle relativement mineur dans le traitement général de la parole. Par contre, si les indices prosodiques peuvent influencer l'interprétation d'ambiguïtés beaucoup plus courantes, c'est-à-dire des ambiguïtés temporaires ou locales, alors la prosodie aurait un rôle plus important dans le traitement de la parole continue. Ce sont donc ces cas d'ambiguïtés temporaires qui ont été largement étudiés.

Une phrase est temporairement ambiguë si, en traitant uniquement le début de cette phrase (donc jusqu'au point d'ambiguïté, avant que toutes les informations ne soient disponibles), il est impossible de prédire quelle sera la structure syntaxique de la phrase complète ; par contre, à la fin de la phrase, une seule interprétation reste possible parmi toutes les alternatives.

Considérons les phrases suivantes, utilisées par Beach (1991) :

*(17a) Jay believed the gossip about the neighbors right away.*

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

*(17b) Jay believed the gossip about the neighbors wasn't true.*

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

Si l'on considère chacune de ces phrases dans leur globalité, une seule interprétation est possible. Par contre si l'on entend uniquement le début de ces phrases, jusqu'à la fin du segment souligné, alors il y a une ambiguïté temporaire : dans la première phrase (17a), ce groupe nominal est le complément d'objet direct du verbe « believed », alors qu'il est le groupe nominal sujet d'une deuxième proposition relative dans la deuxième phrase (17b).

L'autre différence entre ces deux phrases est une différence au niveau des indices prosodiques produits (stimuli synthétisés) : dans la phrase 17b, la dernière voyelle du verbe « believed » est plus longue (275 ms versus 75 ms dans la première phrase) et les variations d'intonation sont beaucoup plus marquées (une baisse plus importante de la fréquence fondamentale au niveau du verbe suivie par une élévation de la fréquence fondamentale au niveau du groupe nominal ambigu).

L'auteur voulait savoir si des auditeurs pouvaient utiliser ces variations prosodiques, caractéristiques d'une phrase à complément d'objet direct ou caractéristiques d'une phrase à proposition relative, pour identifier ou prédire de quelle phrase complète venait un extrait : pour ce faire, elle faisait entendre aux sujets le début de ces phrases test, comme « Jay believed » ou « Jay believed the gossip » et ils devaient choisir de quelle phrase ils avaient été extraits. Les deux débuts de phrases étaient présentés avec les deux prosodies possibles.

Les résultats ont montré que les sujets identifiaient plus fréquemment un début de phrase comme étant le début d'une phrase à complément d'objet quand il contenait le pattern prosodique approprié (55% de réponses « complément d'objet » quand la prosodie était appropriée versus 40% de réponses « complément d'objet » quand la prosodie était inappropriée). Quand les stimuli étaient prononcés avec des durées et des variations de fréquence fondamentale plus élevées au niveau du verbe, les sujets les identifiaient plus comme provenant de phrases à proposition relative. Pour Beach (1991), les auditeurs sont donc capables d'utiliser les indices de durée et de fréquence fondamentale pour identifier la structure syntaxique des phrases qu'ils entendent, et ce même quand les informations sur la phrase sont incomplètes (Beach, 1991, p.650). Mais

cette capacité ne semble pas parfaite, puisque l'effet obtenu dans cette expérience est assez faible: la différence entre les deux pourcentages de réponses « complément d'objet » n'est que de 15%. Ces résultats ont ensuite été répliqués dans d'autres expériences (Stirling & Wales, 1996), et indiquent que la prosodie semble donc pouvoir être informative (mais pas de manière entièrement fiable puisque les effets étaient toujours de petite taille).

#### Expériences « on-line » :

Nagel, Shapiro, Tuller et Naway (1996) ont réalisé à leur tour une expérience où ils manipulaient les mêmes sortes d'ambiguïtés (problèmes d'attachement d'un groupe de mots). La différence fondamentale avec les expériences précédentes (comme celle de Beach, 1991, présentée ci-dessus) était qu'ils essayaient de voir ce qu'il se passait en temps réel. Les autres expériences indiquaient seulement que la prosodie pouvait jouer dans l'interprétation des phrases, mais pas en temps réel, seulement au niveau de « l'interprétation ultime de la phrase » (Nagel et al., 1996, p.320). Les auteurs ont utilisé des phrases comme les exemples suivants :

*(18a) The company owner promised the wage increase to the workers.*

*Le patron de la compagnie a promis une augmentation des salaires à ses employés.*

*(18b) The company owner promised the wage increase would be substantial.*

*Le patron de la compagnie a promis que l'augmentation des salaires serait importante.*

L'ambiguïté locale de ces phrases concerne le site d'attachement du groupe nominal « the wage increase ». La phrase 18a est une phrase à complément d'objet puisque, dans cet exemple, le groupe nominal ambigu est le complément d'objet direct du verbe « promised ». Par contre, la phrase 18b est une phrase à proposition relative puisque le même groupe nominal ambigu est ici le sujet d'une deuxième proposition relative.

Ces deux phrases variaient au niveau prosodique, de manière similaire à ce que nous avons vu dans l'expérience précédente : durée relativement courte du noyau vocalique du verbe et peu de changement de fréquence fondamentale pour les phrases à complément d'objet (18a) versus durée plus longue et variations de fréquence

fondamentale plus importantes au niveau du verbe pour les phrases à proposition relative (18b).

Deux autres phrases ont été construites par la technique de « cross-splicing » : le début des phrases d'une version (jusqu'au groupe nominal ambigu) était attaché à la fin des phrases de l'autre version. Ainsi, il y avait non coïncidence entre la prosodie et le sens véhiculé par la phrase complète (si l'on représente les frontières des groupes intonationnels par des #, on obtiendrait par exemple la phrase suivante : « The company owner promised # the wage increase to the workers # »).

Si la prosodie est utilisée en temps réel pour interpréter la structure syntaxique de la phrase, alors les versions coïncidentes devraient entraîner des interprétations immédiates et correctes de la nature du groupe nominal. Par contre, dans les phrases construites par « cross-splicing », l'interprétation initiale de la nature du groupe nominal devrait être fautive et donc entraîner des temps de traitement plus longs (par exemple, si les sujets utilisent la prosodie, ils devraient conclure, de la durée plus longue du verbe et des variations de fréquence fondamentale importantes, que le groupe nominal est le sujet d'une deuxième proposition et être surpris donc plus lents en entendant la phrase à complément d'objet).

Ces hypothèses ont été confirmées par les résultats de la tâche de décision lexicale que les sujets devaient réaliser. La cible sur laquelle la décision devait se faire était située après le premier mot désambiguïsant de la phrase (c'est-à-dire après « to » pour les phrases à complément d'objet direct et après « would » pour les phrases à proposition relative), mais n'avait aucune relation avec la phrase. Les participants réalisaient donc une tâche de décision lexicale cross-modale, une tâche qui permettait de mesurer la charge de traitement durant la compréhension de la phrase.

Les temps de réaction moyens ont été de 802 ms et 827 ms pour les phrases contrôles à prosodie adéquate (respectivement pour les phrases contrôles à complément d'objet direct et les phrases contrôles à proposition relative) et de 869 ms et 845 ms pour les phrases construites par « cross-splicing », donc avec une prosodie non adéquate par rapport à l'interprétation de la phrase complète. Les sujets utilisaient donc la prosodie pour interpréter la nature du groupe nominal, et celle-ci étant inadéquate, ils devaient effectuer une réanalyse du matériel d'où l'augmentation de leurs temps de réaction. Mais encore une fois, l'effet trouvé par ces auteurs est significatif mais assez

faible : 67 ms pour les phrases à complément d'objet direct et 18 ms pour les phrases à proposition relative, alors que les temps de réaction moyen sont d'environ 800 ms.

Avec ces deux seules expériences, nous pourrions déjà dire que les indices prosodiques marquant certaines structures syntaxiques sont utilisés de manière « off-line » et de manière « on-line ». Une critique qui pourrait être faite est que les stimuli utilisés dans ces expériences n'étaient pas des stimuli naturels puisqu'ils étaient synthétisés, donc construits expérimentalement. Les auditeurs semblent donc pouvoir les utiliser (même si les effets ne sont pas très massifs). Mais une question demeure : est-ce que les locuteurs produisent ces mêmes indices, ou du moins produisent-ils assez d'indices désambiguïsants pour aider un auditeur à réaliser son analyse de la phrase ?

#### 1.2.2.2 Résolution d'ambiguïtés syntaxiques temporaires : stimuli naturels

##### Expériences « off-line » :

Schepman et Rodway (2000) ont utilisé des phrases temporairement ambiguës, comme dans les exemples suivants :

*(19a) The lawyer greeted the powerful barrister and the wise judge who was walking to the courtroom. (Le juriste salua le puissant avocat et le sage juge qui traversait le tribunal)*

*(19b) The lawyer greeted the powerful barrister and the wise judge who were walking to the courtroom. (Le juriste salua le puissant avocat et le sage juge qui traversaient le tribunal)*

Deux lectures et deux interprétations de ces phrases sont possibles au moment de traiter le groupe nominal ambigu et la proposition relative : la première (représentée dans l'exemple 19a et que nous appellerons la condition singulier) est la suivante : « [The lawyer] [greeted the powerful barrister] [*and the wise judge*] [who was walking to the courtroom] ». La seconde, représentée dans l'exemple 19b, correspond à la condition pluriel, soit « [The lawyer] [greeted the powerful barrister *and the wise judge*] [who were walking to the courtroom] » (les crochets indiquent les frontières

prosodiques). L'ambiguïté réside dans l'attachement du deuxième groupe nominal, et donc par la suite, dans l'attachement de la proposition relative : soit cette proposition relative se rapporte uniquement au deuxième groupe nominal (dans la condition singulier), soit elle modifie les deux groupes nominaux coordonnés (dans la condition pluriel).

Ces phrases ont été lues par des locuteurs naïfs et les analyses acoustiques ont montré que des différences existaient au niveau des indices utilisés : il y a une pause plus longue entre les deux groupes nominaux (entre « barrister » et « and ») dans la condition singulier comparée à la condition pluriel ; de plus, la baisse de fréquence fondamentale qui a lieu au niveau du premier groupe nominal (sur « barrister ») est significativement plus importante dans la condition singulier que dans la condition pluriel. Ainsi, même des locuteurs naïfs produisent spontanément des indices prosodiques pour désambiguïser des phrases syntaxiquement ambiguës.

Des auditeurs peuvent-ils utiliser ces indices pour guider leurs analyses syntaxiques des phrases ? Les indices présents sont-ils suffisamment fiables pour être utilisés ? Une première expérience a été réalisée, dans laquelle les sujets entendaient des fragments de phrases (les phrases étaient coupées juste avant l'auxiliaire désambiguïisant comme dans « The lawyer greeted the powerful barrister and the wise judge who... »). Ils devaient alors choisir qu'elle était pour eux la meilleure continuation : « was » ou « were ».

Notons que l'interprétation préférentielle, à l'écrit, est la continuation pluriel : cette préférence d'interprétation, que nous avons décrite dès le début de ce chapitre et qui repose sur l'application des principes d'attachement minimal et de clôture tardive, a été mise en évidence par les auteurs dans une tâche de lecture. Les sujets lisaient le début de la phrase, présentée sans l'auxiliaire, et ceux-ci devaient choisir entre la continuation « was » ou « were ». La continuation pluriel a été choisie dans 87% des cas.

Est-ce que ce biais d'interprétation préférentielle peut être diminué par la prosodie ? Les résultats ont pu apporter une réponse positive à cette question : en effet, les sujets ont choisi la continuation pluriel dans 76,9% des cas quand la prosodie correspondait à la version pluriel ; mais quand ils entendaient la phrase dans la version singulier, ils ne choisissaient la continuation pluriel que dans 16,1% des cas. Contrairement aux expériences présentées ci-dessus, les auteurs ont obtenu cette fois-ci

un effet massif de 61%. Il est donc clair que la désambiguïisation syntaxique est très fortement affectée par les indices prosodiques employés (Schepman et Rodway, 2000, p.386). Mais à nouveau se pose la question de l'utilisation de la prosodie en temps réel.

#### Expériences « on-line » :

Schepman et Rodway (2000) ont réalisé une tâche de dénomination cross-modale. Les sujets entendaient le début des phrases (coupées avant l'auxiliaire désambiguïisant) ; à ce moment, une cible apparaissait, et ils devaient la dénommer le plus rapidement possible (la cible était l'auxiliaire, au singulier ou au pluriel). Dans la condition coïncidente, l'auxiliaire à nommer était approprié par rapport à la prosodie (si le début de la phrase était réalisée avec la prosodie singulier, la cible à nommer était « was »). Dans la condition non coïncidente, la cible était inappropriée par rapport à la prosodie utilisée (le début de la phrase était prononcé avec la prosodie singulier, et les sujets devaient nommer « were »). Dans une condition contrôle, les sujets devaient nommer une cible sans aucun lien sémantique et syntaxique : « web ».

Les résultats ont montré que les sujets nommaient plus rapidement les cibles quand elles apparaissaient dans le bon contexte, que quand elles apparaissaient dans le contexte non coïncident. Les auteurs ont donc pu conclure que des cibles acceptables syntaxiquement sont nommées significativement plus vite quand elles apparaissent avec la prosodie appropriée que quand elles sont dans un contexte prosodique conflictuel (Schepman et Rodway, 2000, p.390).

De la même manière, d'autres études ont démontré que des locuteurs produisaient des indices prosodiques qui permettaient à des auditeurs de désambiguïser des phrases syntaxiquement ambiguës (Kjelgaard & Speer, 1999; Lehiste, 1972; Marslen-Wilson, Tyler, Warren, Grenier, & Lee, 1992; Price, Ostendorf, Shattuck-Hufnagel, & Fong, 1991; Schafer & Speer, 1998; Schepman & Rodway, 2000; Warren, Grabe, & Nolan, 1995). Par exemple, Schafer et al. (2000) ont trouvé que des locuteurs produisaient des indices différents pour distinguer des phrases comme « [When that moves *the square*] [it should land in a good spot] » versus « [When that moves] [*the square* will encounter a cookie] ». Par la suite, les auteurs ont montré que des auditeurs

pouvaient utiliser ces indices différents pour guider leur analyse de la phrase, résoudre ces ambiguïtés et réaliser correctement une tâche de catégorisation à choix forcé.

Kjelgaard et Speer (1999) ont montré que les adultes pouvaient utiliser des indices prosodiques pour effectuer avec succès des tâches « off-line » (jugement de grammaticalité, tâche de compréhension) et des tâches « on-line » (dénomination cross-modale) sur des phrases ambiguës d'un point de vue syntaxique (par exemple, les sujets arrivent à différencier sur la base de la prosodie et avant le contexte ou le mot désambiguïsant « [When Roger leaves] [*the house* is dark] » versus « [When Roger leaves *the house*] [it's dark] »).

Récemment, Kraljic et Brennan (2005) ont aussi trouvé que des locuteurs naïfs pouvaient produire des indices prosodiques différents pour distinguer les deux interprétations de la phrase suivante, qui contient une ambiguïté syntaxique globale : « Put the dog *in the basket* on the star ». En effet, le groupe nominal « in the basket » peut soit se rapporter au verbe (« Put the dog # *in the basket* on the star »), soit modifier le nom précédant (« Put the dog *in the basket* # on the star »). Des auditeurs ont ensuite pu utiliser ces informations prosodiques pour désambiguïser la phrase et choisir très rapidement l'interprétation correcte (ce qui a été mesuré par l'enregistrement des mouvements oculaires des participants, qui avaient face à eux différentes images dont deux qui soutenaient les deux interprétations possibles, à savoir « un chien dans un panier », « une étoile seule » et « un panier sur une étoile »).

Grâce à toutes ces expériences, il semble évident que des indices prosodiques sont produits de manière claire, puis utilisés pour résoudre des ambiguïtés syntaxiques et effectuer une analyse rapide des phrases entendues. L'utilisation de ces indices de frontières prosodiques peuvent ainsi prédominer et annuler les fortes préférences d'interprétation, qui avaient été mises en évidence initialement dans des tâches de lecture, et qui avaient conduits certains auteurs à postuler que seules des stratégies syntaxiques (comme la clôture tardive et l'attachement minimal) guidaient les premiers choix des adultes lors de l'analyse syntaxique des phrases.

Ce que nous devons noter, c'est le fait que les frontières d'unités prosodiques étudiées sont généralement des limites de groupes intonationnels : en effet, dans

chacune de ces expériences, les indices distinguant les phrases ambiguës étaient fortement marqués, les unités prosodiques se terminaient en général par une pause, plus ou moins longue (et les pauses sont des marqueurs acoustiques particulièrement saillants), les variations de fréquence fondamentale étaient très importantes... Même si les analyses acoustiques et prosodiques des stimuli ne sont pas toujours présentes, les auteurs parlaient généralement de frontières intonationnelles ou de frontières majeures. Donc, ces expériences ont montré que des indices fortement marqués sont perçus et utilisés pour réaliser une analyse syntaxique des phrases.

Les frontières intonationnelles se trouvent typiquement dans des phrases compliquées, contenant plusieurs propositions, autant dire des phrases très éloignées de ce qu'entendent les jeunes enfants. Si l'on s'intéresse à la question de l'acquisition, on se rend compte que des frontières plus petites, comme celles des groupes phonologiques, seraient particulièrement utiles pour pouvoir traiter des phrases plus courtes, comme « tu fais un sourire à maman » ou « il a mangé un gâteau au chocolat », c'est-à-dire des phrases dont la longueur correspond plus à ce qu'entendent les enfants.

Il serait donc intéressant d'étudier la perception et l'utilisation de ces frontières de groupes phonologiques, puisqu'elles pourraient permettre un découpage plus fin de la phrase à analyser. De plus, nous avons vu que ces frontières de groupes phonologiques correspondaient à des frontières de constituants syntaxiques : elles pourraient donc être particulièrement importantes pour l'analyse syntaxique des phrases. C'est cette question que nous avons abordée dans cette première étude que nous allons maintenant présenter.

## CHAPITRE 2

### UTILISATION DES FRONTIÈRES DE GROUPES PHONOLOGIQUES DANS LA RESOLUTION D'AMBIGUITÉS SYNTAXIQUES (ETUDE 1)

De nombreuses études, que nous avons présentées et détaillées dans la partie précédente, ont montré que les frontières de groupes intonationnels étaient exploitées en temps réel lors de l'analyse syntaxique des phrases par des auditeurs (Beach, 1991; Kjelgaard & Speer, 1999; Marslen-Wilson et al., 1992; Nagel et al., 1996; Schafer et al., 2000; Schepman & Rodway, 2000; Stirling & Wales, 1996). Ces frontières prosodiques sont très saillantes et très fortement marquées par des allongements des segments, une baisse de la fréquence fondamentale et la présence de pauses en fin d'unités (Beckman & Pierrehumbert, 1986; Nespor & Vogel, 1986). Le résultat majeur de ces expériences est que les auditeurs peuvent utiliser ces frontières de groupes intonationnels pour segmenter les phrases en constituants syntaxiques et résoudre ainsi des ambiguïtés syntaxiques.

Les groupes intonationnels peuvent être divisés en unités plus petites : les groupes phonologiques. Ces unités sont moins fortement marquées mais sont néanmoins contraintes par la syntaxe des phrases (Nespor & Vogel, 1986; Selkirk, 1982; Shattuck-Hufnagel & Turk, 1996). Les phrases possèdent donc une structure prosodique qui dépend en partie de leur structure syntaxique. Percevoir cette structure prosodique pourrait donc guider l'analyse syntaxique des phrases. Mais actuellement, le rôle des frontières de groupes phonologiques dans la résolution d'ambiguïtés syntaxiques n'est pas aussi clair que celui des groupes intonationnels. Par exemple, Kjelgaard et Speer (1999), ainsi que Schafer (1997), proposent que les frontières de groupes phonologiques, de même que les frontières de groupes intonationnels, aident à l'interprétation de la structure syntaxique des phrases. Par contre, des auteurs comme Price et al. (1991) suggèrent que seuls les groupes intonationnels ont un effet sur l'analyse syntaxique des phrases.

Cette première étude a donc été réalisée pour approfondir les connaissances que nous avons sur le rôle des frontières de groupes phonologiques dans l'analyse syntaxique des phrases. Nous avons voulu examiner si ces indices prosodiques plus fins et moins saillants (comparés aux indices de frontières de groupes intonationnels) pouvaient eux aussi aider les auditeurs à résoudre des ambiguïtés syntaxiques.

Pour ce faire, nous avons utilisé des phrases simples de structure « Sujet-Verbe-Objet » contenant une ambiguïté syntaxique locale : celle-ci reposait sur le fait que deux homophones pouvaient appartenir à des catégories syntaxiques différentes. Considérons l'exemple suivant :

*Le petit chien mord la laisse qui le retient.*

*Le petit chien mort sera enterré demain.*

Dans chacune de ces phrases, les quatre premiers mots ont la même prononciation ([lɛ pə ti ʃ jɛ mɔʀ]) : les débuts de ces phrases sont donc identiques d'un point de vue segmental. Mais le quatrième mot est un verbe dans la première phrase, alors que c'est un adjectif dans la deuxième. Les débuts de ces deux phrases diffèrent donc du point de vue de leur structure syntaxique et de leur structure prosodique, comme les transcriptions suivantes le montrent :

Découpage syntaxique<sup>9</sup> :

[ [Le petit chien]<sub>GN</sub> [mord [la laisse]<sub>GN</sub>[qui le retient]<sub>PR</sub>]<sub>GV</sub>]<sub>S</sub>

[ [Le petit chien mort]<sub>GN</sub> [sera enterré demain]<sub>GV</sub>]<sub>S</sub>

Découpage prosodique<sup>10</sup> :

[ [Le petit chien]<sub>GPh</sub> [mord la laisse]<sub>GPh</sub>[qui le retient]<sub>GPh</sub>]<sub>GI</sub>

[ [Le petit chien mort]<sub>GPh</sub> [sera enterré demain]<sub>GPh</sub>]<sub>GI</sub>

Dans la première phrase, on trouve une frontière syntaxique et une frontière prosodique avant le mot ambigu, alors que ces frontières sont placées après le mot ambigu dans la deuxième phrase. Donc si des auditeurs doivent traiter ce mot ambigu

---

<sup>9</sup> Un groupe nominal est codé « GN », un groupe verbal « GV », une proposition relative « PR » et la phrase complète « S » (pour « Sentence »).

<sup>10</sup> Les groupes intonationnels sont codés « GI » et les groupes phonologiques « GPh ».

dans ces phrases, son contenu phonémique ne donnera aucun indice quant à sa catégorie syntaxique, alors que le contexte prosodique pourrait apporter cette information. Nous avons voulu vérifier si la prosodie des phrases pouvait effectivement permettre à des auditeurs de distinguer ces deux structures et d'assigner au mot ambigu sa bonne catégorie syntaxique.

## 2.1 EXPERIENCE 1 : COMPLETION DE PHRASES

Dans cette première expérience, nous voulions vérifier qu'il existait bien une différence prosodique entre les paires de phrases ambiguës, différence qui serait perceptible et utilisable par des auditeurs français. Pour ce faire, nous avons réalisé une expérience de complétion de phrases dans laquelle les sujets entendaient le début des phrases ambiguës, coupées juste après le mot ambigu (pour les exemples ci-dessus, les sujets entendaient le début des phrases jusqu'au mot ambigu « mord/mort » inclus). Les sujets devaient ensuite compléter ces phrases à l'écrit. Si des différences prosodiques entre ces deux débuts de phrases sont perceptibles, et si les sujets peuvent les utiliser pour assigner correctement au mot ambigu sa catégorie syntaxique, alors nous nous attendons à des complétions différentes pour chaque début de phrase, complétions basées sur les propriétés prosodiques des stimuli.

### 2.1.1 *Méthode*

#### 2.1.1.1 Matériel

Nous avons utilisé vingt couples adjectif / verbe ambigus (voir Annexe A.1. pour le détail des stimuli utilisés). Pour chaque couple, deux phrases expérimentales ambiguës ont été construites : la première contenait le mot ambigu dans sa forme adjectivale (cette phrase sera appelée « phrase ambiguë adjectif » par la suite), tandis que l'autre contenait la forme verbale (« phrase ambiguë verbe »). Chaque paire de phrases était identique d'un point de vue phonémique jusqu'au mot ambigu inclus. La seule différence était une différence de structure syntaxique, et donc de structure prosodique. En effet, une frontière de groupe phonologique était placée juste avant le mot ambigu dans les phrases contenant le verbe, alors qu'elle était située juste après le mot ambigu

dans les phrases contenant l'adjectif (la frontière de groupe phonologique se situait toujours entre le groupe nominal sujet et le groupe verbal). Rappelons l'exemple de la paire de phrases ambiguës pour l'item /mɔʀ/ (les frontières de groupes phonologiques seront indiquées par des crochets) :

Phrase ambiguë verbe : [Le petit chien] [mord la laisse] [qui le retient]

Phrase ambiguë adjectif : [Le petit chien mort] [sera enterré demain]

En plus de ces quarante phrases expérimentales, dix phrases distractrices ne contenant aucune ambiguïté ont été construites (par exemple, « le nouveau professeur est très compétent »).

Toutes ces phrases ont ensuite été enregistrées par une locutrice française dans une cabine insonorisée. Elles ont été digitalisées à 16000 Hz et à 16 bits, sur une table de mixage OROSAU22. Elles ont ensuite été éditées grâce aux programmes Manitou et Praat. Toutes les phrases expérimentales ont été coupées à la fin du mot ambigu, à un passage à zéro de la courbe d'amplitude. Les phrases distractrices ont elles aussi été coupées à la fin d'un mot, placé à n'importe quel endroit de la phrase (pour l'exemple cité ci-dessus, les sujets entendaient le début de phrase : « le nouveau professeur »).

Deux conditions de prosodie ont été réalisées pendant les enregistrements des stimuli. La locutrice était une locutrice experte et non-naïve, qui avait pleinement conscience de ce problème d'ambiguïté que nous testions. Elle enregistrait les phrases deux par deux (prononçant ainsi à la suite la phrase ambiguë adjectif et la phrase ambiguë verbe relative à un item). Elle essayait soit de produire des indices prosodiques forts pour désambiguïser le mieux possible ces phrases expérimentales (condition de « prosodie informative »), soit de minimiser les indices prosodiques entre ces couples de phrases ambiguës (condition de « prosodie neutre »). Dans tous les cas, les stimuli résultants étaient possibles : aucune prosodie n'était agrammaticale. Notre hypothèse était que si les indices prosodiques étaient des indices forts pour les sujets, alors ils devraient obtenir de meilleurs résultats quand la prosodie est la plus informative possible.

La moitié des sujets ont traité les stimuli enregistrés avec une prosodie informative, alors que l'autre moitié n'entendaient que les stimuli enregistrés avec une prosodie neutre. Dans chaque condition de prosodie, deux blocs de stimuli ont été créés de manière à ce que chaque membre d'une paire de phrases ambiguës apparaisse dans un bloc différent. Pour chaque prosodie, nous avons donc testé deux groupes de sujets (ceux qui traitaient les items du bloc 1 et ceux qui traitaient les items du bloc 2). Ainsi, chaque bloc contenait dix phrases ambiguës verbe, dix phrases ambiguës adjectif et les dix phrases distractrices. Ces 30 phrases ont été séparées en deux sous-blocs de manière à ce que le sujet puisse faire une pause à la moitié de l'expérience (chaque sous-bloc comprenant 5 phrases ambiguës verbe, 5 phrases ambiguës adjectif et 5 phrases distractrices). La présentation des stimuli à l'intérieur des sous-blocs était aléatoire ; par contre, l'ordre de présentation des sous-blocs était fixe pour un groupe de sujets.

En plus de ces phrases expérimentales et distractrices, deux phrases non ambiguës (similaires aux phrases distractrices) ont été créées et utilisées au début de l'expérience comme entraînement.

#### 2.1.1.2 Procédure

L'expérience a été réalisée en utilisant le programme Expe6 de présentation des stimuli (Pallier, Dupoux, & Jeannin, 1997)<sup>11</sup>. Tous les participants étaient testés individuellement, dans une pièce isolée, équipée d'un ordinateur, d'un casque et d'une feuille réponse. Avant de commencer, chaque participant était assigné à un groupe, puis remplissait un formulaire indiquant sa date de naissance, un éventuel bilinguisme, et s'il avait des problèmes auditifs ou des problèmes de langage. Il découvrait ensuite les consignes de l'expérience : les instructions étaient données oralement, en complément des instructions écrites à l'écran (voir Annexe A.2.1 pour les instructions écrites détaillées). L'expérimentateur s'assurait que le sujet avait bien compris la tâche en restant près de lui pendant les deux essais d'entraînement, ce qui permettait aussi de clarifier toutes les questions éventuelles avant de commencer le test.

Les instructions expliquaient aux participants qu'ils allaient entendre des débuts de phrases, coupées à la fin d'un mot, et que leur tâche serait de trouver une suite à ces phrases. Ils devaient pour cela écrire sur la feuille réponse le début de la phrase qu'ils avaient entendue, puis la compléter en mettant la première idée qui leur venait à l'esprit.

Un essai commençait avec la présentation auditive du début de la phrase (les stimuli auditifs étaient présentés à travers une carte son ProAudioSpectrum 16-bit). Les participants pouvaient écouter ce début de phrase autant de fois qu'ils le désiraient : il leur suffisait d'appuyer sur la barre d'espace de l'ordinateur pour rejouer le stimulus en cours. Ils devaient ensuite écrire la phrase complète (avec la fin imaginée) sur la feuille réponse. Une fois leur tâche réalisée, ils pouvaient lancer l'essai suivant, en appuyant sur la touche entrée. Les sujets devaient traiter les 30 phrases de l'expérience pour que l'expérience se termine. Une pause était proposée aux sujets à la moitié de l'expérience (entre les deux sous-blocs). L'expérience complète durait environ 15 minutes.

### 2.1.1.3 Participants

Trente-six personnes, natives du français, ont participé à cette expérience (dix-huit dans chaque condition de prosodie : informative versus neutre). Aucun d'eux n'a rapporté de problèmes auditifs ou de problèmes liés au langage.

Un participant supplémentaire a été testé dans la condition de prosodie neutre. Mais à la fin de l'expérience, cette personne nous a expliqué qu'elle s'était rendue compte du problème d'ambiguïté dans ces phrases, et qu'elle les avait alors complétées en se basant sur son estimation de la fréquence des verbes et des adjectifs concernés. Elle a de plus utilisé un critère inconsistant, en sélectionnant parfois la forme la plus fréquente, alors qu'à d'autres moments elle préférait choisir la forme la moins fréquente. Ses résultats n'ont donc pas été intégrés dans les analyses que nous allons maintenant présenter.

Quant aux autres, à la fin de l'expérience, aucun n'avait compris que nous testions le rôle de la prosodie dans la résolution d'ambiguïtés syntaxiques (nous leur demandions s'ils avaient compris le but de notre expérience). Certains avaient repéré que certains items étaient ambigus mais ils n'ont pas généralisé ces exemples à l'expérience globale (dans ces cas là, les participants ne semblaient pas avoir mis en place des stratégies de réponse : ils disaient avoir utilisé le premier sens qui leur était venu à l'esprit, ou avoir mis celui qui semblait aller le mieux).

---

<sup>11</sup> Voir le site internet : <http://www.lscp.net/expe/>

## 2.1.2 Résultats

Avant de tester les sujets, nous avons réalisé des analyses acoustiques des stimuli (grâce au logiciel Praat), pour nous assurer que des différences prosodiques existaient bien entre les débuts de paires de phrases ambiguës.

### 2.1.2.1 Analyses acoustiques des stimuli

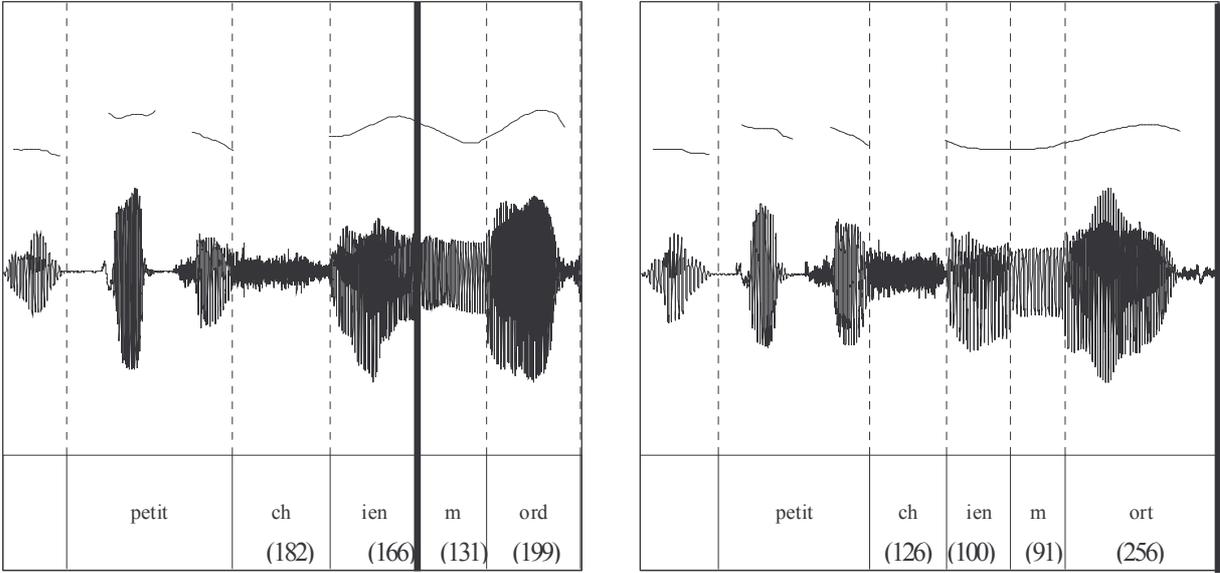
Nous avons centré notre attention sur les segments situés en fin d'unité prosodique dans la région ambiguë que nous analysions. Deux frontières de groupes phonologiques devaient donc être étudiées : la première était placée juste avant le mot ambigu (dans les phrases ambiguës verbe), et la deuxième était placée après le mot ambigu (dans les phrases ambiguës adjectif).

L'étude de la première position possible de la frontière de groupe phonologique concerne les phrases verbe, comme « [le petit chien] [mord la laisse] [qui le retient] » (frontière avant le mot ambigu). Nous avons analysé les caractéristiques acoustiques du mot situé à la fin du premier groupe phonologique, c'est-à-dire « chien », ainsi que la courbe intonative à la fin de cette unité prosodique (entre « petit » et « chien »). Nous avons comparé ces résultats avec les analyses des segments correspondant dans la phrase ambiguë adjectif (comparaison de la durée de « chien » et de la courbe de fréquence fondamentale entre « petit » et « chien » dans les deux phrases ambiguës).

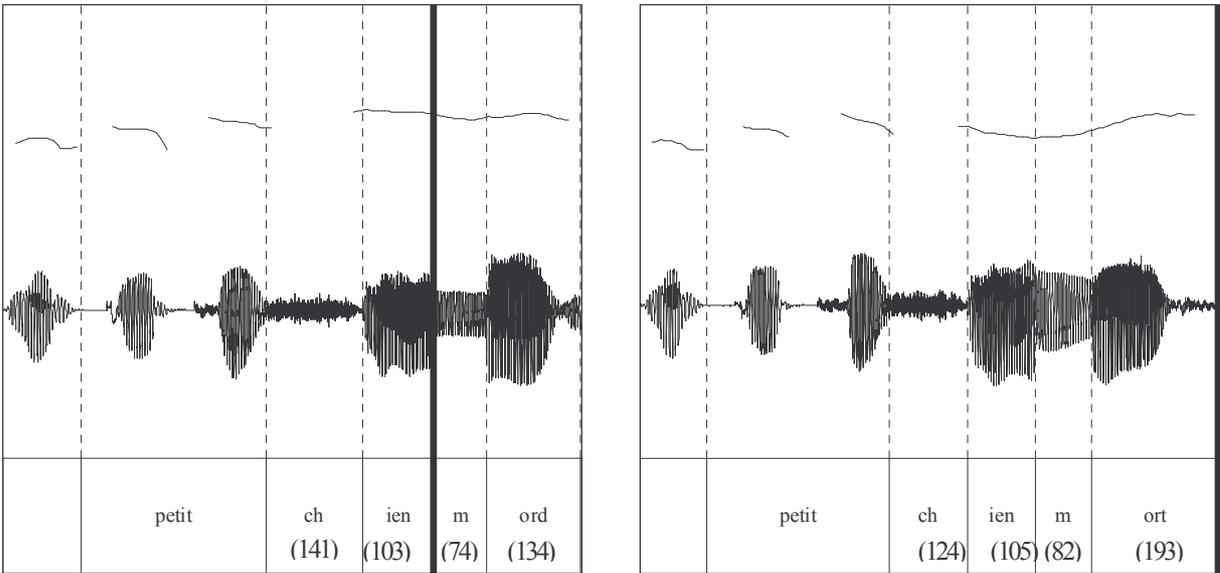
L'étude de la deuxième position possible de la frontière de groupe phonologique concerne les phrases adjectif, comme « [le petit chien mort] [sera enterré demain] » (frontière après le mot ambigu). Dans ce cas, nous avons mesuré la durée du mot « mort », ainsi que la courbe intonative en fin de groupe phonologique (c'est-à-dire entre « chien » et « mort »). Nous avons comparé ces résultats avec ceux obtenus sur les mêmes mots prononcés dans la phrase ambiguë verbe.

La figure suivante (Figure 1) représente la durée des différents éléments (adjectif « petit », attaque et rime du nom « chien », attaque et rime du mot ambigu) ainsi que la courbe intonative de la région ambiguë étudiée, pour les paires de phrases ambiguës de l'item « mord / mort » prononcées avec une prosodie informative (en haut) et une prosodie neutre (en bas).

**PROSODIE INFORMATIVE**



**PROSODIE NEUTRE**



**Figure 1 : (analyses acoustiques – étude 1) : Durée des segments et des mots (en ms) et courbe intonative de la région ambiguë étudiée. Les phrases enregistrées avec une prosodie informative sont en haut de la page, celles enregistrées avec une prosodie neutre en bas de la page. Les phrases ambiguës verbe sont représentées à gauche, les phrases ambiguës adjectif à droite. Les lignes noires en gras indiquent la position des frontières de groupes phonologiques.**

A première vue, pour la condition de prosodie informative, il semble évident que la longueur des segments et la courbe de fréquence fondamentale soient différentes selon que l'on considère la phrase ambiguë verbe (en haut à gauche) ou la phrase ambiguë adjectif (en haut à droite). Ces différences semblent moins saillantes pour les phrases prononcées avec une prosodie plus neutre. Des analyses plus détaillées vont nous permettre de confirmer cette impression.

#### Résultats des analyses de durée :

D'après les résultats trouvés dans la littérature (Delais-Roussarie, 1995, entre autres), nous nous attendions à trouver un allongement des segments en fin d'unité prosodique. Par exemple, la rime du nom « **chien** » devrait être plus longue dans la phrase verbe « [le petit **chien**] [mord...] » de par sa position en fin de groupe phonologique, que dans la phrase adjectif « [le petit **chien** mort]... » où ce mot apparaît en milieu de groupe phonologique.

De la même manière, la rime du mot ambigu devrait être plus longue dans les phrases adjectif puisqu'il est situé en fin de groupe phonologique, que dans les phrases verbe où le mot ambigu apparaît en début d'unité prosodique (dans notre exemple, « **mort** » dans « [le petit chien **mort**]... » devrait être plus long que « **mord** » dans « [le petit chien] [**mord**] »)

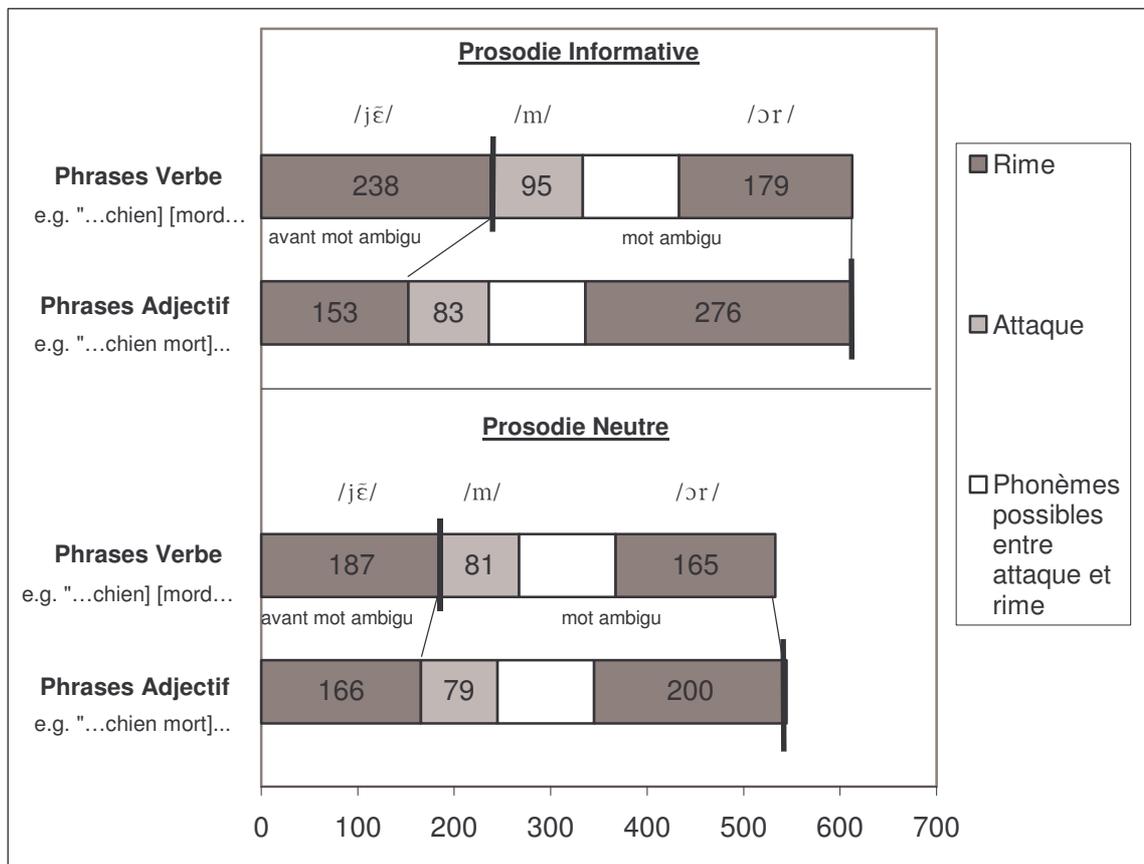
De plus, nous attendions aussi un allongement de l'attaque des mots situés en début d'unité prosodique, donc un allongement de l'attaque du mot ambigu dans les phrases verbes comparé aux phrases adjectif (Fougeron & Keating, 1997)<sup>12</sup>.

Nous avons donc mesuré la durée de la rime du mot ambigu, la durée de la rime du mot précédant le mot ambigu, ainsi que la durée de l'attaque du mot ambigu. La Figure 2 représente les durées moyennes de chacun de ces segments. Dans cette figure, la partie supérieure concerne la prosodie informative (quand la locutrice essayait de

---

<sup>12</sup> Nous n'avons pas mesuré l'attaque du mot suivant la deuxième frontière de groupe phonologique (le mot après le mot ambigu) car nous n'avons pas apparié ce premier phonème dans les paires de phrases ambiguës : dans notre exemple, cela reviendrait à comparer le /s/ de « sera » (dans « [le petit chien mort] [sera...] ») avec le /l/ de « la » (dans « [le petit chien] [mord la...] »). Une différence de durée entre ces phonèmes pourrait être due aux caractéristiques intrinsèques des phonèmes eux-mêmes, plutôt qu'à leur position dans la phrase.

maximiser les différences prosodiques entre les paires de phrases ambiguës), tandis que la partie inférieure représente la condition de prosodie neutre (quand la locutrice essayait de minimiser les différences entre les paires de phrases ambiguës).



**Figure 2 (analyses acoustiques – étude 1) : Durée moyenne des différents segments situés en bordure d'unités prosodiques, pour la prosodie informative (en haut de la figure) et pour la prosodie neutre (en bas de la figure).** Les lignes dessinées en gras représentent les frontières de groupes phonologiques. La catégorie des phonèmes possibles entre l'attaque et la rime du mot ambigu concerne les stimuli bi- et trisyllabiques. En effet, pour ces items, nous n'avons pas analysé les phonèmes qui se situaient entre l'attaque et la rime du mot (par exemple, pour « réfléchi » - /r e f l e ʃ i/ - l'attaque est /r/, la rime est /i/ et les autres phonèmes possibles mais non analysés sont /e f l e ʃ/).

Comme nous l'attendions, nous avons observé un allongement final en fin de groupe phonologique. Dans la condition de prosodie informative, nous avons obtenu un allongement de la rime de 56% avant la première frontière de groupe phonologique (de 153 ms à 238 ms,  $t(19)=10.0$ ,  $p<.001$ ), et de 54% avant la deuxième frontière de groupe phonologique (de 179 ms à 276 ms,  $t(19)=8.9$ ,  $p<.001$ ). Dans la condition de prosodie neutre, cet allongement de la rime est significatif mais d'une moins grande ampleur :

13% avant la première frontière (166 ms versus 187 ms,  $t(19)=4.6$ ,  $p<.001$ ) et 21% avant la deuxième frontière (165 ms versus 200 ms,  $t(19)=4.3$ ,  $p<.001$ ).

Nous avons aussi observé un allongement de l'attaque du mot situé en début d'unité prosodique, mais uniquement dans la condition de prosodie informative. Dans cette condition, l'attaque des mots ambigus était plus longue quand ils étaient situés en début de groupe phonologique (dans les phrases verbe), que quand ils étaient en fin de groupe phonologique (dans les phrases adjectif) : nous avons obtenu un allongement de 14% dans la condition de prosodie informative (83 ms versus 95 ms,  $t(19)=3.0$ ,  $p<.001$ ), alors que la différence n'est que de 2% dans la prosodie neutre (79 ms versus 81 ms,  $t(19)<1$ ).

Enfin, comme nous l'avions attendu, les allongements sont significativement plus importants dans la condition de prosodie informative que dans la condition de prosodie neutre. Ceci est vrai pour les différences de durée des segments autour de la première frontière de groupe phonologique (pour la rime :  $t(39)=5.6$ ,  $p<.001$  ; pour l'attaque :  $t(39)=2.6$ ,  $p<.01$ ), ainsi qu'autour de la deuxième frontière de groupe phonologique (pour la rime :  $t(39)=5.0$ ,  $p<.001$ ).

#### Résultats des analyses sur la courbe de fréquence fondamentale :

Nous avons mesuré la différence de fréquence fondamentale entre les deux voyelles précédant immédiatement une frontière de groupe phonologique, puis nous avons comparé cette mesure avec la différence de fréquence fondamentale existant entre ces deux mêmes voyelles quand elles étaient placées au milieu d'un groupe phonologique. D'après la littérature (Di Cristo, 2000b; Welby, 2003b), nous attendions une montée de l'intonation à la fin des unités prosodiques

Par exemple, pour la première frontière de groupe phonologique, nous avons comparé la différence entre le /jɛ/ de « chien » (/ʃ jɛ/) et le /i/ de « petit » (/p ɛ t i/) dans les deux phrases ambiguës. Nous attendions une courbe ascendante de la fréquence fondamentale sur « petit chien » dans la phrase ambiguë verbe « [le petit chien] [mord...] », comparé à la phrase ambiguë adjectif « [le petit chien mort]... ». Les résultats obtenus autour de cette première frontière de groupe phonologique sont indiqués dans le Tableau 1.

De la même manière, pour la deuxième frontière de groupe phonologique à étudier, nous avons comparé la différence entre le /ɔ/ de « mort/mord » (/mɔʀ/) et le /jɛ̃/ de « chien » (/ʃjɛ̃/). Nous attendions une montée de la courbe de fréquence fondamentale sur le groupe « chien mort » qui est situé en fin de groupe phonologique (dans la phrase adjectif), comparé à la courbe de fréquence fondamentale sur le groupe « chien mord » (dans la phrase verbe) puisque ces deux mots sont situés de part et d'autre d'une frontière de groupe phonologique. Les résultats obtenus autour de cette deuxième frontière de groupe phonologique sont indiqués dans le Tableau 2.

<b>Courbe intonative avant la première frontière de groupe phonologique</b>						
Comparaison entre « [le petit <b>chien</b> ] [mord...] » versus « [le petit <b>chien</b> mort]... »						
		Première Voyelle pet(i)t	Deuxième Voyelle ch(ien)	Position des voyelles dans le groupe phonologique	Différence (erreur standard) entre les voyelles	t-test (phrases verbe - adjectif)
Prosodie Informative	Phrases Adjectif	255	222	Milieu	-33 (7,5)	t(19)=4.3 p<.001
	Phrases Verbe	257	287	Fin	30 (15,4)	
Prosodie Neutre	Phrases Adjectif	277	274	Milieu	-3 (7,7)	t(19)=6.6 p<.001
	Phrases Verbe	270	299	Fin	30 (7,1)	

**Tableau 1 : Valeur moyenne de la fréquence fondamentale (en Hz) des deux voyelles placées avant la première position de frontière de groupe phonologique (frontière avant le mot ambigu).** Nous avons comparé ces segments quand ils étaient dans les phrases ambiguës verbe, donc en fin de groupe phonologique (par exemple, « [le petit **chien**] [mord...] »), et quand ils étaient dans les phrases ambiguës adjectif, donc au milieu d'un groupe phonologique (par exemple, « [le petit **chien** mort]... ». Les résultats obtenus pour les deux prosodies sont indiqués dans ce tableau (prosodie informative en haut, prosodie neutre en bas).

<b>Courbe intonative avant la deuxième frontière de groupe phonologique</b>						
Comparaison entre « [le petit <b>chien</b> mort]... » versus « [le petit <b>chien</b> ] [mord...] »						
		Première Voyelle ch( <b>ien</b> )	Deuxième Voyelle m( <b>o</b> )rt	Position des voyelles dans le groupe phonologique	Différence (erreur standard) entre les voyelles	t-test (phrases adjectif - verbe)
Prosodie Informatrice	Phrases Adjectif	223	291	Fin	69 (10,3)	t(19)=3.5 p=.003
	Phrases Verbe	280	293	de chaque côté	13 (12,4)	
Prosodie Neutre	Phrases Adjectif	260	284	Fin	24 (7,6)	t(19)=1.6 p=0.1
	Phrases Verbe	282	294	de chaque côté	13 (8,3)	

**Tableau 2 : Valeur moyenne de la fréquence fondamentale (en Hz) des deux voyelles placées avant la deuxième position de frontière de groupe phonologique (frontière après le mot ambigu).** Nous avons comparé ces segments quand ils étaient dans les phrases ambiguës adjectif, donc en fin de groupe phonologique (par exemple, « [le petit **chien** mort]... »), et quand ils étaient dans les phrases ambiguës verbe, donc de chaque côté d'une frontière de groupe phonologique (par exemple, « [le petit **chien**] [mord...] »). Les résultats obtenus pour les deux prosodies sont indiqués dans ce tableau (prosodie informative en haut, prosodie neutre en bas).

Nous avons observé des contours intonatifs différents entre les phrases ambiguës verbe et adjectif de la condition de prosodie informative. Avant la première frontière de groupe phonologique, on constate que la courbe intonative tend à être ascendante en fin d'unité prosodique, avec une montée de 30 Hz marginalement différente de 0 (hausse de la fréquence fondamentale de 30 Hz entre « petit » et chien » dans la phrase ambiguë verbe « [le petit chien] [mord...] »,  $t(19)=1.96$ ,  $p=.06$ )<sup>13</sup>. Quand les mêmes voyelles sont situées au milieu d'un groupe phonologique, la courbe intonative entre ces deux voyelles est significativement descendante (-33 Hz entre « petit » et « chien » dans la phrase adjectif « [le petit chien mort...],  $t(19)=4.5$ ,  $p<.001$ ). La courbe de fréquence fondamentale est donc différente selon que les segments considérés sont situés en fin d'unité prosodique ou au milieu de celle-ci (+30 Hz versus -33 Hz,  $t(19)=4.3$ ,  $p<.001$ ). En ce qui concerne la deuxième frontière de groupe phonologique, nous avons aussi mesuré une hausse de la fréquence fondamentale en fin d'unité (+69 Hz entre « chien » et « mort » dans la phrase ambiguë adjectif, significativement différente de 0,  $t(19)=6.7$ ,

<sup>13</sup> Pour faciliter la compréhension des résultats, nous garderons les exemples des phrases ambiguës correspondant à l'item « mort / mord » pour illustrer les résultats trouvés. Mais les chiffres indiqués sont ceux trouvés en moyenne entre toutes les phrases (par exemple, entre « petit » et « chien », la montée de fréquence fondamentale est de 65 Hz. Le chiffre de 30 Hz correspond à la moyenne de toutes les différences mesurées entre les deux voyelles précédant la première frontière de groupe phonologique)

$p < .001$ ), comparé à une différence de 13 Hz dans l'autre phrase (+13 Hz entre « chien » et « mord » dans la phrase ambiguë verbe, ce qui n'est pas significativement différent de 0,  $t(19)=1.0$ ,  $p=0.3$ ). Ces courbes intonatives sont significativement différentes entre elles (+69 Hz versus +13 Hz,  $t(19)=3.5$ ,  $p < .001$ ).

Dans la condition de prosodie neutre, nous avons obtenu un pattern similaire pour la première frontière de groupe phonologique, avec la mise en évidence d'un contour ascendant de 30 Hz quand les segments sont situés en fin d'unité (significativement différent de 0,  $t(19)=4.2$ ,  $p < .001$ ), tandis que la différence entre ces deux segments est de -3 Hz quand ils sont situés au milieu d'un groupe phonologique (ce qui n'est pas différent de 0,  $t(19) < 1$ ). Les mêmes mots ne sont donc pas prononcés avec le même contour intonatif, selon qu'ils sont situés au milieu ou en fin d'unité prosodique (-3 Hz versus + 30 Hz,  $t(19)=6.6$ ,  $p < .001$ ). Au niveau de la deuxième frontière de groupe phonologique, nous avons une tendance similaire à ce qui avait été trouvé avec la prosodie informative : nous obtenons une différence de 24 Hz en fin d'unité prosodique (significativement différente de 0,  $t(19)=3.2$ ,  $p=.005$ ), comparée à une différence de 13 Hz quand les segments sont situés de part et d'autre d'une frontière de groupe phonologique (non significativement différent de 0,  $t(19)=1.5$ ,  $p=0.15$ ). Mais ces valeurs ne sont pas significativement différentes l'une de l'autre (+24 Hz versus + 13 Hz,  $t(19)=1.6$ ,  $p=0.1$ ).

Ces analyses nous permettent donc de conclure que la même suite de mots n'est pas prononcée avec le même contour intonatif quand elle est placée en fin de groupe phonologique ou en milieu de groupe phonologique. De manière consistante, nous avons mis en évidence un contour ascendant en fin d'unité prosodique, comparé à une baisse ou une « stagnation » de la fréquence fondamentale à d'autres endroits. Ces différences sont beaucoup plus importantes dans la condition de prosodie informative que dans la condition de prosodie neutre (pour la première frontière :  $t(39)=2.2$ ,  $p < .03$  ; pour la deuxième frontière :  $t(39)=2.8$ ,  $p < .01$ ).

Les frontières de groupes phonologiques sont donc clairement marquées par des allongements et des variations d'intonation, ces indices prosodiques étant beaucoup plus

saillants quand les phrases sont prononcées avec une prosodie informative<sup>14</sup>. Quand les sujets entendaient le début des phrases ambiguës verbe et adjectif, ils étaient donc confrontés à des séquences similaires d'un point de vue phonémique, mais bien différentes du point de vue de leur structure prosodique. Nous avons donc voulu savoir (1) si les auditeurs français pouvaient utiliser ces différences prosodiques pour distinguer le début des phrases ambiguës et assigner la bonne catégorie syntaxique au mot cible ambigu, et (2) si la performance des sujets allait être facilitée avec de meilleurs indices prosodiques (vont-ils mieux traiter les phrases de la condition « prosodie informative » que les phrases de la condition « prosodie neutre » ?).

### 2.1.2.2 Résultats de la tâche de complétion de phrases

D'après les phrases inventées par les participants, nous avons pu évaluer l'interprétation qui avait été donnée au mot cible. Nous avons donc codé la réponse donnée par les sujets en « interprétation verbe » ou « interprétation adjectif ».

Dans certains cas, les sujets ont donné une réponse qui ne correspondait pas exactement au verbe ou à l'adjectif concerné. Quand leur réponse pouvait quand même être interprétée sans ambiguïté comme un verbe ou un adjectif, nous avons codé ce résultat : par exemple, une des phrases test était « les enfants salent... » et certains participants ont complété la phrase en mettant « les enfants salaient... » ; nous avons alors décidé que l'interprétation qu'ils avaient donnée au dernier mot entendu était bien le sens « verbe » et nous avons donc codé cette réponse comme une interprétation verbe. Ces réponses représentent 1,1% du nombre total de réponses, soit 8 réponses sur 720<sup>15</sup>.

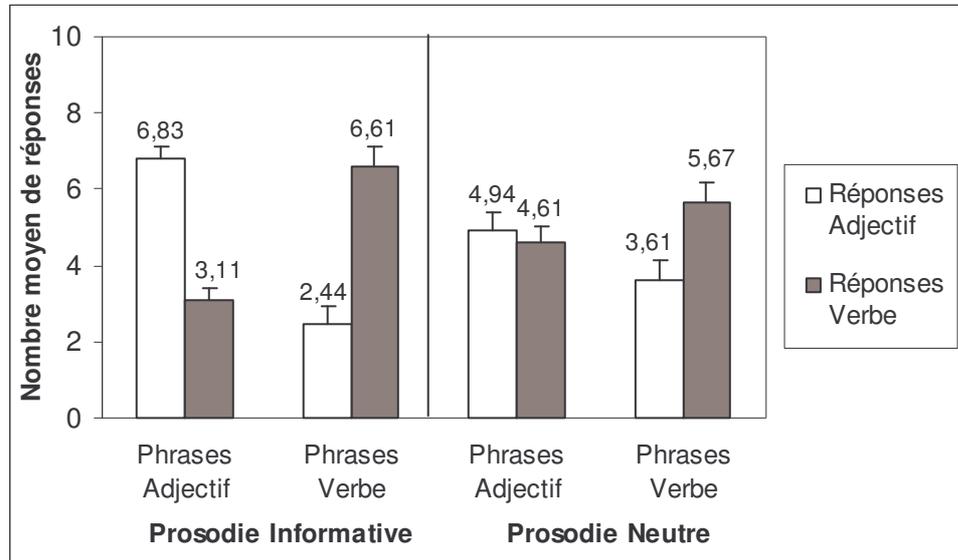
Dans d'autres cas, la réponse donnée par les sujets était trop éloignée du mot test et ne pouvait donc pas être codée en réponse verbe ou adjectif (par exemple, le début de phrase « la fumée dense / danse » a souvent été complétée en « la fumée dans ce lieu... »). Ces réponses ont été exclues des analyses : cela représente 39 réponses sur 720 soit 5,4% du nombre total de réponses.

---

<sup>14</sup> Des analyses similaires ont été menées sur l'énergie des voyelles mais aucune différence n'a été mise en évidence, ni entre les phrases ambiguës verbe et les phrases ambiguës adjectif de chaque condition prosodique, ni entre les deux conditions de prosodie.

<sup>15</sup> Une analyse conduite sans ces huit réponses a donné exactement les mêmes résultats.

La Figure 3 présente le nombre moyen de réponses adjectif et verbe pour les phrases ambiguës adjectif et verbe, dans les deux conditions de prosodie utilisée (voir l'annexe A.2.2 pour les résultats individuels par sujets, et l'annexe A.2.3 pour les résultats individuels par items).



**Figure 3 (résultats expérience 1 – étude 1) : Nombre moyen de réponses adjectif (en blanc) et verbe (en gris) données aux phrases adjectif et verbe dans les deux conditions prosodiques, informative et neutre (sur 10 réponses possibles pour chaque phrase). Les barres d'erreur représentent les erreurs standard de la moyenne.**

Le nombre moyen de réponses adjectif et verbe a été soumis à deux analyses de variance (ANOVA), l'une avec les participants et l'autre avec les items comme variable aléatoire. Dans l'analyse par sujet, nous avons déclaré deux facteurs intra-sujets : la Réponse donnée (adjectif versus verbe) et le type de Phrase entendue (phrase ambiguë adjectif versus phrase ambiguë verbe), ainsi que deux facteurs inter-sujets : la Prosodie (informative versus neutre) et le Contrebalancement (bloc 1 versus bloc 2). L'analyse par item incluait trois facteurs intra-items : la Prosodie, la Réponse donnée et le type de Phrase entendue.

Nous avons observé un effet principal de « type de Phrase » ( $F_1(1,34)=17.8$ ,  $p<.001$ ;  $F_2(1,19)=4.8$ ,  $p=.04$ ;  $\text{min}F'(1,29)=3.8$ ,  $p=.06$ ), mais aucun effet principal significatif de « Prosodie » ( $F_s<1$ ), ni de « Réponse donnée » ( $F_1(1,34)=2.3$ ,  $p=0.1$ ;  $F_2<1$ ).

L'interaction cruciale pour notre propos entre la « réponse donnée » et le « type de phrase entendue » est significative ( $F_1(1,34)=25.82, p<.001; F_2(1,19)=23.1, p<.001; \text{min}F'(1,47)=12.2, p=.001$ ). Elle reflète le fait que les participants ont donné plus de réponses verbe que de réponses adjectif quand ils entendaient les phrases ambiguës verbe ( $F_1(1,34)=20.5, p<.001, F_2(1,19)=5.7, p=0.03; \text{min}F'(1,30)=4.5, p=.04$ ), alors que le pattern de réponses inverse se produit pour les phrases ambiguës adjectif (plus de réponses adjectif que de réponses verbe,  $F_1(1,34)=13.8, p<.001, F_2(1,19)=1.6, p=0.2$ ).

La triple interaction entre la « réponse donnée », le « type de phrase entendue » et la « prosodie » est aussi significative ( $F_1(1,34)=7.4, p=.01; F_2(1,19)=10.3, p=.004; \text{min}F'(1,52)=4.3, p=.04$ ). Ceci reflète le fait que l'interaction mentionnée ci-dessus est plus forte dans la condition de prosodie informative ( $F_1(1,17)=34.1, p<0.001, F_2(1,19)=35.5, p<0.001$ ) que dans la condition de prosodie neutre où elle n'atteint pas le seuil de significativité ( $F_1(1,17)=2.5, p=0.13, F_2(1,19)=2.9, p=0.1$ ).

Aucune autre interaction significative n'est apparue dans ces analyses. Finalement, le facteur « contrebalancement » ne présente pas d'effet principal significatif et n'interagit pas avec les autres facteurs expérimentaux.

### 2.1.3 Discussion

Cette expérience indique que des locuteurs français ont été capables d'interpréter correctement (et différemment) deux débuts de phrases ambiguës qui ne différaient que par leur structure prosodique et syntaxique. En effet, des paires de phrases qui contenaient les mêmes phonèmes, mais qui avaient des structures syntaxiques différentes (et donc des structures prosodiques différentes), n'ont pas reçu la même analyse syntaxique de la part des participants. Ceux-ci ont assigné des catégories syntaxiques différentes aux mots ambigus, en fonction du contexte prosodique dans lequel ils apparaissaient : ils ont donné plus de réponses adjectif que de réponses verbe quand ils traitaient les phrases ambiguës adjectif, et plus de réponses verbe que de réponses adjectif quand ils entendaient les phrases ambiguës verbe. Ceci a été particulièrement fort dans la condition de prosodie informative.

L'interaction entre la prosodie et l'effet expérimental (interaction entre le type de phrase entendue et le type de réponse donnée) est très importante pour renforcer notre

interprétation que les sujets ont bien exploité la prosodie pour guider leurs réponses. En effet, supposons qu'il ait existé des différences artefactuelles dans nos paires de phrases : les homonymes adjectif – verbe auraient pu parfois être de faux homonymes. Par exemple, il serait possible que le /a/ de l'adjectif « lasse » (dans « [la jeune femme lasse]... ») soit différent du /a/ du verbe « lace » (dans « [la jeune femme] [lace...] »), de manière similaire à ce que l'on trouve dans la paire « pâte / patte ». On voit bien comment de telles différences artefactuelles auraient pu créer notre effet expérimental sans qu'il y ait eu exploitation de la prosodie. Le fait que les sujets échouent dans la condition de prosodie neutre nous prouve que notre matériel ne contient pas de tels artefacts : les phrases expérimentales sont bel et bien localement ambiguës, et seule la prosodie permet de les désambiguïser.

On peut donc conclure de cette première expérience que, quand les indices de frontières de groupes phonologiques sont bien marqués et saillants, les adultes français sont capables de les utiliser pour guider leur analyse syntaxique des phrases et résoudre des ambiguïtés syntaxiques locales.

A quel moment de l'analyse syntaxique les auditeurs utilisent-ils ces indices prosodiques ? Est-ce que cela se passe dès les premières étapes du traitement, ou au contraire sont-ils utilisés en dernier recours, comme un moyen de résoudre l'ambiguïté syntaxique quand celle-ci a été détectée par l'analyseur ? En effet, il est possible d'imaginer que, dès que l'analyseur syntaxique rencontre des indices de frontières de groupes phonologiques, ceux-ci soient directement utilisés pour activer plus fortement la structure syntaxique appropriée. D'un autre côté, au moment où l'analyseur syntaxique rencontre le mot ambigu (et en l'absence d'informations désambiguïsantes puisque la fin de la phrase a été coupée), il est possible qu'il soit alors confronté aux différentes interprétations possibles de ce mot ; il utiliserait alors les informations disponibles (dans notre cas, les indices prosodiques) pour favoriser une interprétation par rapport aux autres. Donc, la prosodie pourrait être (1) soit un indice utilisé très rapidement pour guider les premières étapes et les premiers choix de l'analyse syntaxique, (2) soit un indice utilisé plus tardivement pour permettre de faire un choix entre les différentes alternatives qui auraient été générées au point d'ambiguïté. Notre expérience n'imposait pas aux sujets de répondre en temps réel (expérience « off-line »), et ne nous permet donc pas de distinguer entre ces deux hypothèses. En effet, les participants n'avaient aucune limite de temps pour répondre, ils pouvaient réfléchir à la

phrase qu'ils avaient entendue et à la réponse qu'ils allaient donner, ils pouvaient aussi écouter les phrases test autant de fois qu'ils le désiraient avant de répondre... Pour étudier plus précisément le déroulement de l'analyse prosodique des phrases et son influence sur l'analyse syntaxique, nous avons donc décidé d'utiliser les mêmes phrases ambiguës dans une tâche réalisée en temps réel : une tâche de détection de mot.

## 2.2 EXPERIENCE 2 : DETECTION DE MOT

Nous avons réalisé cette expérience de détection de mot pour étudier la possibilité de l'utilisation en temps réel des frontières de groupes phonologiques pour réaliser une analyse syntaxique des phrases. Dans cette expérience, les participants devaient répondre soit au verbe, soit à l'adjectif, dans les phrases ambiguës présentées dans l'expérience précédente. Puisque ces mots étaient des homophones, nous avons demandé aux sujets de détecter une entrée lexicale abstraite. Ainsi, pour détecter les verbes, la cible était présentée visuellement dans la forme infinitive (par exemple « mordre »). Les participants étaient informés qu'ils devaient répondre à ce verbe, quelle que soit sa forme dans les phrases qu'ils allaient entendre (par exemple, détecter « mordre » impliquait de répondre pour « mordra », « mordons », « mordaient »...). Quant aux adjectifs, ils étaient présentés dans des phrases courtes, de structure « pronom – auxiliaire – adjectif » (par exemple « il est mort »). Cette expérience est donc originale puisqu'elle mettait en jeu des mots spécifiés dans leur forme abstraite, et qu'elle se différencie des expériences plus « classiques » de détection de mot où les sujets doivent réagir à un mot présenté dans une forme spécifique. Même si intuitivement cette tâche paraît difficile, il est important de noter qu'en pratique, les participants l'ont trouvée relativement simple à réaliser.

Dans cette expérience, les sujets étaient confrontés à quatre possibilités, qui sont résumées dans le tableau suivant (Tableau 3).

MOT CIBLE	PHRASE	REPONSE ATTENDUE	
Cible Verbe	Phrase Verbe	Oui	Versions
Cible Adjectif	Phrase Adjectif	Oui	Coïncidentes
Cible Verbe	Phrase Adjectif	Non	Versions Non-
Cible Adjectif	Phrase Verbe	Non	Coïncidentes

**Tableau 3 : Récapitulatif des quatre possibilités que rencontraient les participants de l'expérience 2, selon la cible qu'ils devaient détecter (verbe ou adjectif) et la phrase qu'ils entendaient (verbe ou adjectif).** Les sujets devaient répondre quand la cible et la phrase apparaissaient dans la même catégorie syntaxique (versions coïncidentes), tandis qu'ils ne devaient pas répondre quand la cible et la phrase étaient présentées dans des catégories syntaxiques différentes (versions non-coïncidentes).

Dans les deux premiers cas, la cible à détecter était réellement présente dans la phrase que les sujets entendaient, et ceux-ci devaient donc répondre. Ces situations étaient appelées « versions coïncidentes » (le mot cible présenté visuellement coïncide avec le mot présent dans les phrases présentées auditivement). Par contre, dans les deux dernières possibilités, quand les sujets devaient détecter un mot cible spécifique, ils entendaient ensuite des phrases contenant le « compétiteur syntaxique », c'est-à-dire l'homonyme de l'autre catégorie, et devaient donc s'abstenir de répondre. Ces situations étaient appelées « versions non-coïncidentes » (le mot cible ne coïncide pas avec le mot présent dans les phrases).

### 2.2.1 Méthode

#### 2.2.1.1 Matériel

Nous avons utilisé les mêmes phrases expérimentales que celles de l'expérience 1, enregistrées avec les deux prosodies (informative versus neutre). Cette fois-ci, les participants entendaient les phrases complètes (et non les phrases coupées après le mot ambigu). Pour chaque paire de phrases ambiguës, nous avons créé une paire de phrases contrôles non-ambiguës (une contenant la cible adjectif, l'autre contenant la cible verbe). Dans ces phrases contrôles, les mots cibles étaient les mêmes que dans les phrases ambiguës, mais ils apparaissaient dans des positions où une seule des deux formes syntaxiques était possible, et ce pour des raisons sémantiques ou syntaxiques (voir l'annexe A.1 pour le détail des stimuli utilisés). Par exemple, pour l'item « mord / mort », nous avons utilisé les quatre phrases suivantes :

### Phrases Verbe :

*Ambiguë* : [Le petit chien] [mord la laisse] [qui le retient]...

*Contrôle* : [Parfois] [on se mord la langue] [quand on mange] [trop vite].

### Phrases Adjectif :

*Ambiguë* : [Le petit chien mort] [sera enterré demain]...

*Contrôle* : [Maintenant qu'il est mort] [les batailles d'héritage] [vont commencer].

Pour chaque item, un sujet entendait les quatre phrases (les deux phrases ambiguës et les deux phrases contrôles) seulement une fois. Puisque chaque phrase était présentée en version coïncidente (cible appropriée) et en version non-coïncidente (cible non appropriée), chaque item apparaissait dans 8 situations différentes, sur lesquelles chaque sujet n'en traitait que 4 : ils entendaient les deux phrases ambiguës (phrase ambiguë verbe et phrase ambiguë adjectif), chacune présentée dans une version différente (une des phrases ambiguës est présentée en version coïncidente, l'autre en version non-coïncidente). Les deux phrases contrôles étaient présentées dans les autres versions (si un sujet avait la phrase ambiguë adjectif de l'item 1 en version coïncidente, il avait aussi la phrase ambiguë verbe de cet item mais en version non-coïncidente ; il entendait aussi la phrase contrôle adjectif de l'item 1 en version non-coïncidente, et la phrase contrôle verbe en version coïncidente). Donc pour chaque catégorie de phrases (ambiguë et contrôle) et pour chaque cible (adjectif et verbe), les sujets entendaient une version coïncidente et une version non-coïncidente. Nous avons donc mis en place deux groupes de sujets (un premier groupe qui entend les items du groupe 1 avec les phrases ambiguës verbe en version coïncidente, et les items du groupe 2 avec les phrases ambiguës verbes en version non-coïncidente, et un deuxième groupe qui entend les autres situations, pour qu'à la fin les 8 situations de chaque item aient été traitées).

De plus, les items présentés à un sujet étaient divisés en deux blocs, de manière à ce que chaque membre d'une paire de phrases apparaisse dans un bloc différent (par exemple, la phrase ambiguë adjectif et la phrase contrôle verbe d'un item apparaissent dans le premier bloc, tandis que la phrase ambiguë verbe de ce même item et sa phrase contrôle adjectif apparaissent dans le deuxième bloc). L'ordre de présentation des blocs a été contrebalancé entre les participants. Au total, les participants ont tous entendu 80 phrases expérimentales dont la moitié était des phrases ambiguës et l'autre moitié des

phrases contrôles, dont la moitié contenait des cibles adjectif et l'autre moitié des cibles verbe, et dont la moitié était présentée en version coïncidente et l'autre moitié en version non-coïncidente. Tous les items ont été présentés dans tous les cas possibles à travers tous les participants.

En plus de ces phrases expérimentales, nous avons mis en place 50 phrases distractrices demandant une réponse de la part des sujets. Dix de ces phrases contenaient une cible verbe, dix autres contenaient une cible adjectif, et trente contenaient une cible nom. Les cibles verbe et adjectif étaient des mots potentiellement ambigus (les mêmes que ceux utilisés dans les phrases expérimentales), mais ils étaient placés vers la fin des phrases en position non-ambiguë (par exemple, les participants devaient détecter le verbe « mordre » dans « ne t'approche pas trop de ce chien car il mord »). Les cibles nom étaient des mots bi- ou trisyllabiques non ambigus, placés à différentes positions dans les phrases.

Finalement, le matériel expérimental comportait aussi 30 phrases distractrices auxquelles les sujets ne devaient pas répondre. Les cibles à détecter dans ces phrases étaient des noms, des adjectifs et des verbes. Sur ces 30 phrases, dix ne présentaient aucun piège (puisque aucun mot dans les phrases ne ressemblait au mot cible à détecter) et les vingt autres étaient des phrases pièges qui contenaient un mot commençant par la même syllabe que le mot cible à détecter (ou par les mêmes phonèmes quand le mot cible à détecter était monosyllabique). Dans dix de ces vingt phrases pièges, le mot qui ressemblait au mot cible à détecter était en plus de la même catégorie syntaxique (par exemple, on demandait aux sujets de détecter « réfléchir » dans « cette jeune fille aimerait réussir à contrôler ses émotions » où les sujets doivent éviter de se faire piéger par le mot « réussir »). Pour les dix autres phrases, le mot ressemblant au mot cible à détecter était d'une autre catégorie syntaxique (par exemple, détecter « il est distrait » dans « les personnes crédules n'arrivent pas à distinguer le vrai du faux »).

Au total, chaque participant a entendu 160 phrases (80 phrases expérimentales et 80 phrases distractrices), ainsi que 15 phrases d'entraînement. De plus, il faut noter que chaque sujet était assigné à une condition de prosodie, de sorte qu'il n'entendait que des stimuli enregistrés avec une prosodie informative ou que des phrases enregistrées avec une prosodie neutre.

Les phrases ambiguës expérimentales utilisées dans cette expérience étaient celles qui avaient été enregistrées pour la première expérience. Les phrases distractrices

et les phrases d'entraînement ont été enregistrées par la même locutrice que celle de la première expérience (sans véritable consigne au niveau de la prosodie, hormis celle de prononcer ces phrases le plus distinctement possible et de manière vivante et agréable à écouter). Les enregistrements ont eu lieu dans une pièce insonorisée, et toutes les phrases ont été digitalisées à 16 KHz et 16 bits sur une table de mixage OROSAU22. Elles ont ensuite été éditées grâce aux programmes Manitou et Praat. Le début des mots cibles a été marqué manuellement avec le logiciel Manitou.

#### 2.2.1.2 Procédure

L'expérience a été réalisée en utilisant le programme Expe6 de présentation des stimuli (Pallier et al., 1997)<sup>16</sup>. Un essai commençait par la présentation visuelle du mot cible, au centre de l'écran, pendant 1,5 secondes. Le mot cible était présenté avec une indication sur sa catégorie syntaxique (par exemple, « mordre » pour détecter le verbe, et « il est mort » pour détecter l'adjectif). La cible écrite était ensuite effacée, et l'écran restait noir pendant une seconde avant que la phrase expérimentale soit jouée (les stimuli auditifs étaient présentés à travers une carte son ProAudioSpectrum 16-bit). L'essai se terminait 2,5 secondes après la réponse du sujet ou à la fin de la présentation auditive de la phrase (en effet, les participants devaient soit appuyer sur un bouton réponse, placé devant eux, quand ils entendaient le mot cible dans la phrase, soit ne rien faire et attendre l'essai suivant quand le mot cible n'était pas présent dans la phrase). L'essai suivant commençait immédiatement après. Les temps de réponse des sujets étaient mesurés depuis le début du mot cible. Les instructions indiquaient clairement aux participants qu'ils devaient répondre le plus rapidement possible, mais si possible sans faire d'erreurs.

Durant la phase d'entraînement, on informait les participants de la justesse de leurs réponses (correct ou incorrect), et quand un temps de réaction était enregistré (réponse correcte à la cible présente dans la phrase), celui-ci était présenté sur l'écran à la fin de l'essai. S'il était supérieur à 1 seconde, l'indication « trop lent » apparaissait aussi, pour que les sujets essaient d'être encore plus rapides sur les essais suivants. Par contre, lors de la phase de test, aucune indication n'était donnée aux participants sur leurs réponses.

---

<sup>16</sup> Voir le site internet : <http://www.lscp.net/expe>

L'expérience durait en moyenne 25 minutes (avec une pause entre les deux blocs de stimuli). Tous les participants étaient testés individuellement, dans une pièce isolée, équipée d'un ordinateur, d'un casque et d'un boîtier réponse. Avant de commencer, chaque participant remplissait un formulaire indiquant sa date de naissance, un éventuel bilinguisme, et s'il avait des problèmes auditifs ou des problèmes de langage. Puis, il découvrait les consignes de l'expérience : les instructions étaient données oralement, en complément des instructions écrites à l'écran (voir Annexe A.3.1 pour les consignes écrites détaillées). L'expérimentateur s'assurait que le sujet avait bien compris la tâche en restant près de lui pendant la phase d'entraînement, ce qui permettait aussi de clarifier toutes les questions éventuelles avant de commencer le test.

### 2.2.1.3 Participants

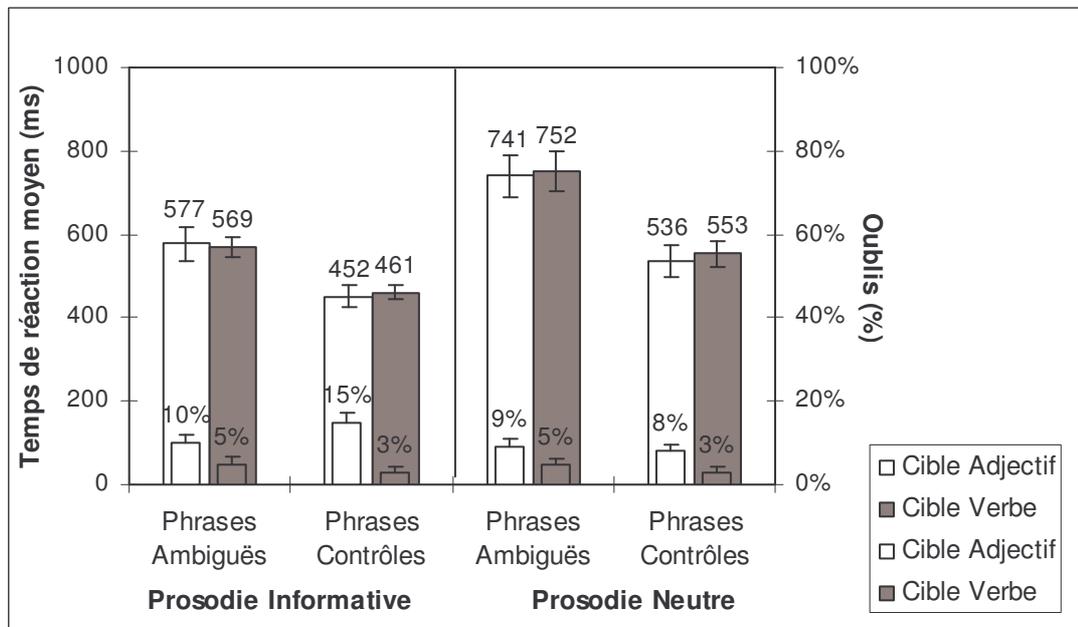
Cinquante-six personnes, de langue maternelle française, ont pris part à cette expérience, vingt-huit dans chaque condition de prosodie (prosodie informative versus prosodie neutre). Aucune d'elles n'a rapporté de problèmes auditifs, ni de problèmes liés au langage. Avant de commencer l'expérience, les participants ont été assignés aléatoirement à l'un des quatre groupes expérimentaux.

Six personnes supplémentaires ont été testées mais leurs résultats n'ont pas été inclus dans les analyses pour les raisons suivantes : non-natives du français (2 participants), problèmes techniques (1 participant) et plus de 70% d'erreurs sur toute l'expérience (3 participants).

## 2.2.2 Résultats

### 2.2.2.1 Analyses des temps de réaction et des oublis

Les temps de réaction moyen et les pourcentage d'oublis (« misses ») sur les phrases ambiguës et contrôles sont représentés dans la Figure 4 (voir les annexes A.3.2.1 et A.3.3.1 pour les résultats individuels par sujets et par items).



**Figure 4 (résultats expérience 2 – étude 1) : Moyenne des temps de réaction et pourcentage d’oublis pour chaque type de phrase expérimentale (phrases ambiguës versus phrases contrôles avec cibles adjectif versus cibles verbe) et dans chaque condition prosodique (prosodie informative versus prosodie neutre). Les barres d’erreur représentent les erreurs standard de la moyenne.**

Nous avons conduit deux analyses de variance (ANOVA) sur les temps de réaction et les oublis, l’une avec les participants, l’autre avec les items comme variable aléatoire. L’analyse par sujet déclare 3 facteurs inter-sujets : la Prosodie (informative versus neutre), ainsi que deux facteurs de contrebalancement : le Contrebalancement des items (groupes de sujets 1 et 2 versus groupes de sujets 3 et 4) et l’Ordre de présentation des stimuli (bloc 1 en premier versus bloc 2 en premier). Cette analyse déclare aussi 2 facteurs intra-sujets : l’Ambiguïté (phrases ambiguës versus phrases contrôles) et le type de Cible (adjectif versus verbe). Quant à l’analyse par item, elle inclut 3 facteurs intra-items : la Prosodie, l’Ambiguïté et le type de Cible.

L’analyse des temps de réaction montre un effet principal d’Ambiguïté ( $F_1(1,48)=88.3, p<.001; F_2(1,19)=17.5, p<.001$ ), ainsi qu’un effet principal de Prosodie ( $F_1(1,48)=9.9, p=.003; F_2(1,19)=40.9, p<.001; \text{min}F'(1,57)=5.2, p=.03$ ). Par contre le type de Cible ne présente pas d’effet principal significatif ( $F_s<1$ ). L’interaction entre Ambiguïté et Prosodie est significative ( $F_1(1,48)=6.3, p=.02; F_2(1,19)=7.7, p=.01; \text{min}F'(1,59)=3.5, p=.07$ ), indiquant que l’effet d’ambiguïté est plus fort dans la condition de prosodie neutre (202 ms,  $F_1(1,24)=63.1, p<0.001, F_2(1,19)=22.0, p<0.001$ )

que dans la condition de prosodie informative (117 ms,  $F_1(1,24)=27$ ,  $p<0.001$ ,  $F_2(1,19)=5.03$ ,  $p=0.04$ ). Aucune autre interaction significative n'est ressortie de cette analyse. De plus les facteurs de Contrebalancement et d'Ordre n'ont pas présenté d'effet principal et n'ont pas interagi de manière significative avec les autres facteurs expérimentaux.

Concernant l'analyse des oublis, nous n'avons pas trouvé d'effet principal d'Ambiguïté ( $F_s<1$ ), ni de Prosodie ( $F_1(1,48)=1.3$ ,  $p=0.3$ ;  $F_2(1,19)=3.0$ ,  $p=0.1$ ). Par contre, l'analyse a révélé un effet principal significatif du type de Cible ( $F_1(1,48)=25.9$ ,  $p<.001$ ;  $F_2(1,19)=8.0$ ,  $p=.01$ ). Il y a plus d'oublis sur les phrases adjectif que sur les phrases verbe (10,5% vs. 4%). Cet effet sur les phrases adjectif pourrait être dû à la structure que nous avons utilisée pour présenter la cible adjectif à détecter : en effet, l'adjectif était utilisé dans sa forme attribut (« il est mort ») alors que pendant le test, l'adjectif était présenté sous une forme d'adjectif qualificatif épithète (« le petit chien mort »). Ce changement dans la structure syntaxique utilisée a pu entraîner ce taux plus important d'oublis de réponse.

L'interaction entre les facteurs Cible et Ambiguïté est marginalement significative dans l'analyse par sujet uniquement ( $F_1(1,48)=3.9$ ,  $p=.055$ ;  $F_2(1,19)<1$ ). L'effet mentionné ci-dessus est en effet particulièrement vrai pour les phrases contrôles (11,5% d'oublis pour les phrases contrôles adjectif). Ceci semble dû à quelques items déviants dans la condition contrôle : par exemple, nous avons trouvé que la phrase contrôle adjectif associée à l'item 9 a généré 71,4% d'oublis, alors que cette phrase n'était pas ambiguë et qu'elle était prononcée avec une prosodie informative. Dans cette phrase, l'adjectif « sale » à détecter était utilisé dans l'expression « faire le sale boulot ». Il se pourrait que les participants aient traité cette expression comme une seule unité, ce qui aurait ralenti la reconnaissance de l'adjectif « sale ». Une autre explication pourrait reposer sur l'antéposition de l'adjectif dans la phrase (adjectif placé avant le nom auquel il se rapporte), qui n'est pas l'utilisation la plus courante pour un adjectif. De manière plus générale, une analyse plus détaillée des résultats par item a mis en évidence 5 items déviants (qui obtenaient plus de 40% d'oublis), et trois d'entre eux étaient dans des phrases contrôles adjectif<sup>17</sup>.

---

<sup>17</sup> Nous avons réalisé des analyses sans ces items déviants : cela n'a pas changé les principaux résultats que nous avons trouvés dans l'analyse des temps de réaction. Dans l'analyse des oublis, l'effet principal de type de cible s'est maintenu, mais tous les autres effets principaux ou interactions ont disparu.

Concernant les autres facteurs expérimentaux, l'interaction entre le type de cible et la prosodie n'est pas significative ( $F_1(1,48)=2.56$ ,  $p=0.12$ ;  $F_2(1,19)=2.2$ ,  $p=0.2$ ), de même que l'interaction entre le facteur ambiguïté et le facteur prosodie ( $F_1(1,48)=2.4$ ,  $p=0.13$ ;  $F_2(1,19)=1.3$ ,  $p=0.3$ ). De la même manière, la triple interaction entre ambiguïté, type de cible et prosodie n'est pas significative ( $F_1(1,48)=3.2$ ,  $p=.08$ ;  $F_2(1,19)=1.8$ ,  $p=0.2$ ). Quant au facteur ordre, il ne présente aucun effet principal et n'interagit pas avec les autres facteurs expérimentaux ( $F_s < 1$ ). Par contre, nous avons trouvé un effet principal du facteur contrebalancement ( $F_1(1,48)=5.3$ ,  $p=.03$ ) qui interagit uniquement avec le facteur ambiguïté ( $F_1(1,48)=3.3$ ,  $p=.03$ ;  $F_s < 1$  pour les autres interactions). Ce résultat semble être dû aux items déviants mentionnés précédemment, puisque les 3 phrases contrôles adjectif qui ont obtenu des taux d'oublis importants appartenaient toutes aux items du groupe 1.

Cette expérience a été réalisée pour étudier l'utilisation en temps réel des indices de frontières de groupes phonologiques pour guider l'analyse syntaxique des phrases et résoudre des ambiguïtés syntaxiques temporaires. Si les frontières de groupes phonologiques étaient (1) des indices parfaitement fiables pour indiquer les frontières des constituants syntaxiques, et (2) des indices utilisés par l'analyseur syntaxique, elles devraient alors permettre aux auditeurs de résoudre directement l'ambiguïté syntaxique étudiée. Si cela était vrai, les auditeurs ne devraient pas percevoir l'ambiguïté, ou ne devraient pas être gênés par elle, et devraient alors répondre aussi vite aux phrases ambiguës qu'aux phrases contrôles non-ambiguës. Or nos résultats ont montré qu'il existait un effet d'ambiguïté : les participants mettent plus de temps pour répondre aux phrases ambiguës qu'aux phrases contrôles. Ceci indique que les indices de frontières prosodiques ne désambiguïsent pas totalement les phrases ambiguës, même dans la condition de prosodie informative où les indices sont particulièrement forts et saillants.

Se pourrait-il que les indices prosodiques ne soient pas du tout utilisés par les auditeurs (ou pas du tout pertinents pour eux) ? Puisque les phrases n'étaient pas coupées après le mot ambigu (comme dans l'expérience de complétion de phrases), les participants ont pu attendre les mots suivants pour les aider à réaliser la tâche que nous leur demandions. Si cela était le cas, nous devrions observer des résultats identiques dans les deux conditions prosodiques (les participants n'auraient pas prêté attention au

fait que la prosodie était plutôt neutre ou particulièrement informative, et auraient utilisé l'information lexicale désambiguïsante faisant suite au mot ambigu). En fait, l'analyse des temps de réaction a montré une interaction significative entre les facteurs ambiguïté et prosodie. Ce résultat reflète le fait que les participants ont mieux réalisé la tâche quand les indices prosodiques étaient plus marqués (temps de réaction plus courts avec la prosodie informative). Les auditeurs ont donc utilisé les indices de frontières de groupes phonologiques pour réaliser cette expérience.

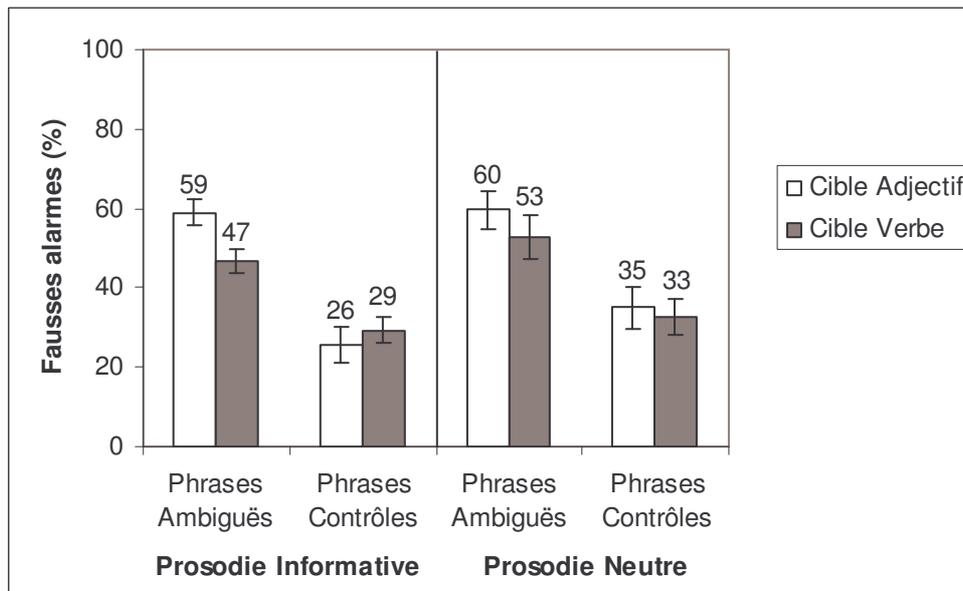
Une autre manière de vérifier que les indices de frontières prosodiques peuvent être utiles aux sujets pour réaliser l'analyse syntaxique de ces phrases ambiguës, est de mesurer le taux de fausses alarmes obtenues pour chacune de ces phrases présentées en version non-coïncidente. En effet, nous avons mis en évidence le fait qu'il y avait une asymétrie entre les phrases ambiguës adjectif et les phrases ambiguës verbe au niveau de la position de la frontière de groupe phonologique. Dans les phrases verbe, cette frontière est placée avant le mot ambigu et les sujets ont donc accès à plus d'informations désambiguïsantes au moment de traiter le mot ambigu (comparé aux phrases adjectif dans lesquelles la frontière prosodique est placée après le mot ambigu). Si les sujets peuvent utiliser le fait que les frontières de groupes phonologiques correspondent à des frontières syntaxiques, alors l'effet d'ambiguïté devrait être moins fort sur les phrases verbe que sur les phrases adjectif : il devrait être plus facile de rejeter l'adjectif « mort » en entendant la phrase verbe « [le petit chien] [mord...] », que de rejeter le verbe « mord » en entendant « [le petit chien mort]... » (en effet, rappelons que l'absence de frontières de groupe phonologique ne signifie pas l'absence de frontières syntaxiques, tandis que la présence d'une frontière de groupe phonologique signale la présence d'une frontière syntaxique).

#### 2.2.2.2 Analyses des fausses alarmes

Les pourcentages de fausses alarmes ont été mesurés sur les phrases pièges (phrases distractrices auxquelles les sujets ne devaient pas répondre) et sur les phrases expérimentales présentées en version non-coïncidente.

Trois sortes de phrases pièges ont été utilisées : certaines ne contenaient aucun mot ressemblant au mot cible à détecter et étaient donc les plus faciles à rejeter (0,9% de fausses alarmes) ; d'autres contenaient un mot commençant par la même syllabe (ou par les mêmes phonèmes) que le mot cible à détecter, et ces mots pièges pouvaient ou non appartenir à la même catégorie que le mot cible. Les participants ont trouvé plus difficile de rejeter les mots pièges quand ils étaient de la même catégorie syntaxique que le mot cible (20,9% de fausses alarmes), et plus facile de rejeter les autres mots pièges (5,7% de fausses alarmes). Ce premier résultat suggère que les auditeurs construisaient en temps réel des attentes quant à la catégorie syntaxique des mots qu'ils allaient entendre, et qu'ils utilisaient ces attentes pour répondre à la tâche demandée. Donc même pour des mots non-ambigus, les participants ont trouvé assez difficile de se retenir de répondre à un mot de la même catégorie syntaxique qui commençait par les phonèmes adéquats, et se sont fait piéger assez fréquemment.

Pour analyser les fausses alarmes faites sur les phrases expérimentales, nous avons conduit deux analyses de variance sur le pourcentage moyen de fausses alarmes, avec les participants et les items comme variables aléatoires, et avec les mêmes facteurs expérimentaux que ceux décrits dans l'analyse précédente. Le pourcentage de fausses alarmes obtenues aux phrases expérimentales est représenté dans la Figure 5 (voir les annexes A.3.2.1 et A.3.3.1 pour les analyses détaillées par sujets et par items).



**Figure 5 (résultats expérience 2 – étude 1) : Pourcentage moyen de fausses alarmes obtenues pour chaque type de phrase expérimentale (phrases ambiguës versus phrases contrôles avec cibles adjectif versus cibles verbe) et dans chaque condition prosodique (prosodie informative versus prosodie neutre). Les barres d’erreur représentent les erreurs standard de la moyenne.**

Nos analyses ont mis en évidence un effet principal d’ambiguïté ( $F_1(1,48)=218$ ,  $p<.001$ ;  $F_2(1,19)=93.5$ ,  $p<.001$ ), avec plus de fausses alarmes sur les phrases ambiguës que sur les phrases contrôles (54.6% versus 30.7%), ce qui montre à nouveau que les indices prosodiques ne sont pas des indices complètement désambiguïsants. L’effet principal du facteur prosodie est significatif mais seulement dans l’analyse par item ( $F_1<1$ ;  $F_2(1,19)=7$ ,  $p=.02$ ) et indique une tendance pour les sujets à faire plus de fausses alarmes dans la condition de prosodie neutre que dans la condition de prosodie informative (45% versus 40%). Nous avons aussi trouvé un effet principal du type de cible, uniquement significatif dans l’analyse par sujet ( $F_1(1,48)=5.6$ ,  $p=.02$ ;  $F_2(1,19)=1.2$ ,  $p=0.3$ ), et une interaction significative entre les facteurs ambiguïté et type de cible ( $F_1(1,48)=5.6$ ,  $p=.02$ ;  $F_2(1,19)=2.2$ ,  $p=0.2$ ). Les sujets ont eu tendance à faire plus de fausses alarmes sur les phrases adjectif que sur les phrases verbe (45% versus 40%), et l’effet d’ambiguïté a été plus important sur les phrases adjectif que sur les phrases verbe (29% versus 19%). Il semble donc que les phrases verbe aient été un peu plus faciles à traiter que les phrases adjectif. Comme nous l’avions supposé, cet effet doit être dû au fait que, dans les phrases verbe, la frontière de groupe phonologique était située avant le mot cible ambigu ; les participants ont donc eu accès à de l’information prosodique pertinente avant de traiter le mot ambigu. Au contraire, dans les phrases adjectif, la frontière de groupe phonologique était située après le mot ambigu (ce qui

pourrait expliquer les moins bons résultats obtenus sur ces phrases puisque les sujets n'avaient pas accès à l'information prosodique désambiguïsante au moment de traiter le mot ambigu).

Jusqu'à présent, les analyses sur les temps de réaction et les analyses sur les taux de fausses alarmes révèlent que les indices de frontières de groupes phonologiques ne sont pas des indices totalement désambiguïsants. Mais ces analyses indiquent aussi que ces indices prosodiques peuvent être utiles pour les auditeurs : ils réussissent mieux la tâche demandée quand les indices sont bien marqués (diminution significative de l'effet d'ambiguïté avec la prosodie informative dans l'analyse des temps de réaction, et tendance à faire moins de fausses alarmes dans la condition de prosodie informative).

Pour évaluer plus précisément le rôle de la prosodie dans les étapes précoces du traitement syntaxique, nous avons analysé les réponses qui avaient été données avant que les participants n'aient accès à l'information lexicale désambiguïsante.

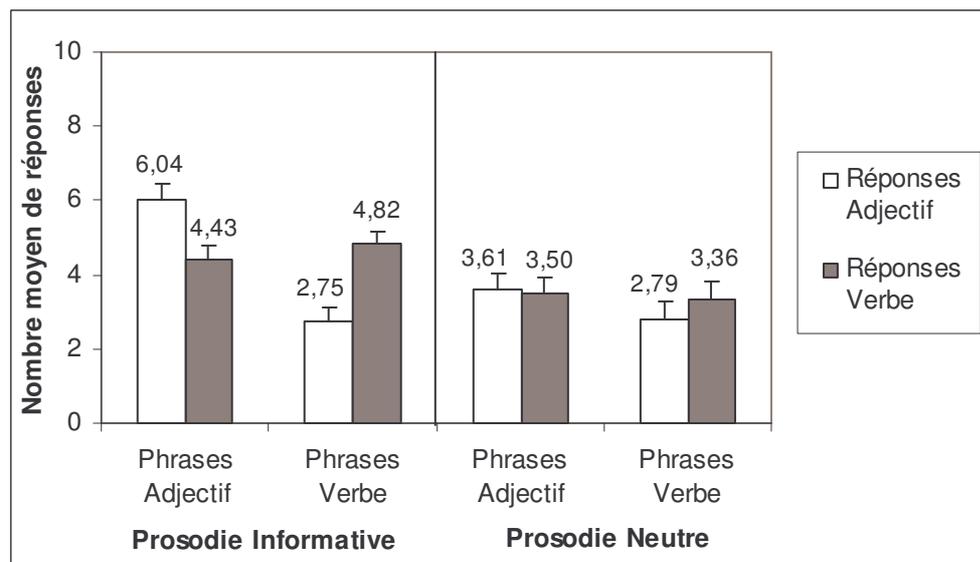
### 2.2.2.3 Analyses des réponses rapides

Cette analyse est restreinte aux réponses données par les sujets avant qu'ils aient eu accès à l'information lexicale désambiguïsante, c'est-à-dire aux réponses données entre le début du mot ambigu et jusqu'à 150 ms après la fin du mot ambigu (nous avons utilisé ce délai de 150 ms car cela correspond au temps de réaction moteur moyen). A ce point de la phrase, les sujets étaient dans une situation similaire à celle de la première expérience : quand ils prenaient leur décision et amorçaient leur réponse, la seule différence entre les débuts de phrases ambiguës était une différence de structure prosodique (et de structure syntaxique).

Comme dans l'expérience de complétion de phrases, nous avons mesuré et analysé le nombre moyen d'interprétations verbe et adjectif données au mot ambigu. Pour ce faire, chaque réponse donnée par les sujets pour une détection de cible adjectif a été codée comme une réponse adjectif (que cette réponse soit exacte ou fausse, donc que cela soit un succès ou « hit » en version coïncidente, ou que cela soit une fausse alarme en version non-coïncidente). De même, quand la cible à détecter était un verbe et que

les sujets appuyaient sur le bouton réponse (quelle que soit la phrase présentée), nous avons codé leur réponse comme une interprétation verbe<sup>18</sup>.

Trente-neuf pour cent des réponses ont été données par les participants dans cet intervalle de temps. Le nombre moyen de réponses verbe et adjectif, données aux phrases ambiguës et dans chaque condition prosodique, est représenté dans la Figure 6 (voir les annexes A.3.2.2 et A.3.3.2 pour les analyses détaillées par sujets et par items).



**Figure 6 (résultats expérience 2 – étude 1) : Nombre moyen de réponses adjectif et verbe données aux phrases ambiguës adjectif et aux phrases ambiguës verbe dans les deux conditions prosodiques, informative et neutre (sur 10 réponses possibles).** Seules les réponses rapides sont prises en compte (réponses données jusqu'à 150 ms avant la fin du mot ambigu). Les barres d'erreur représentent les erreurs standard de la moyenne.

Deux analyses de variance ont été conduites sur ces données, l'une avec les participants, l'autre avec les items comme variable aléatoire. L'analyse par sujet déclarait 3 facteurs inter-sujets : la Prosodie (informative versus neutre), le Contrebalancement des items et le facteur Ordre (bloc 1 en premier versus bloc 2 en premier). Elle déclarait aussi 2 facteurs intra-sujets : le type de Réponse (réponses verbe versus réponses adjectif) et le type de Phrase (phrases ambiguës verbe versus phrases

<sup>18</sup> Dans cet intervalle de temps, nous n'avons pas pu coder les oublis et les rejets corrects. Comme la tâche était une tâche « go – no go » (répondre quand la cible est présente – ne pas répondre quand la cible n'est pas présente), ces réponses n'obtenaient pas de temps de réaction : il nous était donc impossible de savoir si ces décisions avaient été prises avant la fin du mot ambigu ou bien plus tard dans la phrase.

ambiguës adjectif). Quant à l'analyse par item, elle incluait 3 facteurs intra-items : la Prosodie, le type de Réponse et le type de Phrase.

L'analyse a montré un effet principal du type de Phrase ( $F_1(1,48)=40.7, p<.001$ ;  $F_2(1,19)=13.1, p=.002$ ) : les sujets ont donné plus de réponses aux phrases adjectif qu'aux phrases verbe (moyenne de 4.4 réponses pour les phrases adjectif versus 3.4 pour les phrases verbe). Ceci pourrait être expliqué par le fait que les mots ambigus étaient plus longs quand ils étaient prononcés dans les phrases adjectif, de par leur position finale dans le groupe phonologique (on a une durée moyenne des mots ambigus de 410 ms dans les phrases adjectif versus 346 ms dans les phrases verbe, cette différence de 64 ms étant significative, avec  $t(19)=6.9, p<.001$ ). Donc dans l'intervalle de temps considéré (jusqu'à 150 ms après la fin du mot ambigu), les participants ont eu plus de temps pour répondre aux phrases adjectif que pour répondre aux phrases verbe.

L'analyse a aussi mis en évidence un effet principal de la prosodie ( $F_1(1,48)=5.9, p=.02$ ;  $F_2(1,19)=34.0, p<.001$ ;  $\text{min}F'(1,64)=5.0, p=.03$ ), qui illustre le fait que les sujets ont donné plus de réponses dans la condition de prosodie informative que dans la condition de prosodie neutre (moyenne de 4.5 réponses pour la prosodie informative versus 3.3 pour la prosodie neutre). Cet effet est encore probablement lié à la durée des mots ambigus : ceux-ci étaient plus longs quand ils étaient prononcés avec une prosodie informative que quand ils étaient prononcés avec une prosodie neutre (les durées moyennes respectives sont de 404 ms versus 351 ms,  $t(19)=5.9, p<.001$ ). Dans l'intervalle de temps considéré, les sujets ont donc eu plus de temps pour répondre quand ils étaient assignés à la condition de prosodie informative que quand ils traitaient les phrases prononcées avec une prosodie neutre. De plus, les sujets obtiennent des temps de réaction plus lents dans la condition de prosodie neutre car ils sont moins sûrs de leurs réponses et plus hésitants quant à la décision qu'ils doivent prendre.

Le type de réponse donnée par les participants ne présente pas d'effet principal significatif ( $F_1(1,48)=1.3, p=0.3$ ;  $F_2<1$ ) et n'interagit pas avec la prosodie ( $F_s<1$ ). Par contre, l'interaction entre le type de réponse et le type de phrase est significative ( $F_1(1,48)=39.5, p<.001$ ;  $F_2(1,19)=29.2, p<.001$ ;  $\text{min}F'(1,47)=16.8, p<.001$ ), et illustre le fait que les phrases verbe tendent à recevoir plus de réponses verbe que de réponses adjectif, et vice-versa pour les phrases adjectif. La triple interaction entre le type de réponse, le type de phrase et la prosodie est aussi significative ( $F_1(1,48)=18.7, p<.001$ ;

$F_2(1,19)=39.2, p<.001$ ) : cela reflète le fait que l'interaction mentionnée ci-dessus est significative dans la condition de prosodie informative ( $F_1(1,24)=42.7, p<.001$ ;  $F_2(1,19)=48.8, p<.001$ ), alors qu'elle ne l'est pas dans la condition de prosodie neutre ( $F_1(1,24)=2.81, p=0.1$ ;  $F_2(1,19)=2.84, p=0.1$ ).

Finalement, les facteurs contrebalancement et ordre ne présentent pas d'effets principaux significatifs et n'interagissent pas avec les autres facteurs expérimentaux.

### 2.2.3 Discussion

L'analyse faite sur les temps de réaction nous avait indiqué que les indices de frontières de groupes phonologiques ne permettaient pas de désambiguïser complètement les phrases expérimentales : les temps de réponse étaient plus longs pour les phrases ambiguës que pour les phrases contrôles. Cependant, la taille de l'effet d'ambiguïté se trouve modulée en fonction de la qualité des indices prosodiques : l'effet d'ambiguïté a en effet été réduit quand les stimuli ont été enregistrés avec une prosodie bien marquée (prosodie informative). Donc, même si les indices prosodiques ne sont pas complètement désambiguïsants, ils sont quand même clairement exploités par les auditeurs.

L'analyse des réponses données avant l'accès à l'information lexicale désambiguïsante confirme ce résultat. En effet, quand les sujets donnaient leur réponse, ils étaient dans la même situation que pour l'expérience de complétion de phrases (expérience 1) : seule la prosodie permettait de distinguer les deux débuts de phrases ambiguës. Cette analyse a révélé que, quand les indices prosodiques étaient particulièrement saillants, les auditeurs étaient capables de les utiliser en temps réel pour distinguer les deux alternatives possibles : ils ont réussi à assigner la bonne catégorie syntaxique au mot ambigu (plus souvent que l'attribution de la mauvaise catégorie syntaxique, et à un taux de réussite significativement supérieur au hasard), et ce avant même de pouvoir utiliser l'information lexicale désambiguïsante. Ce résultat n'a pas été répliqué dans la condition de prosodie neutre : quand la prosodie a délibérément été rendue neutre, les frontières de groupes phonologiques sont devenues trop peu saillantes pour que les auditeurs puissent les utiliser pour résoudre l'ambiguïté

syntactique étudiée et décider en temps réel si le mot ambigu était un verbe ou un adjectif.

Dans cette expérience de détection de mot abstrait, nous avons montré que plus la prosodie est informative, plus les participants réussissent la tâche qui leur est demandée. Même si la prosodie ne permet pas de désambiguïser complètement les phrases ambiguës, les frontières de groupes phonologiques sont tout de même utilisées en temps réel pour guider l'analyse syntaxique de phrases temporairement ambiguës.

## 2.3 DISCUSSION GENERALE

### *2.3.1 Résultat général*

Les expériences décrites dans ce chapitre suggèrent que les adultes français exploitent les frontières de groupes phonologiques pour réaliser une analyse syntaxique des phrases et résoudre des ambiguïtés syntaxiques temporaires. Dans deux tâches expérimentales différentes (une tâche « off-line » de complétion de phrases et une tâche « on-line » de détection de mot abstrait), les auditeurs français ont été capables de différencier deux débuts de phrases, qui étaient identiques d'un point de vue phonémique, et qui ne se distinguaient que par leurs structures syntaxique et prosodique. Dans certaines situations, ils ont pu donner significativement plus de réponses adjectif que de réponses verbe à un mot ambigu qui était effectivement placé dans une phrase adjectif (et vice-versa pour un mot ambigu placé dans une phrase verbe). Ceci a été particulièrement mis en évidence dans la condition de prosodie informative. Donc, plus les indices prosodiques sont saillants et bien marqués, plus les auditeurs réussissent l'analyse syntaxique des phrases.

### *2.3.2 Modèles de traitement syntaxique : quels indices sont utilisés ?*

Nous ne pouvons décider, sur la base de ces résultats, si l'interprétation donnée par les participants était la seule interprétation activée quand ils ont donné leur réponse

(ce qui serait prédit par les modèles en série que nous avons décrit dans le premier chapitre), ou si elle était l'interprétation la plus activée parmi les différentes alternatives possibles (ce qui serait prédit par les partisans des modèles en parallèle). Par contre, nous pouvons conclure que la prosodie est utilisée en temps réel et qu'elle a un impact sur les étapes précoces de l'analyse syntaxique. De plus, nous pouvons dire que, dans le cas où une seule interprétation est générée par l'analyseur syntaxique, celle-ci n'est alors pas générée sur la seule base de principes syntaxiques. En effet, si ces principes étaient les seules informations utilisées pour attribuer une structure syntaxique aux phrases que nous entendons, alors nous n'aurions pas dû voir un effet de la prosodie émerger dans nos résultats. Par exemple, si l'on suit l'hypothèse de « clôture tardive » (« late closure », Frazier & Fodor, 1978), les participants auraient dû préférer l'interprétation adjectif pour le mot ambigu (puisque l'adjectif restait inclus dans le groupe nominal en train d'être traité). Donc, dans ce cas, on se serait attendu à ce que les auditeurs rencontrent plus de difficultés pour traiter les phrases ambiguës verbe (puisque la catégorie syntaxique du mot ambigu aurait été la catégorie « non-préférée » des auditeurs) par rapport au traitement des phrases ambiguës adjectif. Ils auraient aussi dû donner plus de réponses adjectif que de réponses verbe dans les situations où ils n'avaient pas accès à l'information lexicale désambiguïsante (c'est-à-dire dans l'expérience de complétion de phrases et dans l'analyse des réponses rapides obtenues dans l'expérience de détection de mot abstrait). Et finalement, nous aurions dû trouver des résultats similaires quelle que soit la condition de prosodie utilisée (informative versus neutre). Contrairement à toutes ces prédictions, nous avons mis en évidence le fait que les réponses des participants étaient influencées par la structure prosodique des phrases.

Notre propos n'est pas de dire que les principes syntaxiques ne sont jamais utilisés, ni que seule la prosodie est utilisée pour réaliser une analyse syntaxique des phrases. Notre approche consiste à dire que différentes sortes d'informations peuvent interagir pour arriver à une analyse syntaxique correcte des phrases que nous entendons. Par exemple, il se pourrait que d'autres indices aient parfois guidé l'interprétation des sujets, comme par exemple la fréquence relative des items utilisés (si la forme verbale d'un item est plus fréquente que sa forme adjectif, il se pourrait que les sujets aient plus tendance à donner une interprétation verbe au mot ambigu...), ainsi que la plausibilité des mots utilisés dans les phrases (étant donné le début de la phrase, il se pourrait

qu'une des deux versions du mot ambigu soit plus plausible et donc choisie de manière préférentielle).

Les bases de données actuelles, comme la base Lexique en français<sup>19</sup> (New, Pallier, Ferrand, & Matos, 2001) ne distinguent pas les fréquences de deux homographes ou homophones en fonction de leur catégorie syntaxique. Mais de manière intuitive, on peut dire que certains items ont obtenu des résultats qui pourraient parfois être expliqués en terme de biais de fréquence ou de biais de plausibilité. Par exemple, prenons l'item 10, c'est-à-dire le couple « sombres / sombrent » : que les sujets aient entendu « de gros nuages sombres... » ou « de gros nuages sombrent... », ils ont donné au mot ambigu une interprétation adjectif dans 100% des cas lors de l'expérience de complétion de phrases, et dans presque 70% des cas lors de l'expérience de détection de mot abstrait. Il semble relativement intuitif de dire que l'image « des nuages sombres » est plus fréquente ou plus probable que celle « des nuages qui sombrent ». Nous avons obtenu l'effet inverse pour l'item 11 : que les sujets aient entendu « de drôles de bruits courts... » ou « de drôles de bruits courent... », ils ont donné au mot ambigu une interprétation verbe dans 100% des cas lors de l'expérience de complétion et dans environ 75% des cas lors de l'expérience de détection. Heureusement, ces résultats un peu déviants ne concernent que très peu d'items.

A posteriori, nous avons réalisé deux expériences contrôles auprès de 10 adultes français, pour estimer (1) la fréquence relative de chaque membre d'une paire ambiguë, et (2) la plausibilité de chaque item dans la phrase qui le contenait. Par exemple, ils devaient estimer la fréquence de l'adjectif « mort » dans « il est mort » et la fréquence du verbe « mordre » dans « il mord » sur une échelle allant de 1 (pas fréquent du tout) à 7 (très fréquent). De la même manière, ils devaient estimer la plausibilité (ou la prédictabilité) de l'adjectif « mort » dans « le petit chien mort... » et la plausibilité du verbe « mordre » dans « le petit chien mord... » en utilisant cette même échelle de 1 à 7.

Nous avons ensuite calculé un biais de réponse (par exemple un biais de réponse adjectif) pour chaque item sur les différentes données que nous avons : estimation de fréquence, estimation de plausibilité, résultats en complétion avec prosodie informative, résultats en complétion avec prosodie neutre, résultats en détection avec prosodie informative et résultats en détection avec prosodie neutre. Nous avons ensuite calculé la

---

<sup>19</sup> Site officiel : <http://www.lexique.org>

corrélation qui existait entre les réponses données aux expériences d'estimation de fréquence et de plausibilité, et les résultats obtenus dans les différentes expériences. Ces analyses sont résumées dans le Tableau 4.

	Complétion (prosodie informative)	Complétion (prosodie neutre)	Détection (prosodie informative)	Détection (prosodie neutre)
Estimation de fréquence	0.42	0.59 ***	0.29	0.27
Estimation de plausibilité	0.74 ***	0.66 ***	0.71 ***	0.61 ***

**Tableau 4 : Coefficients de corrélation entre les estimations de fréquence et de plausibilité données pour chaque item ambigu, et les résultats par item obtenus dans les différentes expériences présentées dans ce chapitre.** Les corrélations significatives sont marquées par des astérisques. Une corrélation très élevée indique par exemple que quand un item adjectif a été jugé plus plausible que son homonyme verbe dans les phrases utilisées, cet item a eu tendance à recevoir plus de réponses adjectif que de réponses verbe dans les différentes expériences.

Il semble que la plausibilité des mots dans les phrases soit beaucoup plus corrélée avec les résultats des différentes expériences que la fréquence intrinsèque des mots. Quand à l'indice de fréquence, il semble être un meilleur indice explicatif des résultats obtenus dans les expériences de complétion, que dans les expériences de détection, et en particulier dans la condition où la prosodie utilisée était relativement neutre.

En fait, des analyses plus détaillées de chaque item nous ont permis de trouver les items qui étaient particulièrement déviants (biais de réponses et biais dans les estimations très forts) : ils sont au nombre de 4 (les items « grave / grave », « sombres / sombrent », « courts / courent » et « réfléchi / réfléchit »). Nous avons refait toutes les analyses en retirant ces items déviants et nous avons obtenu des résultats similaires (voire même parfois plus forts) que ceux décrits dans ce chapitre. Les coefficients de corrélation relativement élevés que nous avons trouvés entre les expériences d'estimation et les expériences de complétion et de détection étaient essentiellement dus à ces items déviants. En effet, en retirant ces items, les coefficients de corrélation diminuent largement et les corrélations deviennent non-significatives.

Cette analyse nous a confirmé qu'il faut être très prudent dans le choix des stimuli, ce que nous prendrons en considération dans les expériences suivantes pour la mise en place du matériel expérimental. Quoiqu'il en soit, l'une des caractéristiques de

notre dessin expérimental garantit que nos résultats ne peuvent pas être dus exclusivement à des effets de fréquence et de prédictabilité. Il s'agit de l'utilisation de deux prosodies différentes. En effet, si les réponses des sujets n'avaient été fondées que sur la fréquence des items ou leur plausibilité dans les phrases, nous aurions dû alors obtenir des résultats similaires entre les deux conditions de prosodie (puisque les mêmes mots ambigus et les mêmes phrases étaient utilisés dans ces deux conditions). Or, notre résultat principal est que la saillance des indices prosodiques a eu un effet sur le comportement des sujets : plus les indices prosodiques étaient marqués et plus les auditeurs ont réussi à distinguer les débuts de phrases ambiguës et à attribuer la bonne catégorie syntaxique aux mots ambigus qu'ils rencontraient. Nous pouvons donc conclure que la prosodie, qui avait longtemps été négligée dans les modèles de traitement du langage, semble en fait un indice relativement pertinent et fiable, utilisable par exemple dans les premières étapes précoces du traitement syntaxique des phrases.

### 2.3.3 *Locuteurs experts versus locuteurs naïfs*

Très peu d'expériences avaient jusqu'à présent étudié le rôle des frontières de groupes phonologiques dans l'analyse syntaxique réalisée en temps réel par des auditeurs. Pour essayer de mettre en évidence l'utilisation de ces indices, nous avons décidé de réaliser différentes expériences, mettant en jeu des phrases localement ambiguës. Un point fondamental de notre étude était de faire varier la qualité des indices prosodiques : une modulation des résultats selon la condition prosodique garantit en effet que nos effets sont bien dus à la manipulation de la structure prosodique, plutôt qu'à des artefacts non contrôlés. Pour obtenir ces deux niveaux prosodiques (informative versus neutre) en gardant un signal de parole naturelle, nous avons du faire appel à une « locutrice experte » : celle-ci était d'une part habituée aux séances d'enregistrements de stimuli (être dans une cabine insonorisée, enregistrer des phrases sans donner l'impression de les lire mais en donnant, au contraire, l'impression de converser de manière agréable avec quelqu'un...) ; d'autre part, notre locutrice était pleinement consciente du problème d'ambiguïté que nous testions et du rôle de la prosodie que nous espérions mettre en évidence. Elle a donc pu facilement suivre les

consignes que nous lui donnions, à savoir (1) essayer de différencier les paires de phrases ambiguës en leur assignant des prosodies différentes, et (2) essayer de minimiser le plus possible les différences prosodiques entre les phrases ambiguës mais en produisant toujours quelque chose de naturel. Grâce à cette manipulation, nous avons pu mettre en évidence que quand les indices prosodiques étaient relativement saillants (allongements et variations intonatives importantes), les auditeurs français pouvaient les utiliser pour guider leur analyse syntaxique des phrases en temps réel.

L'étape suivante consiste maintenant à poser la question de la production « spontanée » d'indices prosodiques. Il existe actuellement un débat sur cette question, certains pensant que seuls les locuteurs experts peuvent produire de bons indices prosodiques désambiguïsants (Allbritton, McKoon, & Ratcliff, 1996; Fox Tree & Meijer, 2000), tandis que d'autres ont essayé de démontrer que même des locuteurs naïfs pouvaient produire spontanément des indices de frontières prosodiques utiles pour le traitement syntaxique des phrases (Kraljic & Brennan, 2005; Schepman & Rodway, 2000). Donc, est-ce que des locuteurs naïfs et non-experts produisent spontanément des indices de frontières de groupes phonologiques, suffisamment marqués pour être ensuite utilisés par des auditeurs pour résoudre des ambiguïtés syntaxiques ? L'expérience que nous allons maintenant aborder va essayer d'apporter une réponse à cette question.

## CHAPITRE 3

### UTILISATION D'INDICES PROSODIQUES PRODUITS PAR DES LOCUTEURS NAÏFS (ETUDE 2)

Les expériences que nous avons présentées dans la partie précédente ont mis clairement en évidence le rôle que pouvaient jouer les indices prosodiques de groupes phonologiques dès les premières étapes de l'analyse syntaxique des phrases. Quand ces indices sont assez saillants et bien marqués (comme dans le cas de la prosodie informative), les auditeurs français peuvent les utiliser pour inférer la catégorie syntaxique d'un mot ambigu pouvant appartenir à deux catégories syntaxiques différentes (adjectif versus verbe).

Comme nous l'avons déjà remarqué, les phrases ambiguës utilisées dans les expériences précédentes étaient prononcées par une locutrice experte, à qui nous avons expliqué le problème d'ambiguïté, et qui avait eu comme tâche de produire des indices prosodiques bien différents entre les paires de phrases ambiguës.

Nous voulons maintenant étudier la question de la production plus spontanée et plus naturelle d'indices prosodiques, en regardant si des locuteurs naïfs pourraient eux aussi produire de tels indices, puis en vérifiant si ces informations pourraient être utilisées par des auditeurs français pour résoudre des ambiguïtés syntaxiques temporaires.

#### 3.1 LE DEBAT ACTUEL SUR LA PRODUCTION D'INDICES PROSODIQUES INFORMATIFS PAR DES LOCUTEURS NAÏFS ET NON-EXPERTS

Il existe actuellement un débat sur cette question : certains pensent en effet que seuls les locuteurs experts ou non-naïfs peuvent produire de bons indices désambiguïsants (Allbritton et al., 1996; Fox Tree & Meijer, 2000; Snedeker, Gleitman, Felberbaum, Placa, & Trueswell, 2000; Snedeker & Trueswell, 2003), alors que d'autres tentent de démontrer que même des locuteurs naïfs peuvent produire des indices de frontières prosodiques suffisamment marqués pour être utilisés lors du

traitement syntaxique des phrases (Kraljic & Brennan, 2005; Schafer, Speer, & Warren, 2004; Schafer et al., 2000; Schepman & Rodway, 2000).

Par exemple, Allbritton et al. (1996) ont utilisé des phrases contenant une ambiguïté locale comme « They rose early in May ». Différents locuteurs devaient lire ces phrases : ceux-ci étaient soit des étudiants (locuteurs non experts), soit des acteurs (locuteurs experts). Dans une première partie de cette expérience de production, les participants n'étaient pas mis au courant du but de l'expérience (locuteurs naïfs) ; dans une deuxième partie, on leur expliquait le but de l'expérience, l'importance de la production de bons indices prosodiques pour désambiguïser les phrases, et on leur montrait les deux interprétations possibles de chaque phrase ambiguë, en la retranscrivant par exemple avec une paraphrase (locuteurs non-naïfs).

Les phrases ambiguës étaient présentées à la fin d'un paragraphe qui jouait le rôle de contexte désambiguïsant : en effet, au vu de ce qui précédait, une seule des deux interprétations restait possible à la fin du passage. Par exemple, les locuteurs devaient prononcer les passages suivants :

*Passage 1 : « In spring there was always more work to do on the farm. May was the hardest month. They rose early in May. »*

*Passage 2 : « Bears sleep all winter long, usually coming out of hibernation in late April, but this year they were a little slow. They rose early in May. »*

Dans la phrase ambiguë du passage 1, l'adverbe « early » doit être attaché au verbe, et une frontière prosodique devrait donc apparaître entre « they rose early » et « in May ». Dans la phrase ambiguë du passage 2, « early » doit être attaché au groupe nominal, et une frontière prosodique devrait donc apparaître entre « they rose » et « early in May ».

Les auteurs ont réalisé des analyses acoustiques des phrases ambiguës produites, puis ont conduit différentes expériences, comme des tâches de compréhension, pour voir si des auditeurs pouvaient utiliser la prosodie des phrases pour inférer le sens qu'avaient voulu produire les locuteurs.

Ils ont trouvé que les locuteurs non experts (les étudiants) ne produisaient jamais d'indices prosodiques suffisamment forts pour pouvoir distinguer entre les deux sens

possibles. Quant aux locuteurs experts (les acteurs), ils réussissaient à désambiguïser ces phrases grâce à la prosodie, mais uniquement quand ils étaient mis au courant du but des expériences et quand on leur demandait explicitement de produire des indices prosodiques différents pour les deux interprétations possibles.

Allbritton et al. (1996) avaient alors émis de sérieux doutes quant au rôle des indices prosodiques dans l'analyse syntaxique des phrases, puisqu'il apparaissait que les locuteurs produisaient rarement de tels indices, et que ceux-ci n'étaient pas produits de manière informative spontanément. Fox Tree et Meijer (2000) ont répliqué ces résultats et conclu que la prosodie n'était généralement pas utilisée lors de la production de phrases ambiguës, et qu'elle ne pouvait donc pas être utilisée pour guider l'analyse syntaxique des phrases. Le fait que des locuteurs informés et experts puissent marquer la structure syntaxique des phrases par la prosodie semble bien établi. Mais la généralisation de ce comportement à des locuteurs naïfs a été largement remise en question.

D'autres chercheurs se sont intéressés à ces résultats, mettant en évidence le fait que ces expériences présentaient quelques problèmes, comme par exemple le fait que la tâche d'Allbritton et al. (1996) soit une tâche de lecture, très éloignée d'une situation de communication spontanée (Kraljic & Brennan, 2005; Schafer et al., 2004; Snedeker & Trueswell, 2003).

Kraljic et Brennan (2005) ont décidé d'étudier la production spontanée d'indices prosodiques par des locuteurs naïfs pour résoudre des ambiguïtés syntaxiques globales. Mais leur tâche était beaucoup plus naturelle et avait un but communicatif plus direct que celles utilisées par Allbritton et al. (1996), ou par Fox Tree et Meijer (2000). Les participants étaient testés par deux dans une tâche de jeu coopératif (« cooperative game-playing task ») : l'un des deux était le locuteur (l'instructeur) qui prononçait des phrases indiquant à l'interlocuteur quelles actions réaliser (bouger des objets qui se trouvaient face à lui). Les instructions pouvaient être ambiguës, comme :

(1a) « Put the dog in the basket on the star. »

Cette phrase comporte une ambiguïté au niveau de l'attachement du groupe nominal « in the basket ». Les deux interprétations possibles sont représentées dans les phrases suivantes, dans lesquelles nous avons indiqué par des crochets les positions des

frontières prosodiques (frontières qui devraient être produites par les locuteurs pour permettre de résoudre l'ambiguïté des phrases) :

(1b) « [Put the dog] [in the basket on the star] »

(le groupe nominal ambigu, lié au verbe, indique le lieu de l'action = « goal attachment »)

(1c) « [Put the dog in the basket] [on the star] »

(le groupe nominal ambigu modifie le nom « dog » = « modifier attachment »)

Les auditeurs avaient face à eux une table représentant différents objets. Les locuteurs recevaient des cartes représentant ces objets, et indiquant l'instruction à donner : l'objet à déplacer était marqué d'un astérisque (par exemple, le chien), le lieu où le placer était indiqué par une flèche (par exemple, le panier placé sur l'étoile), et un point représentait un autre objet que le locuteur devait intégrer dans ses consignes (par exemple, l'étoile). Chaque consigne devait commencer par le verbe « put » ou « place ».

Le contexte référentiel (objets face à l'auditeur, carte donnée au locuteur) variait selon les conditions : soit il était ambigu (les deux interprétations de la phrase ambiguë étaient possibles dans ce contexte), soit il était non-ambigu (une seule des deux interprétations était possible).

Les auteurs ont trouvé que les locuteurs produisaient des indices prosodiques fortement marqués, permettant aux auditeurs d'inférer l'interprétation de la phrase et de réaliser correctement l'action demandée. Ceci a été trouvé dans diverses situations : (1) que le contexte référentiel soit ambigu ou non, et (2) que les locuteurs aient conscience ou pas du problème d'ambiguïté. Il semble donc que la production d'indices prosodiques liés à la structure prosodique des phrases soit quelque chose de naturel et spontané. Cette expérience a aussi montré que la production de la prosodie n'était pas uniquement réalisée pour l'interlocuteur (dans le souci de lui apporter le plus d'informations possibles, de manière à ce qu'il comprenne clairement et rapidement le sens de la phrase, ce qui pourrait se rapprocher des maximes de conversation, comme la maxime de manière ou la maxime de quantité, développées par Grice, 1975). Si cela avait été le cas, les locuteurs auraient dû être plus informatifs quand ils étaient conscients de l'ambiguïté syntaxique des phrases, ou quand ils savaient que le contexte référentiel était ambigu pour l'auditeur. Ce résultat va donc à l'encontre de l'hypothèse « audience design » (Kraljic & Brennan, 2005; Lockridge & Brennan, 2002; Schober &

Brennan, 2003). Par contre, cette production d'indices prosodiques qui semble spontanée et naturelle, et qui ne dépend pas du contexte expérimental, valide plutôt l'idée selon laquelle la prosodie est produite spontanément par le locuteur, comme le résultat de la planification et de la prononciation des phrases.

Ainsi, ces deux expériences auraient pu minimiser le débat entre production experte et production naïve de la prosodie. Il semble bien que la tâche expérimentale impliquée joue un rôle important : des tâches de lecture diminuent la production d'indices prosodiques désambiguïsants, alors qu'une tâche plus naturelle de conversation et d'échange permet l'apparition de tels indices. Mais présentons maintenant une dernière expérience, dont les résultats montrent que cette explication n'est pas suffisante.

Snedeker et Trueswell (2003) ont réalisé une expérience similaire à celle de Kraljic et Brennan (2005), à savoir une tâche de jeu coopératif entre un locuteur et un auditeur, impliquant des instructions globalement ambiguës dans des contextes référentiels ambigus ou non. Voici un exemple d'une instruction ambiguë :

(2) « *Tap the frog with the flower* »

En effet, dans cette phrase, le complément « with the flower » peut être considéré soit comme l'outil à utiliser pour tapoter la grenouille (donc cette unité est reliée au verbe), soit comme une indication sur quelle grenouille tapoter (donc cette unité est attachée au nom).

Les auteurs ont trouvé que les locuteurs avaient choisi de produire des indices prosodiques désambiguïsants, uniquement quand le contexte référentiel pouvait supporter les deux interprétations de la phrase ambiguë. A nouveau, ce résultat va plutôt dans le sens des résultats de Allbritton et al. (1996), et semble indiquer qu'il est possible de produire des indices prosodiques seulement dans certaines conditions (si l'on est entraîné, si l'on est informé, et/ou si l'on veut aider notre interlocuteur à comprendre rapidement la phrase prononcée). Cette fois-ci, la tâche expérimentale ne peut pas vraiment être mise en cause : elle est similaire à celle utilisée par Kraljic et Brennan (2005).

Nous pensons qu'une explication supplémentaire pourrait être apportée à ces différences de résultat, explication qui repose sur les stimuli utilisés dans ces différentes expériences. Si l'on regarde plus précisément le matériel ambigu utilisé, on se rend compte qu'il existe une différence au niveau de la longueur des phrases : ceux qui n'ont pas trouvé de production d'indices prosodiques par des locuteurs naïfs ou non-experts ont généralement utilisé des phrases courtes (par exemple, des phrases d'environ 7 syllabes, comprenant deux groupes nominaux, chez Snedeker et Trueswell, 2003) ; ceux qui par contre ont mis en évidence une production naturelle et spontanée d'indices prosodiques désambiguïsants, même par des locuteurs naïfs, ont utilisé des phrases plus longues (par exemple, des phrases d'environ 10 syllabes, comprenant 3 groupes nominaux, chez Kraljic et Brennan, 2005).

Comme nous l'avons vu dans le premier chapitre, la symétrie entre structure syntaxique et structure prosodique n'est pas parfaite : toutes les frontières syntaxiques ne sont pas obligatoirement marquées au niveau prosodique. La réalisation d'une frontière prosodique dépend de plusieurs critères : bien sûr de la structure syntaxique de la phrase, mais aussi de la longueur totale de la phrase, de la longueur des différents constituants résultant du découpage prosodique de la phrase, du débit de parole du locuteur... (Ferreira, 1993; Gee & Grosjean, 1983; Shattuck-Hufnagel & Turk, 1996).

Nous pensons que la manière la plus probable et la plus logique de prononcer les phrases courtes utilisées par Allbritton et al. (1996) ou par Snedeker et Trueswell (2003) est de ne faire qu'une seule unité prosodique. Ainsi, la phrase « They rose early in May » devrait plus logiquement être prononcée en une seule unité intonationnelle (« [They rose early in May] »), plutôt qu'en deux unités (« [They rose] [early in May] » ou « [They rose early] [in May] » selon le sens donné à la phrase), parce qu'elle ne contient que 6 syllabes.

Par contre, nous pensons que les phrases utilisées par Kraljic et Brennan (2005) devraient plus facilement et plus logiquement être prononcées en deux unités prosodiques, de par la longueur plus importante de ces phrases. Si une frontière prosodique doit être insérée dans ces phrases, sa position dépendra alors du sens que veut transmettre le locuteur, et donc de la structure syntaxique de la phrase (ils devraient prononcer « [Put the dog in the basket] [on the star] » si le groupe nominal ambigu modifie le nom, et « [Put the dog] [in the basket on the star] » si le groupe nominal ambigu se rapporte au verbe et indique le but de l'action ; et rappelons que ce sont ces

deux structures prosodiques qui ont été effectivement observées en production spontanée).

Nous avons donc voulu tester cette idée, en utilisant des phrases ambiguës similaires à celles que nous avons étudiées dans les premières expériences (voir chapitre 2), c'est-à-dire des phrases moyennement longues contenant une ambiguïté syntaxique temporaire (comme « Le petit chien mort sera enterré demain dans le jardin de ses maîtres » versus « Le petit chien mord la laisse qui le retient dans l'espoir de se libérer »). Si la longueur des phrases oblige les participants à produire au moins une frontière prosodique, sa position devrait alors dépendre de la structure syntaxique des phrases : une frontière prosodique devrait donc être placée au niveau de la frontière syntaxique entre le groupe nominal sujet et le groupe verbal (donc avant le mot ambigu pour les phrases contenant la cible verbe, après le mot ambigu pour les phrases contenant la cible adjectif).

Nous avons donc demandé à des locuteurs naïfs et non-experts de lire ces phrases, puis nous avons réalisé une expérience de complétion de phrases (similaire à celle présentée dans la partie précédente) pour vérifier que des auditeurs pouvaient utiliser les différences prosodiques (potentiellement produites par les locuteurs naïfs) pour guider leur analyse syntaxique des phrases.

## 3.2 NOTRE ETUDE : PRODUCTION D'INDICES PROSODIQUES PAR DES LOCUTEURS NAÏFS ET UTILISATION DE CES INDICES POUR RESOUDRE DES AMBIGUÏTES SYNTAXIQUES

### 3.2.1 *Production des phrases ambiguës*

Nous voulions comparer la qualité des indices prosodiques produits par une locutrice experte et informée du but de nos recherches (voir la première étude), avec ceux produits par des locuteurs non-experts et naïfs. Pour cela, nous avons donc utilisé le même genre de phrases expérimentales ambiguës, à savoir des paires de phrases

contenant une ambiguïté syntaxique temporaire, reposant sur le fait qu'un des mots de la phrase pouvait appartenir à deux catégories syntaxiques différentes.

Rappelons l'exemple utilisé dans l'étude précédente et repris dans celle-ci :

(3a) *Le petit chien mord la laisse qui le retient.*

(3b) *Le petit chien mort sera enterré demain.*

Dans chacune de ces phrases, les quatre premiers mots ont la même prononciation : les débuts de ces phrases sont donc identiques d'un point de vue segmental. Mais le quatrième mot est un verbe dans la première phrase, alors que c'est un adjectif dans la deuxième. Les débuts de ces deux phrases diffèrent donc du point de vue de leur structure syntaxique, ce qui devrait avoir un impact sur leur structure prosodique (on devrait trouver une frontière de groupe phonologique avant le mot ambigu dans la phrase ambiguë verbe, après le mot ambigu dans la phrase ambiguë adjectif).

### 3.2.1.1 Choix des stimuli à enregistrer

#### Sélection par rapport aux expériences précédentes :

Nous avons repéré que certains items ambigus avaient posé quelques problèmes dans les expériences précédentes : certains avaient montré par exemple un fort biais de réponse dans la première expérience de complétion, même quand la prosodie utilisée était informative (par exemple, l'item « sombres / sombrent » avait été interprété comme un adjectif dans 100% des cas, l'item « courts / courent » n'avait obtenu que des réponses verbe quelle que soit la phrase entendue par les sujets). Nous avons donc décidé de supprimer ces phrases de notre nouveau matériel (ceci concerne 5 items ambigus).

### Estimation de la fréquence et de la plausibilité des mots ambigus dans les phrases :

Suite aux expériences précédentes, nous avons réalisé un test auprès de 10 personnes, pour estimer la fréquence de chaque membre d'une paire ambiguë (pour vérifier qu'une des deux formes n'était pas trop fréquente comparée à l'autre), et pour estimer la plausibilité des items ambigus dans les phrases expérimentales (les paires de phrases commençant de la même manière, il fallait éviter que l'une des deux formes soit beaucoup plus plausible ou prédictible que l'autre). Des analyses post-hoc avaient indiqué que ces informations de fréquence et de plausibilité avaient parfois joué un rôle dans les réponses des sujets (mais ceci ne concernait que quelques items), et que la plausibilité des mots par rapport au début de la phrase porteuse avait plus tendance à influencer les réponses des sujets (comparé à la fréquence).

Pour les 15 items restants de l'expérience précédente, nous avons reconstruit des paires de phrases expérimentales (phrases ambiguës verbe et adjectif). Nous avons aussi recherché d'autres paires de mots homophones pouvant appartenir à deux catégories syntaxiques différentes, puis inventé les phrases ambiguës correspondantes (nous avons ainsi trouvé 16 nouveaux items ambigus).

Toutes ces nouvelles phrases expérimentales (30 phrases adaptées des anciennes et 32 phrases construites pour cette nouvelle expérience) ont été présentées à 10 adultes français pour réaliser deux expériences :

- (1) une estimation de la fréquence de chaque membre d'une paire ambiguë (estimation de la fréquence de l'adjectif et de la fréquence du verbe, sur une échelle allant de 1 à 7, de « pas fréquent du tout » à « très fréquent ») ;
- (2) une estimation de la plausibilité de chaque item dans la phrase qui le contenait (estimation de la plausibilité – ou de la prédictibilité – de l'adjectif et du verbe par rapport au début des phrases porteuses, en utilisant une échelle allant de 1 à 7, de « pas plausible du tout » à « très plausible »).

Nous avons ensuite regardé, pour chaque item, si l'un des deux membres avait un biais de fréquence ou un biais de plausibilité. Nous avons aussi comparé ces résultats avec ceux que nous avons trouvés pour les phrases expérimentales utilisées lors de la première étude (en effet, nous avons construit de nouvelles phrases expérimentales, pour 15 items qui avaient été utilisés dans nos expériences précédentes ; mais il se pouvait que les anciennes phrases soient tout de même meilleures que celles que nous

venions de construire). Nous avons pu ainsi choisir les meilleurs items et les meilleures phrases expérimentales pour cette nouvelle expérience (en évitant les items les plus fortement biaisés, en particulier au niveau de la plausibilité qui semblait être un facteur plus corrélé avec les réponses des sujets, que ne l'était la fréquence).

Au total, nous avons gardé 22 couples d'items ambigus (15 anciens et 7 nouveaux). Sur les 15 couples d'items ambigus utilisés précédemment, peu de phrases expérimentales ont été changées (nous avons choisi les nouvelles phrases expérimentales pour seulement deux items). Au total, nous avons donc 44 phrases ambiguës à faire enregistrer à des locuteurs naïfs.

### 3.2.1.2 Enregistrement des stimuli

Nous voulions que nos locuteurs soient des sujets naïfs et non-experts : nous avons donc contacté des étudiants, qui ne connaissaient pas le principe de nos expériences et qui n'avaient jamais enregistré de stimuli auditifs. Au total, six étudiantes ont répondu et ont enregistré nos stimuli lors d'une séance individuelle durant environ une heure. Pour que les sujets ne se rendent pas compte, au fur et à mesure de la séance d'enregistrement, que les phrases étaient des phrases ambiguës, il nous a fallu compléter notre matériel : en effet, il nous était impossible de seulement présenter les 44 phrases expérimentales aux participants, car ils auraient très rapidement compris quel était le principe de nos expériences.

Nous avons donc construit des phrases contrôles (ou repris certaines de celles que nous avons utilisées dans l'expérience de détection de mot), dans lesquelles se trouvaient les mêmes items test, mais placés dans une position non-ambiguë (seule une des deux interprétations, verbe ou adjectif, était possible pour des raisons sémantiques ou syntaxiques). Ainsi, pour un item, les locuteurs avaient 4 phrases à enregistrer : la phrase ambiguë adjectif, la phrase ambiguë verbe, la phrase contrôle adjectif et la phrase contrôle verbe.

Nous avons réalisé deux blocs de phrases à enregistrer : pour un item donné, les phrases verbe (ambiguë et contrôle) se trouvaient dans le premier bloc, et les phrases adjectif (ambiguë et contrôle) se trouvaient dans le deuxième bloc. Pour un autre item, les phrases adjectif étaient dans le premier bloc alors que les phrases verbe

apparaissaient dans un deuxième bloc. Ainsi, les membres de chaque paire de phrases ambiguës étaient assez éloignés les uns des autres, ce qui permettait d'éviter que les participants réagissent au problème d'ambiguïté, en se rendant compte qu'un mot pouvait avoir deux sens possibles (nous avons pris en compte les effets de mémoire : en effet, si un locuteur avait prononcé une phrase ambiguë verbe relative à un item, et qu'il avait trouvé un peu plus loin une phrase ambiguë adjectif – donc commençant de la même manière mais comportant un mot d'une autre catégorie syntaxique – il aurait plus alors se rappeler de la première phrase produite et trouver étrange la présence de cette deuxième phrase. Si le locuteur avait repéré ces paires de phrases plusieurs fois, il aurait alors sûrement compris le but de nos expériences. De même, si dans un bloc était présentée la phrase ambiguë verbe d'un item, alors la phrase contrôle verbe apparaissait aussi dans ce bloc, et non la phrase contrôle adjectif. Ceci était aussi réalisé pour éviter que les locuteurs réalisent que deux mots se prononçant de la même manière pouvait appartenir à deux catégories syntaxiques différentes). L'ordre de présentation des blocs a été contrebalancé entre les locuteurs (bloc 1 ou bloc 2 enregistré en premier) et l'ordre de présentation des phrases à l'intérieur de chacun des blocs était fixe.

Toutes les phrases étaient présentées à l'écrit. Il faut noter que, pour la majorité des stimuli, l'ambiguïté n'était pas évidente à l'écrit, puisque les items s'écrivaient en général de manière différente entre les phrases ambiguës adjectif et les phrases ambiguës verbe (seuls 7 items sur les 22 utilisés avaient la même orthographe dans les deux phrases). Les sujets devaient donc réaliser une tâche de lecture. Nous avons vu que la lecture des phrases n'était peut-être pas le meilleur moyen d'obtenir de bons indices prosodiques mais, vu les phrases utilisées, il était difficile de les faire produire naturellement et spontanément (même en utilisant un jeu coopératif...).

Donc lors des explications que nous avons données aux locuteurs, nous avons insisté sur plusieurs points, pour pallier ces problèmes : (1) avant de prononcer la phrase, les participants devaient la lire dans leur tête (ceci était fait pour obtenir une production plus fluide, et pour éviter que les sujets ne s'entendent se faire piéger) ; (2) ils devaient prononcer ces phrases de manière vivante, enjouée, relativement rapide, en articulant bien, comme s'ils étaient en train de parler de quelque chose d'intéressant avec des amis...

Les locuteurs ne pensaient pas être testés mais pensaient au contraire participer à la préparation d'une expérience perceptive : nous leur disions uniquement que ces phrases allaient être utilisées dans différentes expériences, et nous insistions donc sur le fait qu'ils devaient produire des phrases bien compréhensibles et agréables à écouter (pour que les futurs participants ne s'ennuient pas trop en les écoutant...).

Les séances d'enregistrement se sont déroulées dans une cabine insonorisée. Toutes les phrases ont été digitalisées à 16000 Hz et 16 bits, sur une table de mixage OROSAU22. Pendant les enregistrements, un expérimentateur, situé dans une pièce adjacente et disposant de la même feuille de stimuli que le locuteur, écoutait les productions des participants grâce à un casque. Il pouvait alors noter les phrases qui nécessiteraient un deuxième enregistrement (dû à des erreurs ou des hésitations que le sujet n'avait pas corrigées par lui-même en répétant la phrase). Il faisait aussi attention aux indices prosodiques produits : en particulier, comme nous voulions tester le rôle des frontières des unités prosodiques intermédiaires (groupes phonologiques), nous ne voulions pas que des indices trop saillants, comme la présence de pauses, apparaissent dans la zone ambiguë étudiée (il ne fallait pas que des pauses soient présentes entre le groupe nominal sujet et le verbe).

Une fois que toutes les phrases avaient été produites par le locuteur, l'expérimentateur demandait au sujet d'en relire certaines (celles qui contenaient des erreurs, des hésitations ou des pauses).

A la fin de la séance, nous discutons avec le locuteur pour savoir ce qu'il avait pensé de ces phrases, s'il avait une idée de ce que nous testions... Heureusement, personne n'avait deviné le but de nos expériences, et n'avait soulevé le problème que les phrases étaient ambiguës et apparaissaient par paires. Il arrivait cependant que certains locuteurs se fassent piéger en prononçant une des phrases ambiguës : par exemple, ils commençaient à lire une phrase avec l'intonation indiquant qu'ils avaient considéré le mot ambigu comme un adjectif, et se rendaient compte en lisant les mots suivants que ce mot ambigu était un verbe. Ceci n'est pas arrivé très fréquemment (1 ou 2 phrases par enregistrement) et les participants n'en ont pas pour autant inféré le but de nos expériences. Au contraire, ils pensaient que ce problème n'avait pas été contrôlé, que nous n'avions pas repéré cette ambiguïté, et nous disaient de faire attention à cette phrase pour la suite de nos expériences. Quand nous leur montrions que la moitié des

phrases étaient des phrases ambiguës, ils étaient alors très surpris. Nous pouvons donc dire que nos locuteurs étaient bien des sujets naïfs et non-experts.

La deuxième étape de notre étude était de vérifier que les locuteurs avaient bien produit des indices prosodiques désambiguïsants : nous avons donc utilisé ces phrases pour réaliser une expérience de complétion de phrases auprès d'adultes français.

### 3.2.2 *Complétion de phrases produites par les locuteurs naïfs*

#### 3.2.2.1 Méthode

##### Matériel :

Nous avons utilisé 22 couples adjectif / verbe ambigus. Pour chaque couple, deux phrases expérimentales ambiguës ont été construites, choisies pour éviter des biais de fréquence et de plausibilité, puis enregistrées par six locutrices naïves et non-expertes (voir l'annexe B.1 pour la liste du matériel expérimental).

Toutes les phrases expérimentales ont été éditées grâce aux programmes Manitu et Praat. Elles ont été coupées à la fin du mot ambigu, à un passage à zéro de la courbe d'amplitude.

Nous avons aussi utilisé des phrases distractrices, en coupant les phrases contrôles (enregistrées aussi par les locuteurs) à un endroit où aucune ambiguïté n'existait (les sujets devaient par exemple compléter « je trouve que mes amies... » ou « la jeune femme... »)

Nous avons les enregistrements de six locutrices. Les sujets qui participaient à cette expérience n'entendaient que les stimuli d'une seule locutrice (nous avons donc 6 groupes de sujets, un par locutrice). Dans chaque groupe, deux blocs de stimuli ont été créés de manière à ce que chaque membre d'une paire de phrases ambiguës apparaisse dans un bloc différent. Pour chaque groupe, nous avons donc testé deux sous-groupes de sujets (ceux qui traitaient les items du bloc 1 et ceux qui traitaient les items du bloc 2).

Ainsi, chaque bloc contenait onze phrases ambiguës verbe, onze phrases ambiguës adjectif et les dix phrases distractrices. Ces 32 phrases ont été séparées en deux sous-blocs (chacun comprenant 5 ou 6 phrases ambiguës verbe, 5 ou 6 phrases ambiguës adjectif et 5 phrases distractrices), de manière à ce que les sujets puissent faire une petite pause à la moitié de l'expérience. La présentation des stimuli à l'intérieur des sous-blocs était aléatoire ; l'ordre de présentation des sous-blocs était fixe pour un groupe de sujets.

En plus de ces phrases ambiguës et distractrices, deux phrases non ambiguës (similaires aux phrases distractrices) ont été créées et utilisées au début de l'expérience comme entraînement.

#### Procédure :

L'expérience a été réalisée en utilisant le programme Expe6 de présentation des stimuli (Pallier et al., 1997)<sup>20</sup>. Tous les participants étaient testés individuellement, dans une pièce isolée, équipée d'un ordinateur, d'un casque et d'une feuille réponse. Avant de commencer, chaque participant était assigné à un groupe, puis remplissait un formulaire indiquant sa date de naissance, un éventuel bilinguisme, et s'il avait des problèmes auditifs ou des problèmes de langage. Il découvrait ensuite les consignes de l'expérience : les instructions étaient données oralement, en complément des instructions écrites à l'écran (voir annexe B.2 pour les instructions écrites détaillées). L'expérimentateur s'assurait que le sujet avait bien compris la tâche en restant près de lui pendant les deux essais d'entraînement, ce qui permettait aussi de clarifier toutes les questions éventuelles avant de commencer le test.

Les instructions expliquaient aux participants qu'ils allaient entendre des débuts de phrases, coupées à la fin d'un mot, et que leur tâche serait de trouver une suite à ces phrases. Ils devaient pour cela écrire sur la feuille réponse le début de la phrase qu'ils avaient entendue, puis la compléter en mettant la première idée qui leur venait à l'esprit.

Un essai commençait avec la présentation auditive du début de la phrase (les stimuli auditifs étaient présentés à travers une carte son ProAudioSpectrum 16-bit). Les participants pouvaient écouter ce début de phrase autant de fois qu'ils le désiraient : il

---

<sup>20</sup> Voir le site internet : <http://www.lscp.net/expe/>

leur suffisait d'appuyer sur la barre d'espace de l'ordinateur pour rejouer le stimulus en cours. Ils devaient ensuite écrire la phrase complète (avec la fin imaginée) sur la feuille réponse. Une fois leur tâche réalisée, ils pouvaient lancer l'essai suivant, en appuyant sur la touche entrée. Les sujets devaient traiter les 32 phrases de l'expérience pour que l'expérience se termine. Une pause était proposée aux sujets à la moitié de l'expérience (entre les deux sous-blocs). L'expérience complète durait environ 15 minutes.

### Participants :

Soixante personnes, de langue maternelle française, ont participé à cette expérience (dix pour chaque locuteur). Aucun d'eux n'a rapporté de problèmes auditifs ou de problèmes liés au langage. A la fin de l'expérience, aucun n'avait compris que nous testions le rôle de la prosodie dans la résolution d'ambiguïtés syntaxiques (nous leur demandions s'ils avaient compris le but de notre expérience). Certains avaient repéré que certains items étaient ambigus mais ils n'ont pas généralisé ces exemples à l'expérience globale (dans ces cas là, les participants ne semblaient pas avoir mis en place des stratégies de réponse : ils disaient avoir utilisé le premier sens qui leur était venu à l'esprit, ou avoir mis celui qui semblait aller le mieux ; cette dernière décision a donc pu être prise sur la base de l'intonation des phrases, ou sur la base de la fréquence ou de la plausibilité des items dans la phrase).

#### 3.2.2.2 Résultats

D'après les phrases inventées par les participants, nous avons évalué l'interprétation qui avait été donnée au mot cible. Nous avons donc codé la réponse donnée par les sujets en « interprétation verbe » ou « interprétation adjectif ».

Deux items ambigus n'ont pas été inclus dans les analyses car ils avaient entraîné de forts biais de réponse : ce sont les items « violet / violait » dans « Il pense que ce grand tableau violet / violait ... » (88% de réponses adjectif quelle que soit la phrase entendue) et « cool / coulent » dans « Il a peur que ses amis cool / coulent ... »

(97% de réponses verbe quelle que soit la phrase entendue). Les analyses concernent donc 20 couples d'items ambigus<sup>21</sup>.

Dans certains cas, les sujets ont donné une réponse qui ne correspondait pas exactement au verbe ou à l'adjectif concerné. Quand leur réponse pouvait quand même être interprétée sans ambiguïté comme un verbe ou un adjectif, nous avons codé ce résultat : par exemple, une des phrases test était « mes amies percent... » et certains participants ont complété la phrase en mettant « mes amis perçoivent... » ; nous avons alors décidé que l'interprétation qu'ils avaient donnée au dernier mot entendu était bien le sens « verbe » et nous avons donc codé cette réponse comme une interprétation verbe. Ces réponses représentent 1,6% du nombre total de réponses, soit 19 réponses sur 1200<sup>22</sup>.

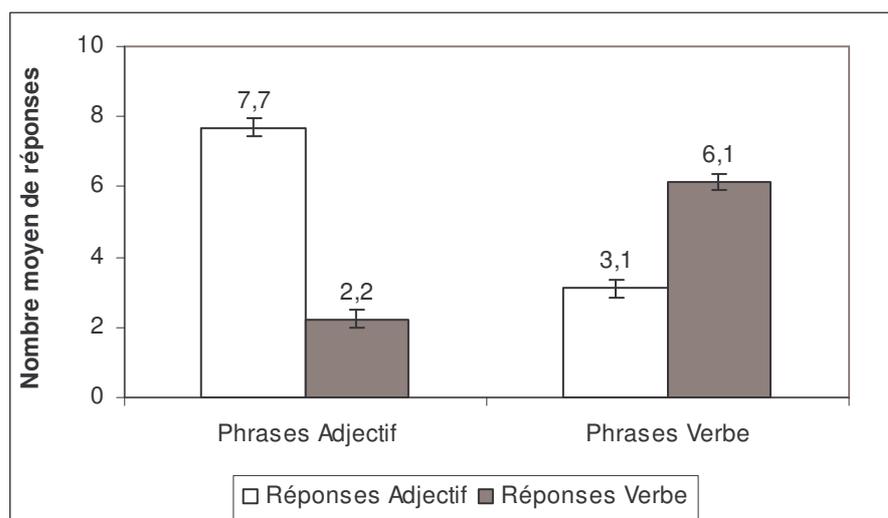
Dans d'autres cas, la réponse donnée par les sujets était trop éloignée du mot test et ne pouvait donc pas être codée en réponse verbe ou adjectif (par exemple, le début de phrase « la fumée dense / danse » a parfois été complétée par « la fumée dans ce lieu... »). Il arrivait aussi que le mot ambigu ne soit pas présent dans la phrase inventée par les sujets. Ces réponses ont toutes été exclues des analyses : cela représente 52 réponses sur 1200 soit 4,3% du nombre total de réponses.

La Figure 7 présente le nombre moyen de réponses adjectif et verbe pour les phrases ambiguës adjectif et verbe, tous locuteurs confondus (voir les annexes B.3.1 et B.4.1 pour les résultats individuels par sujets et par items).

---

<sup>21</sup> Une analyse conduite avec les 22 items ambigus a donné des résultats similaires.

<sup>22</sup> Une analyse conduite sans ces dix-neuf réponses a donné exactement les mêmes résultats.



**Figure 7 (résultats étude 2) : Nombre moyen de réponses adjectif et verbe données aux phrases adjectif et verbe, tous locuteurs confondus (sur 10 réponses possibles pour chaque phrase).** Les barres d'erreurs représentent les erreurs standard de la moyenne.

Le nombre moyen de réponses adjectif et verbe a été soumis à deux analyses de variance (ANOVA), l'une avec les participants et l'autre avec les items comme variable aléatoire. Dans l'analyse par sujet, nous avons déclaré deux facteurs intra-sujets : la Réponse donnée (adjectif versus verbe) et le type de Phrase entendue (phrase ambiguë adjectif versus phrase ambiguë verbe), ainsi que deux facteurs inter-sujets : le Locuteur (du locuteur 1 au locuteur 6) et le Contrebalancement (bloc 1 versus bloc 2). L'analyse par item incluait trois facteurs intra-items : le Locuteur, la Réponse donnée et le type de Phrase entendue.

Nous avons observé un effet principal de « type de Phrase » ( $F_1(1,48)=19.4$ ,  $p<.001$ ;  $F_2(1,19)=11$ ,  $p=.004$ ;  $\text{min}F'(1,41)=7$ ,  $p=.01$ ), qui vient du fait que les participants ont eu tendance à donner plus de réponses aux phrases adjectif qu'aux phrases verbe (9.9 versus 9.2). Ceci peut être expliqué par le fait que la majorité des réponses que nous n'avons pas analysées (à cause des oublis ou des problèmes de compréhension du mot ambigu) concernait essentiellement les phrases verbes (46 réponses à des phrases verbe non analysées versus 6 réponses à des phrases adjectif non analysées). Le facteur « Réponse » a aussi montré un effet principal significatif ( $F_1(1,48)=20.4$ ,  $p<.001$ ;  $F_2(1,19)=5.6$ ,  $p=.03$ ;  $\text{min}F'(1,30)=4.4$ ,  $p=.05$ ), de même que le

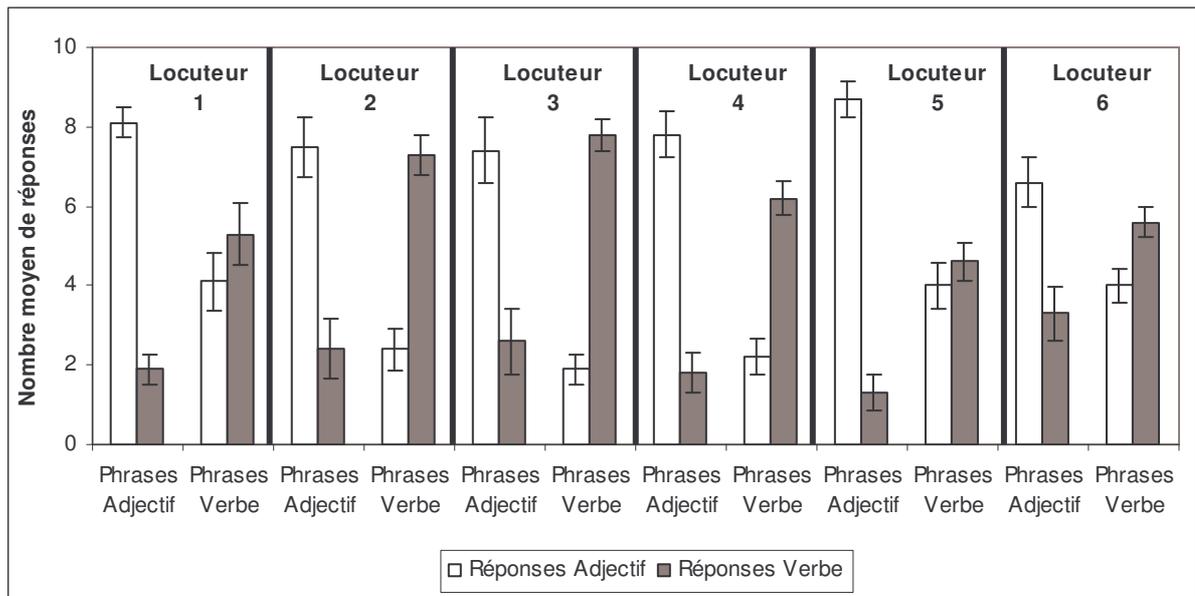
facteur « Locuteur » mais uniquement dans l'analyse par sujet ( $F_1(1,48)=3$ ,  $p=.02$ ;  $F_2(5,95)=1.8$ ,  $p=0.1$ )<sup>23</sup>.

L'interaction cruciale pour notre propos entre la « réponse donnée » et le « type de phrase entendue » est significative ( $F_1(1,48)=179$ ,  $p<.001$ ;  $F_2(1,19)=114$ ,  $p<.001$ ). Elle reflète le fait que les participants ont donné plus de réponses adjectif aux phrases adjectif qu'aux phrases verbe (7.7 versus 3.1,  $F_1(1,48)=198$ ,  $p<.001$ ,  $F_2(1,19)=124$ ,  $p<.001$ ), alors qu'ils ont donné plus de réponses verbe aux phrases verbe qu'aux phrases adjectif (6.1 versus 2.2,  $F_1(1,48)=143$ ,  $p<.001$ ,  $F_2(1,19)=92$ ,  $p<.001$ ).

Ces analyses ont aussi mis en évidence une interaction significative entre « la réponse donnée », « la phrase entendue » et « le locuteur », mais uniquement dans l'analyse par item ( $F_1(5,48)=2$ ,  $p=0.1$  ;  $F_2(5,95)=2.5$ ,  $p=.04$ ). Il semble y avoir une variabilité au niveau des réponses obtenues pour chaque locuteur. Nous avons donc réalisé une analyse plus détaillée des résultats obtenus, en séparant les locuteurs les uns des autres. Ces résultats sont représentés dans la Figure 8 ci-dessous (voir les annexes B.3.2 et B.4.2 pour les résultats individuels par sujets et par items).

---

<sup>23</sup> Il y a logiquement une certaine variabilité entre les locuteurs, qui semble être due à certains items (certains locuteurs ont peut-être été gênés par certaines phrases...). Mais cet effet n'est pas vraiment robuste, et certaines interactions dues à cette variabilité sont apparues significatives, mais ne sont pas vraiment pertinentes pour notre analyse : « Réponse » x « Locuteur » ( $F_1(5,48)=5$ ,  $p<.001$  ;  $F_2(5,95)=5$ ,  $p<.001$ ), « Réponse » x « Locuteur » x « Contrebalancement » ( $F_1(5,48)=4.2$ ,  $p=.003$ ), et « Réponse » x « Phrase » x « Locuteur » x « Contrebalancement » ( $F_1(5,48)=2.4$ ,  $p=.05$ ). Elles illustrent plutôt le fait que certains items ont été marqués différemment entre les locuteurs et que ces items apparaissaient plus fréquemment dans un bloc que dans l'autre.



**Figure 8 (résultats étude 2) : Nombre moyen de réponses adjectif et verbe données aux phrases adjectif et verbe, pour les différents locuteurs (sur 10 réponses possibles pour chaque phrase). Les barres d’erreurs représentent les erreurs standard de la moyenne.**

Ce qui est primordial pour notre propos est que, malgré la variabilité existant entre les locuteurs, l’interaction cruciale entre « la réponse donnée » et « le type de phrase entendue » est significative pour chaque locuteur. Les analyses statistiques ont confirmé que cette interaction était significative quel que soit le locuteur concerné : pour le locuteur 1 ( $F_1(1,8)=47.6, p<.001$  ;  $F_2(1,19)=15.3, p<.001$ ), pour le locuteur 2 ( $F_1(1,8)=57.2, p<.001$  ;  $F_2(1,19)=72, p<.001$ ), pour le locuteur 3 ( $F_1(1,8)=25.6, p=.001$  ;  $F_2(1,19)=82, p<.001$ ), pour le locuteur 4 ( $F_1(1,8)=34.4, p<.001$  ;  $F_2(1,19)=58.6, p<.001$ ), pour le locuteur 5 ( $F_1(1,8)=25.8, p=.001$  ;  $F_2(1,19)=42.2, p<.001$ ), et pour le locuteur 6 ( $F_1(1,8)=13.5, p=.006$  ;  $F_2(1,19)=6.5, p=.02$ ).

### 3.2.2.3 Discussion

Les résultats de cette expérience indiquent que des adultes français ont été capables d’interpréter correctement deux débuts de phrases ambiguës qui ne différaient que par leur structure prosodique et syntaxique. En effet, les paires de phrases expérimentales qui, jusqu’au mot ambigu inclus, contenaient les mêmes phonèmes, mais qui avaient des structures syntaxiques différentes (et donc des structures

prosodiques différentes), n'ont pas reçu la même interprétation, et n'ont donc pas généré la même analyse syntaxique de la part des participants. Ceux-ci ont assigné des catégories syntaxiques différentes aux mots ambigus, en adéquation avec le contexte prosodique dans lequel ils apparaissaient : ils ont donné plus de réponses adjectif que de réponses verbe quand ils traitaient les phrases ambiguës adjectif, et plus de réponses verbe que de réponses adjectif quand ils traitaient les phrases ambiguës verbe.

Le point crucial de cette expérience est que ce résultat a été répliqué pour chaque locuteur : quel que soit le locuteur qui ait produit les phrases qu'ils traitaient, les participants ont toujours réussi à utiliser les seules informations prosodiques pour désambiguïser les débuts des phrases ambiguës.

Ces locuteurs étaient des sujets naïfs et non-experts : ils ont spontanément et naturellement produit des indices de frontières de groupes phonologiques particulièrement bien marqués et assez saillants pour que des auditeurs puissent ensuite les utiliser pour guider leur analyse des phrases. Nous avons comparé les résultats obtenus dans cette expérience avec ceux obtenus dans la première expérience de complétion utilisant une prosodie informative (expérience 1 de l'étude 1) : il n'y aucune différence significative entre ces résultats. L'interaction cruciale entre la réponse donnée par les sujets et la phrase qu'ils entendaient est aussi forte dans les deux expériences ( $F_s < 1$ ). Il apparaît donc que nos locuteurs naïfs et non-experts ont été aussi informatifs et ont produit d'aussi bons indices prosodiques que notre première locutrice (qui pourtant était parfaitement au courant du problème d'ambiguïté et qui avait comme tâche de produire les meilleurs indices prosodiques désambiguïsants).

Certains pourraient peut-être émettre l'hypothèse que nos locuteurs n'étaient pas aussi naïfs qu'ils le prétendaient, et qu'ils s'étaient peu à peu rendus compte de la présence de ces phrases ambiguës lors de leur séance d'enregistrement. Pour répondre à cette critique, nous avons réalisé une dernière analyse. Nous avons séparé les phrases ambiguës en deux groupes : celles qui avaient été produites pendant la première moitié des enregistrements, et celles qui avaient été produites pendant la deuxième moitié des enregistrements. Si les locuteurs s'étaient rendus compte du problème d'ambiguïté syntaxique des phrases, alors ils devraient avoir produit de meilleurs indices prosodiques dans la deuxième partie des enregistrements. L'interaction cruciale de nos

analyses devraient dans ce cas être plus importante pour les phrases produites dans la deuxième partie des séances d'enregistrement.

Nous avons donc repris nos analyses en ajoutant un nouveau facteur : le moment d'enregistrement des phrases. L'interaction entre « la réponse donnée » et « la phrase entendue » s'est révélée significative quel que soit le moment où ont été enregistrées les phrases (phrases enregistrées pendant la première moitié des séances :  $F_1(1,59)=127$ ,  $p<.001$  ;  $F_2(1,9)=12.5$ ,  $p=.007$  ;  $\text{min}F'(1,11)=11.4$ ,  $p=.006$  / phrases enregistrées pendant la deuxième moitié des séances :  $F_1(1,59)=113$ ,  $p<.001$  ;  $F_2(1,9)=8.9$ ,  $p=.02$ ,  $\text{min}F'(1,10)=8.3$ ,  $p=.02$ ). Par contre, cette interaction ne s'est pas révélée différente, que les phrases aient été produites au début ou à la fin des séances d'enregistrement ( $F_1<1$  ;  $F_2(1,9)=1.2$ ,  $p=0.3$ ).

Nous pouvons donc conclure que nos locuteurs étaient complètement naïfs quant au but de nos expériences, et qu'ils ont produit spontanément des indices prosodiques particulièrement désambiguïsants tout au long de leur séance d'enregistrement.

Les résultats que nous avons trouvés dans cette expérience sont donc en faveur des idées développées, par exemple, par Kraljic et Brennan (2005), et vont à l'encontre de celles défendues par Allbritton et al. (1996). Il semble bien que même des locuteurs non-experts et non-informés puissent produire suffisamment d'indices prosodiques pour que des auditeurs puissent inférer correctement le sens des phrases qu'ils entendent.

Pourtant, on aurait pu penser que nous ne nous étions pas mis dans les meilleures conditions pour faire émerger ces indices de frontières de groupes phonologiques. En effet, la tâche utilisée pour faire produire nos phrases expérimentales était une tâche de lecture. Et une des explications qui avait été fournie pour expliquer les résultats divergents sur cette question de la production spontanée d'indices prosodiques, portait sur la tâche de production utilisée (avec l'idée qu'il était nécessaire de ne pas lire les phrases, et d'avoir un interlocuteur ainsi qu'un but communicatif évident pour que ces indices puissent être produits de manière importante). Or, même lors d'une tâche de lecture, nos locuteurs ont produit de très bons indices prosodiques.

Notre hypothèse, selon laquelle les stimuli utilisés dans ces différentes expériences étaient une des explications à ces résultats divergents, semble être validée. Ainsi, des phrases suffisamment longues pour être découpées en au moins deux unités (c'est-à-dire plus de sept syllabes) vont favoriser la production de frontières prosodiques. La position de ces frontières dans les phrases dépendra alors du sens à

communiquer, et donc de la structure syntaxique des phrases. Les locuteurs placeront donc naturellement une frontière prosodique au niveau de la frontière syntaxique entre le groupe nominal sujet et le verbe. Ces frontières pourront être utilisées par les auditeurs pour guider leur analyse syntaxique des phrases.

Nous avons montré que des locuteurs naïfs produisaient spontanément et naturellement des indices prosodiques assez saillants pour être utilisés de manière à résoudre des ambiguïtés syntaxiques temporaires, et à trouver le sens des phrases entendues. Il nous semble important de noter que ce ne sont pas uniquement les indices très marqués qui peuvent être utilisés, comme la présence de pauses plus ou moins longues dans le signal de parole. Les pauses sont toujours présentes dans les expériences qui étudient le rôle des frontières prosodiques dans l'analyse syntaxique des phrases. Mais, nous avons justement contrôlé qu'aucune pause n'apparaissait au niveau de la zone ambiguë étudiée, donc entre le groupe nominal sujet et le verbe. Les allongements et les variations de fréquence fondamentale ont donc été suffisants pour guider l'analyse syntaxique de nos phrases ambiguës.

Enfin, nous voulons aussi rappeler qu'un grand soin avait été pris pour choisir les stimuli de nos expériences, de manière à éviter le plus possible la présence de biais de fréquence et de plausibilité. Une analyse a été réalisée pour vérifier le rôle qu'auraient pu jouer ces indices dans les réponses des sujets (en utilisant les questionnaires remplis par dix adultes français qui évaluaient la fréquence intrinsèque de chaque item, ainsi que sa plausibilité dans la phrase) : aucune corrélation significative n'est ressortie, que ce soit entre l'estimation de la fréquence des items et les résultats obtenus dans l'expérience de complétion ( $r = 0.18$ ,  $p = 0.3$ ), ou entre l'estimation de la plausibilité des items dans les phrases et les résultats obtenus dans l'expérience de complétion ( $r = 0.21$ ,  $p = 0.3$ ).

Cette première partie de notre recherche a donc mis en évidence le fait que, quand des indices prosodiques de frontières de groupes phonologiques sont bien marqués et présents dans le signal de parole, des auditeurs français peuvent les utiliser pour guider leur analyse syntaxique des phrases et résoudre des ambiguïtés syntaxiques temporaires. Nous l'avons aussi bien montré dans des tâches « off-line » (complétion de phrases – expérience 1 de l'étude 1 et étude 2), que dans une tâche réalisée en temps

réel (détection de mot abstrait – expérience 2 de l'étude 1). Cette dernière indique aussi que les indices prosodiques semblent être utilisés dès les premières étapes de l'analyse syntaxique, avant que d'autres informations ne soient disponibles pour désambiguïser les phrases étudiées. Nous avons aussi montré que la production de ces indices se faisait spontanément, même par des locuteurs naïfs et non-experts, et était étroitement liée à la structure syntaxique des phrases.

Ces résultats, en plus d'apporter de nouvelles données à intégrer aux modèles de traitement du langage par les adultes, constituent une base de départ pour de nouvelles recherches, concernant maintenant les enfants. En effet, nous avons démontré que les adultes, qui peuvent pourtant mobiliser leurs différentes connaissances (phonologiques, lexicales, sémantiques, syntaxiques, pragmatiques...) pour analyser les phrases qu'ils entendent, utilisent aussi, de manière importante, les indices prosodiques présents dans les phrases. Puisque les frontières de groupes phonologiques sont calculées en temps réel par les adultes pour contraindre l'analyse syntaxique des phrases qu'ils entendent, il semble plausible de postuler que les enfants qui acquièrent le langage pourraient eux aussi les utiliser pour acquérir, par exemple, certaines notions sur la syntaxe de leur langue maternelle. C'est ce que nous allons maintenant aborder dans la deuxième partie de cette recherche.

## CHAPITRE 4

# ACQUISITION DE LA SYNTAXE DE SA LANGUE MATERNELLE

Acquérir le langage n'est pas une tâche facile. Les enfants doivent en effet apprendre la phonologie de leur langue maternelle (c'est-à-dire découvrir quels sont les sons pertinents dans cette langue), trouver les mots dans les phrases qu'ils entendent (c'est-à-dire réussir à segmenter le signal continu de parole en mots), découvrir le sens de ces mots (pour construire leur lexique mental), et comprendre la syntaxe de cette langue (c'est-à-dire trouver les règles qui permettent de calculer le sens des phrases à partir du sens des mots). L'apprentissage de chacune de ces composantes du langage peut être facilité si l'on suppose des connaissances préalables au niveau des autres composantes. Par exemple, Lila Gleitman (1990) a montré que découvrir le sens des mots pouvait être facilité si l'on avait accès à certains aspects de la structure syntaxique des phrases ; mais puisque la syntaxe définit les relations entre les mots dans les phrases, il semble logique de postuler qu'il est nécessaire d'avoir accès aux mots et à leur signification pour pouvoir apprendre la syntaxe de sa langue maternelle. Ceci met en évidence un problème de circularité dans l'acquisition du langage : les enfants auraient besoin du lexique pour apprendre la syntaxe, tandis que la syntaxe serait nécessaire pour construire le lexique. Ce phénomène est connu sous le nom de problème d'initialisation.

Nous allons centrer notre attention sur une des composantes du langage que les enfants doivent acquérir : la syntaxe. Pour ce faire, les enfants doivent entre autres (1) trouver quelles sont les catégories syntaxiques existant dans leur langue, (2) associer les mots de leur langue maternelle à ces catégories syntaxiques, et (3) comprendre comment sont organisées les catégories syntaxiques dans les phrases.

Comment les enfants découvrent-ils et apprennent-ils quelles sont les catégories grammaticales de leur langue maternelle ? Le premier postulat, proposé pour répondre à cette question, était que cette connaissance faisait partie d'un savoir linguistique inné, ce qui implique que les enfants savent qu'il y a des noms et des verbes dans le langage

qu'ils entendent. En effet, selon Chomsky (1957), toutes les langues humaines reposeraient sur un ensemble de règles de base, dont les principes seraient transmis génétiquement, et qui prédisposeraient l'enfant à acquérir le langage. Ces règles représenteraient des universaux, communs à toutes les langues du monde. D'après la théorie des principes et des paramètres (Chomsky, 1981), la grammaire d'une langue est définie par (1) un ensemble fixe de principes universels et innés et (2) un ensemble de paramètres dont la valeur dépend de la langue. Les principes font donc partie de l'héritage commun de l'espèce, alors que les paramètres doivent être réglés à la valeur qui convient pour la langue d'exposition au cours du développement du langage.

On pourrait postuler, comme Pinker (1984), que les grandes catégories syntaxiques font partie de ces connaissances universelles. Cependant, même en admettant ce postulat de base, il reste le problème de savoir comment les enfants font pour assigner une de ces catégories syntaxiques innées aux mots qu'ils découvrent dans les phrases. Pinker (1984; 1987) a proposé que les enfants utiliseraient le sens des mots pour inférer leur catégorie syntaxique (les objets étant généralement désignés par des noms, et les actions par des verbes). Cette proposition se heurte à la difficulté d'acquisition du sens des mots mentionnée plus haut (Gleitman, 1990). Mais comme nous allons le voir, il existe maintenant d'autres propositions, plus récentes et plus viables, pour résoudre ce problème.

## 4.1 INITIALISATION DISTRIBUTIONNELLE POUR L'ACQUISITION DES CATEGORIES GRAMMATICALES

### 4.1.1 *Approche théorique*

Cette approche postule que les catégories grammaticales pourraient être découvertes sur la base d'informations distributionnelles, que celles-ci soient d'ordre lexical (distribution des mots entre eux) ou d'ordre morphologique (répartition des morphèmes grammaticaux, comme les marqueurs de conjugaison). Maratsos (1982; 1980) a proposé que les catégories grammaticales pouvaient être définies et découvertes grâce à des similarités au niveau des contextes d'occurrence des mots. Un nom serait

défini par le fait qu'il apparaît avec des articles ou des déterminants, tandis qu'un verbe serait défini par le fait qu'il est en co-occurrence avec des pronoms ou des auxiliaires. Maratsos et Chalkley (1980) ont noté que les racines des mots anglais, qui peuvent prendre comme suffixe « -ed », peuvent en général avoir comme suffixe « -s », et sont alors des verbes. A l'inverse, les mots qui peuvent avoir « -s » en suffixe mais pas « -ed » sont en général des noms.

Pinker (1987) s'était élevé contre cette idée en mettant en avant diverses situations où cette stratégie entraînerait des erreurs d'interprétations. Dans certains cas, cette approche pourrait par exemple entraîner des généralisations erronées. Par exemple, étant donné les phrases (1a) et (1b) ci-dessous, une analyse distributionnelle pourrait entraîner à catégoriser « fish » et « rabbits » comme appartenant à la même catégorie syntaxique ; un enfant entendant ensuite la phrase (1c) pourrait en déduire, de manière erronée, que la phrase (1d) est aussi correcte.

- (1) a. John ate fish.  
b. John ate rabbits.  
c. John can fish.  
d. \* John can rabbits.

Cette situation est due au fait que certains mots sont homophones dans une langue, et qu'ils peuvent appartenir à différentes catégories syntaxiques et donc être utilisés dans des contextes syntaxiques différents : c'est le cas de « fish » en anglais, qui peut être soit un nom comme dans la phrase (1a), soit un verbe comme dans la phrase (1c). Ces exemples ne sont pas rares et existent dans différentes langues (comme par exemple en français où le mot « porte » peut être un nom ou un verbe, où le mot « distrait » peut être un adjectif ou un verbe, et où le mot « ferme » peut être à la fois un verbe, un nom et un adjectif). Selon Pinker, ces généralisations erronées seraient très communes et seraient donc un argument en défaveur de cette approche. Une autre difficulté pour cette approche est que d'importantes régularités distributionnelles sont non-adjacentes, donc ne se produisent pas localement (Chomsky, 1965; Pinker, 1987).

Ces contre-exemples existent mais ne sont peut-être pas suffisants pour infirmer complètement cette hypothèse d'initialisation distributionnelle. La meilleure manière de

le prouver était d'apporter une validation expérimentale à ces réflexions. Plusieurs études, en général des expériences de faisabilité avec des analyses de corpus de parole, ont en effet montré que ces problèmes potentiels ne réduisaient pas à néant le fait que les indices distributionnels pouvaient être informatifs pour l'acquisition de la syntaxe et la mise en place des catégories syntaxiques (Finch & Chater, 1992, 1994; Mintz, 2002, 2003; Mintz, Newport, & Bever, 2002; Redington, Chater, & Finch, 1998).

#### 4.1.2 *Analyses de corpus*

##### 4.1.2.1 Rôle des mots adjacents

Redington, Chater et Finch (1998) ont démontré que la distribution des mots entre eux pouvait être très informative pour trouver la catégorie syntaxique des mots. Leur analyse d'un corpus de parole adressée aux enfants (extrait de la base de données CHILDES, Mac Whinney & Snow, 1985) reposait sur trois étapes : (1) mesurer les contextes d'apparition des mots cibles, (2) comparer les contextes d'apparition de paires de mots, et (3) regrouper les mots qui ont des contextes d'apparition similaires. L'idée de base est que les mots qui apparaissent dans des contextes linguistiques similaires (qui tendent à être précédés et suivis par les mêmes mots) ont une grande probabilité d'appartenir à la même catégorie syntaxique. Les auteurs ont donc évalué les contextes d'apparition des 1000 mots les plus fréquents dans leur corpus. Les mots de contexte considérés étaient les 150 mots les plus fréquents du corpus et la fenêtre de contexte considérée était une fenêtre de quatre mots autour du mot test (deux avant, un avant, un après et deux après le mot cible). Pour chaque mot, sa co-occurrence avec les 150 mots de contexte a été comptabilisée, et ce pour les quatre positions possibles considérées (les auteurs ont donc réalisé des statistiques sur des bigrammes, c'est-à-dire des mesures de co-occurrences entre des paires de mots dont la position est fixe). Les quatre vecteurs obtenus pour chaque mot test (un par contexte) ont été combinés pour produire un vecteur à 600 dimensions. Les vecteurs de co-occurrence résultant ont été comparés et groupés ensemble selon leur similarité.

Deux mesures étaient ensuite calculées pour évaluer la performance du modèle :

(1) l'« accuracy » (que l'on pourrait traduire par « homogénéité ») mesure le succès de la catégorisation, et représente la proportion de paires de mots qui sont groupées ensemble et qui appartiennent à la même catégorie syntaxique. Elle est calculée par le nombre de succès (« hits » = les deux items d'une paire de mots sont de la même catégorie syntaxique) divisé par le nombre total de paires groupées par la simulation (somme des « hits » et des « fausses alarmes », ces dernières étant caractérisées par le mauvais regroupement de deux items qui en fait n'appartiennent pas à la même catégorie syntaxique). C'est cette mesure que nous prendrons comme référence pour comparer les différents modèles.

(2) la « completeness » (ou « complétude ») représente la proportion de mots qui appartiennent à la même catégorie et qui sont effectivement regroupés ensemble dans l'analyse. Elle est calculée en divisant le nombre de succès (« hits ») par la somme des succès et des « oublis » (une paire est comptée comme « oubli » si ses membres appartiennent à la même catégorie syntaxique mais ne sont pas regroupés ensemble par l'analyse).

Les résultats de la classification effectuée par Redington et al. (1998) ont été relativement bons puisque la mesure d'« accuracy » atteint 72%, ce qui est significativement supérieur à ce qui aurait été obtenu par un classement aléatoire, soit 27%. Des résultats similaires ont été trouvés par Mintz, Newport et Bever (2002), qui avaient conduit leurs analyses sur des corpus de parole plus petits.

#### 4.1.2.2 Comparaison de paires minimales de phrases

Dans le modèle proposé par Cartwright et Brent (1997), les phrases complètes étaient prises comme le contexte d'un mot cible. Ils recherchaient des paires minimales de phrases, qui différaient entre elles au niveau d'un seul mot ; ceci leur permettait de regrouper des mots dans différentes catégories syntaxiques, et de représenter la structure syntaxique abstraite dans laquelle ces mots apparaissaient. Par exemple, les phrases « I saw the cat » et « I saw the dog », permettaient de regrouper les mots « cat » et « dog » dans la même catégorie syntaxique (catégorie nom « N ») et d'isoler le cadre syntaxique correspondant « I saw the N ». Leur modèle a obtenu 68% d'« accuracy » (comparé à 23% pour un classement aléatoire).

#### 4.1.2.3 Rôle des cadres lexicaux

Mintz (2003) a décidé de regarder d'autres contextes et de ne plus prendre en compte des bigrammes, mais des trigrammes ou des « cadres fréquents » (Mintz les appelle des « frequent frames »). Par exemple, si l'on considère le mot W dans l'environnement « ... X W Y... », la probabilité de co-occurrence qui sera calculée sera celle que le mot W apparaisse en étant à la fois précédé de X et suivi par Y. Les autres approches présentées ci-dessus (Mintz et al., 2002; Redington et al., 1998) ne calculaient que des co-occurrences indépendantes, à savoir celle que W suive X, puis celle que W précède Y. Les cadres qui sont donc considérés dans cette étude peuvent être définis de la manière suivante : deux mots qui sont placés dans un ordre fixe, qui apparaissent ensemble fréquemment, et qui sont séparés par un emplacement libre pouvant être occupés par un autre mot. Pour un cadre considéré, tous les mots qui apparaissent dans cette place libre sont catégorisés ensemble.

L'analyse de Mintz a été réalisée sur six corpus anglais de la base de données Childes (Mac Whinney, 2000) et n'a pris en compte que les 45 cadres les plus fréquents. La classification des mots apparaissant au centre de ces cadres fréquents a été réalisée avec une « accuracy » moyenne de 93% (comparé à 47% pour une classification au hasard). Il s'agit là d'une excellente performance, qui montre que les distributions de trigrammes dans le langage, quand le mot cible est en position centrale, semblent être très pertinentes pour la catégorisation syntaxique des mots. Par exemple, un cadre fréquent très informatif en anglais est le cadre « you X it » dans lequel 100% des X sont des verbes. Ces cadres lexicaux semblent présents et informatifs dans plusieurs langues, puisque les résultats de Mintz ont par exemple été répliqués en français par Chemla (2005). Et par exemple, un cadre fréquent très informatif en français pour la catégorie verbe est le cadre « ne X pas » dans lequel X est toujours un verbe.

Les résultats de toutes ces études indiquent que les patterns distributionnels, présents dans la parole adressée aux enfants, sont des bases potentiellement informatives pour faire une première catégorisation syntaxique des mots. Il semble en effet plausible et possible que les enfants puissent utiliser les relations plus ou moins

adjacentes entre les mots (informations apportées par les bigrammes, les cadres lexicaux...) pour inférer la catégorie grammaticale de certains mots.

Les différentes théories que nous avons développées jusqu'à présent, expliquant l'acquisition des catégories grammaticales, reposent donc sur des informations lexicales (informations apportées par les mots eux-mêmes, que ce soit au niveau de leur sens ou de leur position et distribution dans les phrases). Il semble important de noter qu'un des pré-requis de ces hypothèses de travail est que les enfants aient déjà accès aux mots (ou tout du moins à certains mots), et qu'ils aient donc déjà résolu le problème de la segmentation de la parole continue. Or trouver les mots dans les phrases est une tâche difficile que les enfants ne commencent à réaliser que vers l'âge de 7.5 mois (en ce qui concerne les données des enfants américains, Jusczyk, 1999; Jusczyk & Aslin, 1995).

Pour les chercheurs défendant ces approches distributionnelles, les propriétés physiques de la parole, qui ne concernent pas les séquences de mots, n'ont que peu de choses à voir avec la question de l'acquisition de la grammaire. Or, nous allons maintenant aborder d'autres courants de pensée, dans lesquels sont mises en avant ces informations de plus bas niveau, directement accessibles depuis le signal de parole : les modèles d'initialisation phonologique (dont un sous-ensemble est formé par les modèles d'initialisation prosodique). Les partisans de ces modèles défendent l'idée que les propriétés physiques de la parole (indices acoustiques, phonologiques, prosodiques) ne sont pas seulement du bruit, non pertinent pour les enfants, et qu'elles ne doivent pas être ignorées dans les modèles d'acquisition du lexique et de la syntaxe. L'hypothèse d'initialisation phonologique a donc été développée (Morgan, 1986 ; Morgan & Demuth, 1996a), et propose qu'une analyse purement phonologique du signal de parole pourrait aider les enfants à acquérir le lexique et la syntaxe de leur langue maternelle.

## 4.2 INITIALISATION PHONOLOGIQUE POUR L'ACQUISITION DES CATEGORIES SYNTAXIQUES

### 4.2.1 *Différence entre les mots de contenu et les mots grammaticaux*

Toutes les langues font une distinction entre les mots de contenu et les mots grammaticaux (Abney, 1987 ; cité par Morgan & Demuth, 1996a). Les mots de contenu (aussi appelés « mots de la classe ouverte » ou « open-class words ») sont constitués des noms, des verbes, des adjectifs, des adverbes... Ce sont les mots qui transmettent le sens. Les mots grammaticaux (aussi appelés « mots de la classe fermée » ou « closed-class words ») sont les articles, les pronoms, les auxiliaires, les prépositions... Ils sont considérés comme apportant la structure des phrases. Ces mots grammaticaux semblent importants dans la catégorisation syntaxique des items lexicaux : intuitivement, si l'on entend « une jaurime », on pensera que « jaurime » est un nom.

Les analyses distributionnelles présentées dans la partie précédente ont mis en évidence l'importance des relations existant entre les mots grammaticaux et les mots de contenu (comme le fait que les noms soient généralement précédés d'articles...). De plus, les études sur les cadres lexicaux les plus fréquents impliquaient généralement des mots grammaticaux (par exemple, le cadre « the X is » dans lequel X est toujours un nom, ou « you X it » dans lequel X est toujours un verbe, Mintz, 2003). Il serait donc très intéressant pour les jeunes enfants de disposer d'informations pertinentes pour extraire et apprendre ces mots grammaticaux. Certains auteurs ont émis l'hypothèse qu'ils pourraient effectivement être appris de manière précoce (Christophe, Guasti, Nespor, Dupoux, & van Ooyen, 1997), car ils ont des propriétés particulières qui les différencient des mots de contenu.

#### 4.2.1.1 Analyses de corpus

Différents auteurs ont noté une différence entre la catégorie des mots grammaticaux et celles des mots de contenu au niveau accentuel : les mots de contenu tendent à être des mots accentués, alors que les mots grammaticaux, en général, ne le sont pas (Kelly, 1992; Shi, Morgan, & Allopenna, 1998). Gleitman et Wanner (1982) suggèrent que cette différence d'accentuation pourrait être à la base de l'acquisition de cette distinction par les enfants : les mots qui ne recevraient pas d'accentuation seraient initialement interprétés comme des mots grammaticaux.

Un autre indice qui semble particulièrement important pour distinguer ces deux classes de mots est la qualité des voyelles : les mots de contenu contiennent

généralement des voyelles pleines, en particulier au niveau des syllabes initiales et accentuées, tandis que les mots grammaticaux ont plutôt tendance à contenir des voyelles réduites (Cutler, 1993).

Dans les phrases, les mots qui ont une durée importante, une fréquence fondamentale et une amplitude importantes, ont peu de chances d'être des mots grammaticaux. Au contraire, ces derniers sont relativement courts et sont de basse amplitude.

Il faut aussi noter que ces deux catégories se différencient par leur fréquence : en effet, les mots grammaticaux ont tendance à apparaître très fréquemment, plus que les mots de contenu (par exemple, Morgan et al., 1996, citent la base de données de Kucera et Francis, 1967, dans laquelle les 50 mots anglais les plus fréquents sont des mots grammaticaux).

Morgan et collaborateurs (Morgan et al., 1996) ont analysé un corpus de parole adressée à des enfants américains et mandarins de 12 mois (3 dyades mère-enfant américains, 2 dyades mère-enfant mandarins). Ils ont réalisé diverses analyses pour comparer les mots grammaticaux et les mots de contenu, à différents niveaux :

(1) une première série de mesures, faite sur la retranscription écrite du corpus, concernait les indications de fréquence, la position des items dans les phrases et le nombre de syllabes.

(2) une deuxième série de mesures, faite sur les transcriptions phonétiques, concernait la complexité des syllabes (avec une échelle allant des syllabes les plus simples – sans attaque ni coda – aux syllabes les plus complexes – suite de consonnes en position initiale et finale), la diphtongaison des voyelles, et la prédictabilité des tons (seulement pour le corpus mandarin).

(3) une troisième série de mesures était réalisée sur les représentations acoustiques et concernait les durées des voyelles, l'amplitude des syllabes, les changements de fréquence fondamentale et la qualité des voyelles (ce dernier point ne concerne que le corpus américain).

Les auteurs ont trouvé que les mots de contenu et les mots grammaticaux se différenciaient sur toutes ces mesures pour tous les corpus étudiés : par exemple, la durée moyenne des voyelles des mots de contenu des deux langues étudiées est plus importante que la durée des voyelles des mots grammaticaux. Pour toutes ces mesures, la distribution des valeurs était assez similaire entre les locuteurs et entre les langues.

Par contre, les auteurs ont constaté un chevauchement d'une grande partie des données, entre les mots grammaticaux et les mots de contenu. Chacun de ces indices, pris isolément, ne semble donc pas être un indice complètement informatif. Par contre, une analyse intégrant tous les indices a permis une bonne discrimination entre ces deux catégories de mots (les auteurs ont classifié correctement entre 83 et 91% des mots grammaticaux et des mots de contenu).

#### 4.2.1.2 Données expérimentales sur l'utilisation de ces informations

##### Sensibilité aux mots grammaticaux :

On a longtemps cru que les enfants n'étaient que très peu sensibles aux mots grammaticaux, dans la mesure où ils ne les produisent que très tardivement. Mais de nombreuses recherches ont contredit cette idée, montrant par exemple que les jeunes enfants, et même les nouveau-nés, étaient sensibles à la différence entre les mots grammaticaux et les mots de contenu.

Shi, Werker et Morgan (1999) ont montré que des nouveau-nés pouvaient distinguer la catégorie des mots grammaticaux de la catégorie des mots de contenu. Ils ont réalisé une expérience dans laquelle ils mesuraient le taux de succion des nourrissons<sup>24</sup>. Ces derniers étaient d'abord habitués à écouter une liste de mots grammaticaux ou une liste de mots de contenu. Puis, après habitude, les enfants étaient testés soit sur de nouveaux exemplaires de la catégorie qu'ils venaient d'entendre (groupe contrôle), soit sur des exemplaires de la catégorie opposée (groupe expérimental). Les auteurs ont trouvé que le taux de succion des enfants remontait beaucoup plus quand ils étaient testés sur des mots de la catégorie opposée à celle qu'ils avaient entendue en habitude. Les nourrissons étaient plus sensibles à la différence existant entre des mots grammaticaux et des mots de contenu, qu'à la différence existant entre des exemplaires différents d'une même catégorie. De plus, ces résultats ont été répliqués avec les mêmes stimuli, que les enfants soient de langue maternelle anglaise ou pas. Ces résultats suggèrent donc que les enfants naissent avec une sensibilité

---

<sup>24</sup> Paradigme « High Amplitude Sucking » ou paradigme de succion de haute amplitude

perceptive qui leur permet de détecter les caractéristiques acoustiques et phonologiques distinguant les mots grammaticaux des mots de contenu.

D'autres recherches ont montré que les enfants devenaient rapidement sensibles aux mots grammaticaux de leur langue maternelle, même quand ceux-ci se trouvaient dans des phrases. Par exemple, Shady (1996) a montré que les enfants américains de 10.5 mois préféraient écouter une histoire contenant des morphèmes grammaticaux intacts, plutôt qu'une histoire dans laquelle les mots grammaticaux étaient remplacés par des syllabes n'ayant aucun sens dans la langue de l'enfant (par exemple, les mots grammaticaux « was », « the », « a »... étaient remplacés par les syllabes « ki », « bu », « po »...). Shafer et al. (1998) ont également observé, à l'aide d'un protocole similaire adapté aux potentiels évoqués, que les réponses cérébrales de nourrissons de 11 mois différaient entre les deux histoires.

A l'heure actuelle, de nombreuses recherches ont montré que les enfants reconnaissaient les mots grammaticaux de leur langue maternelle : des résultats existent pour les enfants américains de 10.5 mois et 13 mois (Shady, 1996; Shafer, Shucard, Shucard, & Gerken, 1998; Shi, Werker, & Cutler, 2003), les enfants allemands de 8 mois (Höhle & Weissenborn, 2003), et les enfants français de 8 mois (Shi & Gauthier, 2005).

#### Utilisation des mots grammaticaux pour segmenter les phrases :

En plus de montrer que les enfants connaissaient les mots grammaticaux de leur langue, de nombreuses recherches ont mis en évidence le fait que les enfants utilisaient ces items lexicaux pour traiter le langage qu'ils entendaient, comme par exemple pour identifier et segmenter les mots suivant les mots grammaticaux (Gerken & McIntosh, 1993; Höhle & Weissenborn, 2000, 2003; Shi et al., 1998; Shi et al., 2003; Zangl & Fernald, 2003).

Par exemple, Gerken et McIntosh (1993) ont réalisé une expérience dans laquelle les enfants devaient montrer une image face à eux après avoir entendu un stimulus auditif. Les phrases utilisées étaient des séquences « article + nom », qui contenaient soit un mot grammatical correct (par exemple, « the » dans « show me the car »), soit une syllabe qui n'était pas un article existant dans la langue étudiée (comme

« gub » dans « show me gub car »). Les auteurs ont montré que les enfants de 28 mois étaient sensibles au contexte grammatical dans lequel les noms apparaissaient, leur performance devenant moins bonne quand ils devaient identifier des noms précédés par un mot grammatical anormal (des résultats similaires ont été trouvés à 24 mois par Kedar, Casasola, & Lust, 2004).

Zangl et Fernald (2003) ont enregistré les mouvements oculaires d'enfants de 18 mois, pendant qu'ils regardaient des images face à eux et qu'ils entendaient des phrases nommant ces images. Soit les phrases contenaient un déterminant correct (« look at the shoe »), soit elles contenaient un déterminant incorrect (« look at loo shoe »). Les auteurs ont trouvé que les enfants avaient plus de difficultés pour identifier le mot familier (« shoe ») quand il était précédé par un déterminant n'ayant aucun sens (temps d'orientation vers l'image correspondante plus lent). Dans l'autre situation, ils ont donc pu utiliser leur connaissance de l'article « the » pour segmenter et reconnaître plus facilement le nom suivant.

D'autres expériences ont été réalisées avec des protocoles expérimentaux différents et ont trouvé des résultats similaires. Par exemple, Shi et collaborateurs (Shi, 2005; Shi et al., 2003; Shi, Werker, & Cutler, review) ont familiarisé des enfants américains de 8 mois avec un mot de contenu qui, dans une première condition, était précédé par un vrai déterminant de haute fréquence (« the »); dans une deuxième condition, le mot de contenu était précédé par une syllabe ayant les mêmes caractéristiques phonologiques et prosodiques que le déterminant utilisé dans la première condition, mais n'appartenant pas à la classe des mots grammaticaux (« ke »). Dans une phase de test, les auteurs mesuraient le temps d'écoute des enfants pour le mot de contenu présenté en isolation. Les résultats ont indiqué que les enfants préféraient écouter le mot de contenu qui avait été initialement présenté avec le mot grammatical correct de haute fréquence. Les enfants américains de 8 mois sont donc capables de segmenter et de se rappeler d'un mot nouveau quand celui-ci est précédé par un mot grammatical fréquent comme l'article « the ». Par contre, cet effet n'a été répliqué avec des mots grammaticaux de basse fréquence que vers l'âge de un an (exemple du déterminant « its »).

En français, les mots grammaticaux apparaissent dans des positions relativement constantes dans les phrases (en début d'unité prosodique), et sont des syllabes plus saillantes qu'en anglais. Il se pourrait donc qu'ils soient exploités et reconnus plus précocement. Shi (2005; Shi & Gauthier, 2005) a réalisé une expérience similaire à la précédente, pour tester l'utilisation des mots grammaticaux dans la segmentation lexicale chez des enfants français québécois de 8 mois. Dans une première phase, les enfants étaient habitués à des séquences comme « des preuves » et « kes sangles », puis ils étaient testés sur les mots en isolation « preuves » et « sangles ». Les auteurs ont trouvé que, pendant la phase de test, les enfants écoutaient plus longtemps le nom « preuves » (qui avait été présenté pendant la première phase avec un vrai mot grammatical de haute fréquence) que le nom « sangles » (qui était apparu avec un non-mot grammatical). Les résultats ont été similaires quand ils utilisaient des mots grammaticaux de basse fréquence (comme « mes »). On peut donc conclure de ces résultats que les enfants québécois de langue maternelle française peuvent utiliser les mots grammaticaux de leur langue maternelle pour segmenter et retenir plus facilement des mots nouveaux.

#### Utilisation des mots grammaticaux pour inférer la catégorie grammaticale des mots de contenu :

Des expériences plus récentes se sont intéressées, non plus au rôle des mots grammaticaux dans la segmentation lexicale, mais à leur rôle dans l'attribution des catégories syntaxiques. Les enfants peuvent-ils inférer la catégorie syntaxique « nom » après avoir entendu un article, et inférer la catégorie syntaxique « verbe » après avoir entendu un pronom ou un auxiliaire ? Pour répondre à cette question, Höhle et collaborateurs (2004) ont familiarisé des enfants allemands de 12 et 16 mois avec des pseudo-mots, présentés soit comme des noms (donc avec un article indéfini comme « ein » dans « ein glamm »), soit comme des verbes (donc avec un pronom personnel comme « sie » dans « sie glamm »). Dans une phase de test, les enfants entendaient des passages de 6 phrases dans lesquels les pseudo-mots étaient utilisés soit comme des noms, soit comme des verbes. A 16 mois, les enfants familiarisés avec la séquence « article + pseudo-mot » ont ensuite écouté significativement plus longtemps le passage contenant la cible verbe : ceci s'apparente à un effet de nouveauté (« novelty

preference »). Par contre, aucune préférence d'écoute n'est apparue pour les enfants familiarisés avec la séquence 'pronom + pseudo-mot », ni pour les enfants plus jeunes. Des analyses de corpus de parole adressée aux enfants ont montré qu'il y avait une très forte co-occurrence entre la présence d'un déterminant et celle d'un nom, beaucoup plus importante que la relation entre pronom et verbe (ceci est dû au fait que l'allemand est une langue à ordre libre et que les pronoms peuvent être immédiatement suivis par n'importe quelle catégorie syntaxique). Il semble donc qu'en allemand, les déterminants soient des indices plus pertinents que les pronoms pour inférer la catégorie syntaxique d'un nouveau mot adjacent.

De même, Bernal, Lidz, Waxman, Dutat & Christophe (2005) ont montré que les enfants français de 23 mois pouvaient catégoriser un nouveau mot comme un verbe si celui-ci était présenté précédé par un pronom. Pour ce faire, ils ont familiarisé des enfants avec un non-mot (par exemple « dase »), qui apparaissait dans des phrases courtes, précédé d'un pronom (par exemple, « Regarde : elle dase ! Tu as vu : elle dase ! »), et accompagné de la présentation d'un objet animé sur un écran (par l'exemple, l'image d'une pomme qui tourne sur elle-même). Puis, dans la phase de test, les enfants voyaient deux scènes de chaque côté de l'écran : une qui représentait l'action familière (c'est-à-dire la pomme qui tourne sur elle-même), et une qui représentait le même objet réalisant une nouvelle action (par exemple, une pomme qui faisait des mouvements verticaux, de haut en bas). Au même moment, une consigne orale était présentée aux enfants, qui leur demandait de montrer une des deux scènes (dans cet exemple, les enfants entendaient : « laquelle dase ? », « montre-moi celle qui dase ? »). Les enfants, qui avaient été familiarisés avec ce non-mot précédé d'un pronom, ont pointé significativement plus souvent vers la scène jouant l'action familière. Par contre, les enfants d'un groupe contrôle, qui étaient testés sur les mêmes scènes, mais qui avaient été familiarisés avec ce non-mot dans sa forme nominale (« Regarde : une dase ! Tu as vu : c'est une dase ! »), ont montré un effet de préférence à la nouveauté, en pointant significativement plus souvent vers l'objet réalisant la nouvelle action.

Ces deux expériences indiquent donc que les jeunes enfants sont capables d'inférer la catégorie grammaticale d'un nouveau mot, en utilisant l'information apportée par la présence des mots grammaticaux (un mot précédé d'un article sera

plutôt considéré comme un nom, tandis qu'un mot précédé d'un pronom sera plutôt considéré comme un verbe, donc comme représentant une action).

Avant de conclure cette partie, il nous faut encore noter, qu'en plus d'utiliser les informations de co-occurrence entre les mots grammaticaux et les mots de contenu, les enfants pourraient aussi utiliser les informations phonologiques disponibles dans le signal de parole pour inférer la catégorie grammaticale des mots qu'ils entendent. Par exemple, de nombreuses études se sont intéressées aux différences phonologiques existant entre les noms et les verbes, ce que nous allons maintenant aborder.

#### 4.2.2 *Différence entre les noms et les verbes*

##### 4.2.2.1 Analyses de corpus

De nombreuses études ont montré que les deux grandes classes grammaticales, que sont les noms et les verbes, pouvaient être distinguées sur la base d'indices purement phonologiques. Par exemple, un des traits phonologiques bien connu pour distinguer les noms et les verbes anglais est l'accentuation (Sherman, 1975). Il semble y avoir une corrélation importante entre le pattern d'accentuation des mots bisyllabiques anglais et leur catégorie grammaticale. Kelly et Bock (1988) ont analysé l'accentuation de 3000 noms bisyllabiques et de 1000 verbes bisyllabiques : ils ont trouvé que 94% des noms étaient accentués sur leur première syllabe (pattern d'accent trochaïque), tandis que 69% des verbes étaient accentués sur leur deuxième syllabe (pattern d'accent iambique). De plus, dans l'échantillon analysé, ils ont noté que 90% des mots qui portaient l'accent sur leur première syllabe étaient des noms, alors que 85% des mots ayant une accentuation finale étaient des verbes.

Plusieurs expériences ont démontré que des locuteurs anglais avaient une connaissance implicite de ces différences d'accentuation. Par exemple, Kelly et Bock (1988) demandaient à des sujets de lire à voix haute des phrases contenant un non-mot (par exemple « pernew »). Les sujets avaient plutôt tendance à prononcer ce non-mot avec un accent sur la première syllabe quand il était utilisé comme un nom dans les

phrases porteuses (comme dans « The big *pernew* escaped »), comparé aux phrases où ce non-mot était en position verbale (comme dans « The dogs *pernew* dismay »).

Un autre indice pertinent a été trouvé pour différencier les noms des verbes : le nombre des syllabes (Kelly, 1992). Cassidy et Kelly (1991) ont étudié des corpus de parole adressée aux enfants et ont trouvé que les noms avaient tendance à contenir plus de syllabes que les verbes. Plus le nombre de syllabes d'un mot augmentait, plus la probabilité que ce mot soit un nom augmentait, et plus la probabilité que ce mot soit un verbe diminuait (les probabilités que des mots de une, deux, trois et quatre syllabes soient des noms, plutôt que des verbes, étaient de 38%, 76%, 92% et 100%). On pourrait penser que cette relation entre le nombre de syllabes et la catégorie syntaxique des mots pourrait être caractéristique de la parole adressée aux enfants, et ne pas se retrouver dans le langage adulte (ce qui ferait donc de cet indice un indice non pertinent, voire même gênant, pour l'acquisition des catégories grammaticales). Mais, des données adultes montrent qu'eux aussi utilisent cette corrélation et ont tendance à interpréter des mots longs comme des noms plutôt que comme des verbes (Cassidy & Kelly, 1991).

D'autres variables semblent aussi différencier les noms des verbes, mais pour certains auteurs, ces indices prédisent moins bien la classe grammaticale d'un mot que l'accentuation et le nombre de syllabes (voir Kelly, 1992, pour une revue détaillée). Ces indices sont le type de voyelles (en anglais, les noms ont plus de voyelles basses, alors que les verbes ont plus de voyelles hautes, Soreno & Jongman, 1990), le type de consonnes (les consonnes nasales apparaissent plus dans les noms que dans les verbes anglais), la durée des mots (à nombre de syllabes égal, les noms sont plus longs que les verbes, Sorenson, Cooper, & Paccia, 1978), et le nombre de phonèmes (à nombre de syllabes égal, les noms ont tendance à contenir plus de phonèmes que les verbes anglais).

Des simulations ont été réalisées pour tester la pertinence de cette hypothèse et pour vérifier que les indices phonologiques décrits ci-dessus pouvaient permettre de classer correctement les noms et les verbes de différents corpus. Par exemple, Durieux et Gillis (2001) ont pris en compte un certain nombre d'informations phonologiques, considérées comme étant des bons indices pour distinguer les noms des verbes (pattern d'accentuation, nombre de phonèmes, qualité des voyelles et type de

consonnes). Ils ont fait tourner un algorithme d'apprentissage sur une base de données extraite de CELEX en prenant en considération chacun de ces indices, indépendamment les uns des autres ou combinés les uns aux autres. Les performances de catégorisation ont été globalement bonnes : 62% des noms et des verbes ont été correctement classifiés en prenant en compte le pattern d'accentuation, 67% avec la qualité de la voyelle, 65% avec le type de consonnes et 62% avec le nombre de segments. Il faut noter que la classification des noms était bien meilleure que la classification des verbes (en moyenne, environ 90% des noms ont été bien classifiés, contre 25% des verbes). Les résultats ont été encore meilleurs quand tous ces indices ont été cumulés : ils obtiennent alors une classification correcte globale de 68%, mais avec un taux de réussite assez proche entre la catégorisation des noms (78%) et celle des verbes (61%). Des résultats similaires ont été trouvés par ces mêmes auteurs, avec les mêmes indices phonologiques, sur un corpus hollandais.

Toutes ces analyses montrent que beaucoup d'informations phonologiques existent pour distinguer les deux grandes classes grammaticales que sont les noms et les verbes. Celles-ci existent aussi bien dans la parole adulte que dans la parole adressée aux enfants, et pourraient donc servir aussi bien au traitement syntaxique adulte qu'à l'acquisition de la syntaxe par les enfants. Les simulations réalisées dans différentes études de faisabilité ont montré que ces informations phonologiques (accentuation, qualité des voyelles, nombre de segments...) pouvaient permettre une première classification des mots en noms et verbes (même si ces classifications ne sont pas parfaites, elles n'en demeurent pas moins informatives). Le cumul des différentes informations phonologiques semble être aussi beaucoup plus informatif que la prise en considération d'un seul de ces indices. Bien sûr, les indices phonologiques dépendent des langues étudiées et il serait nécessaire d'obtenir des analyses sur d'autres langues pour valider ce résultat d'une manière plus générale.

Nous venons donc de montrer que les informations phonologiques semblaient pertinentes et pouvaient permettre une catégorisation des mots en noms et verbes ; nous allons maintenant nous intéresser à différentes études qui ont essayé de montrer que ces indices phonologiques étaient effectivement utilisés pour trouver la catégorie grammaticale des mots.

#### 4.2.2.2 Données expérimentales sur l'utilisation de ces informations

Le premier indice à avoir été étudié est le pattern d'accentuation différent entre les noms et les verbes (rappelons que les noms bisyllabiques tendent à avoir un pattern d'accent trochaïque, alors que les verbes ont plutôt tendance à avoir un pattern d'accent iambique).

Kelly et Bock (1988) ont réalisé une expérience pour voir si les locuteurs adultes anglais pouvaient inférer la catégorie grammaticale d'un mot grâce à son pattern d'accentuation. Pour ce faire, les sujets entendaient des non-mots bisyllabiques, accentués sur leur première ou deuxième syllabe, et devaient ensuite inventer des phrases contenant ces non-mots. L'hypothèse était que si les sujets pouvaient classifier les mots dans différentes catégories syntaxiques sur la base de l'accentuation, alors ils devraient plus souvent utiliser les non-mots comme des noms quand ceux-ci portent l'accent sur la première syllabe (que quand ils portent l'accent sur la deuxième syllabe). L'inverse devrait apparaître pour l'utilisation des non-mots comme verbes (les non-mots portant l'accent sur la deuxième syllabe devraient être plus souvent utilisés comme des verbes que les non-mots accentués sur leur première syllabe). Les résultats ont confirmé ces hypothèses (48% des items à pattern d'accent iambique ont été utilisés comme des verbes versus 31% des items à pattern d'accent trochaïque).

Dans une autre expérience, Kelly et Martin (1995, cité dans Kelly, 1996) faisaient entendre des mots bisyllabiques à des sujets, qui devaient ensuite les classifier en noms et en verbes. La moitié des noms et des verbes qu'ils entendaient avaient un pattern d'accent prototypique (pour les noms, la première syllabe était accentuée, alors que pour les verbes, la deuxième syllabe était accentuée). L'autre moitié avait un pattern d'accentuation inverse. Les temps de réponse des sujets ont été influencés par cette information phonologique : les items trochaïques étaient plus rapidement classifiés quand ils étaient des noms que quand ils étaient des verbes (503 ms versus 596 ms). De la même manière, les items iambiques étaient plus vite classifiés quand ils étaient des verbes que quand ils étaient des noms (430 ms versus 526 ms). Ces résultats démontrent que les adultes anglais ont des attentes quant à la catégorie syntaxique des mots selon le pattern accentuel qu'ils présentent, et qu'ils utilisent cette information pour inférer la catégorie grammaticale des mots.

Cassidy et Kelly (1991) ont examiné si des enfants anglais de 3 et 4 ans avaient appris et pouvaient utiliser le fait que les noms ont tendance à être formés de plus de syllabes que les verbes. Dans leur étude, les enfants devaient interagir avec une marionnette qui aimait regarder des émissions télévisées. Chaque séquence vidéo présentée mettait en scène un objet réalisant une action (comme une balle qui roule, ou un cochon qui applaudit). Durant chaque séquence, la marionnette répétait un pseudo-mot de une ou trois syllabes (par exemple, « gorp » ou « gorpinklak »). L'enfant devait deviner de quoi parlait la marionnette en choisissant entre deux interprétations. Pour l'exemple de la balle qui roule, les chercheurs demandaient à l'enfant si « gorp » voulait dire « balle » ou « rouler » (un autre groupe d'enfants devait trouver le sens de « gorpinklak » dans la même situation). Le choix des enfants a été significativement affecté par le nombre de syllabes du pseudo-mot : ils choisissaient plus l'interprétation « objet » pour les mots trisyllabiques que pour les monosyllabes.

Nous avons donc vu que plusieurs indices phonologiques, directement disponibles dans le signal de parole que les enfants entendent, pouvaient apporter des informations pertinentes quant à la catégorie grammaticale des mots : ils peuvent permettre de distinguer par exemple les noms des verbes, ou les mots grammaticaux des mots de contenu.

Nous allons maintenant aborder un dernier indice qui permettrait, non plus de catégoriser des mots individuels, mais plutôt d'obtenir des informations sur l'organisation syntaxique des phrases entières, en particulier sur leur organisation en constituants syntaxiques. Il s'agit là de la plus ancienne idée concernant l'initialisation prosodique de l'acquisition de la syntaxe (Gleitman & Wanner, 1982; Morgan, 1986), l'idée selon laquelle la prosodie des phrases pourrait permettre à l'enfant de construire une représentation de la structure syntaxique des phrases qu'il entend.

## 4.3 INDICES PROSODIQUES DE LA STRUCTURE SYNTAXIQUE DES PHRASES

### 4.3.1 *Approche théorique*

Un des arguments qui motive l'idée que la prosodie des phrases pourrait être un indice d'initialisation, est le fait que la structure prosodique reflète souvent la structure syntaxique des phrases (voir le chapitre 1). Les phrases que nous entendons et produisons ne sont pas des suites de mots sans structure interne. Au contraire, les mots sont regroupés en constituants syntaxiques (dont la structure se reflète au niveau prosodique). Ainsi, les mots se combinent pour former des groupes syntaxiques (« phrases » en anglais), qui eux-mêmes vont former des propositions (« clauses » en anglais) puis des phrases. Des analyses phonologiques des phrases ont montré que les unités prosodiques correspondaient généralement à des constituants syntaxiques. Il a donc été suggéré que les propriétés prosodiques de la parole, c'est-à-dire les variations de rythme et d'intonation, pourraient permettre un découpage de la parole en unités pertinentes d'un point de vue linguistique et syntaxique, c'est-à-dire en constituants syntaxiques (Gleitman, Gleitman, Landau, & Wanner, 1988; Gleitman & Wanner, 1982; Morgan, 1986). En prêtant attention aux propriétés acoustico-prosodiques des phrases, les enfants pourraient donc obtenir des informations sur la place des frontières syntaxiques dans les phrases qu'ils entendent.

Il est important de noter que tous les indices, reflétant la structure prosodique des phrases (et donc leur structure syntaxique), sont exagérés dans la parole adressée aux enfants (Fernald & Mazzie, 1991; Gleitman, Newport, & Gleitman, 1984; Morgan, 1986). Par exemple, les allongements que l'on trouve en fin d'unités prosodiques sont plus importants dans la parole adressée aux enfants que dans la parole adressée aux adultes (Bernstein Ratner, 1986; Morgan, 1986). La parole adressée aux enfants est aussi caractérisée par une fréquence fondamentale plus haute et des variations de fréquence fondamentale plus larges (Fernald & Kuhl, 1987; Fernald & Simon, 1984), ainsi que par un débit plus lent et des pauses plus longues dans le discours.

Etablir que ces variations prosodiques sont plus saillantes dans la parole adressée aux enfants, que dans la parole adressée aux adultes, n'est pas suffisant pour dire que les enfants les utilisent pour traiter le langage qu'ils entendent. Il faut d'abord montrer qu'ils sont capables de les percevoir et de les interpréter.

#### 4.3.2 *Sensibilité des jeunes enfants aux indices prosodiques*

Plusieurs études se sont intéressées à la sensibilité des jeunes enfants aux indices de frontières prosodiques (Christophe, Dupoux, Bertoni, & Mehler, 1994; Gerken, Jusczyk, & Mandel, 1994; Hirsh-Pasek et al., 1987; Jusczyk et al., 1992; Kemler-Nelson, Hirsh-Pasek, Jusczyk, & Cassidy, 1989).

Par exemple, Christophe et al. (1994) ont montré que les nouveaux-nés français pouvaient discriminer deux syllabes qui différaient seulement sur la présence ou l'absence d'une frontière prosodique. Pour ce faire, les auteurs ont extrait des stimuli bisyllabiques (par exemple, « mati »), qui étaient initialement contenus soit dans un mot (comme dans « mathématiicien »), soit dans deux mots différents (comme dans « panorama typique »). Avec la technique de succion non-nutritive, ils ont montré que les enfants pouvaient distinguer des exemplaires bisyllabiques extraits de ces deux conditions. Les nouveaux-nés perçoivent donc ces frontières prosodiques.

En plus de les percevoir, d'autres recherches ont montré que les enfants avaient des attentes quant à l'apparition de ces frontières dans les phrases. Hirsh-Pasek et al. (1987) ont étudié la sensibilité d'enfants américains de 6 et 9 mois aux frontières de groupes intonationnels (qui correspondent typiquement aux frontières de propositions). Pour ce faire, ils ont créé deux sortes de phrases, qui variaient quant à la position des pauses dans les phrases (pauses de une seconde) :

- des phrases coïncidentes dans lesquelles les frontières de groupes intonationnels correspondaient aux frontières des propositions, et où l'apparition des pauses était corrélée avec d'autres indices acoustiques comme des variations de durée et de fréquence fondamentale : par exemple, les enfants entendaient « Cinderella lived in a big house (pause) but it was dark (pause) because she had a mean stepmother » ;

- des phrases non-coïncidentes dans lesquelles les frontières de groupes intonationnels étaient insérées à l'intérieur des propositions, et où l'apparition des pauses n'était pas corrélée avec des variations de durée et d'intonation : par exemple, les enfants entendaient « Cinderella lived (pause) in a big house but it was (pause) dark because she had a mean stepmother ».

Les enfants ont préféré écouter les premières phrases, c'est-à-dire celles où les frontières prosodiques (frontières de groupes intonationnels) correspondaient aux frontières de constituants syntaxiques (frontières de propositions). Ces résultats ont été répliqués dans d'autres études (Kemler-Nelson et al., 1989), et même avec des enfants plus jeunes (4.5 mois, Jusczyk, 1997).

Jusczyk et al. (1992) ont utilisé le même paradigme expérimental que celui de l'expérience précédente, pour étudier les attentes des enfants quant à la position des frontières de groupes phonologiques (coïncidant avec les frontières de groupes syntaxiques, comme entre le groupe nominal et le groupe verbal). Des pauses de une seconde ont été insérées soit entre le groupe nominal et le groupe verbal (versions coïncidentes), soit après le verbe donc au milieu du groupe verbal (versions non coïncidentes). Comme pour l'expérience précédente, les auteurs ont mis en évidence une préférence des enfants de 9 mois pour les phrases coïncidentes. Par contre, l'effet n'a pas été répliqué pour les enfants de 6 mois (des résultats similaires ont été trouvés par Gerken et al., 1994).

De récentes expériences ont étudié la sensibilité des jeunes enfants aux frontières de groupes syntaxiques (Soderstrom, Seidl, Kemler Nelson, & Jusczyk, 2003). Ils ont pu montrer que dès 6 mois, les enfants étaient sensibles aux indices prosodiques marquant les frontières des groupes syntaxiques, et qu'ils pouvaient utiliser cette sensibilité pour reconnaître des unités comme les groupes nominaux et les groupes verbaux, dans les phrases. Pour ce faire, ils ont créé des passages contenant des séquences de mots identiques d'un point de vue phonologique (même prononciation) mais différentes quant à la position des mots dans la structure syntaxique des phrases :

Passage 1 : « At the discount store, **new watches for men** are simple and stylish.

In fact, some **people # buy the whole** supply of them. »

Passage 2 : « In the field, the old frightened **gnu # watches for men** and women seeking trophies. Today, **people by the hole** seem scary. »

Chaque passage contient deux séquences cibles : une séquence formant un groupe nominal, donc une unité syntaxique correcte (« new watches for men » dans le premier passage et « people by the hole » dans le deuxième passage), et une séquence non syntaxique, chevauchant un groupe nominal et un groupe verbal (« people # buy the whole » dans le premier passage et « gnu # watches for men » dans le deuxième passage, le symbole # indiquant la frontière entre le groupe nominal sujet et le groupe verbal). Dans le premier cas, la séquence qui forme une unité syntaxique, forme aussi une unité prosodique ; par contre, dans le deuxième cas, la séquence non syntaxique contient deux morceaux d'unités prosodiques avec une frontière syntaxique, et donc une frontière prosodique, au milieu.

Dans une première phase, les enfants étaient familiarisés avec les deux versions d'une même séquence phonologique (par exemple, ils entendaient des répétitions de « new watches for men » et « gnu # watches for men »). Puis, dans une deuxième phase, ils entendaient les deux passages présentés ci-dessus. Les résultats ont montré que les enfants préféraient écouter le passage contenant la séquence correcte d'un point de vue syntaxique que le passage contenant la séquence incorrecte d'un point de vue syntaxique (les enfants familiarisés avec « new watches for men » et « gnu # watches for men » ont écouté plus longtemps le passage 1, contenant la version correcte d'un point de vue syntaxique, que le passage 2 qui contenait la version non-syntaxique). Les indices prosodiques aideraient donc les enfants à retrouver dans la parole continue des séquences familières, quand celles-ci forment des unités syntaxiques bien formées d'un point de vue prosodique. Cet effet pourrait être dû au fait que les enfants ont segmenté les passages qu'ils entendaient en unités syntaxiques bien formées (grâce aux indices prosodiques de groupes syntaxiques), et qu'il y a donc plus de chances qu'ils détectent l'apparition de la séquence syntaxique entendue en familiarisation, plutôt que l'apparition de la séquence non-syntaxique.

Un point important à noter est que ces résultats n'ont pas été répliqués sur le deuxième groupe d'enfants, qui étaient familiarisés avec les autres séquences

phonologiques (« people # buy the whole » et « people by the hole »). Ils auraient dû préférer écouter le passage 2 par rapport au passage 1 (puisque le passage 2 contient la séquence légale d'un point de vue syntaxique). Les analyses acoustiques des stimuli apportent une explication à cette absence d'effet. Ces deux dernières séquences étaient beaucoup moins différentes d'un point de vue acoustique que les deux premières séquences étudiées. La frontière prosodique entre « gnu » et « watches » était beaucoup plus saillante que celle entre « people » et « buy » (allongements plus grands et plus grande variation dans la courbe intonative). Par exemple, on sait que la durée de la voyelle précédant une frontière prosodique est plus longue que la durée de cette même voyelle placée dans un autre contexte prosodique : dans les exemples utilisés dans cette expérience, la voyelle /u/ était plus longue dans l'extrait « gn # watches for men » (où elle était située avant la frontière prosodique) que dans la séquence « new watches for men » (229 ms versus 89 ms). Cet allongement en fin d'unité prosodique a aussi été trouvé pour l'autre séquence phonologique : le /l/ de « people » est plus long dans « people # buy the whole » que dans « people by the hole » (92 ms versus 74 ms). Mais comme on peut le constater, l'allongement était nettement plus important et plus saillant pour la première séquence phonologique (allongement de 157% versus 24%). Il était donc plus clair que la séquence « gnu # watches for men » chevauchait une frontière syntaxique, comparée à la séquence « people # buy the whole ». Les enfants du deuxième groupe n'ont donc peut-être pas eu assez d'indices pour distinguer la séquence syntaxique de la séquence non-syntaxique (ce qui fait qu'ils ont réagi de la même manière aux deux passages, qui contenaient chacun un exemplaire de la séquence phonologique en question).

Les auteurs ont trouvé ces résultats avec des enfants de 9 mois et des enfants de 6 mois, et les ont répliqués avec d'autres sortes de stimuli (comme en comparant un groupe verbal bien formé et une unité incorrecte chevauchant un groupe nominal et un groupe sujet).

Les résultats trouvés par Soderstrom et al. (2003) avec des enfants américains de 6 mois sont en désaccord avec ceux qui avaient été trouvés par Jusczyk et al. (1992) et qui avaient montré que les enfants de cet âge n'étaient pas sensibles aux indices de groupes phonologiques. L'absence d'effet dans l'étude de Jusczyk et al. (1992) pourrait être rapprochée de l'absence d'effet dans le deuxième groupe d'enfants étudié par Soderstrom et al. (2003). Il se pourrait que les indices prosodiques testés par Jusczyk

n'aient pas été suffisamment saillants pour permettre de voir émerger un effet. Ceci rappellerait une conclusion apportée par Gerken et al. (1994), qui suggéraient que les enfants étaient sensibles à la frontière syntaxique entre groupe nominal sujet et groupe verbal, si celle-ci était marquée prosodiquement de manière saillante (combinaison d'allongements et de variations d'intonation en fin de groupe phonologique).

En fait, Gerken et al. (1994) avaient utilisé des phrases contenant un groupe nominal sujet lexical (comme dans « Joe kicked the ball » ou « The caterpillar kicked the ball ») ou un groupe nominal sujet composé d'un pronom (comme dans « He kicked the ball »). Dans le premier cas, un locuteur aura tendance à produire une frontière prosodique après le groupe nominal sujet, frontière qui est alors congruente avec la frontière syntaxique entre le groupe nominal sujet et le groupe verbal (comme dans « [Joe] # [kicked the ball] » ou « [The caterpillar] # [kicked the ball] », où les crochets indiquent les frontières de groupes phonologiques, et où le signe # indique la position de la frontière syntaxique majeure séparant le groupe nominal sujet du groupe verbal). Les auteurs ont fait écouter aux enfants des phrases dans lesquelles une pause était insérée avant le verbe ou après le verbe. Les enfants américains de 9 mois ont préféré écouter des phrases dans lesquelles une pause avait été insérée entre le sujet et le verbe (« Joe (pause) kicked the ball » ou « The caterpillar (pause) kicked the ball »), comparé à des phrases dans lesquelles une pause avait été insérée après le verbe (« Joe kicked (pause) the dog »).

Dans le deuxième cas, un locuteur ne produira pas de frontière prosodique entre le sujet et le verbe, car le sujet est un pronom qui ne forme donc pas une unité prosodique à lui tout seul, mais qui se trouve rattaché prosodiquement au verbe (phénomène de cliticisation : voir par exemple, Nespor & Vogel, 1986). Par contre, une frontière sera généralement produite entre le verbe et le groupe nominal objet (comme dans « [He # kicked] [the ball] »), ce qui fait que la frontière prosodique marquée ne correspond plus à la frontière syntaxique majeure de la phrase. Dans cette situation, les enfants n'ont pas montré de préférence d'écoute pour les passages contenant une pause insérée après le sujet (temps d'écoute similaire pour « He (pause) kicked the ball » et pour « He kicked (pause) the ball »).

Cette expérience indique donc que les enfants américains de 9 mois perçoivent bien les frontières prosodiques, et qu'ils réagissent quand les groupes phonologiques sont interrompus de manière incorrecte (quand une pause est insérée au milieu). Par contre, les enfants de 9 mois (et probablement aussi les enfants plus âgés) ne semblent

pas percevoir les frontières de groupes syntaxiques qui ne correspondent pas à des frontières de groupes phonologiques.

Toutes ces études semblent montrer un certain accord dans la communauté scientifique pour dire que les jeunes enfants sont sensibles aux frontières de propositions (donc aux frontières de groupes intonationnels). Différents résultats ont été trouvés dès 4.5 mois et jusqu'à 9 mois (Gerken et al., 1994; Jusczyk, 1997; Jusczyk et al., 1992; Nazzi, Nelson, Jusczyk, & Jusczyk, 2000). Par contre, un débat semble subsister quant à l'âge auquel les enfants sont sensibles aux frontières de groupes syntaxiques qui correspondent à des frontières de groupes phonologiques. Certains pensent qu'ils y sont sensibles dès l'âge de 6 mois (Soderstrom et al., 2003), alors que d'autres pensent que qu'ils ne le sont que vers 9 mois (Gerken et al., 1994; Jusczyk et al., 1992). L'explication de ce désaccord provient probablement de la manière dont ces frontières étaient réalisées à travers les différentes expériences.

Quoiqu'il en soit, il est clairement démontré que les enfants sont sensibles à ces frontières de groupes phonologiques avant la fin de la première année de vie. Il reste maintenant à prouver, qu'en plus d'y être sensibles, les enfants peuvent utiliser ces indices pour réaliser différentes tâches linguistiques et guider l'acquisition de leur langue maternelle (voir l'étude 3 dans le chapitre 5 qui apportera quelques réponses à cette question, en étudiant le rôle des frontières de groupes phonologiques dans la segmentation de la parole continue en mots).

#### 4.4 LIENS ENTRE PROSODIE ET MOTS GRAMMATICaux

Avant de conclure cette revue de la littérature sur les différentes hypothèses d'initialisation de la syntaxe, je voudrais revenir sur deux indices traités dans les parties précédentes, qui me semblent être liés, et particulièrement utiles pour l'acquisition de la syntaxe.

D'un côté, nous avons parlé du rôle que pouvaient jouer les mots grammaticaux (articles, pronoms, prépositions...) dans l'acquisition de la syntaxe. Ces mots pourraient

être appris précocement du fait de leurs propriétés particulières qui les différencient des mots de contenu. De plus, les mots grammaticaux sont en général situés en bordure d'unités prosodiques. Par exemple, en anglais, les mots grammaticaux signalent typiquement le début des groupes syntaxiques (Morgan, Meier, & Newport, 1987; Morgan et al., 1996). Ces particularités pourraient être exploitées par les enfants à deux titres : d'abord, elles peuvent permettre aux jeunes enfants de découvrir relativement facilement les mots grammaticaux de leur langue (certains résultats suggèrent que cette catégorie de mots est déjà acquise vers 8 ou 10 mois, Shafer et al., 1998; Shi & Gauthier, 2005). D'autre part, la reconnaissance d'un mot grammatical en début d'unité prosodique pourrait permettre à l'enfant d'inférer que le mot suivant est un mot de contenu (on rejoint aussi à cet endroit les théories d'initialisation distributionnelle dans lesquelles les enfants pourraient utiliser le fait que les noms sont en général précédés d'un article, alors que les verbes sont précédés de pronoms ou d'auxiliaires).

D'un autre côté, nous avons aussi mis en évidence le rôle des frontières prosodiques dans la mise en place de la structure syntaxique. Certes, toutes les unités syntaxiques ne sont pas obligatoirement marquées prosodiquement de manière saillante ; certaines ne sont même pas marquées du tout, comme la frontière syntaxique entre un pronom sujet et un verbe (Beckman & Edwards, 1990; Gerken et al., 1994). Mais nous avons vu que différentes recherches avaient démontré que de nombreux indices prosodiques étaient présents dans la parole adressée aux enfants et que les enfants y étaient sensibles (que ce soit des indices très saillants comme ceux marquant les frontières de propositions, ou des indices moins saillants comme ceux marquant les frontières de groupes phonologiques).

La combinaison de ces deux indices pourrait être exploitée par les enfants lors des premières étapes de l'analyse syntaxique. Prenons par exemple une phrase comme « la petite fille a mangé un gros gâteau ». Les frontières de groupes phonologiques pourraient permettre aux enfants de percevoir les frontières des différents groupes syntaxiques, et de se représenter la phrase sous la structure « [xxxx] [xxx] [xxxx] » (les crochets indiquent les frontières de groupes phonologiques, et les x représentent les syllabes entendues). La reconnaissance des mots grammaticaux en début d'unité pourrait permettre à l'enfant de trouver la catégorie (et donc peut-être la fonction puis le sens) de ces différents constituants. L'enfant arriverait donc à avoir une première

représentation syntaxique de cette phrase sous la forme suivante : « [la xxx]<sub>GN</sub> [a xx]<sub>GV</sub> [un xxx]<sub>GN</sub> » (le symbole GN représentant les groupes nominaux, GV représentant le groupe verbal). Bien sûr, des expériences vont être nécessaires pour affirmer cette hypothèse, que ce soit des expériences bébés, des expériences adultes ou des simulations. L'étude 4 dans le chapitre 6, menée auprès d'adultes français, essaiera d'apporter une réponse à cette hypothèse de travail.

Pour conclure, nous avons vu que, depuis longtemps, les chercheurs ont essayé de comprendre comment les enfants réussissaient à acquérir la syntaxe de leur langue maternelle, et comment ils pouvaient, par exemple, assigner une catégorie syntaxique aux mots qu'ils entendaient. Leurs hypothèses et explications se sont d'abord centrées sur des indices sémantiques. Ces explications sont très attractives de par leur caractère universel, et leur base reposant sur un lien inné entre les catégories conceptuelles (objets, actions,...) et les catégories syntaxiques (noms, verbes,...), mais elles ne permettent pas d'expliquer complètement comment les enfants arrivent à assigner une catégorie syntaxique aux mots qu'ils entendent.

D'autres indices ont été proposés et étudiés de manière convaincante : les indices distributionnels. Il semble en effet possible et plausible que les enfants puissent utiliser les relations de co-occurrence entre les mots des phrases (que ces relations soient adjacentes ou non) pour inférer la catégorie grammaticale d'autres mots. Ainsi, les enfants pourraient utiliser le fait qu'un nom est généralement précédé d'un article, qu'un verbe est généralement précédé d'un pronom, ou que certains cadres lexicaux permettent de trouver de manière relativement fiable la catégorie grammaticale d'un nouveau mot.

Reste un problème important pour ces différentes théories : pour pouvoir appliquer ces règles, les enfants doivent d'abord trouver les mots dans les phrases et donc avoir résolu le problème de la segmentation de la parole continue. D'autres indices ont été proposés, indices qui permettraient à la fois d'aider l'enfant à segmenter l'input continu de parole qu'il reçoit, de l'aider à trouver les catégories syntaxiques des mots qu'il entend, et de lui permettre aussi de découvrir la structure et l'organisation syntaxique des phrases qu'il traite. Ces indices sont les indices phonologiques et prosodiques. Ceux-ci ont longtemps été négligés car considérés comme trop spécifiques aux langues. Or de nombreux arguments peuvent renverser cette critique, comme par

exemple le fait que les enfants, et même les nouveaux-nés, sont très rapidement sensibles aux caractéristiques phonologiques de leur langue maternelle (voir par exemple, Mehler et al., 1988). La possibilité que les enfants puissent apprendre rapidement quelles sont les caractéristiques phonologiques permettant de distinguer les noms des verbes, ou les mots de contenu des mots grammaticaux dans leur langue maternelle, est donc très plausible et a été démontrée dans de nombreuses expériences que nous avons décrites dans ce chapitre. De plus, il faut aussi noter que, même si les indices phonologiques ne sont pas universels, mais plutôt spécifiques aux langues étudiées, certains se retrouvent quand même dans toutes les langues et permettent de distinguer par exemple les mots grammaticaux des mots de contenu : ces indices pertinents sont la fréquence des mots, leur position dans les phrases et probablement leur durée (Morgan et al., 1996).

Bien entendu, personne ne défend l'idée qu'un seul indice est nécessaire et suffisant pour que l'enfant puisse acquérir complètement et rapidement la syntaxe de sa langue maternelle. Au contraire, ce sont plutôt des constellations d'indices qui comptent (voir par exemple les modèles avec intégration d'indices multiples ou « multiple cue integration » proposés entre autres par Christiansen & Dale, 2001; Christiansen & Monaghan, in press; Monaghan, Chater, & Christiansen, 2005; Morgan et al., 1987; Reali, Christiansen, & Monaghan, 2003). De plus, certains auteurs ont montré que différents indices pouvaient jouer des rôles différents selon les mots qui composent les phrases : par exemple, Monaghan et al. (2005) ont montré, dans une analyse de corpus et dans une expérience d'apprentissage d'un langage artificiel, que les indices phonologiques étaient des indices plus pertinents pour les mots peu fréquents, alors que les informations distributionnelles étaient plus utiles pour les mots de haute fréquence.

Les études qui vont être présentées dans les deux prochains chapitres ont été réalisées dans cette optique : nous ne voulons pas dire que seule la prosodie des phrases peut apporter des informations pertinentes aux jeunes enfants qui apprennent la syntaxe de leur langue maternelle. Bien au contraire, nous pensons que les enfants pourraient utiliser, de manière précoce, les informations apportées à la fois par (1) les indices phonologiques (pour découvrir les mots grammaticaux de leur langue), (2) les indices distributionnels (pour inférer la catégorie grammaticale d'un mot de contenu suivant un mot grammatical), et (3) la prosodie des phrases (pour délimiter les constituants syntaxiques composant les phrases).

## CHAPITRE 5

### PERCEPTION ET UTILISATION DES FRONTIERES DE GROUPES PHONOLOGIQUES PAR DES ENFANTS DE 16 MOIS (ETUDE 3)

Notre hypothèse de travail sur l'acquisition de la syntaxe par les jeunes enfants repose sur l'utilisation conjointe de deux indices différents : la prosodie des phrases (qui pourrait permettre aux enfants de trouver les différents constituants syntaxiques dans les phrases qu'ils entendent), et les mots grammaticaux (qui pourraient permettre aux enfants d'assigner une catégorie précise à ces unités syntaxiques, comme de trouver quels sont les groupes nominaux et les groupes verbaux).

Nous nous intéressons plus particulièrement aux frontières de groupes phonologiques, qui correspondent à des frontières syntaxiques internes aux propositions. Nous avons vu, dans la première partie de cette recherche (études 1 et 2), que ces frontières prosodiques étaient calculées en temps réel par les adultes français, et qu'elles contraignaient l'analyse syntaxique des phrases qu'ils entendaient. Notre idée est que les jeunes enfants pourraient eux aussi exploiter ces frontières de groupes phonologiques pour guider l'acquisition de la syntaxe de leur langue maternelle. Pour que cette hypothèse soit plausible, il faut d'abord s'assurer que les enfants sont bien capables de les percevoir et de les utiliser.

Dans cette étude, nous avons voulu vérifier que les jeunes enfants français étaient capables de percevoir et d'utiliser les frontières de groupes phonologiques en temps réel. Nous avons exploité le fait qu'une frontière de groupe phonologique correspond toujours à une frontière de mot.

Considérons les phrases suivantes :

« Le joli balcon cloisonnait la terrasse. »

« Le dernier bal conclura la saison. »

Ces deux phrases comportent les deux mêmes syllabes /bal/ et /con/ : dans la première phrase, ces deux syllabes forment un mot, alors que dans la deuxième phrase, elles appartiennent à deux mots différents. Ces syllabes se différencient aussi au niveau de leur position dans la structure prosodique des phrases, comme le montrent les représentations prosodiques suivantes (dans lesquelles les crochets indiquent les frontières de groupes phonologiques) :

« [Le joli balcon] [cloisonnait la terrasse] »

« [Le dernier bal] [conclura la saison] »

Ainsi, dans la première phrase, le mot constitué des deux syllabes se trouve à la fin du premier groupe phonologique (à la fin du groupe nominal), alors que dans la deuxième phrase, les deux syllabes cibles se trouvent de part et d'autre d'une frontière de groupe phonologique (la première syllabe est la dernière du groupe nominal et la deuxième syllabe est la première du groupe verbal).

Si les enfants perçoivent et peuvent interpréter les frontières des groupes phonologiques comme des frontières de mot, ils devraient alors considérer les syllabes /bal/ et /con/ comme un mot dans la première phrase, mais ne pas identifier le mot « balcon » dans la deuxième phrase, et donc considérer ces deux syllabes comme appartenant à deux mots différents.

Notons que des expériences ont démontré que les adultes français pouvaient effectivement utiliser les frontières de groupes phonologiques pour contraindre en temps réel leur accès au lexique (par exemple, des adultes n'activent pas le mot « chagrin » dans « [son gros chat] [grimpait aux arbres] », où les deux syllabes sont séparées par une frontière de groupe phonologique, alors qu'ils sont ralentis par cette ambiguïté lexicale dans « [son chat grincheux] » où le mot « chagrin » entre en compétition lors de la reconnaissance des mots, et plus particulièrement lors de la reconnaissance du mot « chat », Christophe, Peperkamp, Pallier, Block, & Mehler, 2004).

L'expérience que nous allons présenter a été réalisée dans le cadre d'une collaboration avec le laboratoire de James Morgan à Brown University (Providence, USA). Une étude, similaire à celle que nous allons décrire, avait été menée

précédemment par cette équipe, auprès d'enfants américains de 10 et 13 mois, et avait démontré que les enfants américains interprétaient en temps réel les frontières prosodiques comme des frontières de mot, et qu'ils pouvaient donc grâce à ces indices extraire les mots de la parole continue.

Mais, comme nous l'avons vu dans le chapitre précédent, les indices phonologiques et prosodiques peuvent être différents entre les langues ; il nous semblait donc nécessaire de répliquer en français les résultats trouvés par Gout, Christophe et Morgan (2004), plutôt que de les prendre comme acquis et généralisables à toutes les langues.

Notre expérience concerne des enfants français de 16 mois (voir l'annexe E « Différences inter-langues et méthodologie » pour les résultats des enfants français de 13 mois, et une comparaison inter-langues des résultats). Elle a été réalisée avec une variante de la technique de conditionnement de l'orientation de la tête (« Conditioned Head-Turn procedure »).

Le principe est le suivant : dans une première session (phase d'entraînement), on apprenait aux enfants à tourner la tête pour un mot cible : un groupe d'enfants était entraîné sur un mot bisyllabique (enfants du « groupe bisyllabique » qui doivent répondre à *balcon* ou *vipère*) ; un deuxième groupe était testé sur un mot monosyllabique, qui correspondait à la première syllabe du mot cible utilisé dans la condition bisyllabique (« groupe monosyllabique » qui répond à *bal* ou *vie*). Dans une deuxième session (phase de test), les enfants étaient testés sur des phrases complètes : la moitié des phrases expérimentales contenait réellement le mot cible (ce sont les « phrases bisyllabiques » comme, par exemple, « [le joli *balcon*] [cloisonnait la terrasse] »), tandis que l'autre moitié contenait les deux syllabes du mot cible séparées par une frontière de groupe phonologique (comme dans « [le dernier *bal*] [conclura la saison] » qui fait partie des « phrases monosyllabiques »).

Si les frontières de groupes phonologiques sont utilisées en temps réel comme un indice de fin de mot, nous nous attendons à ce que les enfants, habitués à répondre au mot bisyllabique, tournent plus la tête sur les phrases contenant le mot bisyllabique, et moins sur les phrases contenant les deux syllabes du mot cible séparées par cette frontière prosodique. Quant aux enfants du groupe monosyllabique, nous attendons l'effet inverse, c'est-à-dire qu'ils tournent plus leur tête sur les phrases contenant le mot

monosyllabique que sur les phrases contenant une syllabe homophone en début de mot. Mais pour ce dernier groupe, il faut noter que leur tâche est plus difficile : en effet, alors que la présence d'une frontière de groupe phonologique peut indiquer avec certitude une frontière de mot, l'inverse n'est pas vrai puisque toutes les fins de mot ne coïncident pas obligatoirement avec une frontière de groupe phonologique. Ainsi, les enfants devant tourner la tête pour le mot « bal » devraient pouvoir détecter le mot cible dans les phrases contenant la séquence « [... bal] [con...], mais pourraient aussi se faire piéger en entendant des phrases contenant le mot « balcon », ce qui entraînerait un effet plus petit pour ce groupe d'enfants (comparés aux enfants du groupe bisyllabique).

## 5.1 METHODE

### 5.1.1 Matériel

Les stimuli utilisés pendant la phase d'entraînement, donc les stimuli auxquels les enfants devaient répondre, étaient des mots isolés : soit différents exemplaires de « balcon » ou « vipère » pour les enfants du groupe bisyllabique, soit différents exemplaires de « bal » ou « vie » pour les enfants du groupe monosyllabique. Ces mots cibles ont été choisis car la première syllabe des mots bisyllabiques (/bal/, /vi/) étaient des mots existants, et car leur deuxième syllabe (/kɔ̃/, /pɛr/) apparaissait au début de plusieurs verbes. Ceci nous a donc permis de construire de nombreuses phrases expérimentales contenant soit les mots bisyllabiques, soit ces deux syllabes séparées par une frontière de groupe phonologique.

Pendant la phase de test, les stimuli étaient des phrases entières. Pour chaque mot test (« balcon » et « vipère »), nous avons construit douze paires de phrases expérimentales : un membre de cette paire contenait réellement le mot cible bisyllabique, tandis que l'autre membre contenait les deux syllabes du mot bisyllabique mais dans deux mots différents et séparées par une frontière de groupe phonologique. Des exemples de paires de phrases sont indiqués ci-dessous (voir l'annexe C.1 pour une liste complète du matériel) :

- (1a) « [Le joli balcon] [cloisonnait la terrasse] » (phrase balcon, bisyllabique)  
(1b) « [Le dernier bal] [conclura la saison] » (phrase bal, monosyllabique)  
(2a) « [Une grosse vipère] [pénétra dans l'entrée] » (phrase vipère, bisyllabique)  
(2b) « [Une longue vie] [perpétue les traditions] » (phrase vie, monosyllabique)

Les phrases de chaque paire étaient appariées au niveau de leur structure prosodique avant le mot cible, et au niveau du nombre de syllabes. Elles ont toutes été enregistrées par une locutrice française, qui était au courant du principe de l'expérience, et qui a produit ces phrases avec une intonation très marquée (proche du « motherese »). La locutrice a aussi enregistré différents exemplaires de chaque mot prononcé en isolation avec diverses intonations, ainsi que des phrases courtes contenant le mot cible en fin de phrase (comme « c'est le balcon ») qui étaient présentées lors d'une petite phase de rappel au début de la deuxième session (avant de commencer le test proprement dit).

L'enregistrement s'est déroulé dans une pièce insonorisée, et les stimuli ont été digitalisés à 16 KHz et 16 bits, sur une table de mixage OROSAU22. Elles ont ensuite été éditées grâce aux programmes de préparation du son, Manitou et Praat. Le début du mot cible a ensuite été marqué manuellement avec le programme Bliss.

### 5.1.2 *Equipement*

L'expérience s'est déroulée dans une cabine de test insonorisée. Au centre de cette pièce, se trouvait une table avec deux chaises placées de chaque côté : une pour l'assistant de recherche, l'autre pour le parent qui s'asseyait avec son enfant sur ses genoux. Du côté de l'assistant, à sa gauche, un petit coffre à jouets comportait différents objets servant à attirer l'attention de l'enfant face à lui. Ces jouets étaient classés avec une échelle croissante d'intérêt, allant des moins intéressants (anneaux en plastique monocolores) aux plus intéressants (marionnettes), en passant par une grande variété de jouets (éléments de construction en bois, toupies, lapin en peluche...).

Sur le mur à gauche de l'enfant, deux hauts parleurs encadraient quatre boîtes en plexiglas fumé. Dans chacune de ces boîtes, se trouvait un jouet (un chien, un chat, un lapin et un panda) pouvant s'animer et qui servait de renforcement pendant

l'expérience. La pièce contenait aussi deux casques anti-bruits (un pour chaque adulte) qui permettaient de délivrer un bruit de fond pendant toute la durée de l'expérience, couvrant les stimuli auditifs entendus par l'enfant, de manière à ce que les adultes ne puissent pas influencer le comportement de l'enfant.

Un deuxième expérimentateur était installé à l'extérieur de la cabine insonorisée : c'est lui qui contrôlait le bon déroulement de l'expérience et indiquait à l'expérimentateur dans la cabine de test quel comportement avoir (grâce à un micro relié au casque que l'expérimentateur portait pendant tout le test). Il observait l'enfant grâce à système télévisuel Panasonic à circuit fermé (une caméra filmant le comportement de l'enfant était placée au dessus des boîtes en plexiglas dans la cabine de test), et renseignait l'ordinateur sur le comportement de l'enfant (en indiquant, par un clic de souris, à quel moment l'enfant tournait la tête).

La durée des essais, la présentation des stimuli, et le lancement des renforcements étaient contrôlés par un programme spécial (écrit par James Morgan et créé pour faire tourner les expériences en « Conditioned Head-Turn ») installé sur un ordinateur PC.

### *5.1.3 Procédure*

L'expérience consistait en trois phases principales : l'apprentissage de la tâche et la phase de critère (qui avaient lieu pendant la première session), ainsi que la phase de test (qui avait lieu durant la deuxième session). Ces deux sessions étaient espacées de sept jours en moyenne.

Pendant les deux sessions, l'enfant était assis sur les genoux d'un de ses parents, dans la cabine de test, face à la table et à l'assistant de recherche. Ce dernier utilisait différents jouets (en commençant par les moins intéressants) pour attirer et maintenir l'attention de l'enfant face à lui. Le lancement de chaque essai était décidé par l'expérimentateur hors de la cabine, qui pressait sur le bouton gauche de la souris quand l'enfant était bien calme et attentif à ce que faisait l'assistant de recherche. Si pendant la présentation de l'essai, l'enfant tournait la tête vers les haut-parleurs, l'expérimentateur pressait le bouton droit de la souris pour le signaler à l'ordinateur. Si l'enfant avait

tourné la tête de manière correcte (en entendant le mot cible), celui-ci délivrait alors un renforcement (allumage d'une ou plusieurs boîtes en plexiglas, avec ou sans animation du jouet situé à l'intérieur). Le choix du renforcement était décidé par l'expérimentateur qui pouvait ainsi faire varier de manière croissante la qualité du renforcement (pour maintenir l'intérêt de l'enfant pendant toute la durée de l'expérience), et adapter son choix par rapport à ce qui se trouvait sur la table (de manière à garder un équilibre entre ce que l'enfant voyait sur la table et ce que l'enfant voyait comme renforcement).

#### Première session :

Durant cette session, les enfants entendent un mot en fond sonore, qui est joué de manière continue à un niveau sonore confortable de 68 dB, avec un intervalle d'une seconde entre chaque exemplaire de ce mot. Quand l'expérimentateur décide de lancer un essai, le fond sonore est alors remplacé par le mot cible, répété trois fois, ce qui ouvre une fenêtre de réponse de quatre secondes (pendant laquelle le comportement de l'enfant est codé en temps réel). Seuls les essais où les enfants tournent leur tête de manière adéquate (c'est-à-dire quand ils ont entendu le mot cible) sont renforcés.

Les enfants sont entraînés à tourner la tête vers les haut-parleurs (et donc vers les boîtes en plexiglas) à chaque fois qu'ils entendent un mot cible précis : les enfants du groupe bisyllabique entendent les mots « balcon » et « vipère », tandis que les enfants du groupe monosyllabique entendent les mots « bal » et « vie ». Pour chaque enfant, un de ces mots sert de mot test (ou mot cible), et l'autre sert alors de bruit de fond (la moitié des enfants du groupe bisyllabique a entendu « balcon » en mot cible et « vipère » en fond sonore, alors que l'autre moitié a entendu « vipère » en mot cible et « balcon » en fond sonore ; de la même manière, la moitié des enfants du groupe monosyllabique a entendu « bal » en mot cible, tandis que l'autre moitié avait « vie » comme mot cible).

Durant la phase d'apprentissage (première phase de la première session), le mot cible est joué à chaque essai lancé par l'expérimentateur. Au premier essai, donc la première fois que le mot cible apparaît, celui-ci est joué avec un volume sonore plus élevé de 12 dB comparé au fond sonore, ceci pour déclencher un mouvement de tête vers les haut-parleurs de la part de l'enfant (ce mouvement de tête est probablement dû

à la surprise d'entendre quelque chose de plus fort et au réflexe de s'orienter vers la source sonore élevée, mais il permet de montrer rapidement à l'enfant l'existence des renforcements puisqu'une des boîtes s'allume). La différence d'intensité entre le mot cible et les stimuli de bruit de fond diminue de 4 dB à chaque fois que l'enfant répond correctement à un essai, jusqu'à atteindre le même volume de présentation. Si l'enfant échoue à trois essais consécutifs, le niveau sonore remonte de 4 dB. Cette phase continue jusqu'à ce que l'ordinateur ait joué trente essais (dans ce cas, l'expérience s'arrête), ou jusqu'à ce que l'enfant ait réussi deux essais consécutifs d'intensité égale à celle du fond sonore (et dans ce cas, l'expérience continue avec la phase de critère).

La phase de critère (deuxième phase de la première session) est similaire à la phase d'apprentissage, sauf que cette fois-ci, quand l'expérimentateur décide de lancer un essai, deux situations sont possibles : soit l'essai consiste en l'apparition du mot cible (et on vérifie que les enfants tournent la tête vers les haut-parleurs), soit l'ordinateur continue à jouer le fond sonore (et on vérifie alors que, pendant la fenêtre de quatre secondes qui s'est ouverte, les enfants ne tournent pas la tête par hasard vers les haut-parleurs). L'apparition de ces essais est contrôlée par l'ordinateur (50% de chaque, présentés de manière aléatoire). Les enfants sont testés jusqu'à ce qu'ils aient atteint un critère de réussite prédéterminé qui montre que leur performance a une faible chance de se produire au hasard (ce critère peut être atteint dès sept réponses correctes sur huit essais successifs, ou avec huit réponses correctes sur dix réponses consécutives, etc...). Quand un enfant échoue à trois essais consécutifs, il repasse en phase d'apprentissage jusqu'à ce qu'il obtienne à nouveau deux essais corrects consécutifs présentés au même niveau d'intensité que le fond sonore. Les enfants qui n'atteignent pas le seuil du critère au bout de quarante essais ne continuent pas l'expérience, tandis que ceux qui passent le seuil du critère reviennent une semaine plus tard pour passer la deuxième session.

#### Deuxième session :

La deuxième session commence avec deux phases de rappel : dans la première phase, les mots test et les mots du fond sonore apparaissent à la fin d'une petite phrase (comme dans « c'est le balcon ! » ou « regarde la vipère ! »). Dans la deuxième phase de rappel, les enfants entendent des phrases complètes, au milieu desquelles

apparaissent le mot cible ou le mot du fond sonore (comme dans « j'aimerais que mon balcon soit tout en bois » ou « la petite vipère se prélassait au soleil »). Quand un essai est lancé par l'expérimentateur, une fenêtre de 3.5 secondes s'ouvre à partir du mot cible, et la phrase cible est répétée deux fois (dans la première phase de rappel) ou une fois (dans la deuxième phase de rappel). Le comportement de l'enfant est codé par l'expérimentateur. Les premiers essais de chaque phase sont joués avec un volume sonore, plus élevé de 8 dB comparé au fond sonore. Cette intensité est diminuée progressivement, par palier de 4 dB, à chaque fois que l'enfant tourne correctement la tête vers les haut-parleurs, et ce jusqu'à atteindre le niveau sonore du fond sonore.

Après ces phases de rappel, la phase de test commence. Les enfants entendent les 24 phrases expérimentales de leur groupe (c'est-à-dire, soit les 12 phrases bisyllabiques « balcon » et les 12 phrases monosyllabiques « bal », soit les 12 phrases « vipère » et les 12 phrases « vie »). En fond sonore, certaines phrases de l'autre groupe (les phrases « vipère » pour les enfants du groupe « balcon », les phrases « bal » pour les enfants du groupe « vie »...) sont jouées de manière continue, jusqu'à ce qu'un essai soit lancé par l'expérimentateur. Au début du mot test (ou au début du mot piège), une fenêtre de 2.5 secondes s'ouvre, pendant laquelle le comportement de l'enfant est codé (s'il tourne correctement la tête sur le mot test pendant cette période, une des boîtes s'allumera pendant 2.5 secondes). Si un enfant ne tourne pas la tête sur trois essais tests successifs, il repasse alors automatiquement en deuxième phase de rappel jusqu'à ce qu'il réussisse à nouveau la tâche demandée (l'expérimentateur peut aussi décider de relancer lui-même cette phase de rappel, par exemple, si l'enfant n'a pas eu d'essai test depuis longtemps et semble s'être désintéressé totalement des boîtes de renforcement). Si une pause est nécessaire pendant le test (par exemple si l'enfant est trop énervé ou semble trop se désintéresser du test), celle-ci a lieu hors de la cabine de test pendant le temps nécessaire, et l'expérience reprend ensuite avec la deuxième phase de rappel.

#### *5.1.4 Participants*

Trente-six enfants français, âgés de 16 mois, ont participé à cette expérience (20 dans le groupe bisyllabique et 16 dans le groupe monosyllabique). Cent quatre enfants supplémentaires ont été testés mais n'ont pas été intégrés dans les analyses : une partie de ces enfants a été rejetée au moment de la première session (71 enfants qui n'étaient pas intéressés par l'expérience, qui avaient peur, qui n'étaient pas coopératifs car fatigués ou malades, ou qui ne sont pas revenus pour la deuxième session alors qu'ils avaient bien réussi la première). Le reste des enfants (33) a été rejeté au moment de la deuxième session pour différentes raisons (fatigué, pas intéressé, pas coopératif...). Le taux de rejet est donc important (près de 74%), mais il faut noter que cette expérience était relativement longue (environ 10-15 minutes pour la première session et 20 minutes pour la deuxième) et assez délicate à contrôler.

## 5.2 RESULTATS

### 5.2.1 *Analyses acoustiques des stimuli*

Nous avons réalisé des analyses acoustiques des différentes phrases expérimentales, grâce au logiciel Praat, pour évaluer la réalisation de la frontière de groupe phonologique. Nous nous sommes centrés sur les analyses de durée, de fréquence fondamentale et d'énergie des différents segments des syllabes cibles.

#### 5.2.1.1 Analyse de la durée des segments

Les résultats obtenus pour les valeurs de durée sont indiqués dans le Tableau 5 ci-dessous :

		Frontière de groupe phonologique ...bal] [con... ...vie] [per...	Pas de frontière balcon vipère	Difference	t-test	Allongement (%)	
		Moyenne (ms)	Moyenne (ms)	Moyenne (ms)	Erreur standard	<i>t</i> (23) <i>p</i>	
Attaque 1	b v	102	73	29	7.1	4.1 <.001	39.8
Voyelle 1	a i	306	129	177	19.8	9.0 <10 <sup>-9</sup>	137.4
Coda 1	l	285	124	161	6.2	18.4 <10 <sup>-9</sup>	129.3
Attaque 2	k p	124	120	4	8.2	<1 0.7	3
Voyelle 2	õ ε	124	303	-180	18.8	9.6 <10 <sup>-9</sup>	-144.9
Coda 2	r	88	132	-44	4.8	5.0 <10 <sup>-5</sup>	-49.5

**Tableau 5 (analyses acoustiques – étude 3) : Durées moyennes des différents segments des syllabes cibles (attaque, voyelle et coda) quand ces syllabes formaient un mot (donc quand il n’y avait pas de frontière de groupe phonologique entre ces deux syllabes, comme dans « [un grand balcon]... ») versus quand ces syllabes étaient séparées par une frontière de groupe phonologique (comme dans « [un grand bal] [convenait...] »).**

Les analyses montrent des patterns de durée différents selon que les segments considérés sont en fin de groupe phonologique (et aussi en position finale de mot) ou qu’ils sont en début de groupe phonologique (et en début de mot).

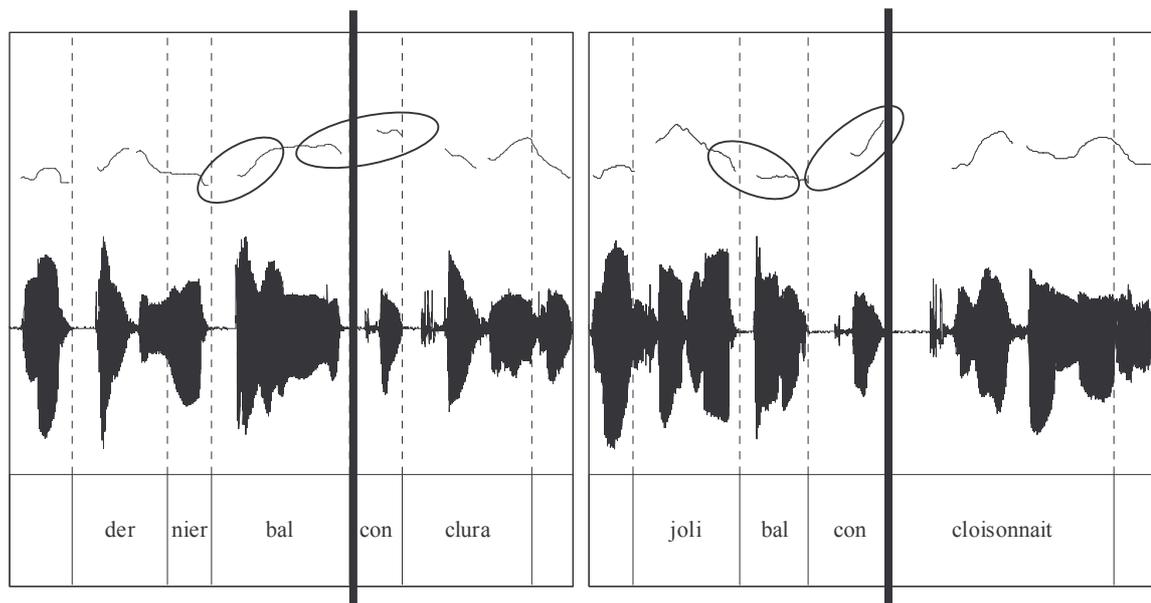
Comme il a été décrit dans la littérature (Delais-Roussarie, 1995 ; Gussenhoven & Rietveld, 1992 ; Wightman et al., 1992), nous avons trouvé un allongement des segments en fin de groupe phonologique, qui concerne en particulier la voyelle et la coda de la syllabe. Par exemple, le /a/ de « **ba**l » (en fin de groupe phonologique) est plus long de 137% que le /a/ de « **ba**lcon » (en milieu de mot). Cet allongement est de presque 130% quand on regarde la coda (le /l/ de « **ba**l » est plus long que le /l/ de « **ba**lcon »). De même, le /õ/ de « **bal**con » est beaucoup plus long que le /õ/ situé dans les phrases comprenant le mot « bal » suivi par un verbe commençant par « **con**... » (145% d’allongement)

Que ce soit les phrases pour les items « balcon » ou celles pour les items « vipère », on constate le même résultat : les segments situés en fin de groupe phonologique sont beaucoup plus longs que les mêmes segments situés au milieu d’un mot. Les segments constitutifs des syllabes /bal/ et /vi/ sont plus longs quand ils

forment les mots cibles monosyllabiques « bal » et « vie » que quand ils font partie des mots cibles bisyllabiques « balcon » et « vipère ». A l'inverse, les syllabes /kɔ/ et /pɛr/ sont plus longues quand elles sont les syllabes finales des mots « balcon » et « vipère » que quand elles forment le début du verbe comme dans « convenait » et « perpétue ». Ceci est dû à la position des segments dans la structure prosodique (plus longs en fin de groupe phonologique).

### 5.2.1.2 Analyse de la fréquence fondamentale

La Figure 9 ci-dessous représente le signal acoustique de deux paires de phrases expérimentales (extraits d'une phrase « bal » à gauche, et d'une phrase « balcon » à droite, centrés autour des syllabes cibles que nous avons étudiées), ainsi que la courbe de la fréquence fondamentale de cette zone.



**Figure 9 (analyses acoustiques – étude 3) : Signal acoustique et courbe de fréquence fondamentale au niveau des zones ambiguës, contenant les syllabes /bal/ et /con/. La figure de gauche représente une phrase monosyllabique : les deux syllabes sont séparées par une frontière prosodique (« [le dernier bal] [conclura la saison] »). La figure de droite représente une phrase bisyllabique : les deux syllabes forment un seul mot (« [le joli balcon] [cloisonnait la terrasse] »). Les barres noires en gras représentent la position de la frontière de groupe phonologique dans les phrases. Les parties entourées indiquent les zones qui vont être étudiées pour comparer les courbes de fréquence fondamentale.**

En ce qui concerne les courbes de fréquence fondamentale, nous avons d'abord mesuré la première zone indiquée sur chaque figure, c'est-à-dire la courbe arrivant à la première voyelle du mot cible (la voyelle /a/ de « **bal** » et de « **balcon** »). Pour cela, nous avons calculé la différence de fréquence entre cette syllabe cible et la voyelle précédente. Dans les phrases monosyllabiques, nous avons par exemple mesuré la différence entre le /a/ et le /e/ dans « [le dernier **bal**] [conclura la saison] »<sup>25</sup>. Dans cette condition, nous avons trouvé que la courbe de fréquence est ascendante. La fréquence monte significativement de 78 Hz entre ces deux voyelles (donc en fin de groupe phonologique,  $t(23)=2.1$ ,  $p<10^{-7}$ , avec un écart-type de 49 Hz). Par contre, dans la condition bisyllabique (quand le mot cible est « balcon »), la fréquence fondamentale tend à diminuer de 34 Hz ( $t(23)=2.1$ ,  $p=.05$ , écart-type de 80 Hz). La différence entre ces deux conditions est évidemment significative ( $t(23)=6.2$ ,  $p<.001$ ).

Nous avons ensuite mesuré la courbe de fréquence fondamentale au niveau de la deuxième zone indiquée sur les figures, c'est-à-dire la courbe partant de la première voyelle du mot cible et allant à la voyelle suivante. Dans la condition monosyllabique, nous avons par exemple mesuré la différence entre le /ɔ/ et le /a/ dans « [le dernier **bal**] [**con**clura la saison] »). Quand les deux voyelles considérées se trouvent de chaque côté de la frontière de groupe phonologique, la courbe ne diffère pas entre ces deux voyelles et reste stable (différence non significative de 7 Hz,  $t(23)<1$ , mais il y a une grande variabilité entre les stimuli, avec un écart-type de la différence de 60 Hz, et des valeurs allant de -165 Hz à +87 Hz). Par contre, dans la condition bisyllabique (on mesure alors la différence entre le /ɔ/ et le /a/ de « balcon » dans « [le joli **balcon**] [cloisonnait la terrasse] »), on observe une différence de 119 Hz, donc à nouveau une augmentation significative de la fréquence fondamentale en fin de groupe phonologique ( $t(23)=2.1$ ,  $p<10^{-9}$ , écart-type de 52 Hz).

Ces résultats sont en adéquation avec ce qui avait été trouvé dans la littérature, à savoir une montée de la fréquence fondamentale à la fin d'un groupe phonologique, en français (Di Cristo, 2000b; Welby, 2003b).

---

<sup>25</sup> Par souci de compréhension, nous garderons les mêmes exemples de phrases expérimentales pour expliciter nos analyses. Mais les chiffres indiqués sont ceux trouvés en moyenne entre toutes les phrases (les courbes de fréquence fondamentale indiquées sur les figures pourront donc être parfois un peu différentes de ce que nous avons trouvé en moyenne)

En résumé, on trouve que dans les phrases « balcon » (à droite dans la figure 1), la courbe de fréquence fondamentale est d'abord descendante (sur la première syllabe du mot cible) puis ascendante (sur la dernière syllabe du mot cible, en fin de groupe phonologique) ; dans les phrases « bal » (à gauche de la figure 1), la courbe est ascendante sur le mot cible (donc en fin de groupe phonologique) puis elle reste au même niveau jusqu'à la syllabe suivante. Ainsi, quand les enfants entendent la syllabe /bal/, le pattern intonatif n'est pas le même selon que cette syllabe est au début d'un mot ou à la fin d'un groupe phonologique.

### 5.2.1.3 Analyse de l'énergie des voyelles

Les analyses d'énergie ont aussi mis en évidence des différences entre les stimuli des paires de phrases expérimentales, mais de manière moins marquée. En fait, les analyses ont montré une différence entre les syllabes /kɔ̃/ (et les syllabes /pɛr/) qui étaient soit la dernière du mot cible (/kɔ̃/ dans « balcon »), soit la première du verbe (/kɔ̃/ dans « convenait »). En effet, dans la première situation, nous avons trouvé une énergie de 0.104, alors que dans la deuxième la valeur atteint 0.148. Cette différence de 0.044 est significative ( $t(23)=4.5$ ,  $p<.001$ ). La syllabe semble donc être produite avec une énergie plus forte quand elle est au début d'un mot (et en début de groupe phonologique) que quand elle est en fin de mot (et en fin de groupe phonologique).

Toutes ces analyses acoustiques indiquent clairement que les mots cibles sont réalisés de manière différente dans les paires de phrases expérimentales. Leur durée, leur courbe de fréquence fondamentale et leur énergie, sont différentes selon que les deux syllabes cibles forment réellement un mot ou qu'elles appartiennent à deux mots différents et sont alors séparées par une frontière de groupe phonologique.

Nous allons maintenant examiner si ces différences prosodiques ont influencé le comportement des bébés.

### 5.2.2 Résultats comportementaux

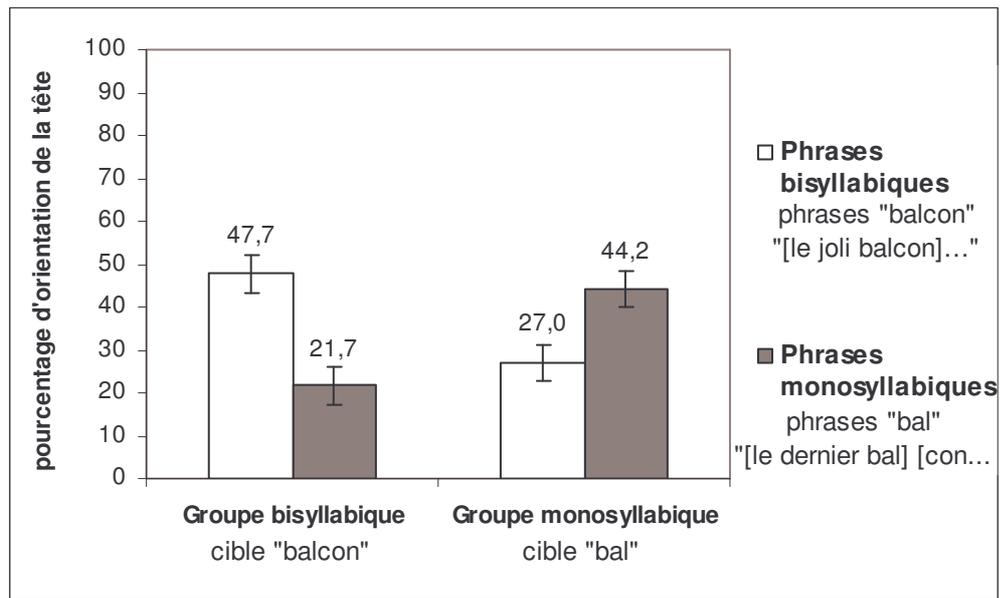
L'orientation de la tête des enfants était codée en temps réel pendant les expériences. Ceci pouvait parfois poser problèmes : par exemple, il y avait un délai variable entre le moment où l'expérimentateur demandait le lancement d'un essai et le moment où cet essai était effectivement lancé (ceci était dû au fait que la phrase test ne pouvait être jouée qu'une fois la phrase de fond sonore terminée). Pendant ce délai, l'enfant pouvait complètement changer d'attitude et passer d'un état très calme à un état très agité, mais surtout l'enfant pouvait s'orienter vers les haut-parleurs pendant ce délai et être donc déjà en train de regarder dans cette direction au moment où l'essai était joué. De plus, le codage en temps réel n'était pas le plus précis possible au niveau des temps de réaction, puisqu'il faut bien sûr prendre en compte le temps nécessaire pour que l'expérimentateur décide que l'enfant regarde bien vers les boîtes, et le temps pris pour appuyer sur le bouton de la souris.

Nous avons donc recodé toutes les deuxièmes sessions des enfants gardés dans les analyses, grâce à l'enregistrement vidéo de l'expérience. Ce codage se déroulait en aveugle (sans savoir ce que l'enfant entendait comme stimulus). La vidéo se positionnait au début de chaque essai, et l'expérimentateur vérifiait alors que l'enfant était bien concentré face à l'assistant, et non déjà tourné vers les haut-parleurs. Dans le cas contraire, l'essai n'était pas intégré dans les analyses (ceci concerne 36 essais, 21 dans le groupe bisyllabique et 15 dans le groupe monosyllabique, soit 4.2% du nombre total d'essais).

Pour les essais valides (enfant non tourné vers les haut-parleurs au début de l'essai), l'expérimentateur vérifiait ensuite le comportement de l'enfant pendant une fenêtre d'environ 2.5 secondes. Certaines décisions et interprétations initiales, faites en temps réel, ont été modifiées par ce codage, comme des essais codés « tourné » (l'enfant regarde vers les boîtes) qui sont devenus des essais « non tourné » (l'enfant se tournait pour regarder sa mère...). Ceci concerne 9 essais (soit 1% du nombre total d'essais). Dans d'autres cas, l'enfant tournait la tête vers les boîtes juste à la fin de la fenêtre de 2.5 secondes dans laquelle on pouvait coder le comportement de l'enfant. Le temps que l'expérimentateur prenne sa décision, plus le temps de réaction moteur, ont fait que quelques essais ont été codés comme « non tourné » alors que l'enfant s'était orienté

vers les boîtes dans le temps imparti (ceci concerne 12 essais, soit 1.4% du nombre total d'essais).

Après cette étape de recodage, nous avons mesuré le pourcentage d'orientation de la tête des enfants sur les phrases expérimentales. Ces résultats sont représentés dans la Figure 10 ci-dessous (voir les annexes C.2 et C.3 pour les résultats individuels par sujets et par items).



**Figure 10 (résultats de l'étude 3) : Pourcentage moyen d'orientation de la tête des enfants de 16 mois du groupe bisyllabique (barres à gauche, mot cible « balcon » ou « vipère ») et du groupe monosyllabique (barres à droite, mot cible « bal » ou « vie »). Les enfants entendaient des phrases contenant le mot bisyllabique (phrases « balcon » ou « vipère », indiquées par les barres blanches) et des phrases contenant les deux syllabes cibles séparées par une frontière de groupe phonologique (phrases « bal » ou « vie », indiquées par les barres grises). Les barres d'erreur représentent les erreurs standard de la différence des moyennes.**

Deux analyses de variance ont été conduites sur ces données, une avec les participants, l'autre avec les items comme variable aléatoire. Dans l'analyse par sujet, nous avons déclaré deux facteurs inter-sujets : le Groupe (bisyllabique versus monosyllabique) et le Matériel (c'est-à-dire un facteur de contrebalancement, entre les sujets devant répondre aux mots vipère/vie, et ceux ayant été habitués aux mots balcon/bal). Nous avons aussi déclaré un facteur intra-sujets : le type de Phrase (phrases bisyllabiques versus phrases monosyllabiques). L'analyse par item déclare trois

facteurs : un facteur inter-items (Matériel) et deux facteurs intra-items (le Groupe et le type de Phrase).

Nous avons observé un effet principal du type de Phrase entendue uniquement dans l'analyse par sujet ( $F_1(1,32)=4.8$ ,  $p=.04$  ;  $F_2(1,22)=2.1$ ,  $p=0.2$ ), de même qu'un effet principal de Matériel ( $F_1(1,32)=5.1$ ,  $p=.03$  ;  $F_2(1,22)=6.2$ ,  $p=.02$  ;  $\min F'(1,54)=2.8$ ,  $p=0.1$ ). L'interaction cruciale entre le facteur Groupe et le facteur Phrase s'est révélée significative ( $F_1(1,32)=45.3$ ,  $p<.001$  ;  $F_2(1,22)=37.9$ ,  $p<.001$ ), reflétant le fait que les enfants du groupe bisyllabique tournent plus leur tête sur les phrases bisyllabiques que sur les phrases monosyllabiques (48,2% versus 21,7%,  $F_1(1,18)=32.6$ ,  $p<.0001$  ;  $F_2(1,22)=21.8$ ,  $p=.0001$ ), tandis que les enfants du groupe monosyllabique montrent le pattern de réponse inverse (27% versus 44,2%,  $F_1(1,14)=15.7$ ,  $p=.001$  ;  $F_2(1,22)=20.7$ ,  $p=.0002$ ).

Nous avons ensuite calculé la valeur du A' pour chaque enfant (version non-paramétrique du d') en utilisant le nombre de succès (hits), d'oublis (misses), de fausses alarmes et de rejets corrects (Grier, 1971). Pour les enfants du groupe bisyllabique, la valeur du A' moyen est de 0.72 (l'erreur standard est de .03), ce qui est supérieur au niveau du hasard représenté par la valeur 0.5 ( $t(19)=6.6$ ,  $p<10^{-5}$ ). Pour les enfants du groupe monosyllabique, cette valeur atteint une moyenne de 0.64 (erreur standard de .03), significativement différente du hasard ( $t(15)=4.2$ ,  $p=.001$ ), mais un peu plus faible que celle obtenue par le groupe bisyllabique ( $t(35)=2.09$ ,  $p=.04$ ).

Une analyse des résultats individuels a aussi été réalisée. Pour déterminer quel score de A' reflète une performance significativement différente du hasard, nous avons réalisé une simulation de Monte Carlo. Un million de séquences de 24 essais (12 essais où il fallait répondre et 12 essais où il ne fallait pas répondre) a été générée, basée sur des réponses aléatoires. Un A' a été calculé pour chaque séquence, et la distribution de ces scores a été évaluée. Cette simulation montre que des scores de A' supérieurs ou égaux à .755 apparaissent par hasard dans moins de 5% des séquences. Nous avons comparé les résultats individuels des enfants à cette valeur. En ce qui concerne le groupe bisyllabique, nous trouvons que 9 enfants sur 20 obtiennent une performance significativement supérieure au hasard. Dix-huit enfants sur 20 obtiennent une différence dans leur taux de réponse qui va dans la direction attendue (à savoir plus de succès sur les phrases bisyllabiques que de fausses alarmes sur les phrases

monosyllabiques, ce qui donne un score de A' supérieur à .5). Pour le groupe monosyllabique, 3 enfants sur 16 montrent une performance supérieure au hasard (deux autres enfants obtiennent un A' de .754). Quatorze enfants sur 16 ont un comportement qui va dans la direction attendue (plus de succès sur les phrases monosyllabiques que de fausses alarmes sur les phrases bisyllabiques).

Nous avons également analysé les latences d'orientation sur les phrases tests et les phrases pièges. Nous avons comparé les temps de réaction sur les succès à ceux enregistrés pour les fausses alarmes. Dans l'analyse par sujet, nous avons déclaré deux facteurs inter-sujets : Groupe (bisyllabique versus monosyllabique) et Matériel (balcon/bal versus vipère/vie) ; un facteur intra-sujets a été déclaré : le type de Phrase entendue (phrases tests versus phrases pièges). Dans l'analyse par item, nous avons déclaré les facteurs Groupe et Phrase comme des facteurs intra-items alors que le facteur Matériel a été déclaré en inter-items.

Les analyses n'ont indiqué aucun effet principal significatif ni aucune interaction significative. En particulier, les enfants ont mis autant de temps, en moyenne, pour tourner la tête quand ils répondent correctement aux mots tests que quand ils se font piéger (pour les succès : 1327 ms, écart-type de 249 ms / pour les fausses alarmes : 1324 ms, écart-type de 342 ms).

Nous avons finalement réalisé des analyses de corrélation pour essayer de voir s'il y avait un lien entre les réponses des sujets et les indices prosodiques (de durée et de fréquence fondamentale) contenus dans les stimuli. Nous avons centré notre attention sur les résultats du groupe bisyllabique, c'est-à-dire ceux qui devaient répondre au mot bisyllabique (« balcon ») et qui pouvaient clairement utiliser la présence d'une frontière de groupe phonologique, pour distinguer les phrases bisyllabiques (« balcon ») et les phrases monosyllabiques (« bal »).

Pour chaque item, nous avons mesuré un score de réussite global, en calculant la différence entre le pourcentage de réussite (regards vers les phrases bisyllabiques) et le pourcentage de fausses alarmes (regards vers les phrases monosyllabiques). Nous avons analysé la corrélation qu'il y avait entre ce score de réussite et le pourcentage d'allongement de la première syllabe /bal/ et /vi/. Notre hypothèse était la suivante :

plus l'allongement de la première syllabe est grand (dans les phrases monosyllabiques, comparées aux phrases bisyllabiques), plus les enfants devraient pouvoir postuler la présence d'une frontière de groupe phonologique entre les deux syllabes cibles dans les phrases monosyllabiques, et plus ils devraient pouvoir se retenir de répondre aux phrases monosyllabiques (le taux de fausses alarmes devrait baisser au fur et à mesure que l'allongement de la syllabe augmente, ce qui devrait augmenter le score global de réussite calculé). Mais les analyses n'ont montré aucune corrélation significative entre ces variables ( $r = -0.15$ ,  $p=0.47$ ).

Nous avons ensuite mesuré si une corrélation existait entre les courbes de fréquence fondamentale et les résultats des enfants du groupe bisyllabique. Le résultat principal qui était ressorti des analyses acoustiques sur les valeurs de fréquence fondamentale était qu'il y avait une montée de la fréquence fondamentale en fin de groupe phonologique (montée de la fréquence fondamentale sur les syllabes /bal/ et /vi/ dans les phrases monosyllabiques, comparée à une baisse de la fréquence fondamentale sur ces mêmes syllabes dans les phrases bisyllabiques). Notre hypothèse était que les enfants du groupe bisyllabique pouvaient peut-être utiliser la montée de la fréquence fondamentale sur les syllabes /bal/ et /vi/ dans les phrases monosyllabiques, pour inférer la présence d'une frontière prosodique derrière cette syllabe, et donc pour ne pas répondre à ces phrases monosyllabiques (et faire moins de fausses alarmes). Nous avons donc mesuré le coefficient de corrélation entre le pourcentage de fausses alarmes sur les phrases monosyllabiques et la montée de la fréquence fondamentale sur les premières syllabes cibles : celui-ci s'est avéré non significatif ( $r = -0.22$ ,  $p=0.29$ ).

Ces analyses ne nous ont donc pas permis de mettre en évidence une corrélation significative entre les indices prosodiques marquant une frontière de groupe phonologique et les réponses des sujets. Cela ne veut pas dire bien sûr qu'aucune corrélation n'existe : il faut peut-être prendre d'autres facteurs en compte pour expliquer tous les résultats des enfants (ce serait une constellation d'indices qui guiderait les réponses des sujets). De même, un autre point à noter est la grande variabilité qui existe entre les stimuli au niveau des indices prosodiques produits dans la zone ambiguë, sur les syllabes cibles.

Nous avons également regardé quelles étaient les phrases qui avaient été le mieux réussies et celles qui avaient obtenu les moins bons résultats (en prenant en compte tous les résultats de l'expérience, donc aussi bien les réponses des enfants du groupe bisyllabique que celles des enfants du groupe monosyllabique ; voir l'annexe C.3 pour les résultats détaillés par item).

Parmi les phrases les mieux réussies, on trouve les phrases 02 et 12 de l'item « bal ». Toutes les deux ont obtenu des taux de réponses importants de la part des enfants du groupe monosyllabique (ils ont correctement répondu à ces phrases, donc correctement identifié le mot « bal », dans 63% des cas pour chacune de ces phrases). Par contre, aucun des enfants du groupe bisyllabique n'a tourné la tête en entendant ces phrases (0% de fausses alarmes). Il semble donc que ces phrases aient plutôt été interprétées comme contenant le mot « bal », et non le mot « balcon ». Est-ce que les indices prosodiques compris dans ces phrases ont pu aider les enfants à correctement les segmenter ? Il y a en effet une des mesures acoustiques qui semble obtenir des résultats bien supérieurs à la moyenne : la longueur de la première syllabe, donc la longueur de la syllabe /bal/ dans ces phrases monosyllabiques. Alors que l'allongement de ces syllabes ne semble pas différent de la moyenne, la longueur intrinsèque de ces syllabes apparaît comme relativement élevée : elles font en effet partie des plus longues syllabes /bal/ produites dans les phrases monosyllabiques (respectivement 669 ms et 696 ms, pour une moyenne de 621 ms). Par contre, les autres données (que ces données concernent les valeurs de fréquence fondamentale ou les valeurs d'énergie) ne semblent pas obtenir des résultats déviants par rapport à la moyenne<sup>26</sup>.

De l'autre côté, parmi les phrases les moins bien réussies, on trouve la phrase 12 de l'item « vipère » et la phrase 10 de l'item « bal ». Cette dernière a entraîné un fort taux de réponse de la part des enfants, quel que soit leur groupe expérimental : on mesure en effet 50% de fausses alarmes (réponses du groupe bisyllabique) et 63% de « hits » (réponses du groupe monosyllabique). Si l'on regarde les analyses acoustiques réalisées sur cette phrase, on constate que, contrairement à la moyenne, la courbe de fréquence fondamentale n'a pas diminué sur la syllabe /bal/. Ceci est dû au fait que la syllabe précédente avait une fréquence très élevée. Cette fréquence élevée a pu peut-être

---

<sup>26</sup> Nous ne disons pas qu'une syllabe très longue sera toujours interprétée comme un seul mot, ni qu'elle ne sera jamais considérée comme le début potentiel d'un mot. De toute façon, un contre-exemple existe en effet dans ces résultats : la phrase bal06, dont la syllabe /bal/ a une durée de 676 ms, mais qui n'entraîne pas une grande réussite de la part des enfants : elle obtient 38% de « hits » (réponses du groupe monosyllabique) et 20% de fausses alarmes (réponses du groupe bisyllabique).

attirer l'attention des enfants, qui ont eu alors tendance à tourner la tête pour ces phrases-là. Une autre mesure pourrait peut-être expliquer pourquoi les enfants ont eu tendance à toujours tourner la tête en entendant cette phrase : en effet, nous avons aussi remarqué que l'énergie de la voyelle de cette syllabe /bal/ était relativement importante (0.250) comparée aux autres voyelles /a/ (moyenne de 0.155 avec un écart-type de 0.052). Peut-être que la perception de cette voyelle produite avec beaucoup d'énergie a attiré l'attention des enfants.

Quant à la phrase 12 de l'item « vipère », elle a obtenu un pattern de réponse inverse à ce que l'on attendait, à savoir plus de fausses alarmes que de « hits » (respectivement, 38% de réponses pour les enfants du groupe monosyllabique, qui ont donc pensé, dans ces cas-là, entendre le mot « vie », versus 22% de réponses pour les enfants du groupe bisyllabique, qui ont donc plus souvent considéré entendre le mot « vie » que le mot « vipère »). Il semblerait donc que cette phrase ait plutôt été interprétée comme une phrase « vie » que comme une phrase « vipère ». En fait, une des explications qui pourrait être donnée pour rendre compte du résultat obtenu se rapporte à la courbe de la fréquence fondamentale : on trouve en effet une montée de la fréquence fondamentale sur la première syllabe de « vipère » (montée importante de 148 Hz), suivie d'une baisse de 39 Hz sur la deuxième syllabe. Ce contour intonatif est typiquement ce qui avait été décrit pour les phrases monosyllabiques (et contraire à ce qui avait été trouvé en moyenne pour les phrases bisyllabiques), et c'est ce qui a pu entraîner une plus forte proportion d'interprétation monosyllabique.

Ainsi, si l'on regarde chaque phrase isolément, on peut parfois trouver des liens entre les indices prosodiques (de durée, de fréquence fondamentale et d'énergie) et les résultats obtenus.

### 5.3 DISCUSSION

Cette expérience indique que les enfants français de 16 mois peuvent utiliser les frontières de groupes phonologiques pour segmenter la parole continue et retrouver les mots dans les phrases. En effet, quand les enfants sont habitués à détecter un mot bisyllabique, ils tournent ensuite significativement plus la tête quand ils entendent réellement ce mot cible dans des phrases, que quand ils entendent ses deux syllabes

constitutives séparées par une frontière de groupe phonologique. A l'inverse, des enfants habitués à répondre à un mot monosyllabique, tournent significativement plus la tête quand ils entendent réellement ce mot cible dans des phrases que quand cette syllabe fait partie d'un autre mot.

Le comportement des enfants des deux groupes est significativement différent alors qu'ils entendent les mêmes phrases expérimentales. L'interaction significative entre le Groupe expérimental et le type de Phrase test indique que les enfants du groupe bisyllabique ont plus répondu aux phrases bisyllabiques « balcon » qu'aux phrases monosyllabiques « bal », non pas pour des raisons artefactuelles (par exemple, les propriétés acoustiques des phrases « balcon » auraient pu être différentes et plus attractives que celles des phrases « bal »), mais bien car la présence de la frontière de groupe phonologique entre les syllabes /bal/ et /con/ a réduit la détection du mot « balcon » dans les phrases monosyllabiques. Ainsi, nous pouvons conclure que la présence ou l'absence d'une frontière de groupe phonologique contraint l'accès au lexique des enfants français de 16 mois.

Il faut noter que nous avons fait l'hypothèse d'une différence de résultats entre les enfants du groupe bisyllabique et ceux du groupe monosyllabique. En effet, puisque les frontières de groupe phonologique coïncident toujours avec des frontières de mot, il devrait être très peu probable d'accéder à un « mot » dont les syllabes chevauchent une frontière de groupe phonologique (il devrait être très rare d'activer le mot « balcon » dans une phrase « ...bal ] [con... ») et c'est ce que nous avons trouvé avec les résultats du groupe bisyllabique. Par contre, l'absence d'une frontière de groupe phonologique ne prédit pas une absence de frontière de mot (tous les mots ne sont pas délimités par des frontières de groupe phonologique), et en ce sens, la tâche des enfants du groupe monosyllabique était plus difficile : pour eux, détecter « bal » dans des phrases « [...bal ] [con... ] » était supposé être relativement facile (puisque la syllabe cible est suivie par une frontière de groupe phonologique, donc relativement facile à segmenter). Mais ne pas répondre à la syllabe « bal » dans des phrases « [...balcon ] » est un comportement plus difficile à avoir puisque cette syllabe est effectivement présente dans le signal de parole. On s'attendait donc à une moins bonne performance des enfants du groupe monosyllabique (due à une plus grande proportion de fausses alarmes sur les phrases bisyllabiques). Les pourcentages de réponses aux deux types de phrases ne semblent pas refléter cette hypothèse. En effet, la taille de l'effet (différence entre le pourcentage de

succès et le pourcentage de fausses alarmes) est de 26.5% pour le groupe bisyllabique, donc plus importante que pour le groupe monosyllabique (17.2%), mais cette différence n'est pas significative ( $F_1(1,32)=2, p=0.2$ ).

Pour tester si la tâche du groupe monosyllabique est plus difficile que celle du groupe bisyllabique, nous pouvons également comparer les scores de A'. Le score obtenu en moyenne par le groupe bisyllabique est significativement supérieur à celui obtenu par le groupe monosyllabique (0.72 versus 0.64,  $t(35)=2.1, p=.04$ ), suggérant effectivement que les enfants du groupe bisyllabique ont plus de facilités à rejeter les phrases monosyllabiques que l'inverse. De même, nous avons vu que plus d'enfants du groupe bisyllabique réussissent à obtenir une performance significativement supérieure à la chance, comparé aux enfants du groupe monosyllabique (45% des enfants du groupe bisyllabique versus 19% des enfants du groupe monosyllabique).

En conclusion, les résultats de cette expérience indiquent que les enfants français de 16 mois peuvent reconnaître un mot bisyllabique au milieu d'une phrase, et qu'ils peuvent aussi le distinguer de la suite de ses deux syllabes constitutives séparées par une frontière de groupe phonologique. Ils peuvent donc exploiter la présence d'une frontière de groupe phonologique pour inférer la présence d'une frontière de mot, et extraire ainsi les mots de la parole continue (rappelons que l'annexe E contient les résultats de deux expériences pilotes auprès d'enfants français de 13 mois, ainsi qu'une comparaison inter-langues entre les résultats américains et les résultats français).

Une des conditions nécessaire à l'hypothèse que nous avons développée sur la question de l'acquisition de la syntaxe par les enfants, vient donc d'être démontrée. Les jeunes enfants perçoivent et peuvent utiliser en temps réel les frontières des groupes phonologiques pour analyser les phrases qu'ils entendent.

Nous savons maintenant que les adultes et les enfants peuvent interpréter les frontières de groupes phonologiques comme des frontières de mot ; nous avons aussi démontré que les adultes pouvaient utiliser le fait que ces frontières prosodiques correspondent toujours à des frontières syntaxiques, pour guider en temps réel leur analyse syntaxique des phrases. Il semble donc plausible de postuler que les enfants pourraient eux aussi utiliser ces mêmes indices prosodiques pour guider l'acquisition de la syntaxe de leur langue maternelle et découvrir la structure syntaxique des phrases

qu'ils entendent. La dernière étude que nous allons maintenant présenter s'est intéressée à cette idée.

## CHAPITRE 6

# ROLE DES MOTS GRAMMATICaux ET DES FRONTIERES PROSODIQUES DANS L'ACQUISITION DE LA SYNTAXE (ETUDE 4)

Nous avons abordé, dans le chapitre 3, de nombreux indices qui pourraient être utilisés par les jeunes enfants pour acquérir la syntaxe de leur langue maternelle. Nous avons ensuite mis en évidence le rôle de deux indices particuliers qui nous paraissent fondamentaux :

- d'une part les mots grammaticaux : de par leurs caractéristiques phonologiques particulières et leur fréquente position en bordure d'unité, il a été montré que les enfants pouvaient acquérir assez rapidement ces items lexicaux. Ils pourraient ensuite permettre aux enfants, par exemple, d'inférer la catégorie syntaxique des mots de contenu qui les suivent (comme l'ont montré différentes expériences et analyses de corpus) ;

- d'autre part les informations prosodiques apportées par les frontières de groupes phonologiques : comme ces frontières prosodiques correspondent toujours à des frontières de constituants syntaxiques, elles pourraient permettre un découpage de la parole en unités pertinentes d'un point de vue syntaxique. Cette hypothèse a été vérifiée, chez des adultes français, dans les deux premières études de cette recherche. Et nous venons de voir (dans l'étude 3) que de jeunes enfants français de 16 mois percevaient eux aussi ces frontières prosodiques et pouvaient les utiliser pour analyser et segmenter les phrases qu'ils entendaient.

Le but de cette expérience est de vérifier si les informations apportées par les mots grammaticaux et les frontières prosodiques peuvent aider à la construction d'une représentation syntaxique des phrases, et ainsi aider à la catégorisation des mots en grandes catégories syntaxiques (comme les noms et les verbes).

L'idée de départ est que les frontières prosodiques pourraient aider au pré-découpage des phrases, en regroupant les mots en constituants syntaxiques. La catégorisation de ces constituants syntaxiques (c'est-à-dire trouver le label que l'on pourrait apposer à ces constituants, comme « groupe nominal » ou « groupe verbal ») serait possible grâce à l'utilisation des mots grammaticaux.

Par exemple, un enfant qui entend une phrase comme « le chien a mordu le ballon » pourrait avoir une première représentation basée sur la prosodie, comme « [xx] [xxx] [xxx] », où les crochets indiquent les frontières des groupes phonologiques (et donc les frontières de constituants syntaxiques). La reconnaissance des mots grammaticaux pourrait permettre d'étoffer cette représentation, qui deviendrait alors « [le x] [a xx] [le xx] ». Ces deux types d'informations, lexicales et prosodiques, permettraient alors à l'enfant de construire un squelette de la structure syntaxique de cette phrase, telle que « [le x]<sub>GN</sub> [a xx]<sub>GV</sub> [le xx]<sub>GN</sub> », où les frontières des constituants sont données par la prosodie, et où les labels de ces constituants peuvent être déduits des mots grammaticaux (un article est généralement suivi d'un nom, alors qu'un auxiliaire est généralement suivi d'un verbe).

Pour vérifier cette hypothèse, nous avons créé une expérience de détection de mot (détection d'une entité lexicale abstraite, similaire à ce que nous avons présenté dans la deuxième expérience de la première étude - chapitre 2). Les mots à détecter sont présentés aux sujets avec leur catégorie syntaxique (les sujets doivent détecter des noms et des verbes).

Pour construire le matériel, nous avons utilisé une langue artificielle : le Jabberwocky (d'après « Alice au pays des merveilles » de Lewis Carroll). La particularité de cette langue est qu'elle conserve certaines informations de la langue que nous voulons étudier : les informations prosodiques et les mots grammaticaux (articles, pronoms, auxiliaires...). Par contre, tous les mots de contenu (noms, verbes, adjectifs...) sont remplacés par des non-mots (qui pourraient être possibles dans la langue étudiée mais qui n'existent pas). Nous avons ainsi pu faire varier le contexte d'apparition des mots cibles que les sujets devaient détecter (présence ou non des mots grammaticaux et/ou des frontières prosodiques) dans des phrases qu'ils ne comprenaient pas.

Nous avons créé deux conditions expérimentales : dans la première condition, que nous appellerons « avec clitique », tous les mots cibles étaient directement précédés par un mot grammatical. La position de cet ensemble « clitique – mot cible » dans les constituants prosodiques des phrases a été variée. Nous avons ainsi mis en place trois sous-conditions :

(1) Une sous-condition où le couple « clitique – mot cible » formait à lui seul une unité prosodique, comme dans : (1a) « [tu **bamoules**] [saman ti] [à mon ada] » où « bamoules » est un verbe, et dont la traduction française pourrait être « [tu **travailles**] [souvent trop] [à mon avis] » ; et (1b) « [une **bamoule**] [dri se froliter] [dagou] » où « bamoule » est un nom, et dont la traduction pourrait être « [une **expo**] [doit se dérouler demain] » ;

(2) Une sous-condition où le couple « clitique – mot cible » se situe au début d'une unité prosodique mais dans laquelle le mot cible est suivi par un autre non-mot, comme dans : (2a) « [tu **bamoules** mirdan] [jali tru la sti] » pour la phrase verbe (« [tu **travailles** souvent] [jusque tard la nuit] »), et (2b) « [une **bamoule** ferdine] [pitora des zures meldunes] » pour la phrase nom (« [une **expo** nouvelle] [présentera des oeuvres modernes] ») ;

Dans ces deux sous-conditions, le mot cible est donc précédé des deux types d'informations testées dans cette expérience : une frontière prosodique directement suivie d'un mot grammatical. La différence entre ces deux sous-conditions vient de la présence ou de l'absence d'une frontière de groupe phonologique à la fin du mot-cible (on pourrait s'attendre à ce que le mot cible soit plus facilement extrait du signal de parole dans la première sous-condition, car mieux délimité, ce qui faciliterait la réponse des sujets).

(3) Une troisième sous-condition a ensuite été construite, dans laquelle le couple « clitique – mot cible » se trouve au milieu d'une unité prosodique, comme dans : (3a) « [croume tu **bamoules** bine] [ce nachi calune] » (« [comme tu **portes** bien] [ce petit tailleur] ») et « [tu as une **bamoule** rube] [pour goder dimare] » (« [tu as une **cagoule** rouge] [pour aller dehors] »). Dans ce dernier cas, la seule information pertinente qui précède directement le mot cible est la présence du mot grammatical. On peut s'attendre à ce que les sujets réussissent aussi à catégoriser correctement le mot cible dans cette condition. Par contre, la segmentation et la reconnaissance du mot grammatical pourrait être plus difficile dans cette sous-condition, de par l'absence de la frontière de groupe phonologique juste avant lui, ce qui pourrait entraîner une moins bonne performance des sujets dans cette sous-condition que dans les deux précédentes.

Dans une deuxième condition, condition que nous avons appelée « sans clitique », les mots cibles ne sont jamais précédés par un mot grammatical. Cette condition a été réalisée pour tester le rôle direct des frontières prosodiques, en comparant les réponses aux phrases nom et aux phrases verbe. En effet, les phrases ont été élaborées par paires de manière à obtenir la sous-condition suivante :

(4) Les sujets devront traiter des phrases verbe comme « [sa cramona] [**bamoule** muche] [le mirtou] » (« [sa camarade] [**travaille** mieux] [le matin] »), et des phrases nom comme « [sa cramona **bamoule**] [camiche dabou] » (« [sa formidable **expo**] [commence demain] »). Jusqu'au mot ambigu inclus, ces deux phrases sont identiques d'un point de vue phonémique. Par contre, elles diffèrent d'un point de vue syntaxique et donc prosodique. Dans la phrase verbe, on trouve une frontière de groupe phonologique juste avant le mot cible. Si cette information est perçue et utilisée par les sujets, on s'attend donc à ce que la reconnaissance du verbe (et le rejet de l'interprétation nom) soit possible, grâce à la présence de cette frontière de groupe phonologique entre le groupe nominal et le mot cible (encore faut-il que la première unité prosodique soit reconnue comme un groupe nominal). Par contre, dans la phrase nom, la frontière de groupe phonologique n'arrive qu'après le mot cible. On peut donc s'attendre à ce que les sujets aient plus de difficultés à correctement interpréter le mot cible comme un nom (et à ne pas le considérer comme un verbe) puisqu'ils n'auront pas eu accès à l'information désambiguïsante au moment de traiter le mot cible.

(5) Dans une dernière sous-condition, nous avons repris le principe des paires de phrases mentionné ci-dessus, mais nous avons fait précéder le mot cible d'un non-mot dont la dernière syllabe (ou la rime) était ambiguë car homophone à un mot grammatical. L'ambiguïté résidait dans le choix de cette homophonie : quand les mots cibles étaient des noms, la syllabe ou la rime précédente était homophone à un pronom ; quand les mots cibles étaient des verbes, la syllabe ou la rime précédente était homophone à un article. Nous avons par exemple utilisé des phrases verbe comme « [ma jine chalune] [**bamoule** muche] [le mirtou] » (« [ma jeune cousine] [**travaille** mieux] [le matin] »), et des phrases nom comme « [une cramotu **bamoule**] [camiche dabou] » (« [une incroyable **expo**] [commence demain] »). Si toutes les syllabes homophones à des mots grammaticaux sont considérées (même potentiellement) comme des mots grammaticaux, alors les sujets devraient soit être gênés dans cette condition

(temps de réaction plus longs), soit faire beaucoup d'erreurs (mal interpréter le mot cible). A nouveau, la présence de la frontière prosodique avant le mot cible dans les phrases verbe pourrait faciliter la tâche des sujets et entraîner des patterns de résultats différents pour les phrases nom et les phrases verbe.

Ainsi, toutes ces sous-conditions vont nous permettre de tester le rôle des mots grammaticaux et le rôle des frontières de groupes phonologiques dans la représentation syntaxique des phrases et la catégorisation syntaxique des mots. La dernière condition nous permettra d'élargir notre questionnement à la question de savoir ce qui est considéré comme un mot grammatical dans les phrases (est-ce que toutes les syllabes homophones à un mot grammatical sont considérées comme un mot grammatical potentiel ?)

## 6.1 METHODE

### 6.1.1 Matériel

Nous avons choisi 16 non-mots comme mots cibles, chacun de ces non-mots étant utilisé à la fois comme un nom et comme un verbe à travers toutes les conditions. Dans chaque sous-condition, nous avons créé 32 phrases expérimentales, 16 contenant le mot cible dans sa forme nominale, 16 contenant la forme verbale (voir annexe D.1 pour le détail des stimuli utilisés). Reprenons l'exemple « bamoule » pour illustrer les phrases construites dans chaque condition et sous-condition :

#### *CONDITION « AVEC CLITIQUE »*

##### (1) sous-condition 1

phrase verbe : « [tu **bamoules**] [saman ti] [à mon ada] »

phrase nom : « [une **bamoule**] [dri se froliter] [dagou] »

##### (2) sous-condition 2

phrase verbe : « [tu **bamoules** mirdan] [jali tru la sti] »

phrase nom : « [une **bamoule** ferdine] [pritora des zures meldunes] »

(3) sous-condition 3

phrase verbe : « [croume tu **bamoules** bine] [ce nachi calune] »

phrase nom : « [tu as une **bamoule** rube] [pour goder dimare] »

CONDITION « SANS CLITIQUE »

(4) sous-condition 4

phrase verbe : « [sa cramona] [**bamoule** muche] [le mirtou] »

phrase nom : « [sa cramona **bamoule**] [camiche dabou] »

(5) sous-condition 5

phrase verbe : « [ma jine chalune] [**bamoule** muche] [le mirtou] »

phrase nom : « [une cramotu **bamoule**] [camiche dabou] »

La position du mot cible dans la phrase (et donc la position du constituant prosodique contenant le mot cible) a été variée ; ainsi, la moitié des items se trouve en début de phrase (il y a entre 1 et 5 syllabes avant le mot cible, avec une moyenne de 2,6 syllabes), tandis que l'autre moitié se trouve en milieu de phrase (en moyenne 6,5 syllabes avant le mot cible, soit entre 4 et 9 syllabes).

Les phrases sont présentées aux sujets dans deux versions : (1) une version coïncidente dans laquelle la cible à détecter est présente dans la phrase (par exemple, cible nom à détecter – phrase nom entendue), et (2) une version non-coïncidente où la cible à détecter n'est pas présente dans la phrase que les sujets entendent puisqu'elle contient, au contraire, l'item homophone appartenant à l'autre catégorie (par exemple, cible nom à détecter – phrase verbe entendue).

Dans chaque sous-condition, un sujet entend les deux phrases expérimentales correspondant à chaque item (la phrase verbe et la phrase nom). Puisque chaque phrase est présentée en version coïncidente (cible appropriée) et en version non-coïncidente (cible non appropriée), chaque item apparaît dans 4 situations différentes, sur lesquelles chaque sujet n'en traitera que deux : pour chaque item, ils entendront la phrase nom et la phrase verbe ; une de ces phrases sera présentée en version coïncidente tandis que l'autre apparaîtra en version non-coïncidente (on a donc deux groupes de sujets pour qu'à la fin les quatre situations aient été traitées). De plus, les items présentés aux

participants ont été divisés en deux blocs, de manière à ce que les deux phrases correspondant à un item apparaissent dans des blocs différents (par exemple, phrase nom dans le premier bloc – phrase verbe dans le deuxième). L'ordre de présentation des blocs a été contrebalancé entre les participants (on arrive donc à 4 groupes expérimentaux : le groupe 1 qui entend bloc 1 puis bloc 2, le groupe 2 qui entend bloc 2 puis bloc 1, le groupe 3 qui entend bloc 3 puis bloc 4, et le groupe 4 qui entend bloc 4 puis bloc 3).

Au total, chaque sujet entend donc 160 phrases expérimentales, soit 32 dans chaque sous-condition. Ils doivent traiter 80 phrases nom (40 en version coïncidente et 40 en version non coïncidente) et 80 phrases verbe (la moitié en version coïncidente et l'autre moitié en non-coïncidente).

Les 160 phrases test et les 10 phrases d'entraînement ont été enregistrées par l'auteur. Elles ont été digitalisées à 22050 Hz et 16 bits, avec le logiciel Adobe Audition, dans une cabine insonorisée. Elles ont ensuite été éditées grâce aux programmes CoolEdit Pro et Praat. Le début du mot cible a été marqué manuellement dans le logiciel Praat.

Lors de l'enregistrement, la locutrice avait les phrases écrites sous les yeux. Pour chaque phrase en jabberwocky, un modèle en français avait été construit et indiqué au dessus de la phrase test. Chaque couple de phrases (en jabberwocky et en français) était apparié au niveau de la structure prosodique, de la structure syntaxique ainsi qu'au niveau du nombre de syllabes. Ceci était fait pour que la lecture des phrases en jabberwocky soit plus facile, et que les phrases expérimentales aient une intonation et un rythme naturels.

### 6.1.2 Procédure

L'expérience a été réalisée en utilisant le programme Expe6 de présentation des stimuli (Pallier et al., 1997)<sup>27</sup>. Un essai commençait par la présentation visuelle du mot cible, au centre de l'écran, pendant 1.5 secondes. Le mot cible était présenté avec une

---

<sup>27</sup> Voir le site internet : <http://www.lscp.net/expe/>

indication sur sa catégorie syntaxique (soit à l’infinitif pour détecter le verbe – par exemple « bamouler » – soit avec un article pour détecter le nom – par exemple « une bamoule »). Après la présentation de la cible, la phrase expérimentale était jouée (les stimuli auditifs étaient présentés à travers une carte son ProAudioSpectrum 16-bit). L’essai se terminait 2.5 secondes après la réponse du sujet ou à la fin de la présentation auditive de la phrase (en effet, les participants devaient soit appuyer sur un bouton réponse, placé devant eux, quand ils entendaient le mot cible dans la phrase, soit ne rien faire et attendre l’essai suivant quand le mot cible n’était pas présent dans la phrase). L’essai suivant commençait immédiatement après.

Les temps de réponse des sujets étaient mesurés depuis le début du mot cible. Les instructions indiquaient clairement aux participants qu’ils devaient répondre le plus rapidement possible, mais le plus précisément possible.

Durant la phase d’entraînement, on informait les participants de la justesse de leurs réponses (correct ou incorrect) : s’ils avaient donné une mauvaise réponse (c’est-à-dire s’ils avaient manqué le mot cible, ou s’ils avaient répondu alors que ce n’était pas le mot cible qui était présenté dans la phrase), on leur rejouait l’essai en question, de manière à ce qu’ils corrigent leur erreur. Quand un temps de réaction était enregistré (réponse correcte à la cible présente dans la phrase), celui-ci était présenté sur l’écran à la fin de l’essai. S’il était supérieur à 1 seconde, l’indication « trop lent » apparaissait, pour que les sujets essaient d’être encore plus rapides sur les essais suivants. Par contre, lors de la phase de test, aucune indication n’était donnée aux participants sur leurs réponses.

L’expérience durait en moyenne 25 minutes (avec une pause entre les deux blocs de stimuli). Tous les participants étaient testés individuellement, dans une pièce isolée, équipée d’un ordinateur, d’un casque et d’un boîtier réponse. Avant de commencer, chaque participant remplissait un formulaire indiquant sa date de naissance, un éventuel bilinguisme, et s’il avait des problèmes auditifs ou des problèmes de langage. Puis, il découvrait les consignes de l’expérience : les instructions étaient données oralement, en complément des instructions écrites à l’écran (voir annexe D.2 pour les instructions écrites détaillées). L’expérimentateur s’assurait que le sujet avait bien compris la tâche en restant près de lui pendant la phase d’entraînement, ce qui permettait aussi de clarifier toutes les questions éventuelles avant de commencer le test.

### 6.1.3 *Participants*

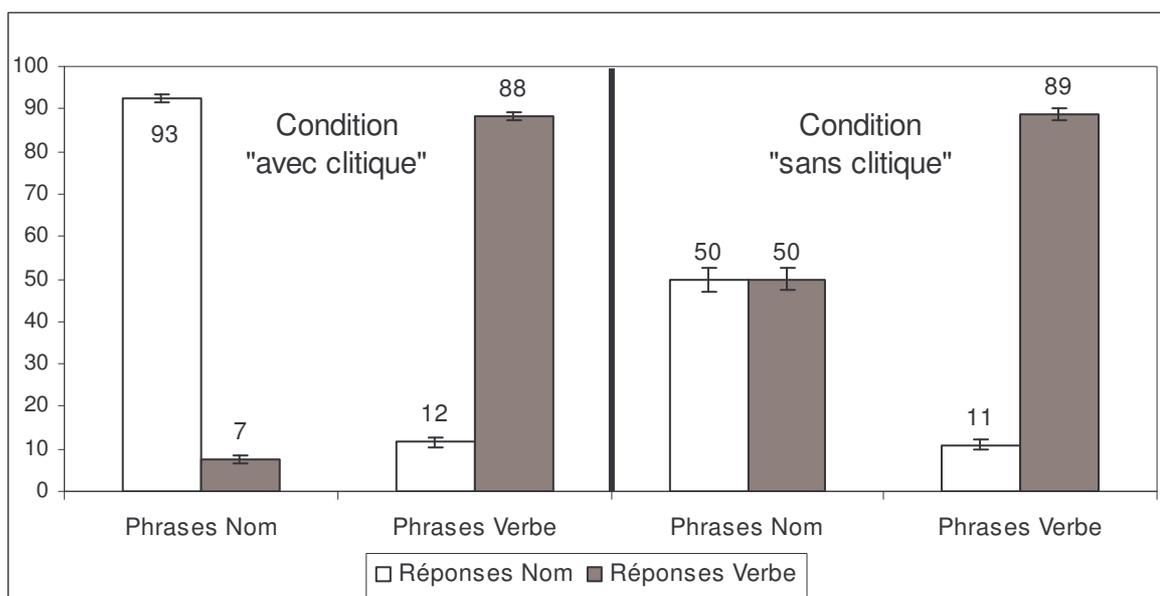
Cinquante-deux personnes ont participé à cette expérience et ont été dédommagées pour leur participation. Toutes étaient natives du français. Aucun d'eux n'a rapporté de problèmes auditifs ou de problèmes liés au langage. Avant de commencer l'expérience, ils ont été assignés aléatoirement à l'un des quatre groupes expérimentaux. Deux autres personnes ont été testées mais leurs réponses n'ont pas été incluses dans les analyses à cause de leurs temps de réponse trop lents (temps de réaction moyen supérieur à 1 seconde).

## 6.2 RESULTATS

### 6.2.1 *Analyses des réponses noms et verbe*

Nous avons analysé l'interprétation que les sujets avaient donnée au mot cible dans chaque phrase. Quand la cible à détecter était un nom et que les sujets appuyaient sur le bouton réponse (quelle que soit la phrase présentée, donc que ce soit une bonne réponse sur une phrase nom ou une fausse alarme sur une phrase verbe), nous avons codé leur réponse comme une « interprétation nom ». Si par contre, ils ne répondaient pas après avoir vu cette cible nom (donc que ce soit un manqué sur une phrase nom ou un rejet correct sur une phrase verbe), nous avons codé cette réponse comme une « interprétation verbe ». Le même raisonnement a bien sûr été suivi pour les cibles verbe (une réponse à cette cible a été codée comme une « interprétation verbe », tandis qu'une absence de réponse a été codée comme une « interprétation nom »).

Le pourcentage moyen de réponses verbe et nom pour chaque phrase expérimentale (phrases nom et verbe) dans chaque condition (avec et sans clitique) est représenté sur la Figure 11.



**Figure 11 (résultats de l'étude 4) : Pourcentage moyen de réponses nom (en blanc) et de réponses verbe (en gris), données aux phrases nom et aux phrases verbe dans chaque condition. La condition « avec clitique » se trouve à dans la partie gauche du graphique, la condition « sans clitique » à droite. Les barres d'erreurs représentent les erreurs standard de la moyenne.**

Le pourcentage moyen de réponses verbe et nom a été soumis à deux analyses de variance (ANOVA), l'une avec les participants et l'autre avec les items comme variable aléatoire. Dans l'analyse par sujet, nous avons déclaré quatre facteurs intra-sujets : la Condition (« avec clitique » versus « sans clitique »), la Position du mot cible (début de phrase versus milieu de phrase), le Type de Phrase entendue (phrase verbe ou phrase nom) et la Réponse Donnée (réponse verbe ou réponse nom)<sup>28</sup>. De plus, nous avons déclaré deux facteurs inter-sujets binaires, pour rendre compte du contrebalancement des groupes de sujets : le facteur Contrebalancement (Groupe 1 versus Groupe 2) et le facteur Ordre (bloc 1 ou bloc 2 en premier). L'analyse par item déclare trois facteurs intra-items (Condition, Phrase entendue et Réponse Donnée) et un facteur inter-items (Position).

Nous avons observé un effet principal de Condition uniquement par sujet ( $F_1(1,48)=4.7, p=.04$  ;  $F_2(1,14)=2.8, p=0.1$ ), ainsi qu'un effet principal de Position dans

<sup>28</sup> Les réponses nom et verbe sont complémentaires (donc inversement reliées). Nous avons mené d'autres analyses en ne prenant comme variable dépendante qu'un seul type de réponse (les réponses verbe), pour s'assurer que la puissance statistique du modèle n'était pas augmentée de manière artificielle. Les résultats obtenus ont été similaires à ceux présentés ici (modulo les réponses non intégrées aux analyses).

l'analyse par sujet ( $F_1(1,48)=8.4$ ,  $p=.006$  ;  $F_2(1,14)=2.1$ ,  $p=0.2$ )<sup>29</sup>. Nous avons aussi observé un effet principal de la Réponse Donnée ( $F_1(1,48)=120$ ,  $p<.001$  ;  $F_2(1,14)=144.7$ ,  $p<.001$ ) dans le sens où les participants ont donné plus de réponses verbe que de réponses nom (58,5 % de réponses verbe et 41,4 % de réponses nom). Ceci est vrai pour les deux groupes de sujets, mais la taille de l'effet est plus importante pour les sujets du groupe 2 que pour ceux du groupe 1 (d'où l'interaction significative entre les facteurs Contrebalancement et Réponse,  $F_1(1,48)=4.8$ ,  $p=.04$ ). De plus, cet effet ne se trouve que dans la condition « sans clitique » (interaction entre les facteurs Réponse et Condition,  $F_1(1,48)=235$ ,  $p<.001$  ;  $F_2(1,14)=182.6$ ,  $p<.001$ , et interaction significative entre les facteurs Groupe, Réponse donnée et Condition,  $F_1(1,48)=4.8$ ,  $p=.03$ )<sup>30</sup>. Par contre, le facteur de Type de Phrase entendue ne présente pas d'effet principal significatif, de même que les facteurs de Contrebalancement et d'Ordre ( $F_s < 1$ ). Ce dernier n'interagit pas non plus avec les autres facteurs expérimentaux.

Les effets expérimentaux cruciaux pour notre propos concernent l'interaction du facteur « Réponse » avec le facteur « Type de Phrase ». En effet, nous voulions savoir si les participants allaient réussir à associer le mot cible à la bonne catégorie syntaxique, en fonction des indices présents dans le signal. Cette interaction est significative ( $F_1(1,48)=1173.6$ ,  $p<.001$  ;  $F_2(1,14)=1671.1$ ,  $p<.001$ ) et indique que les sujets donnent plus de réponses nom que de réponses verbe quand ils traitent les phrases nom (71,5 % versus 28,5 %,  $F_1(1,48)=239.1$ ,  $p<.001$  ;  $F_2(1,14)=267.2$ ,  $p<.001$ ); inversement, ils donnent plus de réponses verbe que de réponses nom quand ils traitent les phrases verbe (88,5 % versus 11,4 %,  $F_1(1,48)=1828.4$ ,  $p<.001$  ;  $F_2(1,14)=4036.3$ ,  $p<.001$ ).

Cet effet interagit avec le facteur « Position » ( $F_1(1,48)=44.7$ ,  $p<.001$  ;  $F_2(1,14)=9.3$ ,  $p=.009$ ). Quelle que soit la position de la cible dans la phrase, l'interaction cruciale entre la Réponse donnée et la Phrase entendue est significative (en début de phrase : effet de 55.7%,  $F_1(1,48)=932.9$ ,  $p<.001$  ;  $F_2(1,7)=808.4$ ,  $p<.001$  ; en milieu de phrase : effet de 64.5%,  $F_1(1,48)=1125.1$ ,  $p<.001$  ;  $F_2(1,7)=865.5$ ,  $p<.001$ ). Par contre,

---

<sup>29</sup> Ces effets sont dus aux réponses données avant la cible que nous n'avons donc pas analysées. Ils sont relativement marginaux et ne concernent que 18 réponses, soit 0.2% des réponses totales. La répartition de ces réponses explique aussi l'interaction significative entre le facteur Contrebalancement et le facteur Position ( $F_1(1,48)=5.2$ ,  $p=.03$ ).

<sup>30</sup> Les facteurs Réponse et Position interagissent aussi ( $F_1(1,48)=32.7$ ,  $p<.001$  ;  $F_2(1,14)=8.7$ ,  $p=.01$ ) et l'interaction entre les facteurs Réponse, Position et Contrebalancement est significative ( $F_1(1,48)=42.1$ ,  $p<.001$  ;  $F_2(1,14)=11.2$ ,  $p=.005$ ) : il y a plus de réponses nom que de réponses verbe dans la condition « avec clitique », et l'inverse dans la condition « sans clitique » (quelle que soit la position de la cible).

cette interaction est plus forte quand la cible est située au milieu de la phrase que quand elle est en début de phrase.

D'une manière globale, les participants arrivent donc à trouver la catégorie syntaxique du mot cible et à répondre de manière correcte à la tâche demandée : quand ils entendent des phrases nom, ils répondent de manière beaucoup plus importante que le mot cible est un nom (plutôt que de se tromper en estimant que la cible entendue est un verbe), et vice-versa pour les phrases verbe.

L'interaction cruciale entre « Type de Phrase » et « Réponse » interagit également avec le facteur Condition ( $F_1(1,48)=295.7$ ,  $p<.001$  ;  $F_2(1,14)=212.6$ ,  $p<.001$ ) : ceci reflète le fait que dans la condition « avec clitique », l'interaction cruciale est plus importante (81.2%,  $F_1(1,48)=3919.8$ ,  $p<.001$  ;  $F_2(1,14)=2935.3$ ,  $p<.001$ ) que dans la condition « sans clitique » (39%,  $F_1(1,48)=203.5$ ,  $p<.001$  ;  $F_2(1,14)=241.5$ ,  $p<.001$ ).

L'observation des données montre que, dans la condition « avec clitique », l'interaction est symétrique : les phrases nom obtiennent plus de réponses nom que de réponses verbe (92.9% versus 7.2%,  $F_1(1,48)=3493.1$ ,  $p<.001$  ;  $F_2(1,14)=4170.4$ ,  $p<.001$ ), alors que les phrases verbe obtiennent plus de réponses verbe que de réponses nom (88.4% versus 11.7%,  $F_1(1,48)=1983$ ,  $p<.001$  ;  $F_2(1,14)=1120.2$ ,  $p<.001$ ). Par contre, dans la condition « sans clitique », seules les phrases verbe montrent un effet de « Réponse donnée » (88,7 % versus 11,1 %,  $F_1(1,48)=1098.7$ ,  $p<.001$  ;  $F_2(1,14)=1800.7$ ,  $p<.001$ ). Les sujets ont un comportement différent seulement sur les phrases nom de la condition « sans clitique » (les phrases verbe entraînant des réponses similaires, que le mot cible soit précédé ou non d'un clitique). Dans cette situation, les sujets répondent au hasard (autant d'interprétations verbe que d'interprétations nom).

Comme nous l'avons noté précédemment, les phrases verbe et nom de la condition « sans clitique » diffèrent grandement entre elles, au niveau de la position des frontières de groupes phonologiques, et donc au niveau de la quantité d'informations disponibles au moment de traiter le mot cible. Rappelons que dans les phrases verbe, on trouve une frontière de groupe phonologique juste avant le mot cible (frontière marquant la séparation entre le groupe nominal et le groupe verbal, comme dans « [sa cramona] [**bamoule** muche]... »), alors que cette frontière de groupe phonologique n'arrive qu'après le mot cible dans la phrase nom (comme dans « [sa cramona **bamoule**] [camiche dabou] »). Ainsi quand les sujets entendent un mot grammatical ou

une frontière prosodique avant de traiter le mot cible, ils arrivent à inférer la catégorie syntaxique du mot cible. Si par contre, aucune information n'est directement présente avant le mot cible, les sujets sont au hasard et ne peuvent trouver la bonne catégorie syntaxique.

Le fait que les sujets réussissent aussi bien à utiliser les informations sur les mots grammaticaux que les informations prosodiques pour réaliser leur tâche peut sembler étonnant. Nous avons conduit de nouvelles analyses, sur les temps de réaction, pour affiner nos conclusions. En effet, il se pourrait que la performance des participants soit identique, mais qu'il leur faille plus de temps pour prendre leur décision (on peut penser que déduire qu'un pronom est suivi par un verbe serait plus facile et plus automatique, et donc que les phrases verbe obtiendraient des temps de réaction plus rapides dans la condition avec clitique que dans la condition sans clitique).

#### 6.2.2 Analyses des temps de réaction

Le temps de réaction moyen pour chaque phrase expérimentale dans chaque condition est représenté dans la Figure 12.

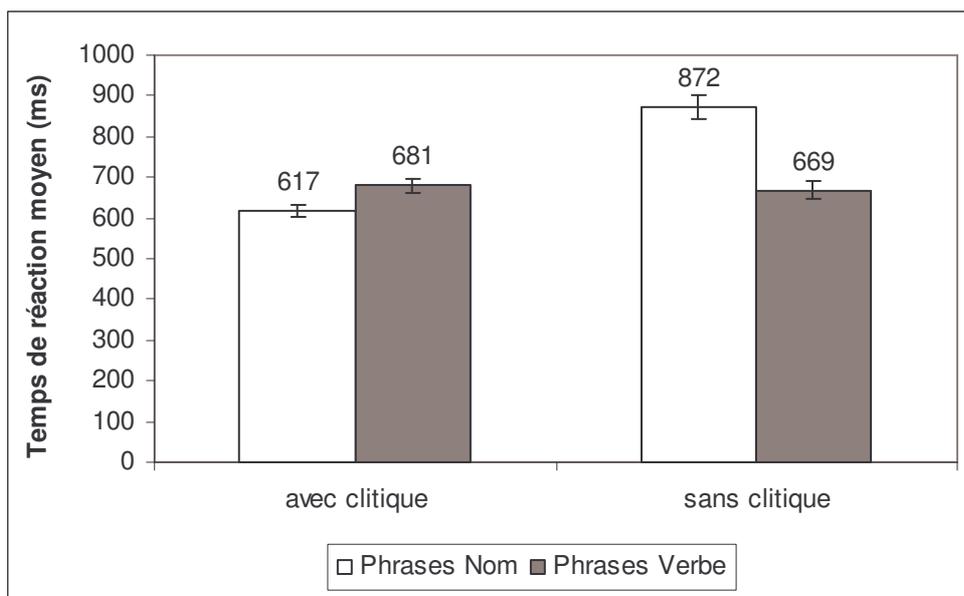


Figure 12 (résultats de l'étude 4) : Temps de réaction moyen pour répondre aux phrases nom (en blanc) et aux phrases verbe (en gris) dans les deux conditions expérimentales. La condition « avec clitique » se trouve à gauche de la figure, la condition « sans clitique » à droite. Les barres d'erreurs représentent les erreurs standard de la moyenne.

Les temps de réaction moyen ont été soumis à deux analyses de variance, l'une avec les participants, l'autre avec les items comme variable aléatoire. L'analyse par sujet déclare trois facteurs intra-sujets : la Condition (« avec clitique » versus « sans clitique »), la Position du mot cible (début de phrase versus milieu de phrase), et le Type de Phrase entendue (phrases verbe versus phrases nom) ; elle déclare aussi deux facteurs inter-sujets : le Contrebalancement (groupe 1 versus groupe 2) et le facteur Ordre (bloc 1 versus bloc 2 en premier). L'analyse par item déclare deux facteurs intra-items (Condition et Type de Phrase) et un facteur inter-items (Position)

Le résultat crucial pour notre propos est l'interaction significative entre le facteur Phrase et le facteur Condition ( $F_1(1,48)=113.8$ ,  $p<.001$ ;  $F_2(1,14)=207.3$ ,  $p<.001$ ). En effet, cette interaction vient du fait que les temps de réaction pour les phrases nom sont beaucoup plus lents dans la condition sans clitique que dans la condition avec clitique (872 versus 617 ms,  $F_1(1,48)=116.5$ ,  $p<10^{-13}$ ;  $F_2(1,14)=304$ ,  $p<10^{-10}$ ), alors que les phrases verbe obtiennent des temps de réaction similaires dans ces deux conditions (669 versus 681 ms,  $F_s<1$ ). Les conclusions développées ci-dessus semblent se confirmer et s'affiner : dans la condition sans clitique, les sujets qui entendent les phrases nom n'ont aucune information réellement disponible au moment de traiter le mot ambigu (ni de mot grammatical, ni de frontière prosodique). Ce manque d'informations désambiguïsantes, et donc leur indécision quant à la catégorie syntaxique du mot cible, se retrouvent dans ces temps de réaction beaucoup plus lents. Par contre, quand les participants entendent les phrases verbe, ils ne rencontrent pas de difficultés. Ils peuvent utiliser la présence d'une frontière prosodique placée juste avant le mot cible, aussi bien que la présence d'un pronom, pour prendre leur décision.

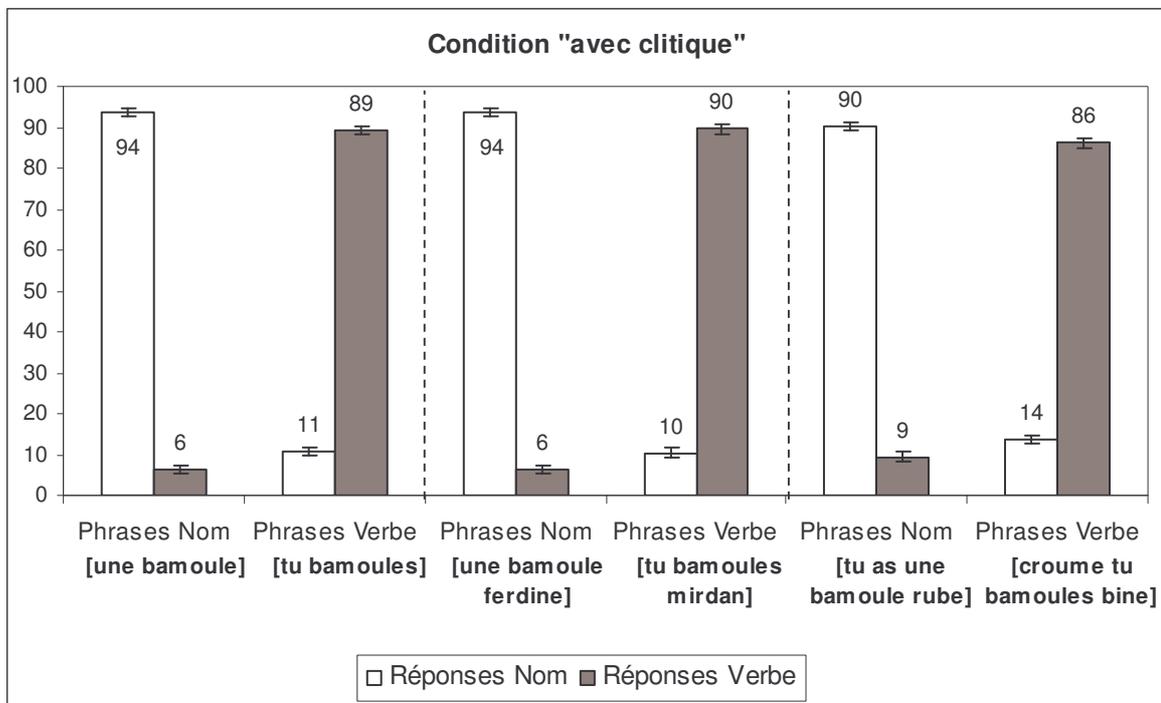
### 6.2.3 *Analyses détaillées par sous-conditions*

Nous venons de voir que d'une manière générale les sujets se comportent de la même manière pour traiter les phrases verbe, que le mot cible soit directement précédé d'un pronom (condition avec clitique) ou d'une frontière de groupe phonologique (condition prosodique). Les informations apportées par les mots grammaticaux et les

frontières prosodiques sont donc autant utilisables par les auditeurs et aussi efficaces pour déduire la catégorie syntaxique du mot suivant. Nous allons maintenant analyser les résultats en comparant certaines sous-conditions pour affiner ces conclusions.

### 6.2.3.1 Comparaison des sous-conditions « avec clitique »

Le pourcentage moyen de réponses verbe et nom pour chaque phrase expérimentale dans les sous-conditions « avec clitique » est représenté dans la Figure 13 (voir les annexes D3 et D4 pour les résultats individuels par sujets et par items).



**Figure 13 (résultats de l'étude 4) : Pourcentage moyen de réponses nom (en blanc) et de réponses verbe (en gris) dans les trois sous-conditions de la condition « avec clitique ».** Les barres d'erreurs représentent les erreurs standard de la moyenne.

L'interaction cruciale entre la Phrase entendue et la Réponse Donnée est significative pour chaque sous-condition (première sous-condition :  $F_1(1,48)=3085$ ,  $p<.001$  ;  $F_2(1,14)=969$ ,  $p<.001$  ; deuxième sous-condition :  $F_1(1,48)=2949$ ,  $p<.001$  ;  $F_2(1,14)=1988$ ,  $p<.001$  ; troisième sous-condition :  $F_1(1,48)=2149$ ,  $p<.001$  ;  $F_2(1,14)=941$ ,  $p<.001$ ).

On observe une tendance à ce que la taille de l'effet soit un peu moins forte quand l'information désambiguïsante est le clitique seul, comparé aux deux premières sous-conditions (donc quand le mot cible est précédé à la fois par un mot grammatical et par une frontière de groupe phonologique). Mais cette différence ne ressort que dans l'analyse par sujet et reste assez faible (environ 2-3 % d'écart pour certaines réponses). Elle pourrait donc être due à certains items de cette sous-condition « clitique seul » qui ont obtenu des taux de réponses un peu déviants. En particulier, les phrases verbe dans lesquelles le mot cible apparaissait en début de phrase ont été traitées plus difficilement. Ceci ressort surtout dans l'analyse des temps de réaction que nous avons menée précédemment : on obtient en effet des temps de réponse aux phrases nom similaires entre ces trois sous-conditions (624, 603 et 623 ms) ; quant aux phrases verbe, on trouve que les participants mettent plus de temps pour répondre aux phrases de la condition « clitique seul » (723 ms) qu'aux phrases verbe des deux autres conditions (645 et 675 ms) ; ceci est surtout vrai quand la cible est au début de la phrase. La majorité des phrases de cette catégorie était construite sur le schéma « adverbe + pronom + verbe cible » regroupés en une seule unité prosodique, ce qui semble plutôt rare dans le langage courant (on aurait par exemple une grande chance d'obtenir une frontière prosodique entre l'adverbe et le pronom plutôt qu'une seule unité prosodique). Par contre, que le mot cible soit suivi ou non par une frontière prosodique ne facilite pas la tâche des sujets.

#### 6.2.3.2 Comparaison des sous-conditions « sans clitique »

Le pourcentage moyen de réponses verbe et nom pour chaque phrase expérimentale dans les sous-conditions « sans clitique » est représenté dans la Figure 14.

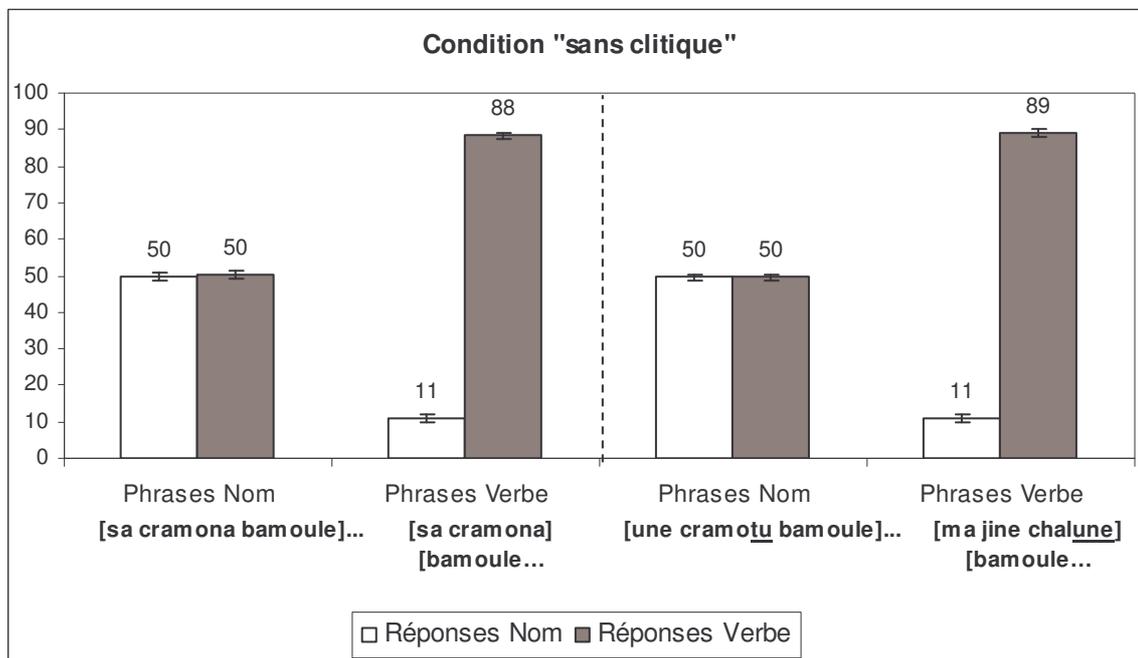


Figure 14 (résultats de l'étude 4) : Pourcentage moyen de réponses nom (en blanc) et de réponses verbe (en gris) dans les deux sous-conditions de la condition « sans clitique ». Les barres d'erreurs représentent les erreurs standard de la moyenne.

L'interaction cruciale entre la Réponse donnée et la Phrase entendue est significative dans ces deux sous-conditions (première sous-condition :  $F_1(1,48)=179$ ,  $p<.001$  ;  $F_2(1,14)=134$ ,  $p<.001$  ; deuxième sous-condition :  $F_1(1,48)=171$ ,  $p<.001$  ;  $F_2(1,14)=247$ ,  $p<.001$ ).

Le comportement des participants est identique entre ces deux situations : les sujets ne sont pas gênés par la présence d'une syllabe ambiguë avant le mot cible (syllabe, ou rime, pouvant entraîner la catégorisation du mot cible dans la mauvaise catégorie syntaxique). Tous réussissent aussi bien à répondre que le mot cible est un verbe quand ils entendent réellement une phrase verbe. Par contre, tous sont perdus quand ils doivent réagir aux phrases nom et ne savent pas quelle réponse donner.

## 6.3 DISCUSSION

Le but de cette expérience était de vérifier si les indices apportés par les mots grammaticaux et les frontières de groupes phonologiques pouvaient permettre de construire une représentation syntaxique des phrases et d'attribuer à des mots inconnus une catégorie syntaxique. Le résultat principal de cette étude confirme cette idée de départ. En effet, nous avons trouvé que, dans des phrases n'ayant aucun sens, et dans lesquelles seules les informations liées aux mots grammaticaux et aux indices prosodiques sont préservées, les adultes français peuvent mettre en place une analyse syntaxique des phrases et trouver la catégorie syntaxique de certains mots cibles. Dans la tâche de détection de mot que nous leur avons proposée, les sujets devaient identifier des noms et des verbes cibles dans des phrases contenant soit la forme nominale soit la forme verbale. Les participants ont réussi à donner plus de réponses nom que de réponses verbe quand ils entendaient les phrases contenant réellement le nom ; à l'inverse, quand ils entendaient les phrases verbes, les sujets ont alors donné plus de réponses verbe que de réponses nom.

### *6.3.1 Comparaison du rôle des indices utilisés*

Dans cette expérience, nous avons fait varier la nature, et parfois même la présence, des indices pertinents pour notre étude. Dans la condition avec clitique, tous les mots cibles étaient immédiatement précédés d'un mot grammatical permettant d'inférer sa catégorie syntaxique (tous les noms étaient précédés d'articles alors que tous les verbes étaient précédés de pronoms). Cette information a été très utile pour les participants qui ont parfaitement réussi à identifier la catégorie syntaxique du mot cible.

Que le couple « mot grammatical – mot cible » soit au début d'une unité prosodique, ou forme une unité prosodique à lui tout seul, n'a pas changé le comportement des participants. Nous pensions que la présence d'une frontière prosodique à la fin du mot cible faciliterait la segmentation de ce mot et donc son traitement. Nous aurions alors dû observer des différences de résultats entre les deux premières sous-conditions. Mais que ce soit au niveau de leur performance dans la tâche de catégorisation, ou au niveau des temps de réaction mis pour prendre leur décision, les

participants ont obtenu exactement les mêmes résultats dans ces deux sous-conditions. De fait, leur performance est excellente, et l'absence de différences entre ces deux sous-conditions peut refléter un effet de « plafond ».

Une légère différence est par contre apparue entre ces deux premières sous-conditions et celle où le couple « mot grammatical – mot cible » se trouvait au milieu d'une unité prosodique. Dans cette dernière situation, le mot cible était toujours précédé d'un mot grammatical mais celui-ci n'apparaissait pas juste après une frontière prosodique. A nouveau, nous pensions que l'absence de cette frontière diminuerait la facilité d'extraction du mot grammatical en début d'unité, ce qui gênerait ou ralentirait la réponse des sujets. En fait, il est difficile d'interpréter la différence obtenue en ces termes. Il est vrai que la taille de l'effet (c'est-à-dire l'interaction entre la réponse donnée par les sujets et la phrase qu'ils entendaient) est moins forte dans cette situation. Mais les tests statistiques ne donnent pas de résultats très probants et ne sont significatifs que dans certaines analyses (les analyses par sujet). Il se pourrait que cette différence soit plutôt due aux réponses obtenues par certains items de cette catégorie, en particulier les phrases verbe dont la cible se situait en début de phrase. Il semblerait que ce cas précis ait particulièrement gêné les sujets : leur performance est toujours bonne mais moins forte (81% de réponses verbe et 19% de réponses nom aux phrases verbe quand la cible se situe en début de phrase), et leurs temps de réaction sont plus lents que pour toutes les autres conditions (785 ms). Même si l'on ne peut pas exclure le fait que les sujets ont plus de mal à réaliser la tâche demandée quand ils ont uniquement accès à l'information apportée par les mots grammaticaux, on ne peut pas non plus exclure le fait que certains items ont posé problème pour d'autres raisons (que nous pourrions mieux contrôler dans nos prochaines expériences).

Toutefois, ce qui ressort de ces premiers résultats est que les adultes français réussissent parfaitement et rapidement à inférer la catégorie d'un mot inconnu si celui-ci est précédé d'un mot grammatical. On peut donc supposer que les jeunes enfants pourraient eux aussi utiliser cette information (au préalable, il faudrait bien entendu que les enfants connaissent certaines associations comme « un article est généralement suivi d'un nom » et « un pronom est généralement suivi par un verbe »).

Le deuxième indice testé était le rôle des frontières prosodiques dans la construction d'une représentation syntaxique des phrases. Nous avons pour cela créé

des paires de phrases, identiques jusqu'au mot cible d'un point de vue phonémique, mais différentes du point de vue syntaxique et prosodique. Ainsi, dans les phrases verbe, comme « [sa cramona] [bamoule...] », on trouve une frontière de groupe phonologique juste avant le mot cible ; par contre, dans les phrases nom, comme « [sa cramona bamoule]... », la frontière prosodique se situe juste après le mot cible. Il y a donc une asymétrie entre ces deux phrases au niveau des informations disponibles au moment de traiter le mot cible, asymétrie qui s'est retrouvée au niveau des résultats des sujets. En effet, les sujets ont parfaitement réussi à interpréter le mot cible comme un verbe quand ils traitaient les phrases verbes (plus de réponses verbe que de réponses nom sur ces phrases). Ils ont donc pu utiliser l'information apportée par les frontières prosodiques, en association avec l'information apportée par les mots grammaticaux. En effet, la frontière prosodique seule, avant le mot cible, ne peut pas permettre de trouver la catégorie syntaxique du mot cible. Par contre, les sujets ont pu utiliser les frontières prosodiques pour délimiter des constituants syntaxiques, puis identifier la nature du premier constituant grâce au mot grammatical placé en début d'unité (identifier que [sa cramona] est un groupe nominal grâce au pronom possessif « sa »). Ensuite, après avoir perçu la frontière prosodique, ils ont pu faire l'hypothèse qu'un nouveau constituant syntaxique allait commencer et que le plus logique était de trouver un groupe verbal après le groupe nominal (d'où une interprétation massive du mot « bamoule » comme un verbe). Ce raisonnement n'est bien sûr qu'une hypothèse de recherche, mais qui semble bien étayée par les résultats obtenus. Il serait maintenant intéressant de mettre en place d'autres stimuli pour tester ceci (comme faire suivre le groupe nominal d'un constituant différent d'un groupe verbal, dans lequel le mot « bamoule » pourrait être un adjectif, et de voir si les adultes se trompent beaucoup en continuant d'interpréter le mot cible comme un verbe : par exemple, « [son milouton] [**bamoule** et joume] [ne la go pas en fiboune] » qui pourrait être l'équivalent de « [son pantalon] [**violet** et jaune] [ne la met pas en valeur] »).

L'asymétrie dont nous parlions plus haut se reflète dans le comportement totalement différent des sujets, lorsqu'ils devaient traiter les phrases nom. Les sujets n'ont pas réussi à les traiter correctement puisqu'ils ont en effet donné autant de réponses nom que de réponses verbe dans cette condition. Cette incertitude dans l'interprétation à donner s'est aussi reflétée dans leur temps de réponse beaucoup plus lents. Il y a deux interprétations possibles à ce résultat :

(1) l'absence d'indices clairs avant le mot ambigu a gêné les participants qui ne savaient pas quelle interprétation choisir, les deux étant possibles (comme toutes les frontières syntaxiques ne sont pas obligatoirement marquées d'un point de vue prosodique). Cette idée s'est retrouvée dans le discours des participants à la fin de l'expérience : certains ont en effet parlé de la difficulté de ces phrases (selon leurs propres termes, « quand il y a un mot entre l'article et le mot sur lequel on doit répondre »). Ils disaient alors qu'il leur était difficile de décider de la catégorie syntaxique du mot cible car les deux étaient possibles : soit ils pouvaient considérer que c'était un nom et que le mot d'avant était donc un adjectif, soit ils pouvaient considérer que le mot cible était un verbe et que le mot précédent était donc un nom ;

(2) la deuxième interprétation repose plus sur la structure syntaxique utilisée : ces phrases nom ambiguës étaient formées d'un groupe nominal de structure « article – adjectif – nom ». L'antéposition de l'adjectif n'est pas la structure la plus courante (elle existe bien évidemment mais apparaît en général avec des adjectifs très fréquents et dont le nombre est limité). Il se pourrait donc que les participants aient interprété le mot suivant l'article comme étant un nom, et le mot suivant (la cible) comme un adjectif ; dans la mesure où les seuls choix qui leur étaient offerts étaient une réponse nom ou verbe, ils ont donc répondu au hasard.

Si les participants ont effectivement pensé que ce mot cible pouvait être un adjectif, cela expliquerait leurs réponses au hasard. Ceci est à nouveau une hypothèse de travail mais qui semble aussi se confirmer par le discours d'une grande partie des participants à la fin de l'expérience. En effet, ils étaient nombreux à dire à l'expérimentateur qu'ils avaient été très dérangés par les phrases où le mot cible était un adjectif car ils étaient surpris de découvrir cette possibilité et ne savaient plus quoi faire. Un moyen de vérifier cette idée serait de réaliser une expérience similaire mais où cette fois-ci les sujets devraient détecter des verbes et des adjectifs (et non plus des verbes et des noms).

Les résultats que nous avons obtenus dans l'expérience 2 de la première étude (voir chapitre 2) apportent un début de réponse : en effet, dans cette expérience, les sujets devaient détecter des verbes et des adjectifs dans des phrases ambiguës de structure prosodique similaire à celles utilisées dans cette expérience. Par exemple, ils devaient détecter l'adjectif « mort » dans « [le petit chien mort] [sera enterré demain] » : dans cette situation, et quand les frontières prosodiques étaient bien marquées, les sujets ont donné plus de réponses adjectif que de réponse verbe (ils ne répondaient pas au

hasard, contrairement à ce que l'on a trouvé dans cette dernière étude, et bien que la frontière de groupe phonologique soit aussi située après le mot cible). Bien sûr, le fait de devoir traiter des vrais mots a dû faciliter leur tâche (en ce qui concerne la reconnaissance des mots précédant le mot cible et donc la construction de la structure syntaxique de ces phrases ambiguës). Mais il est possible d'imaginer que les adultes réussiraient de la même manière à interpréter les phrases jabberwocky s'il leur était de demander de détecter des adjectifs et des verbes.

Un dernier point à noter, concernant la comparaison des deux expériences de détection de mot abstrait réalisées dans cette recherche, est le fait que les performances des sujets semblent meilleures sur les phrases jabberwocky (si l'on fait bien sûr abstraction des réponses au hasard données aux phrases nom de la condition « sans clitique ») que sur les phrases françaises. Quand les sujets traitent les phrases jabberwocky, ils donnent une interprétation correcte aux cibles ambiguës dans environ 90% des cas (et ne se font piéger que dans 10% des cas). Par contre, sur les phrases françaises, le taux de fausses alarmes a été plus important (environ 40%). Il semble donc que les sujets aient plus facilement réussi à analyser les phrases que pourtant ils ne comprenaient pas. Dans cette situation, pour réaliser l'analyse syntaxique des phrases, ils ont utilisé les seules informations qui étaient disponibles, donc les mots grammaticaux et les frontières prosodiques, informations qui semblent être donc particulièrement pertinentes pour guider la construction de la structure syntaxique des phrases. Si les résultats ont été moins bons sur les phrases françaises, cela pourrait être dû à une sorte d'interférence qui aurait été causée par les différentes connaissances des sujets, comme leurs connaissances lexicales et sémantiques, et par les attentes qu'ils pouvaient avoir sur les phrases qu'ils étaient en train d'entendre.

Il ressort de toutes ces analyses que les adultes peuvent utiliser différents indices pour réaliser l'analyse syntaxique des phrases : pour inférer la catégorie syntaxique d'un mot inconnu, ils peuvent utiliser des indices lexicaux (comme la présence directe d'un mot grammatical avant le mot cible), et des indices prosodiques reflétant la structure syntaxique des phrases. Ces deux sortes d'informations semblent aussi pertinentes et aussi utiles pour les sujets. Si l'on regarde les réponses aux phrases verbes dans toutes les sous-conditions, on voit en effet que les participants ont obtenu de très bons résultats, similaires entre toutes les sous-conditions. Cette similitude ne concerne pas seulement leur réussite en terme de catégorisation syntaxique, mais aussi le temps

moyen qu'ils ont mis pour prendre leur décision. Par exemple, on aurait pu penser qu'il serait plus facile et donc plus rapide de décider que « bamoule » est un verbe dans « [tu bamoules] » (condition avec clitique) que dans « [sa cramona] [bamoule...] » (condition sans clitique). Or il n'en est rien puisque les temps de réaction sur les phrases verbe ne sont pas significativement différents entre les sous-conditions. Interpréter les mots grâce à la structure prosodique et syntaxique des phrases que l'on entend semble donc être aussi facile que d'utiliser la simple co-occurrence entre un mot grammatical et un mot de contenu.

### 6.3.2 *Qu'est-ce qui peut être considéré comme un mot grammatical ?*

Le dernier point essentiel à aborder est celui des mots grammaticaux : en effet, qu'est-ce qui est considéré comme un mot grammatical dans la parole orale ? On pourrait faire l'hypothèse que chaque syllabe homophone à un mot grammatical connu est potentiellement activée comme un mot grammatical. Nous avons voulu tester cette idée en créant des phrases ambiguës à ce niveau, c'est-à-dire en faisant précéder un mot cible (d'une catégorie x) par un mot dont la rime ou la dernière syllabe était ambiguë (car homophone à un mot grammatical entraînant une catégorie syntaxique y à sa suite). C'était le cas de la phrase « [une cramotu bamoule] [camiche dabou] », où le nom « bamoule » est précédé de la syllabe /tu/.

Si effectivement toutes les syllabes homophones au pronom « tu » sont considérées, même brièvement, comme le pronom « tu », alors on aurait pu avoir une forte proportion de réponse verbe pour cette phrase-là. En effet, les participants auraient pu reconnaître le pronom potentiel et n'auraient eu que cette information à disposition pour prendre leur décision (puisque la frontière prosodique se situait après le mot cible). On se serait donc attendu à ce que les résultats aillent plutôt dans le sens d'une plus grande proportion de réponses verbe que de réponses nom pour ces phrases là. Or nous avons vu plus haut que les résultats étaient identiques dans les deux sous-conditions « sans clitique », à savoir autant de réponses verbe que nom pour les phrases nom.

De même, on aurait pu s'attendre à un même pattern de réponses entre ces phrases et celle de la condition « clitique seul ». En effet, si la syllabe /tu/ était considérée comme un pronom potentiel, alors les participants auraient dû répondre

autant de fois que « bamoule » est un verbe dans « [une cramotu bamoule]... » que dans « [croume tu bamoules bine] ». Or le pourcentage de réponses verbe pour la première phrase est significativement inférieur (50 % versus 86%). Il semble donc que toutes les syllabes homophones à des mots grammaticaux ne soient pas initialement considérées comme des mots grammaticaux potentiels.

Une des explications pourrait résider dans la forme acoustique et prosodique de cette syllabe qui serait différente selon que cette syllabe est un mot grammatical ou qu'elle est la dernière syllabe d'un mot. Nous avons réalisé quelques analyses acoustiques en comparant la durée, la fréquence fondamentale et l'énergie des pronoms dans ces deux conditions. Nous n'avons pas comparé les articles entre eux (comme comparer la syllabe /une/ dans « [tu as une bamoule rube] » versus « [ma jine chalune] [bamoule...] ») puisque d'autres raisons auraient pu expliquer des différences entre ces deux exemplaires de /une/ : par exemple, dans la deuxième phrase, la syllabe homophone ambiguë est toujours située à la fin d'un groupe phonologique. Elle porte donc des indices acoustiques de fin d'unités prosodiques comme un allongement et une baisse de la fréquence fondamentale. Nous avons donc centré notre attention sur les pronoms qui, dans ces deux conditions, se situaient à des endroits similaires dans la structure prosodique des phrases (au milieu d'une unité prosodique). Rappelons juste les deux situations que nous avons étudiées :

- (a) pronom : « [croume tu bamoules bine] [ce nachi calune] »
- (b) homophone : « [une cramotu tu bamoule] [camiche dabou] »

Les analyses acoustiques ont montré qu'il existait des différences de durée entre ces deux syllabes : elle est plus longue en fin de mot que quand c'est un pronom (153 ms versus 131 ms,  $t(15)=2.1$ ,  $p=.05$ ). De même, la courbe de fréquence fondamentale est différente selon que la syllabe cible est un mot grammatical ou la dernière syllabe d'un mot. Dans les deux cas, la courbe est similaire en arrivant sur la syllabe cible : on constate en effet une baisse significative de la fréquence fondamentale (baisse significative de 81 Hz en fin de mot versus baisse significative de 65 Hz sur le pronom, la différence entre les deux n'étant pas significative :  $t(15)=1.3$ ,  $p=0.2$ ). Par contre, les valeurs se différencient si l'on regarde la courbe de fréquence fondamentale entre la syllabe cible et la syllabe suivante : entre la fin du mot et la première syllabe du mot

suivant, on ne constate pas de différence significative dans les valeurs de fréquence alors qu'on trouve une hausse significative de la fréquence fondamentale entre le pronom et la première syllabe du mot de contenu suivant (hausse de 42 Hz,  $t(15)=4.8$ ,  $p=.0002$ ). Quand la syllabe cible est un pronom, on trouve une courbe « descendante puis ascendante » au niveau de cette syllabe. Ceci pourrait s'apparenter à l'indice de montée initiale, proposé entre autres par Welby (2003a; 2004). Par contre, nous n'avons pas trouvé de différences au niveau de l'énergie de ces syllabes cibles.

Il semble donc que des indices acoustiques fins différencient deux syllabes homophones, quand elles sont des pronoms ou les dernières syllabes d'un mot. Il se pourrait donc que les auditeurs soient sensibles à ces distinctions fines et les utilisent pour traiter le discours (en distinguant par exemple entre les syllabes d'un mot de contenu et les mots grammaticaux). Bien sûr d'autres indices pourraient aussi expliquer ces différences et le fait que les auditeurs n'ont pas été piégés par cette syllabe homophone. Dans les phrases nom de la condition « sans clitique », la syllabe homophone se trouvait toujours dans un groupe nominal, à quelques syllabes de distance d'un article (dans « [une cramotu bamoule] [camiche dabou] », la syllabe homophone au pronom « tu » se trouve à deux syllabes de distance de l'article « une »). Il se pourrait qu'une telle proximité entre deux mots grammaticaux de catégorie différente, à l'intérieur d'une même unité prosodique, soit très rare, et les participants auraient pu utiliser ces informations de prédictabilité ou de plausibilité pour guider leur interprétation. Certes, on peut trouver des phrases comme « [la fille elle pleure] », où un article (« la ») et un pronom (« elle ») sont relativement proches l'un de l'autre, et peuvent être prononcés dans la même unité prosodique, mais il me semble que ces structures sont loin d'être les plus courantes (et certaines semblent impossibles comme avec « [une xx tu xx] »).

De nouvelles expériences devront maintenant être réalisées pour affiner plus précisément les caractéristiques des mots grammaticaux (caractéristiques locales comme leurs propriétés acoustiques, et caractéristiques plus générales comme leur position dans les phrases et dans les structures prosodiques).

Pour conclure, le résultat principal de cette dernière étude est que des adultes français ont réussi à construire une représentation syntaxique de phrases qu'ils ne

comprenaient pas, grâce aux informations apportées conjointement par les mots grammaticaux et les frontières prosodiques. Un non-mot précédé directement d'un article a été interprété comme un nom, tandis qu'un non-mot directement précédé par un pronom a été interprété comme un verbe. Mais la présence conjointe d'un mot grammatical et d'un mot de contenu n'est pas la seule information que les sujets ont utilisée pour inférer la catégorie grammaticale du mot de contenu. Ils ont aussi utilisé la présence des frontières de groupes phonologiques, qu'ils ont interprétées comme des frontières de constituants syntaxiques. Ils ont ainsi pu construire une première ébauche de la structure syntaxique des phrases et attribuer une catégorie syntaxique à certains mots de ces phrases (par exemple, construction de la structure syntaxique « [sa cramona]<sub>GN</sub> [bamoule muche]<sub>GV</sub> », puis interprétation de « cramona » comme un nom et de « bamoule » comme un verbe).

Ainsi il semble relativement facile de faire de la syntaxe même sans avoir un grand lexique à sa disposition. Les mots grammaticaux et les frontières de groupes phonologiques apportent une information utile pour effectuer une première ébauche d'analyse syntaxique des phrases, et pourraient donc être utilisés par les jeunes enfants pour commencer à comprendre et analyser la structure syntaxique de leur langue maternelle (et ce avant même que les enfants ne disposent d'un grand répertoire de mots connus). Certes, la situation n'est pas complètement identique entre nos adultes et les enfants, puisque l'attention des adultes est focalisée pendant toute l'expérience sur la recherche des noms et des verbes qui se trouvent dans les phrases. Leur tâche est beaucoup plus explicite que celle des enfants qui apprennent leur langue maternelle. Néanmoins, les résultats apportés par cette étude offrent des données très encourageantes pour les futures recherches sur l'acquisition de la syntaxe par les jeunes enfants.

## CONCLUSION GENERALE

Dans ce travail de thèse, nous avons développé et testé une hypothèse de recherche permettant de rendre compte à la fois du traitement syntaxique réalisé par les adultes, et de l'acquisition de la syntaxe par les enfants. Cette hypothèse repose sur les modèles d'initialisation phonologique et d'initialisation prosodique proposés par Morgan (1986; 1996b). Il postule qu'une analyse de bas niveau purement phonologique du signal de parole pourrait apporter de nombreuses informations sur, par exemple, les catégories syntaxiques et la structure syntaxique de la langue concernée.

### *1. Rappel du problème d'initialisation et de notre hypothèse de travail*

Comme nous l'avons vu dans le chapitre 1, les adultes ont à leur disposition de nombreuses connaissances qu'ils peuvent utiliser pour traiter le discours qu'ils entendent et guider leur analyse syntaxique des phrases (connaissances phonologiques, lexicales, syntaxiques, pragmatiques...). Nous avons présenté de nombreuses données qui indiquent, par exemple, que les adultes peuvent réaliser une analyse syntaxique des phrases qu'ils entendent sur la base de stratégies syntaxiques (comme la clôture tardive et l'attachement minimal, Frazier, 1978), mais aussi sur la base d'informations lexicales (comme la structure argumentative ou thématique des verbes, Shapiro et al., 1993; Tanenhaus et al., 1989; Trueswell et al., 1994; Trueswell et al., 1993), d'informations contextuelles (Altmann et al., 1994; Altmann & Steedman, 1988; Crain & Steedman, 1985; Tanenhaus et al., 1995) et d'informations statistiques (Jurafsky, 1996; Mitchell & Cuetos, 1991).

Par contre, les enfants qui acquièrent leur langue maternelle n'ont pas accès à toutes ces sources d'informations, puisqu'ils sont en train de les découvrir et de les construire. Pourtant, très rapidement, ils deviennent des locuteurs compétents : pour cela, ils ont dû entre autres découvrir quels étaient les sons pertinents dans leur langue maternelle (acquérir sa phonologie), trouver les mots dans les phrases qu'ils entendaient et découvrir leur signification (dépasser le problème de la segmentation lexicale et mettre en place un lexique mental), et finalement découvrir et comprendre les règles qui

agencent les mots dans les phrases (donc acquérir la syntaxe de leur langue maternelle pour comprendre le sens des phrases à partir du sens des mots).

L'apprentissage de chacune de ces composantes peut être facilité en supposant des connaissances préalables dans les autres. Par exemple, dans la mesure où la syntaxe spécifie les relations entre les mots d'une phrase, on a longtemps supposé que les enfants devaient avoir accès aux mots et à leur signification pour pouvoir apprendre la syntaxe. D'autre part, Gleitman (1990) a démontré, de manière fort convaincante, que l'apprentissage du sens des mots était un problème difficile qui serait considérablement simplifié si les enfants avaient accès à certains aspects de la structure syntaxique. Ceci met en évidence un paradoxe, ou problème d'initialisation : le lexique serait nécessaire à l'acquisition de la syntaxe, et la syntaxe serait nécessaire à l'acquisition du lexique.

Une manière de résoudre ce paradoxe est d'évaluer quelles informations peuvent être obtenues par une analyse « bottom-up » du signal de parole, et de recourir alors aux informations phonologiques et prosodiques directement disponibles et accessibles dans ce signal.

Nous avons d'abord centré notre attention sur les informations prosodiques, c'est-à-dire sur les variations de rythme et d'intonation de la parole. En effet, les mots et les syllabes ne sont pas prononcés d'une manière monotone et monocorde, mais sont regroupés en unités mélodiques, ou constituants prosodiques, qui peuvent être intégrés dans un modèle hiérarchique de la structure prosodique (voir par exemple Shattuck-Hufnagel & Turk, 1996). Nous avons étudié le rôle des unités de taille intermédiaire que sont les groupes phonologiques (alors que la majorité des études qui ont testé le rôle des frontières prosodiques dans l'analyse syntaxique des phrases se sont intéressées aux frontières majeures de groupes intonationnels). Ces groupes phonologiques nous paraissent particulièrement pertinents pour notre propos, puisque leurs frontières sont étroitement liées aux frontières des constituants syntaxiques (Nespor & Vogel, 1986). De plus, on ne trouve qu'un ou deux mots de contenu dans ces unités prosodiques : elles permettraient donc de découper des phrases relativement courtes (c'est-à-dire la majorité des phrases que l'enfant reçoit) en constituants syntaxiques.

Le deuxième indice que nous avons étudié concerne les mots grammaticaux (articles, pronoms, auxiliaires...), qui jouent un rôle central dans l'analyse syntaxique des phrases. Comme l'ont montré différentes expériences présentées dans le chapitre 4, les mots grammaticaux pourraient permettre d'inférer la catégorie syntaxique des mots

de contenu suivants (un mot précédé d'un article est un nom, tandis qu'un mot précédé d'un pronom est un verbe). Il a été proposé que ces items lexicaux pourraient être appris précocement du fait de leurs propriétés particulières (Christophe et al., 1997) : ce sont des mots extrêmement fréquents, situés en général en bordure d'unités prosodiques (début ou fin selon les langues), qui possèdent généralement des caractéristiques phonologiques particulières les distinguant des mots de contenu (Shi et al., 1998; Shi et al., 1999). Ces particularités peuvent être exploitées à deux titres : d'une part, elles peuvent permettre aux enfants de découvrir relativement facilement les mots grammaticaux de leur langue maternelle (à savoir les syllabes les plus fréquentes situées en début ou en fin de constituant prosodique) ; puis, elles peuvent permettre à l'enfant, grâce à la reconnaissance et à l'extraction des mots grammaticaux du signal de parole, d'inférer que le mot suivant est un mot de contenu dont la catégorie syntaxique peut être déduite.

Pris ensemble, ces deux indices (structure prosodique et mots grammaticaux) pourraient aider à l'initialisation de l'acquisition syntaxique. En effet, les enfants pourraient, dans un premier temps, découper le signal de parole en unités syntaxiques grâce à la perception des frontières de groupes phonologiques. L'identification de ces unités syntaxiques (c'est-à-dire pouvoir leur assigner le label « groupe nominal » ou « groupe verbal ») serait rendue possible grâce à la reconnaissance des mots grammaticaux en début d'unités prosodiques (une unité syntaxique commençant par un article serait un groupe nominal...).

## *2. Les adultes français peuvent-ils utiliser des indices prosodiques de frontières de groupes phonologiques pour contraindre leur analyse syntaxique des phrases ?*

Notre hypothèse de recherche suppose que les frontières prosodiques sont calculées en temps réel et exploitées pour guider l'analyse syntaxique des phrases. Si ces frontières sont véritablement disponibles précocement dans le traitement des phrases, alors leur présence ou leur absence devrait contraindre le traitement syntaxique réalisé en temps réel par des adultes. C'est ce que nous avons testé dans notre première étude (chapitre 2).

Une des principales méthodologies utilisées pour répondre à la question du rôle de la prosodie dans le traitement syntaxique des phrases est d'étudier des phrases contenant des ambiguïtés syntaxiques. En effet, de telles phrases permettent aux chercheurs d'examiner comment ces ambiguïtés sont résolues et d'isoler ainsi les mécanismes impliqués plus généralement dans la construction de la structure syntaxique des phrases.

Nous avons donc mis en place des phrases temporairement ambiguës, dans lesquelles un mot pouvait appartenir potentiellement à deux catégories syntaxiques différentes. Par exemple, /mɔʀ/ peut être soit un verbe (comme dans « [le petit chien] [mord la laisse] »), soit un adjectif (comme dans « [le petit chien mort] [sera enterré demain] »). Ces deux phrases commencent par la même suite de phonèmes jusqu'au mot ambigu inclus, mais diffèrent au niveau de leur structure syntaxique et donc de leur structure prosodique. Dans la première phrase, le mot ambigu est un verbe et il est précédé d'une frontière de groupe phonologique (indiquant la frontière syntaxique entre le groupe nominal sujet et le verbe), tandis que c'est un adjectif dans la deuxième phrase et qu'il est suivi par une frontière prosodique.

Cette différence de structure prosodique influence-t-elle l'interprétation du mot ambigu ? Nous avons réalisé une première expérience, afin de tester si les indices prosodiques sont perceptibles par les adultes. Dans cette expérience, les phrases ambiguës étaient coupées juste à la fin du mot ambigu, puis présentées auditivement à des sujets qui devaient alors compléter ces débuts de phrases. Les stimuli utilisés étaient prononcés par une locutrice experte et informée du but de notre expérience, soit avec une prosodie particulièrement informative (allongements importants et grande variation de la fréquence fondamentale en fin d'unité prosodique), soit avec une prosodie beaucoup plus neutre (mais toujours d'une manière naturelle).

Les résultats que nous avons obtenus montrent que les participants ont donné significativement plus de réponses verbe sur les phrases ambiguës verbe que sur les phrases ambiguës adjectif, et vice-versa (plus de réponses adjectif sur les phrases adjectif que sur les phrases verbe), et ce uniquement dans la condition de prosodie informative (pour la prosodie neutre, la tendance va dans la même direction mais n'est pas significative). Cette expérience indique donc que les frontières de groupes phonologiques sont perçues et exploitées par les sujets adultes. Cependant, elle ne nous indique pas quand la prosodie intervient dans le traitement : en effet, elle pourrait soit

contraindre les premières étapes de l'analyse syntaxique (en favorisant dès le départ une analyse syntaxique par rapport à une autre), soit être exploitée comme une stratégie de dernier recours, lorsque l'analyseur syntaxique rencontre une ambiguïté qui ne peut pas être résolue autrement.

Pour répondre à cette question, nous avons réalisé une deuxième expérience dans laquelle les adultes devaient détecter des entrées lexicales abstraites. Dans cette expérience, les mots cibles étaient présentés visuellement avec une indication sur leur catégorie syntaxique (par exemple, « il est mort » pour détecter l'adjectif et « mordre » pour détecter le verbe). Les participants devaient alors répondre à la cible dans les versions coïncidentes (quand la cible était effectivement présente dans la phrase qu'ils entendaient ensuite, par exemple cible verbe – phrase verbe), mais se retenir de répondre dans les versions non-coïncidentes (par exemple, cible verbe – phrase adjectif).

Les premières analyses ont mis en évidence un effet d'ambiguïté : les adultes étaient plus lents pour répondre aux phrases ambiguës que pour répondre à des phrases contrôles qui contenaient les mêmes cibles mais sans aucune ambiguïté possible quant à leur catégorie syntaxique. Ce ralentissement lié à l'ambiguïté était significativement plus important en prosodie neutre qu'en prosodie informative, ce qui indique que les informations prosodiques, même si elles ne sont pas parfaitement fiables, ont été tout de même exploitées par les sujets.

De la même manière, les analyses des fausses alarmes ont montré que la présence d'une frontière prosodique avant le mot cible (dans les phrases verbe) pouvait plus aider les adultes à analyser les phrases entendues, comparée à l'absence de frontière prosodique avant le mot cible (dans les phrases adjectif). L'effet d'ambiguïté était effectivement moins fort sur les phrases verbe que sur les phrases adjectif.

Finalement, une analyse des réponses, restreinte aux interprétations données avant la fin du mot ambigu (donc avant l'accès à l'information lexicale désambiguïsante), a indiqué que les sujets pouvaient utiliser les informations prosodiques dès les premières étapes de leur analyse syntaxique : ils donnaient effectivement plus de réponses verbe aux phrases verbe, et plus de réponses adjectif aux phrases adjectif, alors que les seules informations disponibles pour distinguer les deux phrases au moment de leur réponse étaient les informations apportées par les frontières

de groupes phonologiques. Encore une fois, cet effet n'a été mis en évidence que dans la condition de prosodie informative.

Au total, cette première étude nous a permis de démontrer que, même si la prosodie ne permet pas de désambiguïser totalement les phrases ambiguës, elle joue effectivement un rôle précoce dans l'analyse syntaxique des phrases, quand les indices prosodiques sont bien marqués (prosodie informative). Mais on pouvait nous faire la critique que la prosodie produite par notre locutrice experte et informée n'était peut-être pas une prosodie très naturelle, mais au contraire une prosodie forcée que l'on trouverait uniquement dans ce cadre expérimental. Est-il possible que notre locutrice ait trop accentué les frontières de groupes phonologiques, comparé à ce que feraient des adultes plus naïfs, de sorte que les conclusions que nous avons développées ne seraient valables que dans ce cas particulier (ce qui diminuerait grandement la plausibilité et la viabilité de notre hypothèse sur le rôle des frontières prosodiques dans le traitement syntaxique adulte en temps réel et dans les premières étapes de l'acquisition de la syntaxe) ? Pour répondre à cette question, nous avons donc décidé de faire enregistrer ces phrases ambiguës par des locuteurs « naïfs ».

### *3. Les locuteurs français produisent-ils naturellement des indices prosodiques assez saillants pour être ensuite utilisés par des auditeurs dans l'analyse syntaxique des phrases ?*

Dans une deuxième étude (chapitre 3), nous avons donc utilisé les mêmes phrases ambiguës (certaines avaient été modifiées pour pallier à certains problèmes de biais de fréquence ou de biais de plausibilité des items ambigus dans les phrases). Nous avons demandé à six locutrices françaises de les enregistrer : celles-ci étaient des locutrices non expertes (étudiantes), non informées du but de nos expériences, qui ne se sont pas rendues compte du problème d'ambiguïté contenu dans nos phrases.

Ces phrases enregistrées ont ensuite été coupées à la fin du mot ambigu, puis présentées à des auditeurs français dans une tâche de complétion de phrases (similaire à celle que nous avons réalisée précédemment). Les résultats que nous avons obtenus dans notre première expérience ont été répliqués, et ce de manière générale mais aussi

pour chaque locutrice : en effet, nous avons trouvé que les auditeurs ont été capables d'interpréter correctement (et différemment) deux débuts de phrases qui ne différaient que par leur structure prosodique et syntaxique. Ils ont assigné des catégories syntaxiques différentes aux mots ambigus, selon le contexte prosodique dans lequel ils apparaissaient (plus de réponses verbe que de réponses adjectif aux phrases verbe, et plus de réponses adjectif que de réponses verbe aux phrases adjectif).

Les résultats de cette expérience ont pu permettre d'apporter de nouvelles données au débat qui existe sur la question de la production spontanée et naturelle d'indices prosodiques désambiguïsants par des locuteurs naïfs et non-experts. Contrairement à ce qui avait été initialement proposé, il ne semble pas que le contexte expérimental dans lequel les locuteurs produisent les stimuli soit un facteur déterminant : on croyait initialement que le fait de faire lire des phrases (une des premières techniques utilisées pour générer la production de phrases ambiguës, comme chez Allbritton et al. 1996) empêchait la production de bons indices prosodiques, et qu'il fallait au contraire que les locuteurs soient impliqués dans une relation avec un interlocuteur et dans une situation ayant un véritable but communicationnel (comme par exemple, dans les tâches de jeu coopératif présentées par Kraljic et Brennan, 2005). Or nous avons obtenu des indices particulièrement désambiguïsants et informatifs pour les auditeurs en utilisant une tâche de lecture. Il semble en fait que le vrai facteur impliqué dans les différences de « production spontanée » soit surtout lié à la structure prosodique des phrases. Plus les phrases sont courtes, plus elles auront tendance à être prononcées comme une seule unité prosodique ; plus les phrases sont longues, plus le fait de les diviser en plusieurs unités prosodiques devient obligatoire. Dans les études utilisant des phrases très courtes (7 syllabes ou moins), la production la plus naturelle consistait à ne produire qu'une unité prosodique. Dans nos études, nous avons utilisé des phrases moyennement longues, et nous avons exploité une frontière entre un groupe nominal sujet et un groupe verbal, une frontière que toutes les théories s'accordent à considérer comme bien marquée.

La première partie de notre recherche a donc permis de mettre en évidence que : (1) les indices bien marqués de frontières de groupes phonologiques sont utilisés, même en temps réel, pour contraindre l'analyse syntaxique des phrases, et (2) même des locuteurs naïfs et non-experts produisent naturellement ces indices de frontières prosodiques, qui sont étroitement liées à la structure syntaxique des phrases.

Puisque les frontières de groupes phonologiques sont calculées en temps réel par les adultes pour guider l'analyse syntaxique des phrases qu'ils entendent, il semble plausible de postuler que les enfants qui acquièrent le langage pourraient eux aussi les utiliser pour acquérir, par exemple, certaines notions sur la syntaxe de leur langue maternelle. Mais avant d'étudier ce point précis, nous devons nous assurer que les enfants français pouvaient percevoir et utiliser ces frontières prosodiques pour réaliser des tâches linguistiques.

#### *4. Les enfants français peuvent-ils percevoir et exploiter les frontières de groupes phonologiques pour analyser en temps réel les phrases qu'ils entendent ?*

En plus de toujours correspondre à des frontières de constituants syntaxiques, les frontières de groupes phonologiques coïncident toujours avec des frontières de mot. Nous avons décidé d'étudier dans une troisième étude (chapitre 5) l'utilisation de ces frontières prosodiques dans la segmentation de la parole continue en mots, par des enfants français de 16 mois. Pour ce faire, nous avons utilisé une variante de la technique de conditionnement de l'orientation de la tête : dans une première session, les enfants étaient habitués à répondre à un mot bisyllabique (« balcon ») puis, dans une deuxième session, nous leur présentions des phrases complètes qui soit contenaient le mot cible (« [Le joli balcon] [cloisonnait la terrasse] »), soit contenaient les deux syllabes du mot cible dans deux mots différents et séparées par une frontière de groupe phonologique (« [Le dernier bal] [conclura la saison] »).

Les résultats trouvés indiquent que les enfants français de 16 mois perçoivent et interprètent en temps réel des frontières de groupe phonologique comme des frontières de mot : on observe en effet que les enfants, qui ont appris à tourner la tête pour un mot bisyllabique, répondent ensuite significativement plus aux phrases qui contiennent réellement ce mot bisyllabique, comparées aux phrases dans lesquelles les deux syllabes cibles sont séparées par une frontière de groupe phonologique.

Ainsi, nous avons trouvé que les frontières de groupes phonologiques étaient utilisées en temps réel par les enfants pour contraindre leur accès au lexique ; des résultats similaires ont été trouvés auprès d'adultes français (Christophe et al., 2004). De plus, nous avons mis en évidence le fait que les adultes français utilisaient aussi en

temps réel ces frontières prosodiques pour contraindre leur analyse syntaxique des phrases. Notre hypothèse selon laquelle les enfants pourraient aussi utiliser ces informations prosodiques, conjointement avec l'utilisation des mots grammaticaux, semble donc très plausible, mais reste à tester.

*5. Les mots grammaticaux et les informations prosodiques peuvent-ils être utilisés pour guider les premières étapes de l'acquisition de la syntaxe ?*

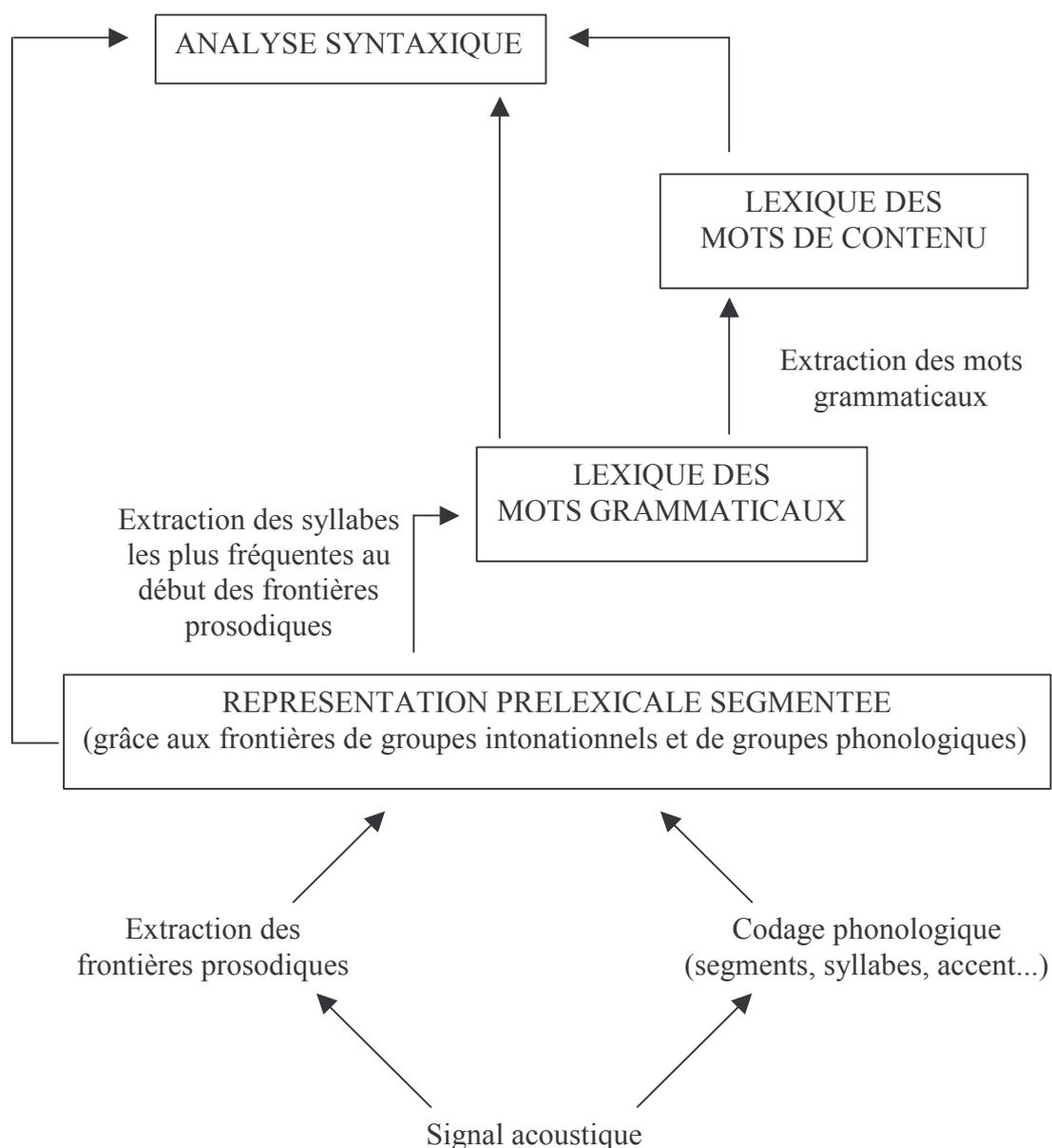
Dans une dernière étude (étude 4 - chapitre 6), nous avons décidé de tester directement la viabilité de notre hypothèse de recherche en réalisant une expérience de faisabilité auprès d'adultes français : ceux-ci étaient placés dans une situation que l'on pourrait rapprocher de celle perçue et vécue par les enfants dans leur deuxième année de vie. En effet, nous avons demandé à des adultes de traiter des phrases dans lesquelles les informations prosodiques et les mots grammaticaux étaient présents (et nous savons maintenant que les enfants, vers 16 mois, ont accès à ces deux sortes d'informations), mais dans lesquelles tous les mots de contenu étaient remplacés par des non-mots (c'est donc une situation un peu extrême puisque nous savons que les enfants de 16 mois connaissent déjà quelques mots de leur langue maternelle). Ils devaient réaliser une tâche de détection de mot abstrait (similaire à ce que nous avons réalisé sur des phrases françaises dans l'étude 1) sur des phrases qu'ils ne comprenaient pas, et ne pouvaient donc pas avoir recours à leurs connaissances lexicales pour réaliser cette tâche.

Le premier résultat trouvé par cette expérience est que les adultes utilisent parfaitement la présence immédiate des mots grammaticaux pour inférer la catégorie syntaxique du mot de contenu suivant (dans plus de 90% des cas, un non-mot précédé d'un article est interprété comme un nom, tandis qu'un non-mot précédé d'un pronom est interprété comme un verbe). Le résultat le plus impressionnant est que la présence d'une frontière de groupe phonologique, utilisée conjointement avec la présence des mots grammaticaux, permet d'inférer aussi facilement et aussi rapidement la catégorie syntaxique de certains mots cibles : en entendant « [sa cramona] [bamoule...], les adultes interprètent le mot cible comme un verbe dans 90% des cas (et ne se font piéger que dans 10% des cas, en répondant que ce mot est un nom).

Pour répondre correctement, les participants ont donc vraisemblablement isolé les constituants syntaxiques de cette phrase (en utilisant les frontières de groupes phonologiques), puis interprété le premier constituant comme un groupe nominal (grâce à la présence du pronom possessif « sa »), et enfin inféré que le plus probable après ce groupe nominal était d'avoir un groupe verbal, ce qui impliquait que « bamoule » était un verbe. Ceci veut dire que, conformément à notre hypothèse de départ, les adultes ont construit une première représentation syntaxique de la phrase qu'ils entendaient, sous la forme suivante : « [sa xxx]<sub>GN</sub> [xx x]<sub>GV...</sub> », où les frontières syntaxiques sont données par les frontières prosodiques, où les noms des groupes sont donnés par les mots grammaticaux, et dans laquelle ils ont pu correctement inférer que le premier mot du groupe verbal était le verbe cible. Il semble parfaitement logique de supposer que les enfants se comportent de manière similaire pour traiter les phrases qu'ils entendent.

## *6. Perspectives*

L'hypothèse de recherche qui a guidé ce travail de thèse pourrait être représentée par le modèle illustré sur la Figure 15.



**Figure 15 : Modèle d'acquisition de la syntaxe dérivant de notre hypothèse de recherche, à savoir le rôle prédominant des indices prosodiques et des mots grammaticaux dans la construction d'une représentation syntaxique des phrases.**

Selon notre modèle, nous partons du signal acoustique des phrases que les enfants entendent. Notre idée est que les enfants pourraient, à partir de ce signal acoustique continu, d'une part réaliser un codage phonologique de ce signal de parole (représentation de ce signal en segments, syllabes, accent...), et d'autre part extraire les frontières prosodiques (en utilisant les variations de durée et de fréquence fondamentale

portées sur certaines syllabes ou certains segments). Ils auraient alors accès à une représentation prélexicale segmentée (grâce aux frontières de groupes intonationnels et frontières de groupes phonologiques). Grâce à cette représentation prélexicale segmentée en unités prosodiques, les enfants pourraient extraire les syllabes les plus fréquentes qui se trouvent aux débuts des unités prosodiques, et ainsi créer un lexique spécifique pour les mots grammaticaux. Par la stratégie d'extraction des mots grammaticaux, les enfants pourraient alors accéder aux mots de contenu (le mot suivant un mot grammatical est considéré comme un mot de contenu) et développer ainsi un lexique des mots de contenu. L'utilisation des indices distributionnels entre articles et noms, et entre pronoms et verbes, permettraient à l'enfant d'accéder à la catégorie syntaxique des mots qu'ils perçoivent. Les informations apportées par les frontières de groupes phonologiques et les mots grammaticaux permettraient à l'enfant de commencer à construire une structure syntaxique des phrases qu'ils entendent et de guider leurs premières analyses syntaxiques. Celles-ci seraient ensuite affinées et se développeraient de plus en plus grâce aux mots de contenu que les enfants intégreraient progressivement dans leur lexique puis dans leur représentation syntaxique.

Cette hypothèse de recherche sur le rôle des frontières de groupes phonologiques, utilisées conjointement avec les informations apportées par les mots grammaticaux, a été confirmée pour le traitement syntaxique en temps réel réalisé par les adultes français. Elle semble aussi parfaitement plausible et viable en ce qui concerne les premières étapes de l'acquisition de la syntaxe. Mais il est évident que la meilleure démonstration serait de montrer cet effet directement auprès d'une population d'enfants, et pour cela différentes expériences devront être réalisées pour tester chaque étape de ce modèle.

### Perception et utilisation des frontières de groupes phonologiques

Une des premières conditions nécessaires à la validité de cette hypothèse est que les enfants puissent percevoir et interpréter les frontières de groupes phonologiques dans des tâches linguistiques. L'expérience que nous avons réalisée et présentée dans le chapitre 5 indique que les enfants sont capables d'interpréter les frontières de groupe phonologique comme des frontières de mot, et ce dès l'âge de 16 mois.

Y a-t-il une quantité d'indices prosodiques nécessaire pour que les enfants perçoivent ces frontières prosodiques ? Pour le vérifier, il faudra réaliser une expérience similaire à celle de l'étude 3 (chapitre 5), dans laquelle les stimuli seront synthétisés (par exemple, sur la base des indices prosodiques produits par notre locutrice), puis dans lesquels nous ferons varier la qualité des indices prosodiques de frontières de groupes phonologiques. Ces indices prosodiques pourront être classés sur une échelle allant de 100% (indices fiables et bien marqués, similaires à ceux produits naturellement dans nos expériences) à 0% (pas d'indices prosodiques disponibles), avec différents paliers de 20% (par exemple, à 80%, les allongements des segments et les variations de la fréquence fondamentale seront diminués de 20% par rapport aux enregistrements originaux...). Nous pourrions alors étudier l'influence de ces indices prosodiques dans le traitement de la parole continue. En effet, on pourrait faire deux hypothèses :

(1) d'une part, un modèle à seuil dans lequel l'information prosodique est utilisée très précocement quand elle signale clairement une frontière de mot (mais dans ce cas, quelle est la quantité nécessaire et suffisante d'information ?). Cette information serait alors utilisée avant toutes les autres et permettrait un découpage initial du signal de parole. Les processus d'accès au lexique se dérouleraient ensuite à l'intérieur de ces unités, avec une activation très forte des candidats lexicaux en bordures d'unités prosodiques et une inhibition très forte des candidats traversant une frontière de groupe phonologique ; de même, on pourrait supposer que l'analyseur syntaxique exploiterait fortement cette frontière prosodique comme une frontière syntaxique (une frontière syntaxique serait postulée entre les syllabes traversant une frontière de groupe phonologique) ;

(2) d'autre part, on peut aussi faire l'hypothèse d'un modèle dans lequel le signal de parole serait analysé sur la base de différents indices prélexicaux (statistiques, allophoniques, prosodiques...), chacun apportant une partie de l'information qui serait ensuite utilisée pour accéder au lexique. L'effet de la prosodie ne serait pas en « tout-ou-rien » comme dans le modèle précédent, mais aurait plutôt un effet graduel et l'activation d'un mot varierait avec la force des indices prosodiques : dans ce cas, plus les indices de frontières prosodiques sont clairs, plus les mots qui commencent ou finissent à cette frontière vont gagner en activation ; de la même manière, plus ces indices de frontières prosodiques sont forts, plus les mots potentiels traversant une frontière prosodique vont être inhibés ; de la même manière, on pourrait postuler que l'activation des structures syntaxiques varierait aussi en fonction de la taille des indices

prosodiques présents dans le signal. De plus, on sait que la taille des indices de frontières prosodiques varient selon la position de ces unités dans la hiérarchie prosodique (on pourrait peut-être trouver que la nature des frontières prosodiques postulées varierait en fonction de la force des indices prosodiques).

Mais notons qu'il se pourrait aussi que les conséquences ne soient pas les mêmes selon que l'on considère l'activation des mots ou l'activation des structures syntaxiques : un modèle à seuil pour les mots (une frontière de mot n'est postulée que si les indices prosodiques l'indiquent fortement) et un modèle graduel pour les structures syntaxiques (où un gradient des indices prosodiques entraînerait soit un gradient dans l'activation des structures syntaxiques, soit un gradient dans la nature des unités prosodiques activées).

Des expériences pourront être réalisées avec des enfants de 16 mois, entendant des phrases synthétiques françaises (comme celles utilisées dans l'étude 3 de cette thèse), ou des phrases synthétisées ne contenant que des non-mots : on comparerait par exemple la détection du non-mot « bronguin » dans « [nur javo bronguin] [clafidon brine fagou] » et dans [nur destole bron] [guinfaloni dro plimuche] » - le fait d'utiliser uniquement des non-mots permettra de mettre en place une expérience de faisabilité avec des adultes français réalisant une tâche de détection de mot, et de pouvoir ainsi comparer les performances des enfants avec celles des adultes.

Finalement, il faudra s'assurer que les indices de frontières de groupes phonologiques sont bien marqués dans la parole adressée aux enfants : en effet, si nos futures expériences montrent par exemple que les enfants peuvent les utiliser, à partir d'un certain niveau, pour segmenter la parole continue, il faudra alors s'assurer que cette quantité d'informations prosodiques nécessaire ne se trouve pas qu'en condition expérimentale dans un laboratoire, mais que les enfants y ont généralement accès dans la vie de tous les jours (donc que les phrases qui leur sont adressées sont bien marquées au niveau prosodique). Pour cela des analyses prosodiques de parole adressée aux enfants devront être mises en place pour s'assurer de la viabilité de cette hypothèse sur la perception et l'utilisation des frontières de groupes phonologiques.

Pour que notre hypothèse de recherche soit plausible, il faut aussi que les enfants réussissent en plus à intégrer l'information apportée par les mots grammaticaux, dans ce découpage en unités prosodiques. En effet le seul découpage de la parole continue en constituants prosodiques (donc en constituants syntaxiques) n'apporte pas en lui-même de véritable information sur la syntaxe des phrases que les enfants entendent. Par exemple, la phrase « le garçon a mangé une pomme » pourrait être représentée sous la forme « [xxx] [xxx] [xx] », c'est-à-dire en groupes de syllabes séparées les unes des autres par des frontières prosodiques. Mais à ce niveau de l'analyse, il n'y a aucun moyen de savoir, par exemple, si le premier groupe est un groupe nominal, ou un groupe verbal, ou un groupe prépositionnel...

Seule l'information apportée par les mots grammaticaux pourrait permettre d'identifier la catégorie de ces unités syntaxiques. Encore faut-il que les enfants puissent apprendre et reconnaître les mots grammaticaux de leur langue maternelle (et construire ainsi un répertoire des mots grammaticaux) ? Mais la simple reconnaissance des mots grammaticaux ne permet pas en elle-même d'inférer la catégorie syntaxique de l'unité porteuse (au mieux, cette reconnaissance permet d'extraire le mot grammatical du signal de parole, ce qui a pour conséquence de pouvoir trouver le début du mot de contenu suivant). Il faut donc s'assurer dans un deuxième temps que les enfants peuvent utiliser ces informations grammaticales pour inférer la catégorie syntaxique du mot suivant (en utilisant par exemple le fait que les noms sont généralement précédés d'articles, tandis que les verbes sont généralement précédés de pronoms).

### Quand et comment les enfants reconnaissent-ils les mots grammaticaux de leur langue maternelle ?

Nous savons depuis récemment, grâce aux travaux de Shi et Gauthier (2005) que les enfants français de 8 mois reconnaissent les mots grammaticaux de leur langue maternelle. Comment ont-ils réussi à découvrir et à catégoriser ces items lexicaux comme un groupe distinct (différent de la catégorie des mots de contenu) ? Plusieurs études ont montré que les mots grammaticaux avaient des propriétés phonologiques particulières, mais ces études ont été réalisées sur d'autres langues que le français. Il sera donc nécessaire d'analyser des corpus de parole adressée aux enfants français afin

de déterminer, puis de tester, quels sont les paramètres les plus pertinents pour découvrir les mots grammaticaux du français.

Quelques propriétés de cette catégorie d'items semblent par contre se retrouver dans toutes les langues, comme le fait qu'ils soient très fréquents et qu'ils apparaissent généralement en bordure d'unités prosodiques. Il sera nécessaire d'enregistrer de la parole adressée aux enfants et de réaliser des analyses prosodiques sur ces corpus (et ne pas se contenter simplement d'une simple retranscription écrite n'intégrant pas les indices de frontières prosodiques). Ceci nous permettra d'ajouter une information supplémentaire dans les analyses et simulations : la position des items par rapport aux frontières prosodiques.

Une des hypothèses possibles, si cette distribution des mots grammaticaux par rapport aux frontières prosodiques se confirme, est que les enfants interpréteraient toutes les syllabes fréquentes en début d'unité prosodique comme des mots grammaticaux. Si cela est le cas, il devraient pouvoir extraire facilement du signal de parole cet item lexical, et interpréter le matériel suivant comme un mot de contenu. Dans cette situation, des enfants entraînés à répondre au non-mot « rion » devraient le détecter plus souvent dans « Marion » (la première syllabe de ce nom propre étant probablement interprétée comme un mot grammatical donc extraite du signal de parole), alors que cela devrait être plus rare dans un contexte où ce non-mot est précédé d'une syllabe non-homophone à un mot grammatical (comme dans « karion »). Des expériences adultes pourront être réalisées avec des stimuli similaires pour tester la plausibilité de cette hypothèse d'extraction des mots grammaticaux : ils devraient être ralentis en rencontrant, au début d'une unité prosodique, une syllabe homophone à un mot grammatical mais qui n'en est pas un

Quand les enfants utilisent-ils les mots grammaticaux pour inférer la catégorie grammaticale des mots de contenu ?

Si les enfants réussissent à extraire les mots grammaticaux du signal de parole, ils pourraient donc les utiliser, dans une première étape, pour inférer que le mot suivant est un mot de contenu. Pour que notre hypothèse sur la mise en route de l'analyse syntaxique des phrases soit confirmée, une deuxième étape est alors nécessaire : les

enfants doivent réussir à inférer la catégorie syntaxique du mot de contenu suivant le mot grammatical, en utilisant par exemple l'association qui existe entre article et nom, ou entre pronom et verbe.

De récentes données ont montré que les enfants français de 23 mois étaient capables d'utiliser la présence d'un mot grammatical pour inférer la catégorie syntaxique du mot de contenu suivant (un non-mot précédé d'un pronom est interprété comme un verbe, Bernal et al., 2005). De nouvelles expériences doivent maintenant être réalisées avec des enfants plus jeunes pour trouver l'âge auquel ils deviennent capables d'inférer la catégorie syntaxique d'un mot de contenu en fonction du mot grammatical qui le précède (des résultats similaires ont été trouvés avec des enfants allemands de 16 mois par Höhle et al., 2004, mais des différences inter-langues peuvent exister entre les performances des enfants et il sera donc nécessaire d'étudier directement les enfants français).

Il sera aussi intéressant d'utiliser l'homonymie qui existe en français entre certains articles et certains pronoms. Par exemple, si l'on considère la phrase « [le jeune homme] [la prend] [par la main] », est-ce que le pronom « la » pourrait être considérée comme un article (de par sa position dans la structure prosodique de la phrase) ou possède-t-il des caractéristiques (phonologiques, prosodiques, position dans la phrase...) qui permettent de le différencier de l'article et d'inférer que ce pronom doit être suivi par un verbe ? Si cette homonymie gêne les enfants, ils pourraient alors initialement penser que « prend » est un nom. Des résultats préliminaires pourront aussi être apportés en réalisant une expérience similaire à notre dernière étude (détection de mot abstrait dans des phrases jabberwocky) : comment les adultes interpréteraient-ils le non-mot « priche » dans « [le jane oure] [*la priche*] [par la mune] » : comme un nom ou comme un verbe ?

Cette technique pourra aussi être utilisée pour tester, non plus la simple co-occurrence existant entre un article et un nom (ou entre un pronom et un verbe), mais le rôle des cadres fréquents dans la catégorisation syntaxique. Par exemple, Chemla (2005) a réalisé une analyse distributionnelle sur des corpus français de parole adressée aux enfants, et trouvé que les cadres fréquents (qui avaient été proposés par Mintz, 2003) semblaient très utiles pour trouver la catégorie syntaxique des mots : par exemple, dans le cadre « la X de », X est un nom dans 100% des cas. Tester des phrases contenant des pronoms homophones à des articles sera un moyen de vérifier l'information apportée

par ces cadres lexicaux et les inférences faites sur l'item situé au centre de ce cadre : il serait possible que le verbe « couvre » soit en fait interprété comme un nom dans « [le jeune homme] [la couvre] [de cadeaux] », de par sa position entre « la » et « de » (mais il se pourrait aussi que la présence de la frontière prosodique ait un impact sur les réponses des sujets : elle pourrait par exemple bloquer la reconnaissance ou la prise en compte du cadre lexical « la X de », et l'on se retrouverait alors dans une situation similaire à l'exemple que nous avons développé juste avant). Cet exemple montre aussi qu'il sera nécessaire d'intégrer les informations apportées par les frontières prosodiques dans les analyses distributionnelles faites sur la parole adressée aux enfants (une autre raison de le faire est que les analyses distributionnelles qui sont réalisées actuellement reposent sur l'hypothèse que les mots sont déjà segmentés : or, accéder aux frontières des mots est quelque chose de relativement difficile. Il serait donc plus logique et plus plausible, au vu des résultats expérimentaux disponibles dans la littérature et présentés dans le chapitre 4 de cette thèse, de supposer que les enfants n'ont pas accès à toutes les frontières des mots mais seulement aux frontières prosodiques).

Les premières idées que nous venons de présenter comme suites expérimentales à donner à nos travaux concernent d'une part la perception des frontières prosodiques, et d'autre part le rôle des mots grammaticaux dans la catégorisation syntaxique. Notre hypothèse de recherche suppose que les enfants auraient besoin de traiter simultanément et en temps réel ces deux indices pour réaliser une première analyse syntaxique des phrases. Nous allons donc maintenant proposer des études qui devront être réalisées pour tester la possibilité que les enfants peuvent utiliser la combinaison des mots grammaticaux et des informations prosodiques pour analyser les phrases.

Quand les enfants utilisent-ils les mots grammaticaux et les indices prosodiques pour assigner une structure syntaxique aux phrases qu'ils entendent ?

Si les expériences futures montrent que vers l'âge de 2 ans, les enfants perçoivent et utilisent les frontières prosodiques, et qu'ils sont aussi capables, par exemple, d'inférer que « jaurime » est un nom en entendant « une jaurime », « la jaurime », « des jaurimes »... alors il nous sera possible d'étudier précisément

l'hypothèse de recherche que nous avons proposée (sur l'utilisation conjointe de ces deux indices dans les premières étapes de l'analyse syntaxique des phrases).

La tâche expérimentale que nous avons inventée pour tester les adultes (détection d'une entrée lexicale abstraite) est trop complexe pour être utilisée avec de jeunes enfants. Mais il est possible de mettre en place des protocoles expérimentaux adaptés, pour tester à la fois chez les adultes et les enfants le rôle des frontières prosodiques et des mots grammaticaux dans la construction en temps réel des structures syntaxiques des phrases.

Je propose de réaliser des tâches de détection de mot, dans des phrases relativement simples, dont nous manipulerons la structure prosodique (de manière à ce qu'il y ait adéquation ou non entre les frontières des constituants syntaxiques et les frontières des constituants prosodiques). Par exemple, les participants entendront des phrases comme « [Le petit ballon] [rebondit bien] » où les frontières de groupes phonologiques sont en adéquation avec les frontières des constituants syntaxiques. Avec la technique de « cross-splicing », il sera possible de créer des phrases chimériques avec une prosodie conflictuelle, comme « [Le petit] [ballon rebondit bien] ». Si les mots sont soigneusement choisis (grâce aux résultats des questionnaires parentaux par exemple) et donc bien connus par les enfants (d'environ 2 ans par exemple), il sera alors possible de tester et de mettre en évidence leurs attentes syntaxiques. L'utilisation de la technique du conditionnement de l'orientation de la tête permettra par exemple d'habituer des enfants à répondre au mot « ballon ». Dans une deuxième session, ils seront confrontés aux phrases contenant soit une prosodie adéquate, soit une prosodie inadéquate. Si les frontières de groupes phonologiques contraignent l'analyse syntaxique des phrases faite par les enfants, ils devraient avoir des attentes concernant la catégorie syntaxique des mots qu'ils vont entendre : en entendant « le » suivi de « petit », les enfants devraient inférer que « petit » est un nom, et que l'unité syntaxique « le petit » (délimitée par des frontières de groupes phonologiques) est donc un groupe nominal ; après avoir entendu une frontière prosodique à la fin du groupe nominal « [le petit] », ils devraient plus s'attendre à entendre ensuite un verbe (c'est le même raisonnement que ce que nous avons proposé pour le traitement des phrases en jabberwocky, comme « [sa cramona] [bamoule...] », où les adultes ont majoritairement interprété « bamoule » comme un verbe). Les enfants devraient donc être plus lents pour répondre au mot cible « ballon » (voire même ne pas le reconnaître) quand les frontières prosodiques sont en inadéquation avec les frontières syntaxiques (des résultats similaires devraient être mis

en évidence auprès d'adultes, avec des temps de réaction plus longs pour détecter le mot cible dans des phrases prononcées avec une prosodie conflictuelle).

Ces mêmes phrases pourront aussi être utilisées dans des expériences mesurant les potentiels évoqués, de manière à vérifier si ces incongruences donnent lieu à l'onde typiquement observée lors d'une incompatibilité syntaxique (LAN, négativité précoce et antérieure à gauche, Friederici, 2002; Osterhout, 1994).

Finalement, nous proposons qu'il est possible de réaliser une expérience similaire à celles que nous avons conduites avec les adultes français dans la première étude, et dans laquelle nous testerons directement notre hypothèse (en testant le rôle des indices prosodiques et des mots grammaticaux dans la résolution d'ambiguïtés syntaxiques temporaires).

Considérons par exemple, les phrases ambiguës suivantes : (1) « [la vieille ferme] [est délabrée] » et (2) « [la vieille] [ferme la fenêtre] ». Comme dans les expériences de détection de mot abstrait que nous avons réalisées, ces paires de phrases commencent par les mêmes phonèmes ; l'ambiguïté réside dans la catégorie syntaxique du troisième mot : dans la première phrase, ce mot ambigu est un nom, tandis que c'est un verbe dans la deuxième. Jusqu'au mot ambigu inclus, la seule différence entre ces deux paires de phrases est une différence de structure syntaxique, donc une différence de structure prosodique (la frontière de groupe phonologique est située avant le mot ambigu quand c'est un verbe, mais après le mot ambigu quand c'est un nom).

Nous proposons de réaliser une expérience en deux phases : dans une première phase, nous familiariserons les enfants avec un mot, défini avec une catégorie syntaxique précise : par exemple, ils entendront « je ferme », « tu fermes », « il ferme », « elle ferme »... jusqu'à habituation (on peut alors supposer que les enfants ont réagi que le mot « ferme » était répété ; mais s'ils utilisent la présence des pronoms avant ce mot, ils devraient plus précisément être familiarisés au verbe « ferme »). Dans une deuxième session, les enfants entendront deux passages différents et nous mesurerons leur préférence d'écoute : le premier passage comportera des phrases temporairement ambiguës contenant la cible verbe (« [la vieille] [ferme la fenêtre] », « [la petite] [ferme le tiroir] », « [la belle] [ferme la grille] [du jardin] »...), tandis que le deuxième passage comportera les phrases temporairement ambiguës avec la cible nom (« [la vieille ferme]

[est délabrée] », « [la petite ferme] [est très jolie] », « [la belle ferme] [sera vendue] [aux enchères] »...).

Si les enfants utilisent, comme les adultes, les informations prosodiques et les mots grammaticaux pour guider leur analyse syntaxique des phrases, ils devraient pouvoir inférer avec certitude que « ferme » est un verbe dans les phrases du premier passage. Prenons par exemple la phrase « la petite ferme le tiroir » : les frontières prosodiques devraient permettre de trouver les frontières des constituants prosodiques (« [xxx] [xxxx] ») ; la reconnaissance des articles, et en particulier celle de l'article « la » au début de la première unité (« [la xx] [x le xx] ») devrait permettre à l'enfant d'inférer que le premier groupe syntaxique est un groupe nominal, et donc que le groupe suivant a de grandes chances d'être un groupe verbal commençant très probablement par un verbe : dans cette situation, il est très plausible que « ferme » soit interprété comme un verbe (alors que les choses sont moins claires pour les phrases du deuxième passage, puisque la frontière de groupe phonologique ne se situe pas avant mais après le mot ambigu ; or l'absence d'une frontière prosodique ne correspond pas avec l'absence d'une frontière syntaxique, et les enfants pourraient donc être gênés dans cette situation).

Toutefois, si notre hypothèse est juste, les enfants d'environ 2 ans devraient se comporter différemment en entendant les deux passages : on s'attendrait plus précisément à ce que les enfants familiarisés avec la répétition du mot « ferme » dans sa forme verbale, préfèrent écouter le passage 1 (qui contient le mot familier) par rapport au passage 2.

Cette dernière expérience serait vraiment une expérience parallèle à celle que nous avons réalisée avec des adultes pour tester la viabilité et la plausibilité de notre hypothèse de recherche (elle utiliserait le même type de stimuli, à savoir des phrases comportant une ambiguïté syntaxique temporaire). Elle nous permettrait d'aller tester directement la pertinence de notre propos auprès d'une population de jeunes enfants.

Pour résumer, nous faisons l'hypothèse que les frontières prosodiques pourraient être utilisées par les enfants pour segmenter le discours en unités prosodiques. Les mots grammaticaux en bordure d'unités prosodiques seraient ensuite extraits du signal de parole, ce qui permettrait aux enfants de segmenter le mot de contenu suivant, puis d'inférer sa catégorie syntaxique grâce à la corrélation existant entre les articles et les

noms, ou entre les pronoms ou auxiliaires et les verbes. Intégrer ces deux indices permettrait à l'enfant d'obtenir une première ébauche de la structure syntaxique des phrases entendues, relativement simple mais néanmoins informative, du genre « [le xx]<sub>GN</sub> [a xx]<sub>GV</sub> [une xx]<sub>GN</sub> » (pour « le garçon a mangé une pomme »).

Notre modèle repose donc essentiellement sur les informations phonologiques et prosodiques du signal de parole, c'est-à-dire les informations qui peuvent être directement extraites du signal de parole. Il permet de proposer des hypothèses sur l'acquisition de la syntaxe qui ne présupposent pas que les enfants connaissent déjà un certain nombre de mots. Il est vrai que de nombreuses expériences restent encore à faire pour démontrer la validité de ces hypothèses. Mais les résultats obtenus et présentés dans cette thèse leur apportent déjà un soutien important et constituent une base fortement encourageante pour les nouvelles études qui seront menées dans ce champ de recherche.

# ANNEXES

## ANNEXES A RELATIVES A L'ETUDE 1 (CHAPITRE 2)

### A.1 STIMULI

Pour chaque item ambigu, écrit en lettres capitales, quatre phrases expérimentales ont été construites : deux phrases ambiguës et deux phrases contrôles. Les phrases ambiguës ont été utilisées dans les deux expériences (elles étaient coupées à la fin du mot ambigu pour l'expérience de complétion de phrases – Expérience 1). Les phrases contrôles n'ont été utilisées que dans l'expérience de détection de mot abstrait – Expérience 2).

Pour chaque item, les phrases expérimentales sont mentionnées dans l'ordre suivant : d'abord les phrases ambiguës (phrase ambiguë verbe puis phrase ambiguë adjectif), puis les phrases contrôles (phrase contrôle verbe puis phrase contrôle adjectif).

#### Item Phrases

- 1 Je trouve que la fumée DANSE joliment en s'échappant du feu de bois.
- 1 Je trouve que la fumée DENSE qui s'échappe de ce bâtiment laisse imaginer le pire.
- 1 Mon cousin DANSE dans le lac des cygnes.
- 1 Ce brouillard DENSE m'empêchait d'avancer sur la route.
  
- 2 Les pommes DURENT plus longtemps que les bananes si on les conserve à l'abri de la lumière.
- 2 Les pommes DURES font de meilleures tartes que les golden, par exemple.
- 2 L'entracte DURE plus longtemps que prévu.
- 2 C'est un coup DUR que vous venez de vivre.
  
- 3 Cet artiste GRAVE le bois et ses œuvres sont très prisées.
- 3 Cet artiste GRAVE médite en contemplant le paysage.
- 3 Quand on GRAVE son nom sur un tronc d'arbre, on l'abîme.
- 3 Ce n'est pas GRAVE donc ne t'inquiète pas.
  
- 4 Assise sur un banc, la jeune femme LACE les souliers de son petit garçon.
- 4 Assise sur un banc, la jeune femme LASSE reprend sa respiration.
- 4 Les enfants sont fiers quand ils LACENT pour la première fois leurs chaussures.
- 4 Je suis vraiment LASSE d'attendre depuis des heures.
  
- 5 Je crois que cet homme LACHE son boulot parce qu'il est trop stressant.
- 5 Je crois que cet homme LACHE refuse de voir la vérité en face.
- 5 Certains enfants ne veulent pas qu'on leur LACHE la main quand il y a des inconnus dans la pièce.
- 5 Il est tellement LACHE qu'il n'osera jamais s'opposer à sa belle-mère.
  
- 6 J'ai appris que cet homme LOUCHE depuis qu'il a eu un accident de voiture.
- 6 J'ai appris que cet homme LOUCHE doit comparaître devant la justice.
- 6 Il existe des personnes qui LOUCHENT depuis leur naissance.
- 6 Je trouve un peu LOUCHE que ma voisine ait toujours les mêmes notes que moi.
  
- 7 Le petit chien MORD la laisse qui le retient dans l'espoir de se libérer.
- 7 Le petit chien MORT sera enterré demain dans le jardin de ses maîtres.
- 7 Parfois on se MORD la langue quand on mange trop vite.
- 7 Maintenant qu'il est MORT, les batailles d'héritage vont commencer.

- 8 Je trouve que le petit tambourin ROMPT l'harmonie de la musique.  
8 Je trouve que le petit tambourin ROND a une très bonne sonorité.  
8 Quand une femme décide de divorcer, elle ROMPT les liens qui l'unissait à son mari.  
8 Au football, on joue avec un ballon ROND alors qu'il est ovale au rugby.
- 9 Elle trouve que les enfants SALENT beaucoup trop leur repas.  
9 Elle trouve que les enfants SALES font la honte de leurs parents.  
9 Je fais partie de ces gens qui SALENT leur repas sans y avoir goûté avant.  
9 Il y a des gens qui veulent qu'on fasse le SALE boulot à leur place car ils se croient supérieurs.
- 10 Ces gros nuages SOMBRENT derrière les montagnes.  
10 Ces gros nuages SOMBRES promettent un orage violent.  
10 Une personne qui SOMBRE dans l'alcoolisme a besoin d'aide.  
10 Il fait tellement SOMBRE qu'on ne sait plus où mettre les pieds.
- 11 De drôles de bruits COURENT dans l'hôtel quant à la venue d'une célébrité.  
11 De drôles de bruits COURTS sont frappés à ma porte comme une sorte de code.  
11 Ce sportif COURT tous les matins pour s'entraîner.  
11 L'entretien fut COURT mais très productif.
- 12 Ce politicien INTEGRE les minorités dans son projet de développement urbain.  
12 Ce politicien INTEGRE sera certainement élu aux prochaines législatives.  
12 Il faut que j'INTEGRE mes nouvelles données dans mon modèle psycholinguistique.  
12 C'était un homme très INTEGRE et tous ses voisins l'admiraient.
- 13 Ce grand écrivain CELEBRE la naissance de Victor Hugo avec son nouveau livre.  
13 Ce grand écrivain CELEBRE fera une dédicace demain dans une grande librairie.  
13 La date qui CELEBRE la prise de la Bastille est le 14 juillet  
13 Mon amie sera CELEBRE si elle continue de travailler dur.
- 14 Manger cette tartelette COMPLETE délicieusement ce repas.  
14 Manger cette tartelette COMPLETE ne me va pas car je n'ai plus faim.  
14 Ce premier timbre en euro COMPLETE ma collection.  
14 Ma sœur a lu les œuvres COMPLETEES de Chateaubriand.
- 15 Cette petite lumière DIFFUSE une agréable sensation de calme.  
15 Cette petite lumière DIFFUSE ne fatigue pas les yeux.  
15 J'écoute une radio qui DIFFUSE des concerts tous les soirs.  
15 Elle éprouve une douleur trop DIFFUSE pour la localiser avec précision.
- 16 Ce petit clown DISTRAIT les enfants malades dans les hôpitaux.  
16 Ce petit clown DISTRAIT n'a pas vu qu'il avait oublié son nez rouge.  
16 La télévision DISTRAIT beaucoup les enfants maintenant.  
16 Ce professeur est tellement DISTRAIT qu'il a perdu les copies de ses étudiants.
- 17 Ce président ILLUSTRÉ la réussite sociale et professionnelle.  
17 Ce président ILLUSTRÉ voyage énormément pour ses affaires.  
17 C'est un tableau qui ILLUSTRÉ bien les paysages de Bretagne.  
17 C'était un ILLUSTRÉ inconnu avant de faire carrière au cinéma.
- 18 Maintenant, certaines montres PRECISENT même la pression atmosphérique.  
18 Maintenant, certaines montres PRECISES permettent les mesures à la milliseconde près.  
18 Des exemples qui PRECISENT notre idée sont très utiles pour se faire comprendre de tous.  
18 J'ai une idée PRECISE de ce que je veux faire plus tard.
- 19 Cette belle femme CAPTIVE l'attention du public.  
19 Cette belle femme CAPTIVE craint pour sa vie.  
19 Cet homme qui CAPTIVE toujours son auditoire m'impressionne beaucoup.  
19 La jeune princesse était retenue CAPTIVE par sa méchante belle-mère.

- 20 Avant d'agir, ce diplomate REFLECHIT peu aux conséquences de ses actes.
- 20 Avant d'agir, ce diplomate REFLECHI pèse toujours le pour et le contre.
- 20 Je sais que tu REFLECHIS beaucoup avant de dire quelque chose.
- 20 Comme je suis quelqu'un de REFLECHI, je ne veux pas m'engager à la légère.

## A.2 EXPERIENCE 1 : COMPLETION DE PHRASES

### A.2.1 Consignes

Bonjour !

Merci de participer à une expérience de perception de la parole.  
Cela vous prendra une petite demi-heure et ne vous réclamera rien de plus qu'un peu d'attention et d'imagination.

L'expérience se déroulera en deux parties de 15 minutes chacune ; entre les deux vous pourrez prendre quelques instants de détente.

Vous allez entendre le début d'une phrase, et vous devrez inventer une suite possible pour cette phrase ; par exemple, si vous entendez « la petite fille... », vous pouvez répondre : « la petite fille va à l'école tous les matins ».

Pour répondre, vous écrirez la phrase entière sur la feuille de réponse qui vous est fournie.  
Pour réécouter le même début de phrase, appuyez sur la barre d'espace ; pour passer à la phrase suivante, appuyez sur « Entrée ».

Vous allez maintenant faire un entraînement de deux phrases pour vous familiariser avec la procédure.

Avant de commencer, vous devez avoir le casque sur la tête et disposer d'un crayon et de la feuille de réponse.

### A.2.2 Résultats individuels par sujets

	<b>PROSODIE INFORMATIVE</b>			
	Phrases Ambiguës Adjectif		Phrases Ambiguës Verbe	
	Réponses Adjectif	Réponses Verbe	Réponses Adjectif	Réponses Verbe
Sujet 01	4	6	4	4
Sujet 02	7	3	0	9
Sujet 03	5	5	2	5
Sujet 04	8	2	1	8
Sujet 05	7	3	7	3
Sujet 06	8	2	0	9
Sujet 07	6	4	4	4
Sujet 08	8	2	1	8
Sujet 09	5	5	1	9
Sujet 10	8	2	0	9
Sujet 11	7	3	6	3
Sujet 12	8	2	1	8
Sujet 13	5	4	3	6
Sujet 14	8	2	2	7
Sujet 15	6	4	5	5
Sujet 16	8	2	2	8
Sujet 17	7	3	3	7
Sujet 18	8	2	2	7
<b>Moyenne</b>	<b>6,83</b>	<b>3,11</b>	<b>2,44</b>	<b>6,61</b>

	<b>PROSODIE NEUTRE</b>			
	Phrases Ambiguës Adjectif		Phrases Ambiguës Verbe	
	Réponses Adjectif	Réponses Verbe	Réponses Adjectif	Réponses Verbe
Sujet 01	4	6	6	4
Sujet 02	5	4	1	9
Sujet 03	6	4	7	3
Sujet 04	8	2	5	5
Sujet 05	6	4	5	5
Sujet 06	6	4	3	7
Sujet 07	5	4	1	9
Sujet 08	2	7	3	4
Sujet 09	7	2	2	6
Sujet 10	2	7	2	7
Sujet 11	7	3	2	7
Sujet 12	2	8	7	3
Sujet 13	8	2	1	8
Sujet 14	4	5	6	3
Sujet 15	6	4	3	6
Sujet 16	2	7	4	4
Sujet 17	4	6	1	8
Sujet 18	5	4	6	4
<b>Moyenne</b>	<b>4,94</b>	<b>4,61</b>	<b>3,61</b>	<b>5,67</b>

Tableau 6 : résultats individuels des 36 adultes ayant passé l'expérience de complétion de phrases (expérience 1 - étude 1). Les résultats pour la condition de « prosodie informative » (18 sujets) se trouvent en haut de la page, ceux de la « prosodie neutre » (18 sujets) en bas de la page. Chaque sujet entendait 10 phrases ambiguës adjectif et 10 phrases ambiguës verbe ; pour chacune, il pouvait donner une réponse adjectif ou une réponse verbe ; le nombre maximum de réponses dans chaque case est donc de 10. Un sujet qui a bien réussi la tâche demandée donne plus de réponses adjectif que de réponses verbe sur les phrases ambiguës adjectif, et plus de réponses verbe que de réponses adjectif sur les phrases ambiguës verbe.

### A.2.3 Résultats individuels par items

	<b>PROSODIE INFORMATIVE</b>			
	Phrases Ambiguës Adjectif		Phrases Ambiguës Verbe	
	Réponses Adjectif	Réponses Verbe	Réponses Adjectif	Réponses Verbe
Item 1	4	5	0	1
Item 2	9	0	4	5
Item 3	0	9	1	7
Item 4	9	0	2	7
Item 5	9	0	2	3
Item 6	9	0	3	6
Item 7	6	3	0	9
Item 8	9	0	5	3
Item 9	9	0	3	4
Item 10	9	0	9	0
Item 11	0	9	0	9
Item 12	6	3	1	8
Item 13	9	0	7	2
Item 14	9	0	2	6
Item 15	8	1	1	8
Item 16	5	4	0	9
Item 17	8	1	4	5
Item 18	2	6	0	9
Item 19	3	6	0	9
Item 20	0	9	0	9
<b>Moyenne</b>	<b>6,15</b>	<b>2,80</b>	<b>2,20</b>	<b>5,95</b>

	<b>PROSODIE NEUTRE</b>			
	Phrases Ambiguës Adjectif		Phrases Ambiguës Verbe	
	Réponses Adjectif	Réponses Verbe	Réponses Adjectif	Réponses Verbe
Item 1	0	4	1	3
Item 2	9	0	2	7
Item 3	2	7	0	9
Item 4	5	4	8	1
Item 5	4	2	3	5
Item 6	8	1	4	5
Item 7	2	7	0	9
Item 8	9	0	5	2
Item 9	9	0	4	2
Item 10	9	0	9	0
Item 11	0	9	0	8
Item 12	5	4	3	6
Item 13	8	1	9	0
Item 14	8	1	3	5
Item 15	4	5	2	7
Item 16	2	7	7	2
Item 17	2	7	5	4
Item 18	1	8	0	9
Item 19	2	7	0	9
Item 20	0	9	0	9
<b>Moyenne</b>	<b>4,45</b>	<b>4,15</b>	<b>3,25</b>	<b>5,10</b>

Tableau 7 : résultats individuels des 20 items ambigus de l'expérience de complétion de phrases (expérience 1 - étude 1). Les résultats pour la condition de « prosodie informative » (18 sujets) se trouvent en haut de la page, ceux de la « prosodie neutre » (18 sujets) en bas de la page. Chaque phrase de chaque item était entendue par la moitié des sujets : le nombre maximum de réponses dans chaque case est donc de 9. Un item bien réussi obtient un nombre élevé de réponses adjectif sur sa phrase ambiguë adjectif et un nombre élevé de réponses verbe sur sa phrase ambiguë verbe.

## A.3 EXPERIENCE 2 : DETECTION DE MOT ABSTRAIT

### A.3.1 Consignes

Bonjour !

Merci de participer à une expérience de perception de la parole.

Cela vous prendra une petite demi-heure et ne vous réclamera rien de plus qu'un peu d'attention.

L'expérience se déroulera en deux parties de 15 minutes chacune ; entre les deux, vous pourrez prendre quelques instants de détente.

Vous allez entendre une série de phrases, à raison d'une toutes les 5 ou 6 secondes. Précédant chacune de ces phrases, un mot sera affiché au centre de l'écran.

Ce mot sera soit un verbe (ex : « manger »), soit un nom (ex : « une fille »), soit un adjectif (ex : « il est gentil »).

Dès que vous entendrez ce mot dans la phrase qui suit, vous devrez appuyer sur le bouton de réponse. Attention, il peut apparaître sous différentes formes ; par exemple, pour « manger », il faudra répondre si vous entendez « la petite fille mange des bonbons » ; pour « il est gentil », il faudra répondre si vous entendez « j'espère que vous serez gentils avec votre grand-mère ».

Certaines phrases ne contiennent pas le mot, dans ce cas vous ne faites rien et vous attendez l'essai suivant.

Nous mesurons votre temps de réaction, c'est pourquoi il faut que vous répondiez **AUSSI VITE** que possible, sans attendre la fin des phrases, et surtout sans faire d'erreurs.

Vous allez maintenant faire un entraînement d'une quinzaine de phrases.

Durant cet entraînement, l'ordinateur corrigera chacune de vos réponses, et vous indiquera votre temps de réaction en millisecondes.

Avant de commencer, vous devez avoir le casque sur la tête, et la main sur le bouton de réponse (à droite si vous êtes droitier, à gauche si vous êtes gaucher)

### A.3.2 Résultats individuels par sujets

#### A.3.2.1 Analyses sur les temps de réaction, oublis et fausses alarmes

	PROSODIE INFORMATIVE											
	Phrases ambiguës Adjectif			Phrases Ambiguës Verbe			Phrases Contrôles Adjectif			Phrases Contrôles Verbe		
	RT	Miss	FA	RT	Miss	FA	RT	Miss	FA	RT	Miss	FA
Sujet 01	1182	10	50	699	0	80	521	30	50	546	0	60
Sujet 02	636	0	60	624	0	60	488	10	50	639	0	20
Sujet 03	363	0	70	419	40	70	257	10	50	450	40	50
Sujet 04	430	20	60	582	20	60	342	10	80	456	0	50
Sujet 05	575	0	70	654	0	30	483	20	40	481	0	40
Sujet 06	482	10	40	341	0	60	528	20	40	338	0	30
Sujet 07	800	10	30	721	10	10	558	20	0	372	0	0
Sujet 08	469	0	50	587	0	40	521	0	50	438	0	20
Sujet 09	362	10	60	466	0	30	359	10	40	449	0	60
Sujet 10	366	10	40	363	0	30	431	30	0	324	0	10
Sujet 11	297	30	50	322	20	50	268	30	20	340	10	30
Sujet 12	987	0	40	543	0	30	462	0	60	605	0	40
Sujet 13	321	10	30	454	10	30	316	30	0	400	0	0
Sujet 14	480	20	40	561	10	40	187	20	20	298	0	20
Sujet 15	702	20	70	776	0	20	479	0	0	503	0	10
Sujet 16	386	0	80	509	0	50	348	10	10	301	0	30
Sujet 17	427	0	60	660	0	50	561	0	10	573	10	20
Sujet 18	421	0	60	388	0	40	390	20	0	439	0	10
Sujet 19	545	30	90	641	0	60	556	20	70	496	0	20
Sujet 20	646	10	50	596	0	50	735	0	0	481	0	30
Sujet 21	672	0	80	544	0	60	453	20	0	508	0	50
Sujet 22	501	0	90	543	10	50	365	20	40	414	0	40
Sujet 23	751	10	60	757	10	60	430	10	40	493	0	20
Sujet 24	771	20	70	669	0	50	313	40	10	493	10	50
Sujet 25	831	20	90	673	0	60	804	0	20	609	0	40
Sujet 26	806	20	70	746	0	40	536	30	10	437	0	30
Sujet 27	515	10	40	515	0	50	451	0	10	475	0	0
Sujet 28	444	0	50	572	0	50	513	0	0	539	0	40
<b>Moyenne</b>	<b>577</b>	<b>10</b>	<b>59</b>	<b>569</b>	<b>5</b>	<b>47</b>	<b>452</b>	<b>15</b>	<b>26</b>	<b>461</b>	<b>3</b>	<b>29</b>

	PROSODIE NEUTRE											
	Phrases ambiguës Adjectif			Phrases Ambiguës Verbe			Phrases Contrôles Adjectif			Phrases Contrôles Verbe		
	RT	Miss	FA	RT	Miss	FA	RT	Miss	FA	RT	Miss	FA
Sujet 01	790	0	30	748	0	50	514	0	20	494	0	20
Sujet 02	622	0	40	526	20	30	268	0	0	363	0	30
Sujet 03	976	40	80	497	20	90	470	20	80	525	20	50
Sujet 04	399	0	70	345	10	80	282	10	50	347	0	50
Sujet 05	1074	0	40	911	0	60	570	10	60	509	10	30
Sujet 06	965	10	50	1042	10	50	694	0	30	690	0	20
Sujet 07	1403	0	20	1073	0	20	910	0	20	919	0	0
Sujet 08	539	0	60	475	0	60	334	20	20	351	0	20
Sujet 09	755	10	0	835	0	20	667	0	0	765	0	0
Sujet 10	881	10	100	1322	20	80	815	10	40	716	0	70
Sujet 11	678	0	90	582	0	90	430	10	80	727	10	90
Sujet 12	462	10	70	641	10	100	285	10	70	384	0	60
Sujet 13	425	10	60	399	0	60	288	10	70	603	10	20
Sujet 14	936	10	50	998	0	20	870	0	20	531	0	0
Sujet 15	1082	0	80	785	0	30	996	10	10	748	0	20
Sujet 16	359	20	30	439	0	20	415	10	10	450	0	30
Sujet 17	420	20	40	414	0	10	253	0	40	350	20	30
Sujet 18	786	10	50	1093	0	40	554	20	10	547	0	30
Sujet 19	897	0	60	1011	0	30	850	0	10	837	0	20
Sujet 20	1099	10	90	779	0	90	438	20	40	504	0	60
Sujet 21	609	10	80	657	10	50	385	10	50	452	0	20
Sujet 22	442	10	100	649	0	100	459	10	80	437	10	90
Sujet 23	515	30	70	536	10	30	466	30	0	400	10	10
Sujet 24	777	10	80	1078	10	80	409	10	70	499	0	40
Sujet 25	850	10	50	666	0	10	644	0	0	471	0	10
Sujet 26	432	0	50	634	0	50	564	0	10	484	0	30
Sujet 27	760	0	90	898	10	90	496	0	70	626	0	40
Sujet 28	806	20	40	1027	0	40	678	0	20	757	0	30
<b>Moyenne</b>	<b>741</b>	<b>9</b>	<b>60</b>	<b>752</b>	<b>5</b>	<b>53</b>	<b>536</b>	<b>8</b>	<b>35</b>	<b>553</b>	<b>3</b>	<b>33</b>

**Tableau 8 : résultats individuels des 28 sujets ayant passé l'expérience de détection de mot abstrait (expérience 2 - étude 1) dans la condition de « prosodie informative » en haut, et la condition de « prosodie neutre » en bas.**

### A.3.2.2 Analyses sur les réponses données avant la fin du mot ambigu

	PROSODIE INFORMATIVE			
	Phrases Ambiguës Adjectif		Phrases Ambiguës Verbe	
	Réponses Adjectif	Réponses Verbe	Réponses Adjectif	Réponses Verbe
Sujet 01	2	5	1	2
Sujet 02	5	4	1	4
Sujet 03	9	3	8	4
Sujet 04	7	5	0	5
Sujet 05	7	5	2	5
Sujet 06	7	5	2	8
Sujet 07	2	3	2	1
Sujet 08	8	6	2	5
Sujet 09	9	7	5	6
Sujet 10	8	4	6	7
Sujet 11	8	5	4	7
Sujet 12	4	5	3	6
Sujet 13	9	6	4	7
Sujet 14	6	5	4	6
Sujet 15	5	2	1	3
Sujet 16	7	9	5	7
Sujet 17	6	2	2	1
Sujet 18	8	6	2	7
Sujet 19	5	1	0	3
Sujet 20	3	5	4	5
Sujet 21	6	8	4	4
Sujet 22	7	4	4	4
Sujet 23	6	3	1	2
Sujet 24	5	7	3	7
Sujet 25	3	2	0	4
Sujet 26	3	2	0	4
Sujet 27	7	3	3	6
Sujet 28	7	2	4	5
<b>Moyenne</b>	<b>6,04</b>	<b>4,43</b>	<b>2,75</b>	<b>4,82</b>

	PROSODIE NEUTRE			
	Phrases Ambiguës Adjectif		Phrases Ambiguës Verbe	
	Réponses Adjectif	Réponses Verbe	Réponses Adjectif	Réponses Verbe
Sujet 01	3	2	2	3
Sujet 02	5	5	5	5
Sujet 03	2	6	1	5
Sujet 04	8	8	9	7
Sujet 05	1	3	1	1
Sujet 06	2	0	1	2
Sujet 07	0	1	0	0
Sujet 08	7	6	3	6
Sujet 09	2	1	2	2
Sujet 10	2	2	0	1
Sujet 11	6	4	4	5
Sujet 12	7	6	7	3
Sujet 13	6	6	5	7
Sujet 14	2	2	0	2
Sujet 15	1	1	0	0
Sujet 16	7	5	6	8
Sujet 17	5	5	8	7
Sujet 18	2	3	3	3
Sujet 19	1	1	0	3
Sujet 20	2	2	2	5
Sujet 21	3	3	2	1
Sujet 22	6	6	6	4
Sujet 23	4	5	1	5
Sujet 24	3	1	0	0
Sujet 25	4	5	2	5
Sujet 26	7	5	5	3
Sujet 27	2	3	1	1
Sujet 28	1	1	2	0
<b>Moyenne</b>	<b>3,61</b>	<b>3,50</b>	<b>2,79</b>	<b>3,36</b>

Tableau 9 : résultats individuels des 28 adultes ayant passé l'expérience de détection de mot abstrait (expérience 2 - étude 1) dans la condition de prosodie informative (en haut) et dans la condition de prosodie neutre. Les réponses analysées sont celles données jusqu'à 150 ms après la fin du mot ambigu. Dans chaque prosodie, chaque sujet entendait 20 phrases ambiguës adjectif (10 en version coïncidente, 10 en version non-coïncidente) et 20 phrases ambiguës verbe (10 en version coïncidente, 10 en version non-coïncidente). Le score maximum dans chaque case est de 10.

### A.3.3 Résultats individuels par items

#### A.3.3.1 Analyses sur les temps de réaction, oublis et fausses alarmes

PROSODIE INFORMATIVE												
	Phrases ambiguës Adjectif			Phrases Ambiguës Verbe			Phrases Contrôles Adjectif			Phrases Contrôles Verbe		
	RT	Miss	FA	RT	Miss	FA	RT	Miss	FA	RT	Miss	FA
Item 1	470	7	29	990	7	29	505	29	0	361	7	7
Item 2	711	14	36	461	0	14	987	43	21	401	7	21
Item 3	837	14	86	342	0	36	636	57	29	707	0	36
Item 4	464	0	71	501	0	36	280	7	21	475	0	14
Item 5	574	0	50	689	7	64	358	0	14	361	0	7
Item 6	629	7	64	469	0	50	452	14	21	365	7	14
Item 7	465	0	79	475	0	21	289	0	7	561	7	7
Item 8	431	7	14	581	7	79	221	0	0	390	7	21
Item 9	407	7	0	579	0	0	1450	71	0	387	0	21
Item 10	451	7	29	582	43	71	451	21	21	501	0	21
Item 11	454	71	100	453	0	21	450	0	21	460	7	21
Item 12	426	7	57	898	0	14	326	0	43	506	0	14
Item 13	324	0	43	523	0	71	435	0	14	645	0	36
Item 14	660	0	43	665	0	64	642	14	21	537	7	57
Item 15	635	7	93	412	0	71	469	7	57	335	0	57
Item 16	522	7	57	547	0	71	311	7	71	470	0	36
Item 17	704	0	64	711	7	64	542	0	7	370	0	43
Item 18	830	7	93	440	0	79	623	7	64	423	0	50
Item 19	616	7	79	705	0	71	583	14	50	460	0	50
Item 20	930	21	93	441	21	7	258	0	29	486	0	50

PROSODIE NEUTRE												
	Phrases ambiguës Adjectif			Phrases Ambiguës Verbe			Phrases Contrôles Adjectif			Phrases Contrôles Verbe		
	RT	Miss	FA	RT	Miss	FA	RT	Miss	FA	RT	Miss	FA
Item 1	1038	29	43	1258	14	36	785	7	21	499	0	0
Item 2	502	0	57	847	0	57	749	14	21	451	7	14
Item 3	1010	14	86	430	0	29	640	21	21	563	0	36
Item 4	591	0	71	546	0	29	404	0	36	698	0	14
Item 5	894	0	36	871	0	29	370	0	36	608	0	21
Item 6	471	0	71	784	0	79	698	21	36	415	0	14
Item 7	440	7	57	522	0	64	410	7	14	578	0	14
Item 8	525	7	29	727	7	50	180	7	0	437	0	29
Item 9	726	7	29	655	7	43	994	36	0	671	14	14
Item 10	677	0	29	908	36	86	338	0	29	705	7	36
Item 11	1635	57	71	649	14	29	488	7	21	408	0	21
Item 12	798	7	50	1025	0	36	536	0	36	545	7	50
Item 13	612	0	43	983	0	50	443	14	50	712	0	43
Item 14	1076	14	64	649	0	64	371	14	29	646	0	50
Item 15	505	0	79	661	0	71	755	0	50	400	0	43
Item 16	595	0	64	1036	7	93	460	0	79	464	7	71
Item 17	1034	7	64	657	0	50	661	0	43	483	7	43
Item 18	838	0	100	628	7	57	912	7	57	544	7	50
Item 19	855	7	57	767	0	57	537	14	64	682	7	50
Item 20	729	21	93	548	0	50	327	0	57	607	0	43

**Tableau 10 : résultats individuels des 20 items de l'expérience de détection de mot abstrait (expérience 2 - étude 1) dans la condition de « prosodie informative » en haut, et la condition de « prosodie neutre » en bas. Pour chaque item sont indiqués son temps de réaction moyen (RT), son pourcentage moyen d'oublis (Miss) et son pourcentage moyen de fausses alarmes (FA) obtenus pour chaque catégorie de phrases expérimentales.**

### A.3.3.2 Analyses sur les réponses données avant la fin du mot ambigu

	<b>PROSODIE INFORMATIVE</b>			
	Phrases Ambiguës Adjectif		Phrases Ambiguës Verbe	
	Réponses Adjectif	Réponses Verbe	Réponses Adjectif	Réponses Verbe
Item 1	9	2	3	1
Item 2	5	3	1	3
Item 3	7	9	2	12
Item 4	10	9	3	8
Item 5	6	3	5	4
Item 6	6	4	4	9
Item 7	12	9	2	9
Item 8	8	0	6	5
Item 9	10	0	0	2
Item 10	10	2	4	4
Item 11	3	6	3	10
Item 12	12	7	2	4
Item 13	13	5	9	8
Item 14	9	4	5	4
Item 15	8	12	8	13
Item 16	8	6	3	5
Item 17	8	8	7	6
Item 18	10	13	2	11
Item 19	10	9	7	7
Item 20	5	13	1	10

	<b>PROSODIE NEUTRE</b>			
	Phrases Ambiguës Adjectif		Phrases Ambiguës Verbe	
	Réponses Adjectif	Réponses Verbe	Réponses Adjectif	Réponses Verbe
Item 1	3	2	2	0
Item 2	3	4	5	1
Item 3	3	7	2	10
Item 4	7	9	4	8
Item 5	3	0	2	3
Item 6	8	4	5	4
Item 7	8	6	4	5
Item 8	3	2	1	1
Item 9	4	1	0	2
Item 10	6	1	6	5
Item 11	0	3	3	6
Item 12	3	6	2	5
Item 13	6	5	5	3
Item 14	4	6	5	6
Item 15	10	8	6	5
Item 16	8	3	8	1
Item 17	4	7	4	7
Item 18	7	8	3	7
Item 19	4	5	6	6
Item 20	7	11	5	9

**Tableau 11 : résultats individuels des 20 items de l'expérience de détection de mot abstrait (expérience 2 - étude 1) dans la condition de prosodie informative (en haut) et dans la condition de prosodie neutre.** Les réponses analysées sont celles données jusqu'à 150 ms après la fin du mot ambigu. Dans chaque prosodie, chaque phrase de chaque item était traitée par la moitié des sujets : le score maximum dans chaque case est de 14.

## ANNEXES B RELATIVES A L'ETUDE 2 (CHAPITRE 3)

### B.1 STIMULI

Pour chaque item ambigu, écrit en lettres capitales, deux phrases expérimentales ont été construites : une phrase ambiguë verbe (indiquée en premier) et une phrase ambiguë adjectif (indiquée en deuxième). Toutes ces phrases ont été coupées à la fin du mot ambigu avant d'être présentées aux participants pour une expérience de complétion de phrases.

#### Item Phrases

- 1 Je trouve que la fumée DANSE joliment en s'échappant du feu de bois.
- 1 Je trouve que la fumée DENSE qui s'échappe de ce bâtiment laisse imaginer le pire.
  
- 2 J'ai vraiment l'impression que les pommes DURENT plus longtemps que les bananes.
- 2 J'ai vraiment l'impression que les pommes DURES font de meilleures tartes que les golden.
  
- 3 La jeune femme LACE les souliers de son petit garçon.
- 3 La jeune femme LASSE lit un livre en attendant de reprendre des forces.
  
- 4 Je crois que cet homme LACHE son boulot parce qu'il est trop stressant.
- 4 Je crois que cet homme LACHE refuse de voir la vérité en face.
  
- 5 J'ai appris que cet homme LOUCHE depuis qu'il a eu un accident de voiture.
- 5 J'ai appris que cet homme LOUCHE doit comparaître devant la justice.
  
- 6 Le petit chien MORD la laisse qui le retient dans l'espoir de se libérer.
- 6 Le petit chien MORT sera enterré demain dans le jardin de ses maîtres.
  
- 7 Le petit tambourin ROMPT l'harmonie de la musique.
- 7 Le petit tambourin ROND a une très bonne sonorité.
  
- 8 Elle trouve que les enfants SALENT beaucoup trop leur repas.
- 8 Elle trouve que les enfants SALES font la honte de leurs parents.
  
- 9 Ce politicien INTEGRE les minorités dans son projet de développement urbain.
- 9 Ce politicien INTEGRE sera certainement élu aux prochaines législatives.
  
- 10 Ce grand écrivain CELEBRE la naissance de Victor Hugo avec son nouveau livre.
- 10 Ce grand écrivain CELEBRE fera une dédicace demain dans une grande librairie.
  
- 11 Manger cette tartelette COMPLETE délicieusement ce repas.
- 11 Manger cette tartelette COMPLETE ne me va pas car je n'ai plus faim.
  
- 12 Ces petites lumières DIFFUSENT d'agréables sensations de calme.
- 12 Ces petites lumières DIFFUSENT ne fatiguent pas les yeux.
  
- 13 Ce petit clown DISTRAIT les enfants malades dans les hôpitaux.
- 13 Ce petit clown DISTRAIT n'a pas vu qu'il avait oublié son nez rouge.
  
- 14 Ce président ILLUSTRE la réussite sociale et professionnelle.
- 14 Ce président ILLUSTRE lit très peu de romans contemporains.
  
- 15 Cette belle femme CAPTIVE l'attention du public.
- 15 Cette belle femme CAPTIVE craint pour sa vie.

- 16 J'ai appris que les adolescents MUAIENT de plus en plus précocement.  
16 J'ai appris que les adolescents MUETS désiraient souvent devenir interprètes en langues des signes.
- 17 Je trouve que mes amies PERCENT rapidement dans le milieu du mannequinat.  
17 Je trouve que mes amies PERSES reçoivent une éducation trop stricte.
- 18 Sa petite amie TAILLE des rosiers dans le jardin.  
18 Sa petite amie THAI découvre Paris pour la première fois.
- 19 On sait que les hommes PARENT s'échapper de ce camp de concentration.  
19 On sait que les hommes PURS sont très rares dans le milieu de la politique.
- 20 Je vous dis que ce numéro PERD toujours à la roulette.  
20 Je vous dis que ce numéro PAIR terminera le tirage du loto de ce soir.
- 21 Il pense que ce grand tableau VIOLAIT délibérément les principes des Impressionnistes .  
21 Il pense que ce grand tableau VIOLET définit parfaitement l'idéal de l'art contemporain.
- 22 Il a peur que ses amis COULENT pendant la tempête.  
22 Il a peur que ses amis COOL puissent gâcher son repas d'affaires.

## B.2 CONSIGNES

Bonjour !

Merci de participer à une expérience de perception de la parole.

Cela vous prendra environ 20 minutes et ne vous réclamera rien de plus qu'un peu d'attention et d'imagination.

L'expérience se déroulera en deux parties de 10 minutes environ chacune ; entre les deux vous pourrez prendre quelques instants de détente.

Vous allez entendre le début d'une phrase. Votre tâche sera d'écrire le début de cette phrases et d'inventer une suite possible. Par exemple, si vous entendez « la petite fille... », vous pouvez répondre : « la petite fille va à l'école en bus ». Les phrases sont toutes coupées à la fin d'un mot (pas au milieu).

Vous écrirez la phrase entière sur la feuille qui vous est fournie.

Pour réécouter le début de la phrase en cours, appuyez sur la barre d'espace ; pour passer à la phrase suivante (après avoir répondu), appuyez sur « Entrée ».

Vous allez maintenant faire un entraînement de deux phrases pour vous familiariser avec la procédure.

Avant de commencer, vous devez avoir le casque sur la tête et disposer d'un crayon et de la feuille de réponse.

## B.3 RESULTATS INDIVIDUELS PAR SUJETS

### B.3.1 Résultats par sujets tous locuteurs confondus

	Phrases Ambiguës Adjectif		Phrases Ambiguës Verbe	
	Réponses Adjectif	Réponses Verbe	Réponses Adjectif	Réponses Verbe
Sujet 01	10	0	2	8
Sujet 02	7	3	5	4
Sujet 03	9	1	2	7
Sujet 04	6	4	7	1
Sujet 05	8	2	2	8
Sujet 06	9	1	7	3
Sujet 07	7	3	2	6
Sujet 08	9	1	6	4
Sujet 09	8	2	2	8
Sujet 10	8	2	6	4
Sujet 11	9	1	5	5
Sujet 12	8	2	1	7
Sujet 13	2	8	1	9
Sujet 14	7	3	2	8
Sujet 15	10	0	2	7
Sujet 16	10	0	3	7
Sujet 17	8	2	5	5
Sujet 18	6	3	3	7
Sujet 19	6	4	0	10
Sujet 20	9	1	2	8
Sujet 21	10	0	1	9
Sujet 22	3	7	3	6
Sujet 23	6	4	0	9
Sujet 24	10	0	4	6
Sujet 25	7	3	3	7
Sujet 26	6	4	2	8
Sujet 27	4	6	2	7
Sujet 28	10	0	1	9
Sujet 29	8	2	1	9
Sujet 30	10	0	2	8
Sujet 31	7	1	2	5
Sujet 32	9	1	2	8
Sujet 33	7	2	1	7
Sujet 34	8	2	1	5
Sujet 35	10	0	2	8
Sujet 36	10	0	1	7
Sujet 37	5	5	5	5
Sujet 38	5	4	3	6
Sujet 39	8	2	4	4
Sujet 40	9	1	1	7
Sujet 41	9	1	3	7
Sujet 42	8	2	1	3
Sujet 43	7	3	4	4
Sujet 44	10	0	5	4
Sujet 45	6	4	6	4
Sujet 46	10	0	2	5
Sujet 47	10	0	4	6
Sujet 48	8	2	7	2
Sujet 49	10	0	5	5
Sujet 50	9	1	3	6
Sujet 51	8	1	6	4
Sujet 52	4	6	5	5
Sujet 53	9	1	3	7
Sujet 54	3	7	4	6
Sujet 55	7	3	6	4
Sujet 56	6	4	2	7
Sujet 57	9	1	3	5
Sujet 58	6	4	3	7
Sujet 59	8	2	5	5
Sujet 60	6	4	3	6
<b>Moyenne</b>	<b>7,7</b>	<b>2,2</b>	<b>3,1</b>	<b>6,1</b>

**Tableau 12 : résultats individuels des 60 sujets ayant passé l'expérience de complétion de phrases dans laquelle les phrases étaient enregistrées par des locuteurs naïfs (étude 2).** Les sujets entendaient 10 phrases ambiguës adjectif et 10 phrases ambiguës verbe (score maximum pour chaque phrase = 10)

B.3.2 Résultats par sujets et par locuteurs

	LOCUTEUR 1			
	Phrases Ambiguës Adjectif		Phrases Ambiguës Verbes	
	Réponses Adjectif	Réponses Verbe	Réponses Adjectif	Réponses Verbe
Sujet 01	10	0	2	8
Sujet 02	9	1	2	7
Sujet 03	8	2	2	8
Sujet 04	7	3	2	6
Sujet 05	8	2	2	8
Sujet 06	7	3	5	4
Sujet 07	6	4	7	1
Sujet 08	9	1	7	3
Sujet 09	9	1	6	4
Sujet 10	8	2	6	4
<b>Moyenne</b>	<b>8,1</b>	<b>1,9</b>	<b>4,1</b>	<b>5,3</b>

	LOCUTEUR 2			
	Phrases Ambiguës Adjectif		Phrases Ambiguës Verbes	
	Réponses Adjectif	Réponses Verbe	Réponses Adjectif	Réponses Verbe
Sujet 01	9	1	5	5
Sujet 02	2	8	1	9
Sujet 03	10	0	2	7
Sujet 04	8	2	5	5
Sujet 05	6	4	0	10
Sujet 06	8	2	1	7
Sujet 07	7	3	2	8
Sujet 08	10	0	3	7
Sujet 09	6	3	3	7
Sujet 10	9	1	2	8
<b>Moyenne</b>	<b>7,5</b>	<b>2,4</b>	<b>2,4</b>	<b>7,3</b>

	LOCUTEUR 3			
	Phrases Ambiguës Adjectif		Phrases Ambiguës Verbes	
	Réponses Adjectif	Réponses Verbe	Réponses Adjectif	Réponses Verbe
Sujet 01	10	0	1	9
Sujet 02	6	4	0	9
Sujet 03	7	3	3	7
Sujet 04	4	6	2	7
Sujet 05	8	2	1	9
Sujet 06	3	7	3	6
Sujet 07	10	0	4	6
Sujet 08	6	4	2	8
Sujet 09	10	0	1	9
Sujet 10	10	0	2	8
<b>Moyenne</b>	<b>7,4</b>	<b>2,6</b>	<b>1,9</b>	<b>7,8</b>

<b>LOCUTEUR 4</b>				
Phrases Ambiguës Adjectif		Phrases Ambiguës Verbes		
Réponses Adjectif	Réponses Verbe	Réponses Adjectif	Réponses Verbe	
Sujet 01	7	1	2	5
Sujet 02	7	2	1	7
Sujet 03	10	0	2	8
Sujet 04	5	5	5	5
Sujet 05	8	2	4	4
Sujet 06	9	1	2	8
Sujet 07	8	2	1	5
Sujet 08	10	0	1	7
Sujet 09	5	4	3	6
Sujet 10	9	1	1	7
<b>Moyenne</b>	<b>7,8</b>	<b>1,8</b>	<b>2,2</b>	<b>6,2</b>

<b>LOCUTEUR 5</b>				
Phrases Ambiguës Adjectif		Phrases Ambiguës Verbes		
Réponses Adjectif	Réponses Verbe	Réponses Adjectif	Réponses Verbe	
Sujet 01	9	1	3	7
Sujet 02	6	4	6	4
Sujet 03	10	0	4	6
Sujet 04	10	0	5	5
Sujet 05	8	2	1	3
Sujet 06	7	3	4	4
Sujet 07	10	0	5	4
Sujet 08	10	0	2	5
Sujet 09	8	2	7	2
Sujet 10	9	1	3	6
<b>Moyenne</b>	<b>8,7</b>	<b>1,3</b>	<b>4</b>	<b>4,6</b>

<b>LOCUTEUR 6</b>				
Phrases Ambiguës Adjectif		Phrases Ambiguës Verbes		
Réponses Adjectif	Réponses Verbe	Réponses Adjectif	Réponses Verbe	
Sujet 01	8	1	6	4
Sujet 02	9	1	3	7
Sujet 03	7	3	6	4
Sujet 04	9	1	3	5
Sujet 05	8	2	5	5
Sujet 06	4	6	5	5
Sujet 07	3	7	4	6
Sujet 08	6	4	2	7
Sujet 09	6	4	3	7
Sujet 10	6	4	3	6
<b>Moyenne</b>	<b>6,6</b>	<b>3,3</b>	<b>4</b>	<b>5,6</b>

**Tableau 13 : résultats individuels des 60 sujets (10 sujets par locuteurs) ayant passé l'expérience de complétion de phrases dans laquelle les phrases étaient enregistrées par des locuteurs naïfs (étude 2).** Les sujets entendaient 10 phrases ambiguës adjectif et 10 phrases ambiguës verbe (score maximum pour chaque phrase = 10)

## B.4 RESULTATS INDIVIDUELS PAR ITEMS

### B.4.1 Résultats par items tous locuteurs confondus

Les résultats des items 21 et 22 sont indiqués en grisé dans le tableau : ce sont les deux items qui, en moyenne, ont montré les plus forts biais de réponse et qui ont été retirés des analyses.

	Phrases Ambiguës Adjectif		Phrases Ambiguës Verbe	
	Réponses Adjectif	Réponses Verbe	Réponses Adjectif	Réponses Verbe
Item 1	22	8	14	14
Item 2	20	9	4	23
Item 3	18	10	14	11
Item 4	22	8	12	16
Item 5	25	5	16	13
Item 6	22	7	2	24
Item 7	27	2	20	8
Item 8	29	1	11	11
Item 9	23	7	10	20
Item 10	25	3	12	18
Item 11	23	7	5	24
Item 12	20	10	1	29
Item 13	23	7	8	22
Item 14	17	13	8	22
Item 15	24	5	15	15
Item 16	28	2	2	28
Item 17	24	5	7	21
Item 18	21	8	4	22
Item 19	23	5	11	16
Item 20	27	3	8	19
Item 21	29	1	22	6
Item 22	2	28	0	30
<b>Moyenne</b>	<b>23,2</b>	<b>6,3</b>	<b>9,2</b>	<b>18,8</b>

**Tableau 14 : Résultats individuels des 20 items de l'expérience de complétion de phrases dans laquelle les stimuli étaient enregistrés par des locuteurs naïfs (étude 2).** Chaque phrase de chaque item était entendue par la moitié des participants (score maximum pour chaque phrase = 30).

B.4.2 Résultats par items et par locuteurs

	LOCUTEUR 1			
	Phrases Ambiguës Adjectif		Phrases Ambiguës Verbes	
	Réponses Adjectif	Réponses Verbe	Réponses Adjectif	Réponses Verbe
Item 1	3	2	3	1
Item 2	5	0	2	3
Item 3	1	4	4	0
Item 4	4	1	2	3
Item 5	2	3	1	3
Item 6	5	0	0	4
Item 7	5	0	5	0
Item 8	5	0	0	4
Item 9	5	0	3	2
Item 10	5	0	4	1
Item 11	4	1	0	5
Item 12	3	2	0	5
Item 13	5	0	2	3
Item 14	3	2	0	5
Item 15	3	2	4	1
Item 16	5	0	1	4
Item 17	5	0	3	1
Item 18	4	1	2	3
Item 19	4	1	4	1
Item 20	5	0	1	4
Item 21	5	0	4	1
Item 22	0	5	0	5
<b>Moyenne</b>	<b>4,1</b>	<b>1,0</b>	<b>2,1</b>	<b>2,7</b>

	LOCUTEUR 2			
	Phrases Ambiguës Adjectif		Phrases Ambiguës Verbes	
	Réponses Adjectif	Réponses Verbe	Réponses Adjectif	Réponses Verbe
Item 1	4	1	2	3
Item 2	3	2	1	4
Item 3	4	1	1	3
Item 4	3	1	1	4
Item 5	5	0	2	3
Item 6	2	3	1	4
Item 7	4	1	3	2
Item 8	5	0	1	3
Item 9	3	2	1	4
Item 10	5	0	0	5
Item 11	5	0	1	4
Item 12	4	1	0	5
Item 13	5	0	2	3
Item 14	1	4	1	4
Item 15	3	2	1	4
Item 16	4	1	0	5
Item 17	4	1	0	5
Item 18	3	2	0	4
Item 19	4	1	3	2
Item 20	4	1	3	2
Item 21	5	0	4	1
Item 22	0	5	0	5
<b>Moyenne</b>	<b>3,8</b>	<b>1,2</b>	<b>1,2</b>	<b>3,7</b>

	<b>LOCUTEUR 3</b>			
	Phrases Ambiguës Adjectif		Phrases Ambiguës Verbes	
	Réponses Adjectif	Réponses Verbe	Réponses Adjectif	Réponses Verbe
Item 1	4	1	2	3
Item 2	4	1	0	5
Item 3	2	3	2	2
Item 4	3	2	1	4
Item 5	4	1	1	4
Item 6	2	3	1	4
Item 7	5	0	3	2
Item 8	5	0	2	1
Item 9	3	2	1	4
Item 10	3	2	1	4
Item 11	5	0	0	5
Item 12	3	2	0	5
Item 13	4	1	0	5
Item 14	3	2	0	5
Item 15	5	0	1	4
Item 16	4	1	0	5
Item 17	3	2	2	3
Item 18	4	1	0	5
Item 19	4	1	1	4
Item 20	4	1	1	4
Item 21	5	0	3	2
Item 22	1	4	0	5
<b>Moyenne</b>	<b>3,7</b>	<b>1,3</b>	<b>1,0</b>	<b>3,9</b>

	<b>LOCUTEUR 4</b>			
	Phrases Ambiguës Adjectif		Phrases Ambiguës Verbes	
	Réponses Adjectif	Réponses Verbe	Réponses Adjectif	Réponses Verbe
Item 1	4	1	2	2
Item 2	2	3	0	2
Item 3	4	0	2	2
Item 4	5	0	1	3
Item 5	4	1	5	0
Item 6	4	1	0	4
Item 7	5	0	3	1
Item 8	5	0	1	1
Item 9	2	3	0	5
Item 10	3	0	0	5
Item 11	3	2	2	2
Item 12	5	0	0	5
Item 13	3	2	1	4
Item 14	3	2	0	5
Item 15	4	1	1	3
Item 16	5	0	1	4
Item 17	4	1	1	3
Item 18	4	1	1	4
Item 19	5	0	0	4
Item 20	4	0	1	3
Item 21	5	0	5	0
Item 22	1	4	0	5
<b>Moyenne</b>	<b>3,9</b>	<b>0,9</b>	<b>1,1</b>	<b>3,1</b>

LOCUTEUR 5				
	Phrases Ambiguës Adjectif		Phrases Ambiguës Verbes	
	Réponses Adjectif	Réponses Verbe	Réponses Adjectif	Réponses Verbe
Item 1	4	1	1	4
Item 2	2	2	1	4
Item 3	4	0	3	1
Item 4	5	0	4	1
Item 5	5	0	5	0
Item 6	4	0	0	3
Item 7	3	1	2	2
Item 8	5	0	5	0
Item 9	5	0	3	2
Item 10	4	1	5	0
Item 11	5	0	1	4
Item 12	4	1	0	5
Item 13	4	1	2	3
Item 14	5	0	2	3
Item 15	5	0	3	2
Item 16	5	0	1	4
Item 17	4	0	0	5
Item 18	3	1	1	1
Item 19	3	1	2	3
Item 20	5	0	2	3
Item 21	5	0	2	1
Item 22	0	5	0	5
<b>Moyenne</b>	<b>4,2</b>	<b>0,5</b>	<b>2,2</b>	<b>2,5</b>

LOCUTEUR 6				
	Phrases Ambiguës Adjectif		Phrases Ambiguës Verbes	
	Réponses Adjectif	Réponses Verbe	Réponses Adjectif	Réponses Verbe
Item 1	3	2	4	1
Item 2	4	1	0	5
Item 3	3	2	2	3
Item 4	1	4	4	1
Item 5	4	1	3	2
Item 6	5	0	0	5
Item 7	5	0	4	1
Item 8	3	2	3	1
Item 9	5	0	2	3
Item 10	4	1	3	2
Item 11	1	4	1	4
Item 12	1	4	1	4
Item 13	2	3	1	4
Item 14	2	3	5	0
Item 15	5	0	4	1
Item 16	4	1	0	5
Item 17	4	1	1	4
Item 18	3	2	0	5
Item 19	3	1	1	2
Item 20	4	1	1	3
Item 21	3	2	5	0
Item 22	0	5	0	5
<b>Moyenne</b>	<b>3,3</b>	<b>1,7</b>	<b>2,0</b>	<b>2,8</b>

Tableau 15 : résultats individuels des 20 items de l'expérience de complétion de phrases, dans laquelle les stimuli étaient enregistrés par 6 locuteurs naïfs (étude 2). Les résultats sont présentés pour chaque locuteur. Chaque phrase de chaque item était entendue par la moitié des participants (score maximum pour chaque phrase =5)

## ANNEXES C RELATIVES A L'ETUDE 3 (CHAPITRE 5)

### C.1 STIMULI

Les stimuli utilisés dans l'étude 3 (expérience avec les enfants français de 16 mois sur le rôle des frontières prosodiques dans la segmentation lexicale) étaient créés par paires : chaque phrase bisyllabique (contenant le mot bisyllabique) était appariée avec une phrase monosyllabique (contenant les deux syllabes cibles séparées par une frontière de groupe phonologique) au niveau du nombre de syllabes et de la structure prosodique.

Paires de phrases « balcon / bal » :

- |    |  |
|----|--|
| 1  | De gracieux <i>balcons</i> courent le long des belles façades.                   |
| 1  | Les immenses <i>bals</i> concourent à la grandeur du roi.                        |
| 2  | Ce spécialiste de grilles de <i>balcons</i> fait fureur auprès des gens.         |
| 2  | Ce spécialiste de robes de <i>bal</i> confectionne aussi des smokings.           |
| 3  | La rangée de <i>balcons</i> fait face au cloître du monastère.                   |
| 3  | La grande salle de <i>bal</i> confère un air solennel au château.                |
| 4  | La fillette qui va au <i>balcon</i> tourne au bord de la table.                  |
| 4  | Le chemin qui mène au <i>bal</i> contourne la tour écroulée.                     |
| 5  | Les portes-fenêtres déglinguées des <i>balcons</i> travaillent par temps humide. |
| 5  | Les plafonds de l'ancienne salle de <i>bal</i> contrastent avec tous les autres. |
| 6  | Les belles enluminures du <i>balcon</i> trompent les spécialistes.               |
| 6  | Les organisateurs du <i>bal</i> contrôlent bien les entrées.                     |
| 7  | Les portes qui mènent au <i>balcon</i> voisinent avec les toilettes.             |
| 7  | Les gens qui donnent des <i>bals</i> convoitent les orchestres renommés.         |
| 8  | Tous les enfants au <i>balcon</i> versent de l'eau en bas.                       |
| 8  | Tous les accès au <i>bal</i> convergent vers le podium.                          |
| 9  | Un grand <i>balcon</i> venait d'être détruit.                                    |
| 9  | Un grand <i>bal</i> convenait à tout le monde.                                   |
| 10 | Le vieux balcon s'appuyait sur le chêne.   |
| 10 | Ce grand <i>bal</i> consacra leur union.   |
| 11 | Le joli <i>balcon</i> cloisonnait la terrasse.                                   |
| 11 | Le dernier <i>bal</i> conclura la saison.  |
| 12 | Les décorations du <i>balcon</i> sauvaient la façade.                            |
| 12 | Le plaisir d'aller au <i>bal</i> consola la jeune fille.                         |

Paires de phrases « vipère / vie » :

- 1 La jolie *vipère* subjuguait les enfants qui l'observaient.  
1 Son étrange *vie persuada* ses amis de sa folie.
- 2 Il pensait que les vieux costumes de *vipères* se louaient bien.  
2 Je prédis que ce nouveau mode de *vie percera* bientôt.
- 3 La plus répandue des *vipères* siffle tout le temps.  
3 Ceux qui en veulent à sa *vie persiflent* en permanence.
- 4 Le sorcier qui parle aux *vipères* faiblissait chaque jour.  
4 Les médecins qui sauvent des *vies perfectionnent* leur technique.
- 5 La proie de la *vipère* sommeillait au soleil.  
5 La femme de sa *vie personnifiait* la beauté.
- 6 Une grosse *vipère* pénétra dans l'entrée.  
6 Une longue *vie perpétue* les traditions.
- 7 L'énorme *vipère* tua la belle souris.  
7 Sa drôle de *vie perturba* son entourage.
- 8 La plus affreuse des *vipères* s'irritait très facilement.  
8 L'espoir d'une meilleure *vie persiste* chez les optimistes.
- 9 Le remède à base de *vipère* soignera bien des maladies.  
9 La grande campagne pour la *vie perçoit* des aides gouvernementales.
- 10 La petite *vipère* mettait des plumes dans son nid.  
10 Une meilleure *vie permettra* leur épanouissement.
- 11 Même le venin des *vipères* servait aux pharmaciens.  
11 Les amis pour la *vie persévèrent* jusqu'à la fin.
- 12 La peau de la *vipère* disparaissait sous les feuilles.  
12 L'histoire de sa *vie perdit* de son charme à leurs yeux.

## C.2 RESULTATS INDIVIDUELS PAR SUJETS

		Pourcentage d'orientation de la tête du groupe bisyllabique		
		Phrases bisyllabiques (hits)	Phrases monosyllabiques (FA)	a'
Condition "balcon"	Sujet 01	25,0	36,4	0,384
	Sujet 02	54,5	25,0	0,734
	Sujet 03	75,0	33,3	0,795
	Sujet 04	45,5	33,3	0,612
	Sujet 05	36,4	27,3	0,594
	Sujet 06	66,7	8,3	0,878
	Sujet 07	90,9	75,0	0,703
	Sujet 08	75,0	8,3	0,904
	Sujet 09	50,0	25,0	0,708
	Sujet 10	41,7	0,0	0,854
Condition "vipère"	Sujet 01	25,0	0,0	0,813
	Sujet 02	33,3	8,3	0,756
	Sujet 03	45,5	9,1	0,800
	Sujet 04	27,3	0,0	0,818
	Sujet 05	58,3	27,3	0,740
	Sujet 06	41,7	8,3	0,791
	Sujet 07	33,3	25,0	0,590
	Sujet 08	36,4	16,7	0,695
	Sujet 09	33,3	50,0	0,354
	Sujet 10	60,0	18,2	0,802
<b>Moyenne</b>		<b>47,7</b>	<b>21,7</b>	<b>0,716</b>

Tableau 16 (résultats de l'étude 3 détaillés par sujets) : Pourcentage moyen d'orientation de la tête de chaque enfant du groupe syllabique sur les phrases bisyllabiques (succès) et sur les phrases monosyllabiques (fausses alarmes). Les enfants habitués à répondre au mot « balcon » sont indiqués dans la moitié supérieure, ceux habitués à répondre au mot « vipère » dans la moitié inférieure. Une valeur du a' supérieure ou égale à .755 représente une performance significativement supérieure au hasard.

		Pourcentage d'orientation de la tête du groupe monosyllabique		
		Phrases bisyllabiques (FA)	Phrases monosyllabiques (hits)	a'
Condition "bal"	Sujet 01	50,0	75,0	0,708
	Sujet 02	33,3	16,7	0,325
	Sujet 03	16,7	66,7	0,838
	Sujet 04	8,3	50,0	0,822
	Sujet 05	27,3	25,0	0,472
	Sujet 06	41,7	58,3	0,643
	Sujet 07	25,0	40,0	0,644
	Sujet 08	25,0	58,3	0,754
Condition "vie"	Sujet 01	18,2	36,4	0,681
	Sujet 02	10,0	36,4	0,754
	Sujet 03	9,1	40,0	0,778
	Sujet 04	25,0	27,3	0,528
	Sujet 05	41,7	45,5	0,537
	Sujet 06	16,7	27,3	0,629
	Sujet 07	50,0	54,5	0,544
	Sujet 08	33,3	50,0	0,646
<b>Moyenne</b>		<b>27,0</b>	<b>44,2</b>	<b>0,644</b>

Tableau 17 (résultats de l'étude 3 détaillés par sujets) : Pourcentage moyen d'orientation de la tête de chaque enfant du groupe monosyllabique sur les phrases bisyllabiques (fausses alarmes) et sur les phrases monosyllabiques (succès). Les enfants habitués à répondre au mot « bal » sont indiqués dans la moitié supérieure, ceux habitués à répondre au mot « vie » dans la moitié inférieure. Une valeur du a' supérieure ou égale à .755 représente une performance significativement supérieure au hasard.

### C.3 RESULTATS INDIVIDUELS PAR ITEMS

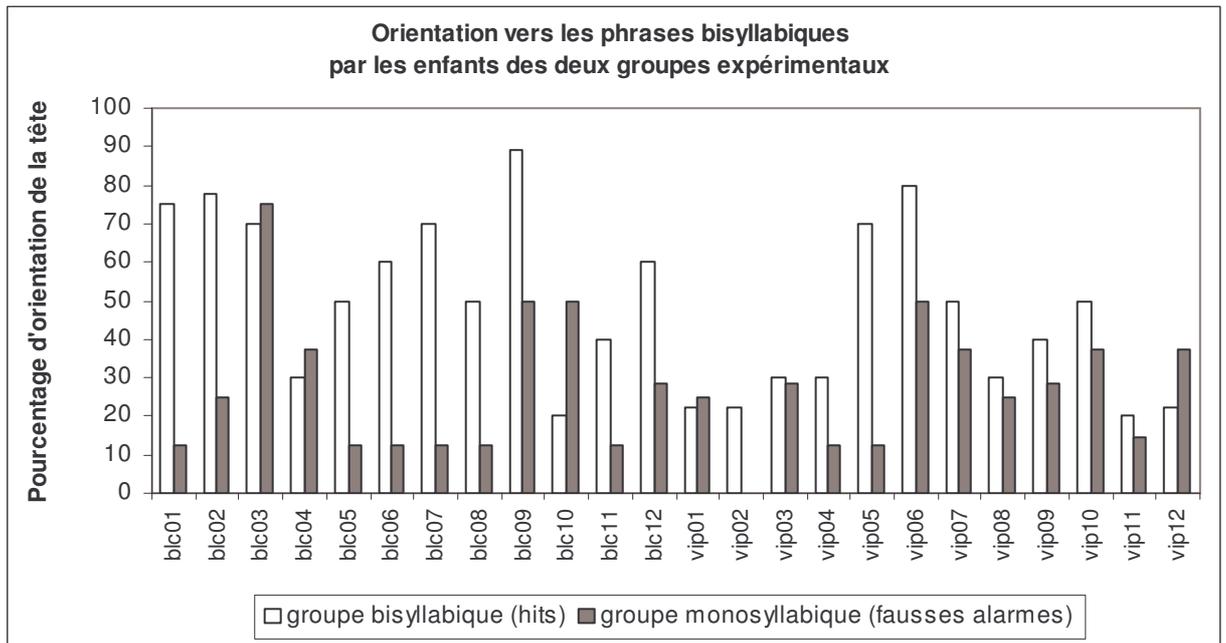


Figure 16 (résultats de l'étude 3 détaillés par items) : Pourcentage moyen d'orientation de la tête vers les phrases bisyllabiques : les réponses des enfants du groupe bisyllabique (en blanc) correspondent à des succès (« hits »), alors que les réponses des enfants du groupe monosyllabique (en gris) correspondent à des fausses alarmes.

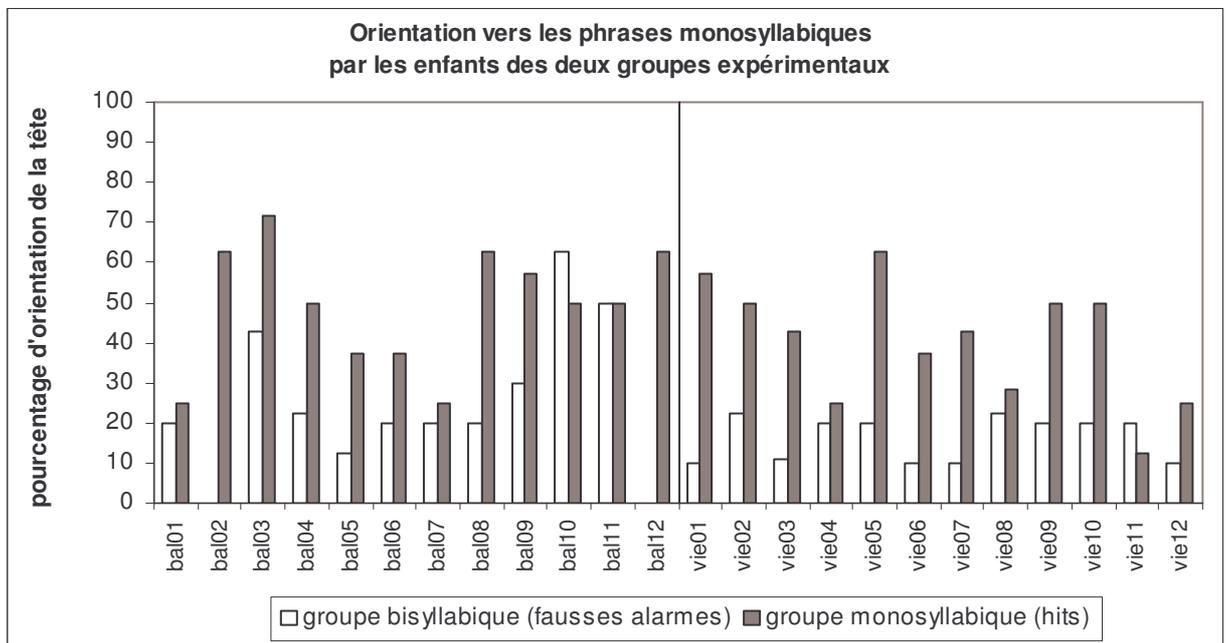


Figure 17 (résultats de l'étude 3 détaillés par item) : Pourcentage moyen d'orientation de la tête vers les phrases monosyllabiques : les réponses des enfants du groupe bisyllabique (en blanc) correspondent à des fausses alarmes, alors que les réponses des enfants du groupe monosyllabique (en gris) correspondent à des succès (« hits »).

## ANNEXES D RELATIVES A L'ETUDE 4 (CHAPITRE 6)

### D.1 STIMULI

Pour chaque item ambigu, écrit en gras, deux phrases expérimentales ont été construites : une phrase nom (indiquée en premier) et une phrase verbe (indiquée en deuxième). Chacune de ces phrases avait un équivalent en français (pour aider à l'enregistrement des stimuli) apparié sur le nombre de syllabes et la structure syntaxique (cette traduction est indiquée sous chacune des phrases). Les frontières de groupes phonologiques sont représentées par des crochets.

Condition « avec clitique » :

Dans cette condition, les mots cibles sont directement précédés d'un mot grammatical (un pronom avant une cible verbe et un article avant une cible nom).

*Sous-condition 1 : le couple « clitique – mot cible » forme une unité prosodique à lui tout seul*

- 1 [Un **piluche**] [doupe dans la palude]  
[Un cheval] [court dans la prairie]
- 1 [Je **piluche**] [tra les pimes] [de trotir ce gadale]  
[Je prévois] [tous les jours] [de finir ce travail]
- 2 [Un **jigoulate**] [s'émunait du ti]  
[Un dromadaire] [s'échappait du zoo]
- 2 [On **jigoulate**] [no s'adiber]  
[On applaudit] [sans s'arrêter]
- 3 [Un **pardale**] [ga tachin proquire]  
[Un cadeau] [fait toujours plaisir]
- 3 [Elle **pardale**] [tru les sbimes] [de grabifouner]  
[Elle promet] [toutes les semaines] [de téléphoner]
- 4 [Une **bamoule**] [dri se froliter dagou]  
[Une expo] [doit se dérouler demain]
- 4 [Tu **bamoules**] [saman ti] [à mon ada]  
[Tu travailles] [souvent trop] [à mon avis]
- 5 [Une **doripe**] [talou poulire] [à mon cabin]  
[Une guitare] [ferait plaisir] [à mon cousin]
- 5 [On **doripe**] [tra les mounes] [la mopine à niler]  
[On répare] [tous les mois] [la machine à laver]
- 6 [Une **goufe**] [pifait d'oguer] [dans le darchin]  
[Une porte] [permet d'aller] [dans le jardin]
- 6 [Elle **goufe**] [dapou ce mitu]  
[Elle pleure] [depuis ce matin]
- 7 [Des **falares**] [choument dans les jotrées]  
[Des renards] [chassent dans les forêts]
- 7 [Elle **falare**] [tra les omules] [avec sopi]  
[Elle franchit] [tous les obstacles] [avec succès]

- 8 [Des **bradoles**] [ralouchent le tro] [du brolimier]  
 [Des coupoles] [recouvrent le toit] [du bâtiment]
- 8 [Il **bradole**] [trachan ce sporale]  
 [Il rigole] [devant ce spectacle]
- 9 [Il racho prito] [des **drilapes**] [pour amuter ses lourtes]  
 [Il cherche partout] [des cartons] [pour emballer ses livres]
- 9 [Pour litou] [je **drilape**] [atrou ramilaton] [la bidunon]  
 [Pour l'instant] [je recherche] [avec détermination] [la solution]
- 10 [Il folichait] [la **lourate**] [rolite et furche]  
 [Il recherchait] [la voiture] [rapide et fiable]
- 10 [Milaton] [je **lourate**] [stire dans ce liquo]  
 [Maintenant] [je travaille] [seule dans ce labo]
- 11 [Nous avons ruchiné] [la **dramiche**] [dabilant la kef]  
 [Nous avons regardé] [la corniche] [surplombant la mer]
- 11 [Ton nurdo faliné] [il **dramiche**] [vutan pave]  
 [Ton nouveau fiancé] [il cuisine] [vraiment peu]
- 12 [Ma livurne] [étilé de fure] [la **mopare**] [crume sa michon]  
 [Ma cousine] [essaie de faire] [la compote] [comme sa maman]
- 12 [C'est larfant indiloume] [tu **mopares**] [tru tes fatideurs]  
 [C'est vraiment incroyable] [tu épates] [tous tes professeurs]
- 13 [J'aidrais blire] [le **maligoure**] [dagoler du birchon]  
 [J'aimerais voir] [le cuisinier] [rechercher du travail]
- 13 [Tes fridons] [ils **maligourent**] [dropiment] [une lature]  
 [Tes enfants] [ils manigacent] [secrètement] [une bêtise]
- 14 [Tu ribaleras] [le **trine**] [jiporant de te nutafer]  
 [Tu amèneras] [le câble] [permettant de te remorquer]
- 14 [La mitu frale] [elle **trine**] [tupe les nades] [qu'elle go briter]  
 [La petite fille] [elle rêve] [toutes les nuits] [qu'elle peut voler]
- 15 [Tu roliras] [les **natoules**] [munés dans la crouge]  
 [Tu trouveras] [les moineaux] [cachés dans la cage]
- 15 [Ces fires manives] [elles **natoulent**] [apou plumine ave d'atrine]  
 [Ces jeunes touristes] [elles repartent] [après plusieurs heures d'attente]
- 16 [Ils inviloutent] [les **pamilabes**] [jilant dans la bruche]  
 [Ils interdisent] [les téléphones] [sonnant dans le train]
- 16 [Ce virno grichon] [il **pamilabe**] [atra joule] [sa lapotine]  
 [Ce petit garçon] [il revendique] [avec force] [sa liberté]

*Sous-condition 2 : le couple « clitique – mot cible » se trouve au début d'une unité prosodique (mais le mot cible n'est pas directement suivi par une frontière prosodique)*

- 1 [Un **piluche** abline] [gaille sur la mastone]  
 [Un cheval alesan] [court sur la montagne]
- 1 [Je **piluche** balti] [tra ma fortille] [par drolimane]  
 [Je préviens souvent] [toute ma famille] [par téléphone]

- 2 [Un **jigoulate** invirunant] [fra mive les drulateurs]  
 [Un dromadaire intelligent] [fait rire les visiteurs]
- 2 [On **jigoulate** arumilement] [son piri atou]  
 [On interroge amicalement] [son petit ami]
- 3 [Un **pardale** gatovinale] [pirna à mon pruche]  
 [Un cadeau original] [plaira à mon père]
- 3 [Elle **pardale** rabivomant] [de vadir nous clure]  
 [Elle promet régulièrement] [de venir nous voir]
- 4 [Une **bamoule** ferdine] [pritora des zures meldunes]  
 [Une expo nouvelle] [présentera des œuvres modernes]
- 4 [Tu **bamoules** mirdan] [jali tru la sti]  
 [Tu travailles souvent] [jusque tard la nuit]
- 5 [Une **doripe** étrulite] [fouchait pitou] [à mion conin]  
 [Une guitare électrique] [ferait plaisir] [à mon cousin]
- 5 [On **doripe** saçant] [pour tribuver] [atrou les migues tadonnes]  
 [On revient souvent] [pour travailler] [avec les mêmes personnes]
- 6 [Une **goufe** luchée] [prato d'aguer] [dans le cirdin]  
 [Une porte vitrée] [permet d'aller] [dans le jardin]
- 6 [Elle **goufe** lapran] [arti de s'agounir]  
 [Elle pleure longtemps] [avant de s'endormir]
- 7 [Des **falares** saltons] [chounent dans les gorpées]  
 [Des renards sauvages] [chassent dans les forêts]
- 7 [Elle **falare** grumon] [tra les émites]  
 [Elle franchit gaiement] [toutes les étapes]
- 8 [Des **bradoles** amilées] [ralichent le tru]  
 [Des coupoles argentées] [recouvrent le toit]
- 8 [Il **bradole** caligant] [ful padore s'amirter]  
 [Il rigole carrément] [sans pouvoir s'arrêter]
- 9 [Il richoule] [des **drilapes** amoules] [raloutant à l'uche]  
 [Il recherche] [des cartons épais] [résistant à l'eau]
- 9 [Marégadant] [je **drilape** tigou] [mave le stu]  
 [Dorénavant] [je travaille beaucoup] [même le soir]
- 10 [Il guivait arder] [la **lourate** bane] [pour son gati]  
 [Il voulait acheter] [la voiture rouge] [pour son amie]
- 10 [Marinon] [je **lourate** muche] [dans ce quito]  
 [Maintenant] [je travaille mieux] [dans ce bureau]
- 11 [Elle cripait toguer] [la **dramiche** adomate] [pour son virdale]  
 [Elle voulait trouver] [la tenue adéquate] [pour son voyage]
- 11 [Ton nachou faliné] [il **dramiche** dalimant] [le prichon aux amutes]  
 [Ton nouveau fiancé] [il cuisine divinement] [le poulet aux amandes]
- 12 [Elle ardile] [la **mopare** navirelle] [cra son éplité]  
 [Elle achète] [la compote naturelle] [chez son épicier]
- 12 [C'est vrami délabine] [tu **mopares** targou] [les jimes que tu da]  
 [C'est vraiment étonnant] [tu épates toujours] [les gens que tu vois]

- 13 [Il a fergalé] [le **maligoure** jirunal] [partin son vunage]  
 [Il a engagé] [le cuisinier japonais] [pendant son voyage]
- 13 [Ces furants] [ils **maligourent** entrane] [une briluse à frale]  
 [Ces enfants] [ils manigacent ensemble] [une bêtise à faire]
- 14 [Je miroune] [le **trine** jafulin] [garmé à la midorémi]  
 [Je recherche] [le film japonais] [primé à la cérémonie]
- 14 [La martu fale] [elle **trine** otamant] [sa prouchon]  
 [La petite fille] [elle vit fortement] [sa passion]
- 15 [J'alogue murner] [les **natoules** bartins] [au proli galuné]  
 [J'adore manger] [les gateaux bretons] [au petit déjeuner]
- 15 [Ces funes trichons] [elles **natoulent** endon] [arpo des aves d'atoche]  
 [Ces jeunes touristes] [elles repartent enfin] [après des heures d'attente]
- 16 [Ils nigatinent] [les **pamilabes** nalades] [perfa le vrou]  
 [Ils interdisent] [les téléphones portables] [pendant le vol]
- 16 [Ce lato garbin] [il **pamilabe** saige] [des frichules amilétones]  
 [Ce petit garçon] [il exécute seul] [des figures acrobatiques]

*Sous-condition 3 : le couple « clitique – mot cible » est au milieu d'une unité prosodique*

- 1 [J'aigrait un **piluche** trupé]  
 [J'aimerais un chaton tigré]
- 1 [Sabant je **piluche** tru] [un muche à l'atri]  
 [Souvent je piluche tout] [un mois à l'avance]
- 2 [J'ai tru un **jigoulate** churton] [dans ce sapoume]  
 [J'ai vu un acrobate chinois] [dans ce spectacle]
- 2 [Il foune qu'on **jigoulate** bame] [sa ramigachilon]  
 [Il faut qu'on applaudisse fort] [sa représentation]
- 3 [J'ai un **pardale** ogranilé] [pour ma fougé]  
 [J'ai un cadeau original] [pour ma soeur]
- 3 [Mirtan elle **pardale** valou] [de ne pas radibouler]  
 [Pourtant elle promet souvent] [de ne pas recommencer]
- 4 [Tu as une **bamoule** rube] [pour goder dimare]  
 [Tu as une cagoule rouge] [pour aller dehors]
- 4 [Croume tu **bamoules** bien] [ce nachi calune]  
 [Comme tu portes bien] [ce petit tailleur]
- 5 [J'ai tra une **doripe** nuse] [dans le fromé]  
 [J'ai vu une souris grise] [dans le grenier]
- 5 [C'est dume qu'on **doripe** tachou] [de lui guirouter le badole]  
 [C'est vrai qu'on essaie toujours] [de lui remonter le moral]
- 6 [Je mide une **goufe** pirlée]  
 [Je veux une douche carrelée]
- 6 [Dabou elle **goufe** sadilousement]  
 [D'ailleurs elle pleure silencieusement]

- 7 [Prin des **falares** chimes] [abou l'irdé]  
[Prends des habits chauds] [avant l'hiver]
- 7 [Il pune qu'elle **falare** endou] [de se pachiner à ce fabou]  
[Il faut qu'elle essaie encore] [de se présenter à ce concours]
- 8 [Il a des **bradoles** bines] [pour nire ses dabons]  
[Il a des cahiers blancs] [pour faire ses dessins]
- 8 [Croume il **bradole** puche] [padou nos sadées]  
[Comme il rigole bien] [pendant nos soirées]
- 9 [Dadin] [j'odrai des **drilapes** sabis] [sur la quiron]  
[Demain] [j'aurais des avis sérieux] [sur la question]
- 9 [Il capite] [quand je **drilape** gaminant] [les fauges]  
[Il accepte] [quand je demande gentiment] [les choses]
- 10 [Je cra bi] [que la **lourate** hitronale] [se trabine épon]  
[Je crois bien] [que la saison hivernale] [se termine enfin]
- 10 [Je vous tru] [que je **lourate** bougue] [en ce paton]  
[Je vous dis] [que je travaille dur] [en ce moment]
- 11 [Je pache que la **dramiche** bloume] [da mite refoute]  
[Je pense que la peinture blanche] [doit être refaite]
- 11 [Pardu] [je dase qu'il **dramiche** touve] [le goine à la michon]  
[Parfois] [je veux qu'il cuisine plus] [le soir à la maison]
- 12 [Dabno] [je crube la **mopare** fite] [pour adinoler ce pline]  
[Maintenant] [je prends la moutarde forte] [pour accompagner ce plat]
- 12 [Je sipe late] [car tu **mopares** vapant] [tru tes radipeurs]  
[Je suis fière] [car tu épates vraiment] [tous tes professeurs]
- 13 [Elle trache] [que le **maligoure** igolin] [a rafougé le rédinant]  
[Elle trouve] [que le cuisinier italien] [a relancé le restaurant]
- 13 [C'est gouve] [mais ils **maligourent** adou] [leur pétille d'agoune]  
[C'est drôle] [mais ils reconnaissent encore] [leur village d'enfance]
- 14 [Amopi] [j'avin le **trine** calique] [de ma patice]  
[Aujourd'hui] [j'atteinds le seuil critique] [de ma patience]
- 14 [C'est abilant qu'elle **trine** angale] [de dapiner gabeuse édole]  
[C'est amusant qu'elle rêve encore] [de devenir danseuse étoile]
- 15 [Dapou des junes] [elle ragoude le **natoule** nouche] [clu en patibon]  
[Depuis des jours] [elle demande le futon noir] [vu en promotion]
- 15 [Je sipe infate] [car tu **natoules** batou] [dapi kaso ti]  
[Je suis inquiète] [car tu faiblis beaucoup] [depuis quelque temps]
- 16 [Il est crouge] [que les **pamilabes** rodals] [se vagent bine]  
[Il est clair] [que les téléphones portables] [se vendent bien]
- 16 [Il est pso] [donc il **pamilabe** ripe] [notre capotaline]  
[Il est psy] [donc il interprète bien] [notre comportement]

Condition « sans clitique » :

Les mots cibles ne sont pas directement précédés par un mot grammatical

*Sous-condition 4 :*

- 1 [Un gorabou **piluche**] [crave dans la fragole]  
[Un minuscule cheval] [court dans la prairie]
- 1 [Un gorabou] [**piluche** patilament] [son enfou du tra]  
[Un diplomate] [prévoit précisément] [son emploi du temps]
- 2 [Ce raduqua **jigoulate**] [agline dans le dasine]  
[Ce ridicule dromadaire] [avance dans le désert]
- 2 [Ce raduqua] [**jigoulate** rafichement] [la vro de sa mope]  
[Ce garnement] [reconnait parfaitement] [la voix de sa mère]
- 3 [Un gouminet **pirdale**] [agoche mon atrulon]  
[Un incroyable cadeau] [attire mon attention]
- 3 [Un gouminet] [**pirdale** tigou] [d'aigo soujé]  
[Un étudiant] [promet toujours] [d'être sérieux]
- 4 [Sa cramona **bamoule**] [camiche dabou]  
[Sa formidable expo] [commence demain]
- 4 [Sa cramona] [**bamoule** muche] [le mirtou]  
[Sa camarade] [travaille mieux] [le matin]
- 5 [Une michorbi **doripe**] [truche dans le dabon]  
[Une magnifique télé] [trône dans le salon]
- 5 [Une michorbi] [**doripe** adou] [le rafore du moune]  
[Une demoiselle] [détient encore] [le record du monde]
- 6 [La jaminu **goufe**] [anope son adilon]  
[La magnifique fille] [attire son attention]
- 6 [La jaminu] [**goufe** sapi] [en rupant]  
[La journaliste] [pleure souvent] [en rentrant]
- 7 [Des maloupets **falares**] [cravent leur titou]  
[Des magnifiques renards] [creusent leur tanière]
- 7 [Des maloupets] [**falarent** au badu de la priche]  
[Des acrobates] [voltigent au dessus de la piste]
- 8 [Ces invilados **bradoles**] [daquent se voune falasement]  
[Ces incroyables terrains] [doivent se vendre facilement]
- 8 [Ces invilados] [**bradolent** bine] [défo ce fouche]  
[Ces adolescents] [rigolent bien] [devant ce film]
- 9 [J'ai l'oubiton] [que ces palifu **drilapes**] [caront l'amuche]  
[J'ai l'impression] [que ces énormes cartons] [feront l'affaire]
- 9 [J'ai l'oubiton] [que ces palifu] [**drilapent** rouchiment]  
[J'ai l'impression] [que ces ouvriers] [travaillent rapidement]
- 10 [Il amodait] [que sa miquou **lourate**] [manole dans sa dilaton]  
[Il espérait] [que sa jolie collègue] [regarde dans sa direction]
- 10 [Il amodait] [que sa miquou] [**lourate** les samineurs]  
[Il espérait] [que sa musique] [saisisse les spectateurs]

- 11 [Cet abou chida] [un mato **dramiche**] [valite notre épule]  
[Cet après-midi] [un acteur célèbre] [visite notre école]
- 11 [Cet abou chida] [un mato] [**dramiche** la nigale] [du cibalo]  
[Cet après-midi] [un acteur] [célèbre la naissance] [du cinéma]
- 12 [Je te cabime] [que ces malinou **mopares**] [pime du mané]  
[Je te confirme] [que ces merveilleux légumes] [viennent du marché]
- 12 [Je te cabime] [que ces malinou] [**moparent** ambinou] [notre prémi]  
[Je te confirme] [que ces diplomates] [visitent aujourd'hui] [notre pays]
- 13 [Il onvait bune] [que ces motina **maligoures**] [dartent cha zou]  
[Il aimerait bien] [que ces valeureux combattants] [rentrent chez eux]
- 13 [Il onvait bune] [que ces motina] [**maligourent** à s'inluter]  
[Il aimerait bien] [que ces artisans] [réussissent à s'implanter]
- 14 [J'espule dabant] [que cette vilucho **trine**] [vanqua bune me ramuler]  
[J'espère vraiment] [que cette merveilleuse femme] [voudra bien me conseiller]
- 14 [J'espule dabant] [que cette vilucho] [**trine** bidou à mon fuche]  
[J'espère vraiment] [que cette récompense] [plaît beaucoup à mon fils]
- 15 [Le machou] [de la péloubi **natoule**] [afon des ramuches]  
[Le montant] [de la magnifique maison] [atteint des records]
- 15 [Le machou] [de la péloubi] [**natoule** varon des sachous]  
[Le montant de la statuette] [atteint vraiment des sommets]
- 16 [J'ai draminé] [que les balines **pamilabes**] [coune fircher les pru jouilles]  
[J'ai découvert] [que les anciennes interprètes] [viennent former les plus jeunes]
- 16 [J'ai draminé] [que les balines] [**pamilabent** encru les faves]  
[J'ai découvert] [que les anciennes] [interprètent encore les rêves]

*Sous-condition 5 : présence d'une syllabe ou d'une rime ambiguë avant le mot cible  
(homophone à un pronom avant un nom et homophone à un article avant un verbe)*

- 1 [Un tra martu **piluche**] [crave dans la fragole]  
[Un tout petit cheval] [court dans la prairie]
- 1 [Un uche quarbin] [**piluche** patilament] [son enfou du tra]  
[Un homme sérieux] [prévoit précisément] [son emploi du temps]
- 2 [Un broulajeux **jigoulate**] [agline dans le dasine]  
[Un gigantesque dromadaire] [avance dans le désert]
- 2 [Le pran falin] [**jigoulate** rafichement] [la vro de sa mope]  
[Le jeune garçon] [reconnait parfaitement] [la voix de sa mère]
- 3 [Un gouratelle **pirdale**] [agoche mon atrulon]  
[Un incroyable cadeau] [attire mon attention]
- 3 [Ce blo kartin] [**pirdale** tigou] [d'aigo bire souge]  
[Ce beau garçon] [promet toujours] [d'être bien sage]
- 4 [Une cramotu **bamoule**] [camiche dabou]  
[Une incroyable expo] [commence demain]
- 4 [Ma jine chalune] [**bamoule** muche] [le mirtou]  
[Ma jeune cousine] [travaille mieux] [le matin]

- 5 [Une michoton **doripe**] [truche dans le dabon]  
 [Une magnifique télé] [trône dans le salon]
- 5 [Ce fale dalune] [**doripe** adou] [le rafore du moune]  
 [Ce jeune athlète] [détient encore] [le record du monde]
- 6 [Une tru matelle **goufe**] [métine son atrito]  
 [Une très jolie fille] [attire son attention]
- 6 [Ma priche carune] [**goufe** sapi] [paton la nuche]  
 [Ma jeune cousine] [pleure souvent] [pendant la nuit]
- 7 [De tru jabelle**s falares**] [crounent leur tamine]  
 [De très jolis renards] [creusent leur tanière]
- 7 [La jibe gadé] [**falare** ditruniment] [tra les igoles]  
 [La jeune sportive] [franchit allègrement] [toutes les étapes]
- 8 [Ces go parati**l bradoles**] [diquent se voune dabilisement]  
 [Ces très minuscules terrains] [doivent se vendre difficilement]
- 8 [Les patou zidé] [**bradolent** bune] [défo ce dabile apoti]  
 [Les petits enfants] [rigolent bien] [devant ce dessin animé]
- 9 [Il poutine tru le tine] [des prijeux drilapes] [pour enviler ses loumes]  
 [Il demande tout le temps] [des petits cartons] [pour emballer ses livres]
- 9 [C'est tru clon] [que ces pritou midé] [**drilapent** soltes] [la samiron]  
 [C'est très bien] [que ces jeunes enfants] [recherchent seuls] [la solution]
- 10 [Il ramodait] [la molijoux lourate] [dapou des multes]  
 [Il recherchait] [la meilleure voiture] [depuis des mois]
- 10 [Je sto] [que ma médile roa] [**lourate** mide] [pour ader ses rétaro]  
 [Je sais] [que ma meilleure amie] [travaille dur] [pour avoir ses examens]
- 11 [Ils ont rolé] [la moti**l dramiche**] [savilant la doufe]  
 [Ils ont aimé] [la jolie corniche] [surplombant la mer]
- 11 [Je frou] [que ton nurdo fala] [**dramiche** bitou] [dapi qu'il est pita]  
 [Je trouve] [que ton petit cousin] [cuisine beaucoup] [depuis qu'il est marié]
- 12 [J'anpais romader] [ta marotu **mopare**] [ébore une frou]  
 [J'aimerais remanger] [ta délicieuse compote] [encore une fois]
- 12 [Amoudi] [les natules ébala] [**moparent** tachou] [pour la pimuté]  
 [Aujourd'hui] [les nouvelles étudiantes] [militent aussi] [pour la parité]
- 13 [Les cafans] [truchent que le cravi**l maligoure**] [pardo des piches infilanes]  
 [Les clients] [trouvent que le nouveau cuisinier] [prépare des plats incroyables]
- 13 [C'est tru caton] [chu les firouleux] [**maligourent** sopiguement] [leurs dafitons]  
 [C'est très étrange] [tous les députés] [ratifient silencieusement] [leurs décisions]
- 14 [Nous avons fraté] [le tru badelle **trine**] [sadourant de gires puches]  
 [Nous avons acheté] [le tout nouveau câble] [supportant de gros poids]
- 14 [Dapi loune] [la foune amochileux] [**trine** dapilouement] [de le cabonne]  
 [Depuis hier] [la jeune adolescente] [tente désespérément] [de le convaincre]
- 15 [Ramivez] [les jabou pinelle**s natoules**] [granalant le milton]  
 [Ecoutez] [les jolis petits moineaux] [gazouillant le matin]
- 15 [C'est ingore] [mon calon pulé] [**natoule** bidou] [aux gapules]  
 [C'est étrange] [mon petit cousin] [parie beaucoup] [aux loteries]

- 16 [Filumant] [les chu drobi] **pamilabes** [plouguent aux cabiluteurs]  
 [Maintenant] [les très anciens téléphones] [plaisent aux collectionneurs]  
 16 [Je brouche] [que la capifulé] **pamilabe** dri [le sidure des binodes]  
 [Je trouve] [que la communauté] [perpétue bien] [le souvenir des anciens]

## D.2 CONSIGNES

Bonjour !

Merci de participer à une expérience de perception de la parole.

Cela vous prendra une vingtaine de minutes et ne vous réclamera rien de plus qu'un peu d'attention.

L'expérience se déroulera en deux parties de 10 minutes environ chacune ; entre les deux vous pourrez prendre quelques instants de détente.

Vous allez entendre une série de phrases d'un langage artificiel. La particularité de ce langage est que tous les mots grammaticaux (articles, pronoms, marques de conjugaison...) sont les mêmes qu'en français. Par contre, tous les mots de contenu (noms, verbes, adjectifs...) sont remplacés par des mots qui n'existent pas en français.

Par exemple, la phrase « la fillette joue à la poupée toute la journée » pourrait être prononcée « la manole june à la dibon trune la javé ».

De même, la phrase « hier il a mangé un gros gâteau au chocolat » pourrait être prononcée « navou il a nuré un fouche griban au valinou ».

Avant chaque phrase, un mot sera affiché au centre de l'écran. Ce mot sera soit un verbe (ex : « nurer » ou « juner »), soit un nom (ex : « une manole », « un griban »).

Dès que vous entendrez ce mot dans la phrase qui suit, vous devrez appuyer sur le bouton réponse le plus vite possible.

Certaines phrases ne contiennent pas le mot, dans ce cas vous ne faites rien et vous attendez l'essai suivant.

Vous allez maintenant faire un entraînement d'une dizaine de phrases. Durant cet entraînement, l'ordinateur corrigera chacune de vos réponses, et vous indiquera votre temps de réaction en millisecondes.

Avant de commencer, vous devez avoir le casque sur la tête, et la main sur le bouton de réponse (à droite si vous êtes droitier, à gauche si vous êtes gaucher).

## D.3 RESULTATS INDIVIDUELS PAR SUJETS

	CONDITIONS AVEC CLITIQUE											
	Sous-Condition 1				Sous-Condition 2				Sous-Condition 3			
	Phrases Nom		Phrases Verbe		Phrases Nom		Phrases Verbe		Phrases Nom		Phrases Verbe	
	Rép. Nom	Rép. Verbe	Rép. Nom	Rép. Verbe	Rép. Nom	Rép. Verbe	Rép. Nom	Rép. Verbe	Rép. Nom	Rép. Verbe	Rép. Nom	Rép. Verbe
Sujet 01	100	0	6	94	100	0	0	100	100	0	6	94
Sujet 02	94	6	6	94	75	25	6	94	75	25	6	94
Sujet 03	100	0	13	88	100	0	19	81	100	0	13	88
Sujet 04	81	19	25	75	81	19	25	75	94	6	38	63
Sujet 05	94	6	25	75	100	0	6	94	88	13	13	88
Sujet 06	88	13	13	88	88	13	13	88	81	19	13	88
Sujet 07	100	0	19	81	100	0	13	88	100	0	38	63
Sujet 08	94	6	6	94	100	0	0	100	94	6	6	94
Sujet 09	100	0	0	100	100	0	0	100	88	13	0	100
Sujet 10	94	6	19	81	94	6	6	94	100	0	6	94
Sujet 11	75	25	19	81	88	13	6	94	100	0	13	88
Sujet 12	94	6	6	94	100	0	0	100	88	13	13	88
Sujet 13	94	6	13	88	88	13	19	81	81	19	6	94
Sujet 14	100	0	19	81	94	6	6	94	94	6	0	100
Sujet 15	100	0	6	94	94	6	13	88	88	13	25	75
Sujet 16	94	6	0	100	88	13	0	100	94	6	13	88
Sujet 17	94	6	19	81	94	6	13	88	81	19	19	81
Sujet 18	100	0	6	94	100	0	6	94	100	0	6	94
Sujet 19	100	0	13	88	94	6	0	100	94	6	0	100
Sujet 20	88	13	13	88	88	13	0	100	88	13	6	94
Sujet 21	88	13	13	88	94	6	13	88	94	6	19	81
Sujet 22	94	6	13	88	88	13	19	81	88	13	13	88
Sujet 23	94	6	19	81	94	6	19	81	94	6	13	88
Sujet 24	94	6	31	69	75	25	0	100	88	13	31	69
Sujet 25	100	0	6	94	100	0	13	88	100	0	13	88
Sujet 26	94	6	13	88	94	6	6	94	100	0	13	88
Sujet 27	94	6	6	94	94	6	0	100	100	0	13	88
Sujet 28	81	19	19	81	88	13	25	75	81	19	25	75
Sujet 29	100	0	6	94	88	13	19	81	88	13	13	88
Sujet 30	100	0	13	88	100	0	19	81	100	0	19	81
Sujet 31	100	0	13	88	94	6	6	94	88	13	6	94
Sujet 32	100	0	6	94	100	0	0	100	94	0	19	81
Sujet 33	100	0	6	94	94	6	13	88	75	25	6	94
Sujet 34	94	6	13	88	100	0	0	100	81	19	13	88
Sujet 35	94	6	6	94	94	6	13	88	88	13	13	88
Sujet 36	100	0	13	88	88	13	13	88	81	19	19	81
Sujet 37	88	13	0	100	100	0	31	69	88	13	25	75
Sujet 38	100	0	6	94	94	6	19	81	100	0	25	75
Sujet 39	94	6	13	88	100	0	6	94	81	19	19	81
Sujet 40	100	0	0	100	88	13	0	100	94	6	6	94
Sujet 41	100	0	0	100	100	0	13	88	94	6	6	88
Sujet 42	88	13	13	88	94	6	13	88	94	6	0	100
Sujet 43	94	6	19	81	100	0	25	75	88	13	13	88
Sujet 44	94	6	13	88	100	0	19	81	88	13	25	75
Sujet 45	94	6	6	94	100	0	6	94	100	0	13	88
Sujet 46	94	6	6	94	100	0	6	94	100	0	6	94
Sujet 47	81	19	6	94	94	6	25	75	88	13	19	75
Sujet 48	100	0	13	88	94	6	13	88	94	6	0	100
Sujet 49	81	19	6	94	88	13	13	88	75	25	19	81
Sujet 50	94	6	0	100	94	6	6	94	100	0	19	81
Sujet 51	81	19	19	81	94	6	25	75	81	19	13	88
Sujet 52	88	13	6	94	100	0	6	94	81	19	19	81
<b>Moyenne</b>	<b>94</b>	<b>6</b>	<b>11</b>	<b>89</b>	<b>94</b>	<b>6</b>	<b>11</b>	<b>89</b>	<b>91</b>	<b>9</b>	<b>14</b>	<b>86</b>

Tableau 18 (résultats de l'étude 4 détaillés par sujets) : Pourcentage de réponses nom et verbe données par chacun des 52 sujets aux phrases nom et verbe dans les trois sous-conditions « avec clitique »

## D.4 RESULTATS INDIVIDUELS PAR ITEMS

	CONDITIONS SANS CLITIQUE							
	Sous-Condition 4				Sous-Condition 5			
	Phrases Nom		Phrases Verbe		Phrases Nom		Phrases Verbe	
	Rép. Nom	Rép. Verbe	Rép. Nom	Rép. Verbe	Rép. Nom	Rép. Verbe	Rép. Nom	Rép. Verbe
Sujet 01	50	50	0	100	38	63	0	100
Sujet 02	50	50	31	69	31	69	25	75
Sujet 03	69	31	6	94	63	38	0	100
Sujet 04	50	50	19	81	38	63	13	88
Sujet 05	63	38	13	88	56	44	6	94
Sujet 06	6	94	13	81	0	94	6	94
Sujet 07	56	44	31	69	75	25	13	88
Sujet 08	75	25	13	88	75	25	0	100
Sujet 09	56	44	13	88	50	50	0	100
Sujet 10	69	31	6	94	75	25	6	94
Sujet 11	63	38	6	88	56	44	25	75
Sujet 12	63	38	19	81	75	25	13	88
Sujet 13	25	75	6	94	44	56	25	75
Sujet 14	63	38	0	100	50	50	6	94
Sujet 15	63	38	13	88	63	38	25	75
Sujet 16	69	31	0	100	38	63	0	100
Sujet 17	38	63	13	88	38	63	25	75
Sujet 18	56	44	6	94	50	50	6	94
Sujet 19	19	81	0	100	44	50	6	94
Sujet 20	69	31	6	94	50	50	0	100
Sujet 21	31	69	6	94	50	50	0	100
Sujet 22	50	50	19	81	50	50	6	94
Sujet 23	81	19	19	81	56	38	6	94
Sujet 24	13	88	19	81	19	81	19	81
Sujet 25	50	50	0	100	31	69	0	100
Sujet 26	75	25	6	94	63	38	6	94
Sujet 27	50	50	0	100	63	38	6	94
Sujet 28	44	56	25	69	38	56	31	69
Sujet 29	50	50	0	100	63	38	0	100
Sujet 30	69	31	19	81	75	25	13	88
Sujet 31	13	88	13	88	6	94	0	100
Sujet 32	81	19	6	94	69	31	0	100
Sujet 33	50	50	6	94	63	38	13	88
Sujet 34	56	44	0	100	50	50	6	94
Sujet 35	81	19	13	88	88	13	6	94
Sujet 36	69	31	31	69	88	13	19	81
Sujet 37	50	50	31	69	50	50	25	75
Sujet 38	69	31	19	81	81	19	13	88
Sujet 39	63	38	6	94	44	56	25	75
Sujet 40	50	50	0	100	63	38	6	94
Sujet 41	25	75	0	100	31	63	13	88
Sujet 42	38	63	6	88	25	75	13	88
Sujet 43	13	88	6	94	19	81	6	94
Sujet 44	63	38	13	88	69	31	13	88
Sujet 45	0	100	6	94	25	75	0	100
Sujet 46	69	31	13	88	50	44	13	88
Sujet 47	38	63	25	75	31	69	19	69
Sujet 48	13	88	0	94	31	69	6	94
Sujet 49	56	44	0	94	50	50	25	75
Sujet 50	31	69	6	94	44	56	13	88
Sujet 51	56	44	19	81	63	38	25	75
Sujet 52	50	50	25	75	38	56	31	69
<b>Moyenne</b>	<b>50</b>	<b>50</b>	<b>11</b>	<b>88</b>	<b>50</b>	<b>49</b>	<b>11</b>	<b>89</b>

Tableau 19 (résultats de l'étude 4 détaillés par items) : Pourcentage de réponses nom et verbe aux phrases nom et verbe des 16 items, dans les trois sous-conditions « avec clitique » (en haut) et les deux sous-conditions « sans clitique » (en bas).

## ANNEXES E

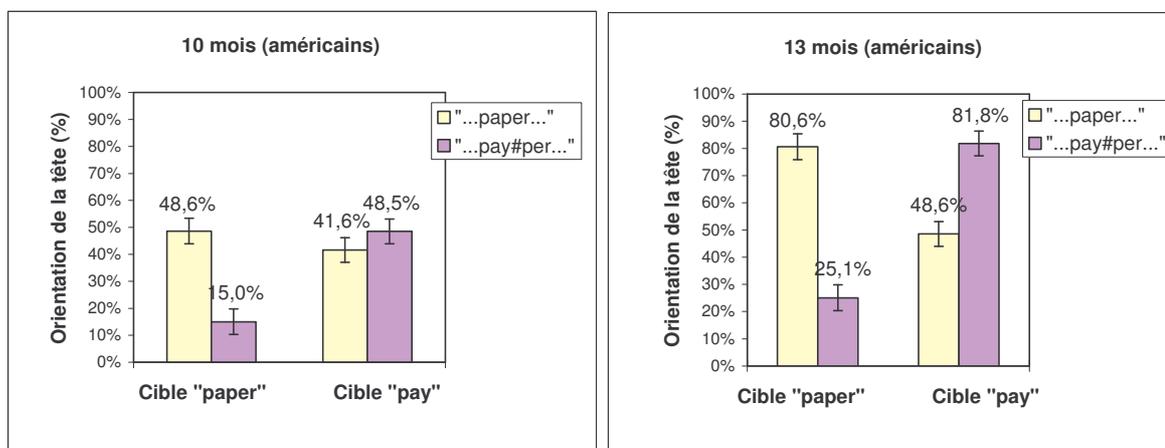
### ROLE DES FRONTIERES PROSODIQUES DANS LA SEGMENTATION LEXICALE : COMPARAISONS INTER-LANGUES ET CONSIDERATIONS METHODOLOGIQUES

Dans le chapitre 5 de cette recherche, nous avons présenté une étude (étude 3) dans laquelle nous étudions le rôle des frontières de groupes phonologiques dans la segmentation lexicale. Ceci était fait pour valider l'hypothèse selon laquelle les jeunes enfants français pouvaient percevoir en temps réel les frontières prosodiques de taille intermédiaire, et les utiliser pour réaliser une tâche linguistique, comme celle de retrouver les mots dans la parole continue. Cette étude faisait suite à une expérience réalisée par Gout, Christophe et Morgan (2004) auprès d'enfants américains de 10 et 13 mois, que nous allons présenter brièvement.

#### E.1 DONNEES AMERICAINES

Les auteurs ont utilisé une variante de la technique de conditionnement d'orientation de la tête pour tester des enfants américains de 10 et 13 mois. Dans une première session, les enfants du groupe bisyllabique étaient habitués à tourner la tête pour un mot bisyllabique, présenté en isolation (par exemple, « paper »), tandis que les enfants du groupe monosyllabique étaient habitués à répondre pour un mot monosyllabique (par exemple, « pay »). Dans une deuxième session, les enfants étaient testés sur des phrases expérimentales complètes contenant la cible bisyllabique (comme dans « the church / with the most *paper* spires / is heavenly »), et sur des phrases contenant la cible monosyllabique, c'est-à-dire que dans ce cas les deux syllabes constitutives du mot bisyllabique étaient séparées par une frontière de groupe phonologique (comme dans « the man / with the least *pay* / *perspires* constantly »). Les enfants entendaient aussi des phrases distractrices, qui ne contenaient pas les syllabes du mot cible, et pour lesquelles les auteurs s'attendaient à que les enfants ne répondent pas.

Les résultats obtenus sont représentés dans la Figure 18 ci-dessous.



**Figure 18 (résultats américains) :** Résultats des enfants américains de 10 mois (à gauche) et des enfants américains de 13 mois (à droite). Ces figures représentent le pourcentage moyen d'orientation de la tête vers les phrases bisyllabiques (« ...paper... », en blanc) et les phrases monosyllabiques (« ...pay#per... », en gris), par les enfants du groupe bisyllabique (cible « paper », barres de gauche) et les enfants du groupe monosyllabique (cible « pay », barres de droite). Les figures sont adaptées de Gout et al. (2004)

Les résultats indiquent que les enfants habitués à tourner la tête pour le mot « paper » tournent plus la tête quand ils entendent les phrases contenant vraiment le mot cible que quand ils entendent les deux syllabes du mot cible séparées par une frontière prosodique (48,6% versus 15% pour les 10 mois et 80,6% versus 25,1% pour les 13 mois). Le comportement inverse a été montré pour les enfants habitués à tourner la tête pour le mot monosyllabique : en effet, ces enfants ont plus tourné la tête quand ils entendaient les phrases contenant réellement le mot « pay » que quand ils entendaient les phrases contenant le mot « paper » (48,5% versus 41,6% pour les enfants de 10 mois, c'est-à-dire un pattern de réponses dans le sens attendu, même si cette différence n'est pas statistiquement significative, et 81,8% versus 48,6% pour les enfants de 13 mois).

Cette expérience démontre donc que les enfants américains pouvaient interpréter en temps réel les frontières de groupe phonologique comme des frontières de mot, et qu'ils pouvaient grâce à ces indices extraire les mots de la parole continue. Nous avons voulu étendre ce résultat aux enfants français en réalisant une expérience similaire auprès d'enfants de 13 mois.

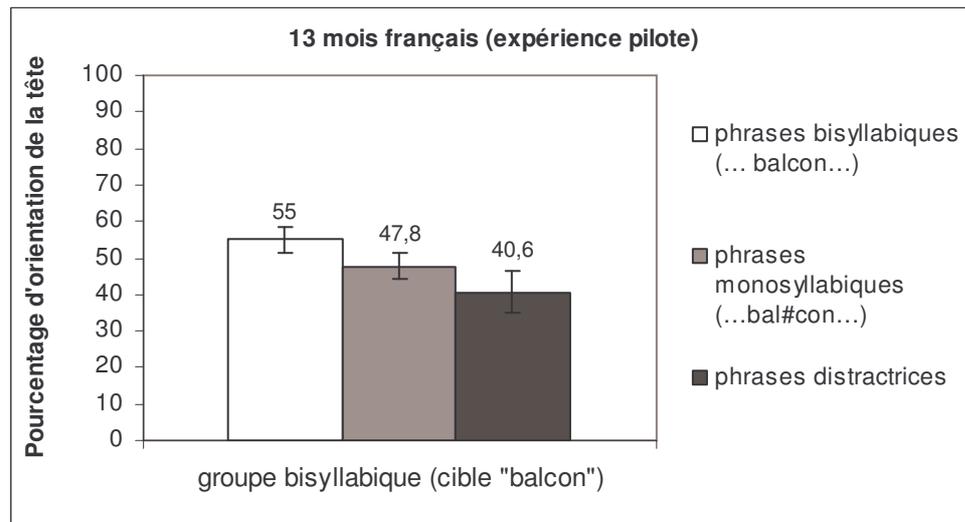
## E.2 DONNEES FRANCAISES

### E.2.1 Résultats d'une expérience pilote avec des enfants de 13mois (n=16)

Nous avons décidé de conduire une expérience similaire à celle réalisée par Gout et al. (2004), en utilisant le même protocole expérimental. Dans une première session, on habitue les enfants à répondre à un mot bisyllabique (« balcon »), puis nous les testions (dans une deuxième

session) sur des phrases expérimentales complètes : soit celles-ci contenaient réellement le mot cible (comme dans « [le joli *balcon*] [cloisonnait la terrasse] »), soit elles contenaient les deux syllabes constitutives du mot cible séparées par une frontière de groupe phonologique (comme dans « [le dernier *bal*] [*con*clura la saison] »). De plus, les enfants entendaient des phrases distractrices (comme « [le petit garçon sage] [regardait les étoiles] ») pour lesquelles on s'attendait à ce que les enfants ne réagissent jamais, puisque aucune syllabe du mot cible n'était présente.

Les résultats obtenus dans cette première expérience sont représentés dans la Figure 19 ci-dessous (seuls les enfants du groupe bisyllabique, habitués à répondre au mot bisyllabique « balcon » ont été testés).



**Figure 19 (résultats 13 mois français – expérience pilote) : pourcentage moyen d’orientation de la tête des enfants du groupe bisyllabique (habitués à répondre au mot bisyllabique « balcon ») sur les phrases bisyllabiques (« ...balcon... », en blanc), les phrases monosyllabiques (« ...bal#con... », en gris clair), et les phrases distractrices (en gris foncé).** Les barres d’erreurs représentent les erreurs standard de la moyenne.

Le résultat principal de cette étude est que les enfants habitués à tourner la tête pour un mot bisyllabique, répondent ensuite aussi souvent pour les phrases bisyllabiques (contenant réellement le mot cible) que pour les phrases monosyllabiques (contenant les deux syllabes du mot cible séparées par une frontière de groupe phonologique). En effet, la différence entre les pourcentage d’orientation de la tête sur les phrases bisyllabiques et les phrases monosyllabiques ne sont pas différents l’un de l’autre (55% versus 47.8%,  $t(15)=1.7$ ,  $p=0.1$ ). Contrairement aux enfants américains du même âge, les enfants français de 13 mois n’arrivent pas à utiliser le fait que les frontières de groupe phonologique correspondent à des frontières de mot. Ils pensent aussi souvent que le mot « balcon » est présent dans les phrases « balcon » et dans les phrases « bal ».

Ce qui était aussi particulièrement étonnant était le nombre important de regards sur les phrases distractrices (qui pourtant ne comportaient aucune ambiguïté), pourcentage qui ne diffère pas non plus significativement de celui obtenu pour les phrases bisyllabiques (55% versus 40.6%,  $t(15)=2.1$ ,  $p=.06$ ) même si la différence va dans le sens attendu.

En fait, il semble que les enfants français aient plus de difficultés, que les enfants américains du même âge, avec la tâche de segmentation en elle-même (et non pas simplement avec l’utilisation des frontières prosodiques), et qu’ils n’arrivent tout simplement pas à trouver

des mots dans les phrases. Leur comportement est identique quel que soit le type de phrase entendue, comme s'ils n'avaient pas réussi à comprendre quelle était la tâche qui leur était demandée.

Nous trouvons ce résultat étonnant, et avant de conclure que les enfants français avaient un retard dans la segmentation de la parole continue en mots, nous voulions vérifier que cet effet n'était pas dû, par exemple, à la moins bonne expertise des expérimentateurs (l'auteur avait été formé un mois à Brown University auprès de Jim Morgan, ce qui paraît relativement peu comparé aux années d'expertise des expérimentateurs américains). Nous avons donc décidé de tester des enfants français de 16 mois avec le même protocole (si l'expertise des expérimentateurs était en cause, aucun résultat significatif ne devrait apparaître à cet âge, alors qu'on s'attend à ce que les enfants français de cet âge soient capables de retrouver des mots dans des phrases).

### E.2.2 Résultats d'une expérience pilote auprès d'enfants de 16 mois (n=11)

Les mêmes stimuli et le même protocole expérimental ont été utilisés dans cette expérience. Les résultats sont indiqués dans la Figure 20 ci-dessous.

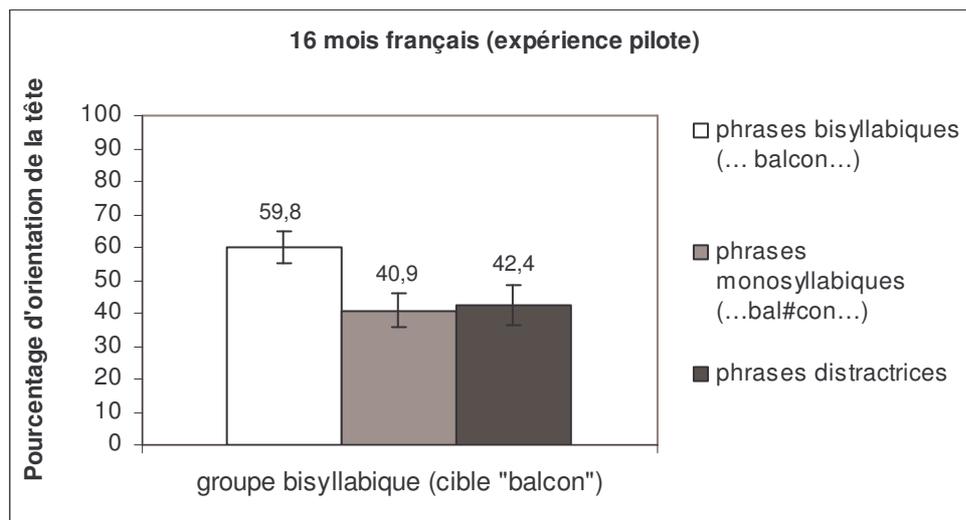


Figure 20 (résultats 16 mois français – expérience pilote) : pourcentage moyen d'orientation de la tête des enfants du groupe bisyllabique (habitué à répondre au mot bisyllabique « balcon ») sur les phrases bisyllabiques (« ...balcon... », en blanc), les phrases monosyllabiques (« ...bal#con... », en gris clair), et les phrases distractrices (en gris foncé). Les barres d'erreurs représentent les erreurs standard de la moyenne.

Dans cette expérience, les enfants habitués à répondre à un mot bisyllabique ont tourné significativement plus la tête sur les phrases bisyllabiques que sur les phrases qui contenaient les deux syllabes cibles séparées par une frontière de groupe phonologique (59.8% versus 40.9%,  $t(10)=6.8$ ,  $p<.001$ ). Par contre, le comportement des enfants a de nouveau été surprenant sur les phrases distractrices : cette fois-ci, ils répondent moins à ces phrases qu'aux phrases bisyllabiques, mais le pourcentage de réponses reste assez élevé (similaire à celui obtenu pour les phrases monosyllabiques, alors qu'elles ne comportent aucune syllabe ambiguë).

Toutefois, l'effet principal de type de phrase est maintenant significatif (les expérimentateurs sont donc capables de réaliser correctement ces expériences). Mais cet effet

reste relativement faible comparé à celui trouvé avec les enfants américains, et semble être beaucoup plus proche de ce qui est ressorti de l'expérience menée avec les enfants américains de 10 mois.

### *E.2.3 Changement de la procédure expérimentale*

Nous pensions que le protocole expérimental était relativement lourd et difficile pour les enfants, et que l'expérience durait peut-être trop longtemps (ce qui pouvait entraîner une diminution progressive des performances des enfants, et donc en moyenne de moins bons résultats). Nous avons donc décidé d'apporter quelques variations à la procédure expérimentale (pour arriver à la procédure décrite dans l'étude 3 du chapitre 5), en espérant faciliter la tâche des enfants, réduire la durée de l'expérience, et ainsi observer de meilleurs résultats de la part des enfants français (plus similaires aux résultats de Gout et al., 2004).

Le changement principal entre les deux protocoles expérimentaux concerne l'utilisation des phrases distractrices dans la deuxième session expérimentale (phase de test). Dans les expériences initiales (Gout et al., 2004, et les expériences pilotes en français), les enfants entendaient pendant la phase de test, les 24 phrases expérimentales de leur groupe (par exemple, les 12 phrases « balcon » et les 12 phrases « bal »), ainsi que 26 phrases distractrices. Il n'y avait pas de fond sonore, c'est-à-dire que des silences séparaient les essais lancés par l'expérimentateur.

Nous pensions que ceci pouvait poser certains problèmes : (1) les enfants pouvaient initialement réagir à l'apparition d'un son et se tourner vers les haut-parleurs, non pas parce qu'ils avaient reconnu le mot cible dans les phrases, mais parce qu'un stimulus auditif était joué après une période de silence (cette réponse de l'enfant pouvait témoigner de sa surprise ou de son intérêt pour les stimuli auditifs après une période de silence d'une durée variable). Ceci pouvait entraîner une forte proportion de codages « tourné vers les haut-parleurs » en temps réel, sans véritable lien avec le type de phrases entendues (ces réactions pourraient bien sûr être recodées après l'expérience, ce qui n'a pas été fait dans les premières expériences, mais pourraient alors entraîner une très forte proportion d'essais non intégrés dans les analyses) ; et (2) la présence des essais composés des phrases distractrices augmentaient considérablement la durée totale de l'expérience (ce qui n'était peut-être pas nécessaire vu que les résultats principaux, présentés suite à ces expériences, ne prenaient en compte que les réponses aux phrases bisyllabiques et aux phrases monosyllabiques).

Nous avons donc décidé : (1) de ne plus laisser de silence entre les essais, mais de jouer certaines phrases de l'autre groupe expérimental en bruit de fond (par exemple, les enfants du groupe « balcon » entendaient en fond sonore les phrases « vipère ») ; et (2) de ne plus utiliser de phrases distractrices (et de ne tester les enfants que sur les phrases bisyllabiques et les phrases monosyllabiques).

D'autres changements, plus minimes, ont aussi été mis en place :

(1) Nous avons mis en place une deuxième phase de rappel (au début de la deuxième session) : initialement une seule phase de rappel était utilisée (les enfants entendaient des phrases courtes, dans lesquelles le mot cible apparaissait en position finale, comme dans « Regarde le balcon ! »). Nous avons construit une deuxième phase de rappel dans laquelle les enfants entendaient des phrases relativement longues (similaires à celles utilisées en phase de test) qui comprenaient le mot cible placé en position médiane. Ces phrases n'étaient pas ambiguës, mais permettaient de créer un gradient de difficulté dans la tâche que les enfants avaient à faire, avant qu'ils entrent réellement en phase de test.

(2) Pendant la première phase de rappel (de la seconde session), nous avons réduit le nombre de répétitions des phrases à chaque essai : dans les expériences initiales, quand l'expérimentateur lançait un essai, une fenêtre de 3.5 secondes s'ouvrait à partir du mot cible, et la phrase cible était alors répétée trois fois. Nous avons réduit le nombre de fois où étaient répétées ces phrases cibles, ne laissant que deux répétitions. Les mêmes décisions ont été prises pour la phase de test : initialement la phrase test était jouée deux fois, mais nous avons décidé de ne la jouer qu'une fois.

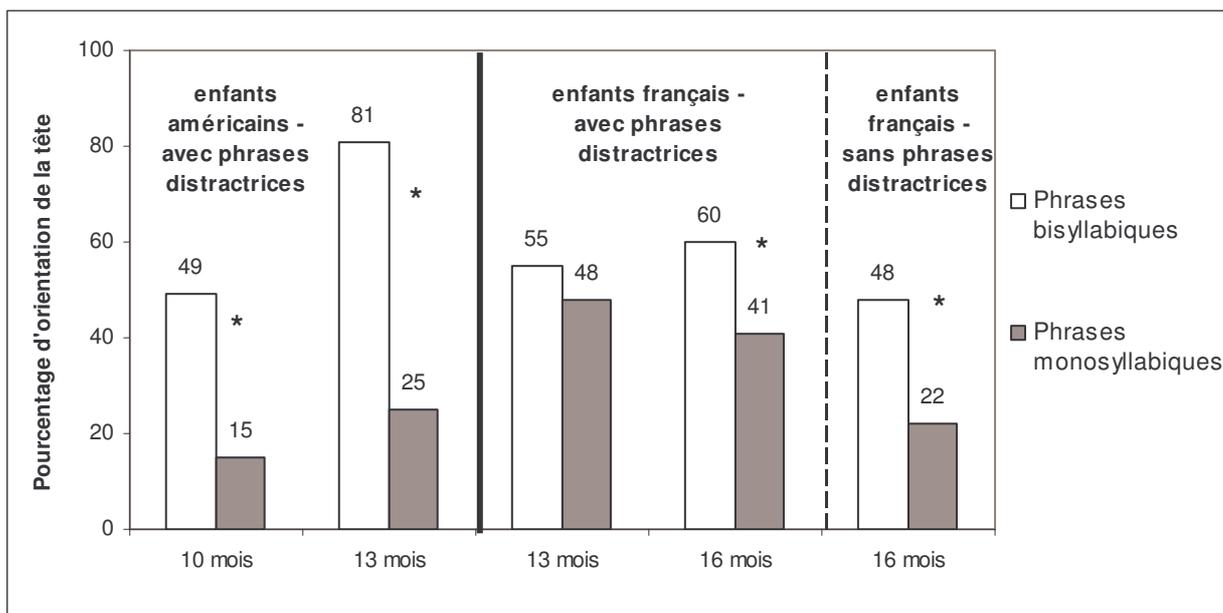
En plus de raccourcir la durée de l'expérience, ces changements nous permettaient de mieux contrôler les renforcements donnés à l'enfant quand il répondait correctement. En effet, certains problèmes apparaissaient quand les phrases étaient répétées trois fois : par exemple, si l'enfant tournait la tête rapidement, sur la première phrase, en entendant le mot cible, il recevait alors un renforcement ; mais celui-ci ne durait que 2.5 secondes. Quand la phrase était répétée une deuxième ou une troisième fois, le renforcement ne pouvait plus être lancé, et ce même si l'enfant tournait à nouveau correctement la tête en entendant le mot cible (ce qui pouvait donc paraître étrange pour l'enfant).

(3) Nous avons aussi décidé d'allonger la durée des renforcements pendant la première phase de rappel : la durée initiale de 2.5 secondes a été allongée à 3.5 secondes ; ainsi, les enfants qui tournaient la tête rapidement en entendant le mot cible, avaient un renforcement qui durait pendant la répétition des deux phrases.

(4) De manière à réduire la durée totale de l'expérience, les volumes sonores utilisés dans la première phase de rappel ont été changés : initialement, les premiers essais de chaque phase étaient joués avec un volume sonore plus élevé de 12 dB comparé au fond sonore (et cette intensité était ensuite diminuée par palier de 4 dB, à chaque fois que l'enfant tournait correctement la tête vers les haut-parleurs). Nous avons décidé de supprimer un palier de niveau sonore, en jouant les premiers essais avec un volume plus élevé de 8 dB par rapport au fond sonore.

Toutes les modifications que nous avons apportées au protocole expérimental étaient réalisées, surtout pour diminuer la durée totale de l'expérience, de manière à diminuer le risque de fatigue ou d'ennui de la part des enfants au fur et à mesure que l'expérience avançait dans le temps : nous espérions ainsi améliorer la performance globale des enfants. Nous avons donc testé de nouveaux enfants français de 16 mois (avec l'idée que si leur performance était bien améliorée, nous irions alors tester des enfants de 13 mois). Ce sont ces résultats que nous avons présentés dans le chapitre 5 (étude3).

De manière à voir si ces changements ont effectivement amélioré la performance des enfants, nous avons résumé les principaux résultats obtenus dans toutes ces expériences, en centrant notre attention sur les enfants des groupes bisyllabiques. La Figure 21 représente ces différentes données.



**Figure 21 (comparaison des résultats américains et français) : Pourcentage moyen d'orientation de la tête des enfants des groupes bisyllabiques, sur les phrases bisyllabiques (en blanc) et les phrases monosyllabiques (en gris).** La ligne en gras sépare les données des enfants américains de 10 et 13 mois (à gauche) des données des enfants français de 13 et 16 mois (à droite). La ligne en pointillés sépare les résultats des enfants français selon le protocole expérimental utilisé : ancien protocole avec les phrases distractrices (à gauche), protocole modifié sans les phrases distractrices (à droite, résultats présentés dans le chapitre 5). Les effets significatifs sont représentés par un astérisque.

Ce que l'on voit immédiatement dans ce graphique est la meilleure performance des enfants américains comparé aux enfants français. Les derniers résultats obtenus avec les français âgés de 16 mois semblent très proches de ceux obtenus par les américains de 10 mois.

Seuls les enfants français de 13 mois ont échoué dans cette tâche : après avoir été habitués à répondre à un mot bisyllabique, les enfants ont répondu autant de fois aux phrases qui contenaient ce mot bisyllabique, qu'aux phrases qui contenaient les deux syllabes du mot cible séparées par une frontière de groupe phonologique.

Pour savoir si les changements apportés au protocole expérimental ont été utiles et ont donné les résultats escomptés, nous avons comparé les résultats des deux groupes d'enfants de 16 mois. Dans les deux situations, ils réussissent à inférer que les frontières de groupe phonologique correspondent à des frontières de mot (ils n'interprètent pas les deux syllabes du mot « balcon » séparées par une frontière prosodique comme le mot « balcon »).

Dans la première version de l'expérience, les essais apparaissaient après des périodes de silence (pas de fond sonore) : nous pensons que cela pouvait entraîner une plus forte proportion de regards vers les haut-parleurs, et que ces regards seraient codés comme des essais où l'enfant répond aux phrases, alors qu'il pouvait en fait ne montrer que sa surprise ou son intérêt pour le fait d'entendre à nouveau des stimuli auditifs. La mise en place d'un fond sonore entre les essais a en effet diminué le nombre de regards vers les haut-parleurs, que ces regards soient des succès (regards sur les phrases bisyllabiques, 60% versus 48%,  $t(29)=1.8$ ,  $p=.07$ ) ou des fausses alarmes (regards vers les phrases monosyllabiques, 41% versus 22%,  $t(29)=2.9$ ,  $p=.01$ ).

Par contre, nous pensons que les changements dans le protocole expérimental amélioreraient la performance des sujets (grâce à la diminution de la durée de l'expérience, grâce à la mise en place de la deuxième phase de rappel qui pouvait habituer les enfants à réaliser des tâches de plus en plus dures avant de commencer véritablement le test...). Ceci n'est pas démontré par cette comparaison : en effet, la taille de l'effet tend à augmenter avec le nouveau protocole mais la différence n'est pas significative (ancien protocole : 17% versus nouveau protocole : 26%,  $t(29)=1.3$ ,  $p=0.3$ ).

## E.3 DISCUSSION

### *E.3.1 Quelle est la méthode la plus adaptée ?*

Les changements opérés entre les deux protocoles expérimentaux n'ont pas eu les effets escomptés : la performance des sujets tend à être légèrement améliorée mais cette différence n'est pas significative. Certes les sujets obtiennent moins de fausses alarmes, mais ils obtiennent aussi moins de succès : ceci doit probablement être dû aux essais qui, avec le premier protocole, avaient dû être codés comme des regards de l'enfant alors que ce dernier avait sûrement réagi à l'apparition d'un stimulus sonore après un silence, plutôt qu'à la phrase expérimentale proprement dite (il devait alors être probablement déjà tourné vers les haut-parleurs au moment de l'apparition du mot cible). Les résultats américains n'ont pas été recodés après les expériences, ni les résultats des expériences pilotes : il est donc difficile de savoir si cette hypothèse est vraie, mais en tout cas elle semble plausible. Nous supposons que le nombre d'essais à rejeter (ceux où l'enfant est déjà tourné vers les haut-parleurs au moment où la fenêtre de mesure du comportement s'ouvre) doit être beaucoup moins important avec le nouveau protocole expérimental qu'avec l'ancien.

Le premier protocole conserve quand même de nombreux avantages : d'abord la précision avec laquelle l'expérimentateur peut décider de lancer un essai. En effet, dès qu'il décide que l'enfant est calme et concentré, et qu'un essai peut donc être lancé, il n'a qu'à appuyer sur le bouton de la souris et la phrase est immédiatement jouée. Avec le deuxième protocole, le problème est que des phrases sont jouées en fond sonore entre les essais : quand l'expérimentateur appuie sur le bouton de la souris pour lancer un essai, le délai avec lequel la phrase est effectivement jouée est très variable (elle peut être jouée immédiatement si la phrase de fond sonore vient de se terminer, mais si par contre elle vient de commencer, l'ordinateur doit alors attendre sa fin avant de lancer la phrase test). Ce délai est parfois gênant car l'enfant peut changer rapidement de comportement, et passer d'un état calme et concentré (au moment où l'expérimentateur demande un essai), à un état beaucoup plus agité ou déjà tourné vers les haut-parleurs (au moment où la phrase est effectivement jouée).

Enfin, la présence des phrases distractrices (même si elles allongent la durée de l'expérience) peut être très informative sur ce que font réellement les enfants pendant l'expérience. Par exemple, le fait que les enfants français de 13 mois n'obtiennent pas de différence significative entre le pourcentage de regards sur les phrases bisyllabiques et le pourcentage de regards sur les phrases distractrices, nous indique que ces enfants n'ont pas seulement des problèmes avec l'utilisation des frontières de groupes phonologiques pour extraire les mots des phrases, mais qu'ils ont aussi des problèmes plus généraux avec la tâche de segmentation lexicale proprement dite. Retirer ces phrases, comme dans le deuxième protocole, nous fait donc perdre certaines informations.

Chacun de ces protocoles expérimentaux a donc ses avantages et ses inconvénients. A chacun de voir ce qu'il veut tester et laquelle de ces deux méthodes est la plus adaptée. Par contre, certaines améliorations ont été apportées avec le deuxième protocole, changements qu'il nous semble nécessaire de garder dans de futures expériences (comme la mise en place d'une deuxième phase de rappel au début de la deuxième session, où l'ajustement de la durée des renforcements aux stimuli entendus...).

### *E.3.2 Différences inter-langues*

Initialement, nous avons mis en place cette expérience de détection de mot pour tester la perception et l'utilisation des frontières de groupes phonologiques dans des tâches

linguistiques, par les enfants français. Ceci était réalisé pour commencer à valider notre hypothèse sur l'acquisition de la syntaxe, qui postule entre autres l'utilisation de ces frontières prosodiques pour segmenter les phrases en constituants syntaxiques.

Au vu des résultats observés par Gout et al. (2004) avec des enfants américains de 10 et 13 mois, nous pensions que les enfants français de 13 mois seraient eux aussi capables d'utiliser la présence d'une frontière de groupe phonologique pour inférer la présence d'une frontière de mot, et extraire ainsi les mots de la parole continue. D'où une première expérience réalisée avec des enfants de 13 mois, grâce à laquelle nous voulions démontrer que les enfants percevaient effectivement ces frontières prosodiques et pouvaient les utiliser en temps réel pour analyser les phrases qu'ils entendaient.

En fait, les résultats de cette expérience pilote ont ouvert la voie à un autre axe de recherche, que nous n'avons pas pu développer ici, mais qui nécessitera de nouveaux travaux. En effet, il semble que les enfants français aient un retard d'environ 6 mois par rapport aux enfants américains pour réussir à trouver des mots dans la parole continue (les enfants américains réussissent cette tâche de détection de mot dès l'âge de 10 mois, alors qu'il faut attendre 16 mois pour que les enfants français obtiennent des résultats similaires).

Quelles peuvent être les interprétations possibles à ce décalage ? Premièrement, les explications pourraient être d'ordre méthodologique : par exemple, si les phrases françaises étaient prononcées plus rapidement que les phrases américaines, cela pourrait expliquer pourquoi les enfants français ont plus de difficultés à identifier les mots dans les phrases. Mais lors de la mise en place du matériel expérimental américain et français, de grandes précautions avaient été prises à ce niveau, le débit syllabique et l'intonation des phrases ayant été parfaitement appariés entre les phrases françaises et les phrases américaines. Il semble donc peu probable que la différence de résultats soit attribuable aux stimuli utilisés. Il se pourrait aussi que la méthode utilisée soit moins sensible en France (ce qui pourrait être dû par exemple à une moins bonne expertise des expérimentateurs français). Une manière de vérifier cette hypothèse est d'utiliser une autre technique expérimentale, comme la technique d'écoute préférentielle (Jusczyk & Aslin, 1995). Et justement, un délai similaire d'environ 6 mois a été mis en évidence, avec cette autre technique expérimentale, entre les performances des enfants américains et celles des enfants français pour segmenter la parole continue et reconnaître des mots dans des phrases (Gout, 2001).

Quelles peuvent être les autres explications à ce retard de performance ? On pourrait par exemple faire l'hypothèse de différences culturelles, dans la manière de s'occuper des enfants et de leur parler. Par exemple, les parents américains sont très réputés pour parler à leurs enfants avec un « motherese » très accentué. Le « motherese », appelé aussi « mamanais », représente la manière qu'ont certains parents de parler à leur(s) enfant(s) avec un débit très lent et des variations d'intonation très exagérées. La prosodie des phrases est donc très accentuée et très perceptible pour les bébés. Des différences acoustiques dans la parole adressée aux enfants américains et français pourraient être une explication possible à cette difficulté que rencontrent les enfants français pour trouver les mots dans les phrases (le « motherese » des parents français pouvant être par exemple moins marqué que celui des parents américains). Il faudrait réaliser des enregistrements de parole adressée aux enfants dans ces deux pays et les comparer pour tester cette hypothèse.

Une autre explication à ce décalage pourrait venir de la langue elle-même, le français pouvant être une langue plus difficile à segmenter que l'américain. Par exemple, contrairement au français, l'américain est une langue accentuée. Les mots sont ainsi composés de syllabes fortes accentuées et de syllabes faibles, et une grande proportion des mots anglais (80%) commencent avec une syllabe forte. Cet indice est donc très pertinent à utiliser pour essayer de trouver le début des mots dans les phrases. Et en effet, il a été montré que les enfants américains de 7,5 mois comme les adultes ont tendance à supposer que les syllabes fortes correspondent au début des mots (Jusczyk, Cutler & Redanz, 1993 ; Jusczyk, Houston & Newsome, 1999). Le français ne comporte pas cet accent de mot et ne présente donc pas cet indice qui semble

particulièrement perceptible et pertinent pour segmenter la parole continue en mots. Quoiqu'il en soit, s'il est vrai que le français est initialement plus difficile à segmenter que l'américain (et ce quelles qu'en soient les raisons), on devrait observer un retard des enfants français dans leur développement lexical précoce : on pourrait ainsi s'attendre à ce que leur vocabulaire soit lui aussi retardé. Pour tenter de quantifier cette possibilité, nous avons examiné les données disponibles dans la littérature, à savoir le développement du vocabulaire des enfants tel qu'il est estimé par les questionnaires parentaux (comme le McArthur Communicative Development Inventory). La comparaison entre l'estimation parentale du nombre de mots compris et/ou produits par les enfants français et par les enfants américains (Fenson et al., 2000; Kern, 2003) met une nouvelle fois en évidence un retard des enfants français : retard allant de 1 à 2 mois pour le nombre de mots compris, et jusqu'à 6 mois pour le nombre de mots produits. Même si l'interprétation du questionnaire parental reste délicate (traduction du questionnaire difficile entre les langues, estimation subjective des parents...), il semble que ces données soient consistantes avec l'hypothèse que le français est une langue plus difficile à segmenter que l'américain (ou aussi avec les différences culturelles).

Pour conclure, nous avons montré que les jeunes enfants français et américains peuvent exploiter l'intonation de la parole pour segmenter la parole continue en mots : ils interprètent une frontière de groupe phonologique comme une frontière de mot. Cette capacité à extraire les mots des phrases pourrait varier d'une langue à l'autre puisqu'un décalage d'environ 6 mois semble se confirmer entre les enfants français et les enfants américains. Ceci nous rappelle qu'il est dangereux de tenir pour acquis que les résultats obtenus dans une langue peuvent se généraliser à toutes les langues du monde.

## REFERENCES BIBLIOGRAPHIQUES

- Abney, S. (1987). *The english noun phrase in its sentential aspect*. Unpublished doctoral dissertation, MIT, Cambridge, MA.
- Allbritton, D. W., McKoon, G., & Ratcliff, R. (1996). Reliability of prosodic cues for resolving syntactic ambiguity. *Journal of Experimental Psychology : Learning, Memory and Cognition*, 22(3), 714-735.
- Altmann, G. (1988). Ambiguity, parsing strategies, and computational models. *Language and Cognitive Processes*, 3(2), 73-97.
- Altmann, G. T. M., Garnham, A., & Dennis, Y. (1992). Avoiding the garden path: eye movements in context. *Journal of Memory and Language*, 31, 685-712.
- Altmann, G. T. M., Garnham, A., & Henstra, J. A. (1994). Effects of syntax in human sentence parsing: evidence against a structure-based proposal mechanism. *Journal of Experimental Psychology : Learning, Memory and Cognition*, 20, 209-216.
- Altmann, G. T. M., & Steedman, M. J. (1988). Interaction with context during human sentence processing. *Cognition*, 30(3), 191-238.
- Beach, C. M. (1991). The interpretation of prosodic patterns at point of syntactic ambiguity: evidence for a cue trading relations. *Journal of Memory and Language*, 30, 644-663.
- Beckman, M., & Pierrehumbert, J. (1986). Intonational structure in Japanese and English. *Phonology Yearbook*, 3, 255-309.
- Beckman, M. E. (1996). The parsing of prosody. In P. Warren (Ed.), *Prosody and parsing*. Hove, UK: Psychology Press.
- Beckman, M. E., & Edwards, J. (1990). Lengthenings and shortenings and the nature of prosodic constituency. In Kingston & Beckman (Eds.), *Papers in Laboratory Phonology: Between the Grammar and the Physics of Speech*.
- Bernal, S., Lidz, J., Waxman, S., Dutat, M., & Christophe, A. (2005). *The role of syntactic cues in verb acquisition : results from 23-month-old French infants*. Paper presented at the 10th International Congress for the Study of Child Language, Berlin, Germany.
- Bernstein Ratner, N. (1986). Durational cues which mark clause boundaries in mother-child speech. *Journal of Phonetics*, 14, 1303-1309.
- Bever, T. (1970). The cognitive basis for linguistic structure. In J. R. Hayes (Ed.), *Cognition and the development of language*. New York: Willey.
- Boland, J. E. (1997). The relationship between syntactic and semantic processes in sentence comprehension. *Language and Cognitive Processes*, 12(4), 423-484.
- Boland, J. E., & Blodgett, A. (2001). Understanding the constraints on syntactic generation: Lexical bias and discourse congruency. *Journal of Memory and Language*, 45, 391-411.
- Britt, M. A. (1994). The interaction of referential ambiguity and argument structure. *Journal of Memory and Language*, 33, 251-283.
- Britt, M. A., Perfetti, C. A., Garrod, S., & Rayner, K. (1992). Parsing in discourse: context effects and their limits. *Journal of Memory and Language*, 31, 293-314.
- Cartwright, T. A., & Brent, M. R. (1997). Syntactic categorization in early language acquisition: formalizing the role of distributional analysis. *Cognition*, 63, 121-170.
- Cassidy, K. W., & Kelly, M. H. (1991). Phonological information for grammatical category assignments. *Journal of Memory and Language*, 30, 348-369.

- Chemla, E. (2005). *Acquisition of syntactic categories: A distributional approach*. Unpublished Mémoire de DEA de Sciences Cognitives, Paris.
- Chomsky, N. (1957). *Syntactic structures*. The Hague: Mouton.
- Chomsky, N. (1965). *Aspects of the Theory of Syntax*.: Cambridge: MIT Press.
- Chomsky, N. (1981). *Lectures on Government and Binding*. Dordrecht: Foris.
- Chomsky, N. A., & Halle, M. (1968). *The Sound Pattern of English*.: New York: Harper and Row.
- Christiansen, M. H., & Dale, R. (2001). *Integrating distributional, prosodic and phonological information in a connectionist model of language acquisition*. Paper presented at the 23rd Annual Conference of the Cognitive Science Society, Mahwah, NJ.
- Christiansen, M. H., & Monaghan, P. (in press). Discovering verbs through multiple-cue integration. In K. Hirsh-Pasek & R. M. Golinkoff (Eds.), *Action meets words: how children learn verbs*.: Oxford University Press.
- Christophe, A., Dupoux, E., Bertoni, J., & Mehler, J. (1994). Do infants perceive word boundaries? An empirical study of the bootstrapping of lexical acquisition. *Journal of Acoustical Society of America*, 95(3), 1570-1580.
- Christophe, A., Guasti, M. T., Nespor, M., Dupoux, E., & van Ooyen, B. (1997). Reflections on phonological bootstrapping: its role for lexical and syntactic acquisition. *Language and Cognitive Processes*, 12, 585-612.
- Christophe, A., Peperkamp, S., Pallier, C., Block, E., & Mehler, J. (2004). Phonological phrase boundaries constrain lexical access: I. Adult data. *Journal of Memory and Language*, 51, 523-547.
- Cooper, W. E., & Paccia-Cooper, J. (1980). *Syntax and Speech*. Cambridge MA: Harvard University Press.
- Crain, S., & Steedman, M. J. (1985). On not being led up the garden path: the use of context by the psychological parser. In D. Dowty, Karttunen, L. and Zwicky, A. (Ed.), *Natural Language Parsing: Psychological, Computational, and Theoretical Perspectives* (pp. 320-358): Cambridge University Press.
- Cuetos, F., & Mitchell, D. C. (1988). Cross-linguistic differences in parsing: Restrictions on the use of late closure strategy in Spanish. *Cognition*, 30, 73-105.
- Cutler, A. (1993). Phonological cues to open-and closed-class words in the processing of spoken sentences. Special Issue: Sentence processing: III. *Journal of Psycholinguistic Research*, 22, 109-131.
- Cutler, A., Dahan, D., & van Donselaar, W. (1997). Prosody in the comprehension of spoken language: a literature review. *Language and Speech*, 40(2), 141-201.
- Delais-Roussarie, E. (1995). *Pour une approche parallèle de la structure prosodique: Etude de l'organisation prosodique et rythmique de la phrase française (For a parallel approach of prosodic structure: A study of the prosodic and rhythmic organisation of the French sentence)*. Unpublished Ph.D. Thesis, Université de Toulouse-Le Mirail, Toulouse, France.
- Di Cristo, A. (1998). Intonation in French. In A. Di Cristo & D. Hirst (Eds.), *Intonation systems: a survey of twenty languages*.: Cambridge University Press.
- Di Cristo, A. (2000a). *Interpréter la prosodie*. Paper presented at the XXIIèmes Journées d'Etudes sur la Parole, Aussois, France.
- Di Cristo, A. (2000b). Vers une modélisation de l'intonation du français: seconde partie. (Towards a model of French intonation: Second part). *French Language Studies*, 9, 143-179.

- Dresher, E. (1996). Introduction to metrical and prosodic phonology. In J. L. Morgan & K. Demuth (Eds.), *Signal to syntax : Bootstrapping from speech to grammar in early acquisition*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Durieux, G., & Gillis, S. (2001). Predicting grammatical classes from phonological cues: An empirical test. In J. Weissenborn & B. Höhle (Eds.), *Approaches to Bootstrapping: Phonological, Lexical, Syntactic and Neurophysiological Aspects of Early Language Acquisition* (Vol. 1). Amsterdam: John Benjamins.
- Fenson, L., Pethick, S., Renda, C., Cox, J. L., Dale, P. S., & Reznick, J. S. (2000). Short-form versions of the MacArthur Communicative Development Inventories. *Applied Psycholinguistics*, 21, 95-116.
- Fernald, A., & Kuhl, P. (1987). Acoustic determinants of infant preference for motherese speech. *Infant Behavior and Development*, 10, 279-293.
- Fernald, A., & Mazzie, C. (1991). Prosody and focus in speech to infants and adults. *Developmental Psychology*, 27, 209-221.
- Fernald, A., & Simon, T. (1984). Expanded intonation contours in mothers' speech to newborns. *Developmental Psychology*, 20, 104-113.
- Ferreira, F. (1993). Creation of prosody during sentence production. *Psychological Review*, 100(2), 233-253.
- Ferreira, F., & Clifton Jr, C. (1986). The independence of syntactic processing. *Journal of Memory and Language*, 25, 348-368.
- Ferreira, F., & Henderson, J. M. (1991). Recovery from misanalyses of garden-path sentences. *Journal of Memory and Language*, 30, 725-745.
- Finch, S., & Chater, N. (1992). *Bootstrapping syntactic categories*. Paper presented at the Fourteenth Annual Conference of the Cognitive Science Society of America.
- Finch, S., & Chater, N. (1994). *Distributional bootstrapping: from word class to proto-sentences*. Paper presented at the 16th annual meeting of the Cognitive Science Society of America.
- Fodor, J. D. (1998). Learning to parse? *Journal of Psycholinguistic Research*, 27(2), 285-319.
- Fodor, J. D. (2002). *Psycholinguistics cannot escape prosody*. Paper presented at the Speech Prosody 2002, Aix-en-Provence, France.
- Ford, M., Bresnan, J., & Kaplan, R. M. (1982). A competence-based theory of syntactic closure. In J. Bresnan (Ed.), *The mental representation of grammatical relations* (pp. 727-796). Cambridge, MA: The MIT Press.
- Fougeron, C., & Keating, P. A. (1997). Articulatory strengthening at edges of prosodic domains. *Journal of Acoustical Society of America*, 101(6), 3728-3740.
- Fox Tree, J. E., & Meijer, P. J. A. (2000). Untrained speakers' use of prosody in syntactic disambiguation and listeners' interpretations. *Psychological Research*, 63, 1-13.
- Frazier, L. (1978). *On comprehending sentences: Syntactic parsing strategies*. Bloomington: Indiana University Linguistics Club.
- Frazier, L. (1987). Sentence processing: A tutorial review. In M. Coltheart (Ed.), *Attention and performance XII : The psychology of reading*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Frazier, L., Clifton, C., & Randall, J. (1983). Filling gaps: decision principles and structure in sentence comprehension. *Cognition*, 13, 187-222.
- Frazier, L., & Fodor, J. D. (1978). The sausage machine: A new two-stage parsing model. *Cognition*, 6, 291-325.

- Frazier, L., & Rayner, K. (1982). Making and correcting errors during sentence comprehension: eye movements in the analysis of structurally ambiguous sentences. *Cognitive Psychology*, *14*, 178-210.
- Frazier, L., & Rayner, K. (1987). Resolution of syntactic category ambiguities: Eye movements in parsing lexically ambiguous sentences. *Journal of Memory and Language*, *26*(505-526).
- Friederici, A. D. (2002). Towards a neural basis of auditory sentence processing. *Trends in Cognitive Sciences*, *6*(2), 78-84.
- Gee, J. P., & Grosjean, F. (1983). Performance structures: a psycholinguistic and linguistic appraisal. *Cognitive Psychology*, *15*, 411-458.
- Gerken, L., Jusczyk, P. W., & Mandel, D. R. (1994). When prosody fails to cue syntactic structure: 9-month-olds' sensitivity to phonological versus syntactic phrases. *Cognition*, *51*, 237-265.
- Gerken, L., & McIntosh, B. (1993). Interplay of function morphemes and prosody in early language. *Developmental Psychology*, *29*, 448-457.
- Gleitman, L. (1990). The structural sources of verb meanings. *Language Acquisition*, *1*, 3-55.
- Gleitman, L., Gleitman, H., Landau, B., & Wanner, B. (1988). Where learning begins: Initial representations for language learning. In F. J. Newmeyer (Ed.), *Language: Psychological and biological aspects* (Vol. III). New York: Cambridge University Press.
- Gleitman, L., & Wanner, E. (1982). The state of the state of the art. In E. Wanner & L. Gleitman (Eds.), *Language acquisition: The state of the art* (pp. 3-48). Cambridge UK: Cambridge University Press.
- Gleitman, L. R., Newport, E., & Gleitman, H. (1984). The current status of the motherese hypothesis. *Journal of Child Language*, *11*, 43-79.
- Goldman-Eisler, F. (1972). Pauses, clauses, sentences. *Language and Speech*, *15*, 103-113.
- Gout, A. (2001). *Etapas précoces de l'acquisition du lexique*. Unpublished PhD thesis, Ecole des Hautes Etudes en Sciences Sociales, Paris.
- Gout, A., Christophe, A., & Morgan, J. (2004). Phonological phrase boundaries constrain lexical access: II. Infant data. *Journal of Memory and Language*, *51*, 548-567.
- Grice, H. P. (1975). Logic and conversation. In P. Cole & J. L. Morgan (Eds.), *Syntax and semantics* (Vol. 3). New York: Seminar Press.
- Grier, J. B. (1971). Nonparametric indexes for sensitivity and bias: Computing formulas. *Psychological Bulletin*, *75*, 424-429.
- Gussenhoven, C., & Rietveld, A. C. M. (1992). Intonation contours, prosodic structure and preboundary lengthening. *Journal of Phonetics*, *20*, 282-303.
- Hirsh-Pasek, K., Nelson, D. G. K., Jusczyk, P. W., Cassidy, K. W., Druss, B., & Kennedy, L. (1987). Clauses are perceptual units for young infants. *Cognition*, *26*, 269-286.
- Hirst, D. (1993). Detaching intonational phrases from syntactic structure. *Linguistic Inquiry*, *24*(4), 781-788.
- Höhle, B., & Weissenborn, J. (2000). *The origins of syntactic knowledge: Recognition of determiners in one year old German children*. Paper presented at the 24th Annual Boston Conference on Language Development, Somerville, Mass.
- Höhle, B., & Weissenborn, J. (2003). German-learning infants' ability to detect unstressed closed-class elements. *Developmental Science*, *6*, 122-127.

- Höhle, B., Weissenborn, J., Kiefer, D., Schulz, A., & Schmitz, M. (2004). Functional elements in infants' speech processing: The role of determiners in the syntactic categorization of lexical elements. *Infancy*, 5, 341-353.
- Jurafsky, D. (1996). A probabilistic model of lexical and syntactic access and disambiguation. *Cognitive Science*, 20, 137-194.
- Jusczyk, P. W. (1997). *The discovery of spoken language*. Cambridge, MA: MIT Press/Bradford Books.
- Jusczyk, P. W. (1999). How infants begin to extract words from speech. *Trends in Cognitive Sciences*, 3(9), 323-328.
- Jusczyk, P. W., & Aslin, R. N. (1995). Infants' detection of the sound patterns of words in fluent speech. *Cognitive Psychology*, 29, 1-23.
- Jusczyk, P. W., Kemler-Nelson, D. G., Hirsh-Pasek, K., Kennedy, L., Woodward, A., & Piwoz, J. (1992). Perception of acoustic correlates of major phrasal units by young infants. *Cognitive Psychology*, 24, 252-293.
- Just, M. A., & Carpenter, P. A. (1992). A capacity theory of comprehension: Individual differences in working memory. *Psychological Review*, 99, 122-149.
- Keating, P., Cho, T., Fougeron, C., & Hsu, C.-S. (2003). Domain-initial articulatory strengthening in four languages. In J. Local & R. Ogden & R. Temple (Eds.), *Papers in Laboratory Phonology, VI: Phonetic Interpretation* (pp. 143-161). Cambridge: Cambridge University Press.
- Kedar, Y., Casasola, M., & Lust, B. (2004). *24-month-olds' sensitivity to the syntactic role of function words in English sentences: noun phrase determiners*. Paper presented at the 32nd Stanford Child Language Research Forum, Palo Alto, CA.
- Kelly, M. (1996). The role of phonology in grammatical category assignments. In J. L. Morgan & K. Demuth (Eds.), *Signal to Syntax: Bootstrapping from speech to grammar in early acquisition* (pp. 249-262). Mahwah, NJ: Lawrence Erlbaum Associates.
- Kelly, M. H. (1992). Using sound to solve syntactic problems: the role of phonology in grammatical category assignments. *Psychological Review*, 99(2), 349-364.
- Kelly, M. H., & Bock, J. K. (1988). Stress in time. *Journal of Experimental Psychology : Human Perception and Performance*, 14, 389-403.
- Kemler-Nelson, D. G., Hirsh-Pasek, K., Jusczyk, P. W., & Cassidy, K. W. (1989). How the prosodic cues in motherese might assist language learning. *Journal of Child Language*, 16, 55-68.
- Kern, S. (2003). Le compte-rendu parental au service de l'évaluation de la production lexicale des enfants français entre 16 et 30 mois. *Glossa*, 85, 48-61.
- Kjelgaard, M. M., & Speer, S. R. (1999). Prosodic facilitation and interference in the resolution of temporary syntactic closure ambiguity. *Journal of Memory and Language*, 40, 153-194.
- Kraljic, T., & Brennan, S. E. (2005). Prosodic disambiguation of syntactic structure: For the speaker or for the addressee? *Cognitive Psychology*, 50(2), 194-231.
- Lacheret-Dujour, A., & Beaugendre, F. (1999). *La prosodie du français*. Paris: CNRS Editions.
- Lehiste, I. (1972). The timing of utterances and linguistic boundaries. *Journal of the Acoustical Society of America*, 51, 2018-2024.
- Lehiste, I. (1973). Rhythmic units and syntactic units in production and perception. *Journal of the Acoustical Society of America*, 54, 1228-1234.
- Lhote, E., Barry, A. O., & Tivane, A. (2000). Acquisition et apprentissage de la prosodie, une double approche, vocale et discursive, *Acquérir, Apprendre*,

- Enseigner : la prosodie au coeur du débat.*: Publications de l'Université de Rouen - CNRS.
- Lockridge, C., & Brennan, S. E. (2002). Addressees' needs influence speakers' early syntactic choices. *Psychonomic Bulletin & Review*, 9(3), 550-557.
- MacWhinney, B. (2000). *The CHILDES project: Tools for analyzing talk. Third Edition*. Mahwah, NJ: Erlbaum.
- MacWhinney, B., & Snow, C. (1985). The child language data exchange system. *Journal of Child Language*, 12, 271-295.
- MacDonald, M. C., Pearlmutter, N. J., & Seidenberg, M. S. (1994). Lexical nature of syntactic ambiguity resolution. *Psychological Review*, 101(4), 676-703.
- MacWhinney, B. (1987). The competition model. In B. MacWhinney (Ed.), *Mechanisms of language acquisition* (pp. 249-308). Hillsdale, New Jersey: Lawrence Erlbaum Associates.
- Maratsos, M. (1982). The child's construction of grammatical categories. In E. Wanner & L. Gleitman (Eds.), *Language acquisition : The state of the art*. Cambridge: Cambridge University Press.
- Maratsos, M., & Chalkley, M. A. (1980). The internal language of children's syntax. In K. Nelson (Ed.), *Children's language, Vol. 2*. New York: Gardner Press.
- Marslen-Wilson, W. D., Tyler, L. K., Warren, P., Grenier, P., & Lee, C. S. (1992). Prosodic effects in minimal attachment. *The Quarterly Journal of Experimental Psychology*, 45A(1), 73-87.
- Martin, P. (1980). Une théorie syntaxique de l'accentuation en français. In P. Léon & M. Rossi (Eds.), *L'accent en français contemporain*.
- McRae, K., Spivey-Knowlton, M. J., & Tanenhaus, M. K. (1998). Modeling the influence of thematic fit (and other constraints) in on-line sentence comprehension. *Journal of Memory and Language*, 38, 283-312.
- Mehler, J., Jusczyk, P. W., Lambertz, G., Halsted, G., Bertocini, J., & Amiel-Tison, C. (1988). A precursor of language acquisition in young infants. *Cognition*, 29, 143-178.
- Mintz, T. H. (2002). Category induction from distributional cues in an artificial language. *Memory and Cognition*, 30, 678-686.
- Mintz, T. H. (2003). Frequent frames as a cue for grammatical categories in child-directed speech. *Cognition*, 90, 91-117.
- Mintz, T. H., Newport, E. L., & Bever, T. G. (2002). The distributional structure of grammatical categories in speech to young children. *Cognitive Science*, 26(4), 393-424.
- Mitchell, D. C. (1989). Verb guidance and lexical effects in ambiguity resolution. *Language and Cognitive Processes*, 4, 123-154.
- Mitchell, D. C. (1994). Sentence Parsing. In M. A. Gernsbacker (Ed.), *Handbook of Psycholinguistics* (pp. 375-409). San Diego: Academic Press, Inc.
- Mitchell, D. C., & Cuetos, F. (1991). *The origins of parsing strategies*. Paper presented at the Current issues on natural language processing, University of Texas, Austin, TX.
- Mitchell, D. C., Cuetos, F., & Corley, M. M. B. (1992). *Statistical versus linguistic determinants of parsing bias: cross-linguistic evidence*. Paper presented at the 5th Annual CUNY Conference on Human Sentence Processing, New York.
- Mitchell, D. C., & Holmes, V. M. (1985). The role of specific information about the verb in parsing sentences with local structural ambiguity. *Journal of Memory and Language*, 24, 542-559.

- Monaghan, P., Chater, N., & Christiansen, M. H. (2005). The differential contribution of phonological and distributional cues in grammatical categorisation. *Cognition*, *96*, 143-182.
- Morgan, J. L. (1986). *From simple input to complex grammar*. Cambridge Mass: MIT Press.
- Morgan, J. L., & Demuth, K. (1996a). Signal to Syntax: an overview. In J. L. Morgan & K. Demuth (Eds.), *Signal to Syntax: Bootstrapping from speech to grammar in early acquisition* (pp. 1-22). Mahwah, NJ: Lawrence Erlbaum Associates.
- Morgan, J. L., & Demuth, K. (1996b). *Signal to Syntax: Bootstrapping from speech to grammar in early acquisition*. Mahwah NJ: Lawrence Erlbaum Associates.
- Morgan, J. L., Meier, R. P., & Newport, E. L. (1987). Structural Packaging in the input to language learning: contributions of prosodic and morphological marking of phrases to the acquisition of language. *Cognitive Psychology*, *19*, 498-550.
- Morgan, J. L., Shi, R., & Allopenna, P. (1996). Perceptual bases of rudimentary grammatical categories: Toward a broader conceptualization of bootstrapping. In J. L. Morgan & K. Demuth (Eds.), *Signal to Syntax: Bootstrapping from speech to grammar in early acquisition* (pp. 263-283). Mahwah, NJ: Lawrence Erlbaum Associates.
- Nagel, H. N., Shapiro, L. P., Tuller, B., & Nawy, R. (1996). Prosodic influences on the resolution of temporary ambiguity during on-line sentence processing. *Journal of Psycholinguistic Research*, *25*(2), 319-344.
- Nazzi, T., Nelson, D. G. K., Jusczyk, P. W., & Jusczyk, A. M. (2000). Six-month-olds' detection of clauses embedded in continuous speech: effects of prosodic well-formedness. *Infancy*, *1*, 123-147.
- Nespor, M., & Vogel, I. (1986). *Prosodic Phonology*. Dordrecht: Foris.
- New, B., Pallier, C., Ferrand, L., & Matos, R. (2001). Une base de données lexicales du français contemporain sur internet: LEXIQUE (A lexical database of contemporary French on internet: LEXIQUE). *L'Année Psychologique*, *101*, 447-462.
- Nicol, J. L. (1996). What can prosody tell a parser? *Journal of Psycholinguistic Research*, *25*(2), 179-192.
- Osterhout, L. (1994). Event-related brain potentials as tools for comprehending language comprehension. In C. Clifton Jr & L. Frazier & K. Rayner (Eds.), *Perspectives on sentence processing*. Hillsdale, NJ: Erlbaum.
- Pallier, C., Dupoux, E., & Jeannin, X. (1997). EXPE: an expandable programming language for on-line psychological experiments. *Behavior Research Methods, Instruments, & Computers*, *29*, 322-327.
- Pinker, S. (1984). *Language learnability and language development*. Cambridge, MA: Harvard University Press.
- Pinker, S. (1987). The bootstrapping problem in language acquisition. In B. MacWhinney (Ed.), *Mechanisms of language acquisition*. Hillsdale, NJ: Erlbaum.
- Price, P. J., Ostendorf, M., Shattuck-Hufnagel, S., & Fong, C. (1991). The use of prosody in syntactic disambiguation. *Journal of Acoustical Society of America*, *90*, 2956-2970.
- Pritchett, B. (1988). Garden path phenomena and the grammatical basis of language processing. *Language*, *64*(3), 539-576.
- Rayner, K., Carlson, M., & Frazier, L. (1983). The interaction of syntax and semantic during sentence processing: Eye movements in the analysis of semantically biased sentences. *Journal of Verbal Learning and Verbal Behavior*, *22*, 358-374.

- Real, F., Christiansen, M. H., & Monaghan, P. (2003). *Phonological and distributional cues in syntax acquisition: scaling up the connectionist approach to multiple-cue integration*. Paper presented at the 25th Annual Conference of the Cognitive Science Society, Mahwah, NJ.
- Redington, M., Chater, N., & Finch, S. (1998). Distributional information: a powerful cue for acquiring syntactic categories. *Cognitive Science*, 22, 435-469.
- Schafer, A. (1997). *Prosodic parsing : The role of prosody in sentence comprehension*. Unpublished Doctoral dissertation, University of Massachusetts.
- Schafer, A. J., & Speer, S. R. (1998). Prosodic influences on the resolution of lexical ambiguity, *Presentation to the Annual Meeting of the linguistic Society of America*. NY, NY.
- Schafer, A. J., Speer, S. R., & Warren, P. (2004). Prosodic influences on the production and comprehension of syntactic ambiguity in a game-based conversation task. In M. K. Tanenhaus & J. Trueswell (Eds.), *World situated language use: Psycholinguistic, linguistic and computational perspectives on bridging the product and action tradition*.
- Schafer, A. J., Speer, S. R., Warren, P., & White, S. D. (2000). Intonational disambiguation in sentence production and comprehension. *Journal of Psycholinguistic Research*, 29(2), 169-182.
- Schepman, A., & Rodway, P. (2000). Prosody and parsing in coordination structures. *The Quarterly Journal of Experimental Psychology*, 53A(2), 377-396.
- Schober, M. F., & Brennan, S. E. (2003). Processes of interactive spoken discourse: The role of the partner. In A. C. Graesser & M. A. Gernsbacher & S. R. Goldman (Eds.), *Handbook of discourse processes*. Hillsdale, NJ: Lawrence Erlbaum.
- Sedivy, J. C., & Spivey-Knowlton, M. J. (1994). The use of structural, lexical and pragmatic information in parsing attachment ambiguities. In C. Clifton & L. Frazier & K. Rayner (Eds.), *Perspectives on sentence processing*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Selkirk, E. (1982). *The syntax of words*. Cambridge, MA: The MIT Press.
- Selkirk, E. (1984). *Phonology and Syntax: the Relation between Sound and Structure*. Cambridge Mass.: MIT Press.
- Selkirk, E. (1986). On derived domains in sentence phonology. *Phonology*, 3, 371-405.
- Selkirk, E. (1996). The prosodic structure of function words. In J. L. Morgan & K. Demuth (Eds.), *Signal to syntax: Bootstrapping from speech to grammar in early acquisition* (pp. 187-213). Mahwah, NJ: Lawrence Erlbaum Associates.
- Selkirk, E. O. (1978). On prosodic structure and its relation to syntactic structure. In T. Fretheim (Ed.), *Nordic Prosody II*. Trondheim: TAPIR.
- Shady, M. (1996). *Infant's sensitivity to function morphemes*. Unpublished PhD Thesis, State University of New York, Buffalo.
- Shafer, V. L., Shucard, D. W., Shucard, J. L., & Gerken, L. (1998). An electrophysiological study of infants' sensitivity to the sound patterns of english speech. *Journal of Speech, Language, and Hearing Research*, 41, 874-886.
- Shapiro, L. P., Nagel, H. N., & Levine, B. E. (1993). Preferences for a verb's complements and their use in sentence processing. *Journal of Memory and Language*, 32, 96-114.
- Shattuck-Hufnagel, S., & Turk, A. E. (1996). A prosody tutorial for investigators of auditory sentence processing. *Journal of Psycholinguistic Research*, 25(2), 193-247.
- Sherman, D. (1975). Noun-verb stress alternation: An example of lexical diffusion of sound change in English. *Linguistics*, 159, 43-71.

- Shi, R. (2005). *Perception of function words in preverbal infants*. Paper presented at the 10th International Congress for the Study of Child Language, Berlin, Germany.
- Shi, R., & Gauthier, B. (2005). Recognition of function words in 8-month-old French-learning infants. *Journal of Acoustical Society of America*, *117*(4), 2426-2427.
- Shi, R., Morgan, J. L., & Allopenna, P. (1998). Phonological and acoustic bases for earliest grammatical category assignment: a cross-linguistic perspective. *Journal of Child Language*, *25*, 169-201.
- Shi, R., Werker, J., & Cutler, A. (2003). *Function words in early speech perception*. Paper presented at the Fifteenth International Congress of Phonetic Sciences, Barcelona, Spain.
- Shi, R., Werker, J., & Cutler, A. (review). Recognition and representation of function words in English-learning infants. *Infancy*.
- Shi, R., Werker, J. F., & Morgan, J. L. (1999). Newborn infants' sensitivity to perceptual cues to lexical and grammatical words. *Cognition*, *72*, B11-B21.
- Snedeker, J., Gleitman, L., Felberbaum, M., Placa, N., & Trueswell, J. (2000). *Prosodic choice: effects of speaker awareness and referential context*. Paper presented at the Twenty-second Annual Conference of the Cognitive Science Society, Mahwah, NJ, USA.
- Snedeker, J., & Trueswell, J. (2003). Using prosody to avoid ambiguity: effects of speaker awareness and referential context. *Journal of Memory and Language*, *48*(103-130).
- Soderstrom, M., Seidl, A., Kemler Nelson, D. G., & Jusczyk, P. (2003). The prosodic bootstrapping of phrases: evidence from prelinguistic infants. *Journal of Memory and Language*, *49*, 249-267.
- Soreno, J. A., & Jongman, A. (1990). Phonological and form class relations in the lexicon. *Journal of Psycholinguistic Research*, *19*, 387-404.
- Sorenson, J. M., Cooper, W. E., & Paccia, J. M. (1978). Speech timing of grammatical categories. *Cognition*, *6*, 135-153.
- Spivey-Knowlton, M., & Tanenhaus, M. K. (1994). Referential context and syntactic ambiguity resolution. In C. Clifton Jr & L. Frazier & K. Rayner (Eds.), *Perspectives in sentence processing* (pp. 415-439): Lawrence Erlbaum.
- Stirling, L., & Wales, R. (1996). Does prosody support or direct sentence processing? *Language and Cognitive Processes*, *11*(1/2), 193-212.
- Tabor, W., Juliano, C., & Tanenhaus, M. K. (1997). Parsing in a dynamical system: an attractor-based account of the interaction of lexical and structural constraints in sentence processing. *Language and Cognitive Processes*, *12*(2/3), 211-271.
- Tanenhaus, M. K., Carlson, G. N., & Trueswell, J. C. (1989). The role of thematic structures in interpretation and parsing. *Language and Cognitive Processes*, *4*, 211-234.
- Tanenhaus, M. K., Spivey-Knowlton, M. J., Eberhard, K. M., & Sedivy, J. E. (1995). Integration of visual and linguistic information in spoken language comprehension. *Science*, *268*, 1632-1634.
- Trueswell, J. C., & Kim, A. E. (1998). How to prune a garden path by nipping it in the bud: fast priming of verb argument structure. *Journal of Memory and Language*, *39*, 102-123.
- Trueswell, J. C., Tanenhaus, M. K., & Garnsey, S. M. (1994). Semantic influences on parsing: Use of thematic role information in syntactic ambiguity resolution. *Journal of Memory and Language*, *33*, 285-318.
- Trueswell, J. C., Tanenhaus, M. K., & Kello, C. (1993). Verb-specific constraints in sentence processing: separating effects of lexical preference from garden-paths.

- Journal of Experimental Psychology : Learning, Memory and Cognition*, 19(3), 528-553.
- Vaissière, J. (1997). Langues, prosodies et syntaxe. *Traitement Automatique des Langues*, 38(1), 53-82.
- Verluyten, S. P. (1982). *Recherches sur la prosodie et la métrique du Français (Investigation on French Prosodics and Metrics)*. Unpublished PhD Thesis, University of Antwerpen.
- Warren, P., Grabe, E., & Nolan, F. (1995). Prosody, phonology and parsing in closure ambiguities. *Language and Cognitive Processes*, 10(5), 457-486.
- Welby, P. (2003a). *French intonational rises and their role in speech segmentation*. Paper presented at the 8th Annual Conference on Speech Communication and Technology (Eurospeech), Geneva, Switzerland.
- Welby, P. (2003b). *The slaying of Lady Mondegreen: Being a study of French tonal association and alignment and their role in speech segmentation*. Unpublished Unpublished PhD thesis, Ohio State University, Columbus.
- Welby, P. (2004). *The structure of French intonational rises: A study of text-to-tune alignment*. Paper presented at the Speech Prosody 2004, Nara, Japan.
- Wightman, C. W., Shattuck-Hufnagel, S., Ostendorf, M., & Price, P. J. (1992). Segmental durations in the vicinity of prosodic phrase boundaries. *Journal of the Acoustical Society of America*, 91, 1707-1717.
- Zangl, R., & Fernald, A. (2003). *Sensitivity to function morphemes in on-line sentence processing: developmental changes from 18 to 36 months*. Paper presented at the biennial meeting of the Boston University Conference on Child Language Development, Boston, MA, USA.