



**HAL**  
open science

# La prégnance perceptive des mots parlés: une réponse au problème de la segmentation lexicale ?

Christian Michel Lachaud

► **To cite this version:**

Christian Michel Lachaud. La prégnance perceptive des mots parlés: une réponse au problème de la segmentation lexicale ?. Psychologie. University of Geneva, 2005. Français. NNT : . tel-00419056

**HAL Id: tel-00419056**

**<https://theses.hal.science/tel-00419056>**

Submitted on 24 Sep 2009

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

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



UNIVERSITÉ DE GENÈVE

FACULTÉ DE PSYCHOLOGIE ET DES  
SCIENCES DE L'ÉDUCATION

Section de Psychologie

---

Sous la direction du Professeur  
Ulrich Hans FRAUENFELDER

# **La prégnance perceptive des mots parlés : une réponse au problème de la segmentation lexicale ?**

## **THÈSE**

Présentée à la  
Faculté de Psychologie et des Sciences de l'Éducation  
de l'Université de Genève  
pour obtenir le grade de Docteur en Psychologie

par

**Christian Michel LACHAUD**



Thèse n° 353

2005 - Genève, GE - SUISSE

# Membres du Jury

*(de gauche à droite sur la photo)*

- ▷ Pr. Alain Content (Université Libre de Bruxelles, Belgique)
- ▷ Dr. Sven Mattys (University of Bristol, United Kingdom)
- ▷ Pr. Ulrich Hans Frauenfelder (Université de Genève, Suisse)
- ▷ *(le candidat)*
- ▷ Pr. Joël Pynte (Université de Provence, France)



# Remerciements

Initiée en 1998 à Aix-en-Provence (France) sous la direction du Pr. Joël Pynte (Directeur de Recherches au CNRS, DR2 HDR - Laboratoire Parole et Langage (LPL), UMR 6057), puis sous la codirection du Pr. Geneviève Caelen-Haumont (Directeur de Recherches au CNRS, DR2 HDR - LPL), l'aventure s'est poursuivie et terminée à Genève sous la direction du Pr. Ulrich Hans Frauenfelder (Professeur Ordinaire - Laboratoire de Psycholinguistique Expérimentale, FPSE, Université de Genève).

Durant ce travail, Chercheurs et Non-chercheurs ont eu l'occasion de m'apporter un soutien matériel, intellectuel, philosophique, ou relationnel et affectif.

J'adresse donc à toutes et tous mes sincères remerciements.

Je souhaite chaleureusement remercier<sup>1</sup> :

- ▷ le Pr. Geneviève Caelen-Haumont, le Pr. Uli Frauenfelder, et le Pr. Joël Pynte, pour m'avoir donné l'opportunité de faire ce travail et pour m'avoir accompagné sur ce parcours. Un grand merci à Geneviève pour son soutien dans des moments difficiles.
- ▷ Les intervenants suivants, qui ont appuyé ce travail avec leurs compétences techniques :
  - ∩ Yoann Bourgeois, du GREQAM (EHESS, Université d'Aix-Marseille), pour sa contribution aux analyses statistiques basées sur le bootstrap (Annexe 11).
  - ∩ Pierre Courrieu (Laboratoire de Psychologie Cognitive, CNRS-UMR 6146, Université de Provence) pour les discussions théoriques concernant les réseaux neuromimétiques, et pour le calcul de l'indice « dtw » utilisé dans certaines analyses (Annexe 11).
  - ∩ Philippe Demierre et Franco Serena (Unité Développements et Maintenances, Université de Genève) pour leur aide à la conception, et pour la réalisation du système de recueil de données par Internet, utilisé pour obtenir les indices d'âge d'acquisition et de familiarité des mots (Annexe 1).
  - ∩ Sophie Dufour (FPSE, Université de Genève), qui a mis à ma disposition ses données expérimentales afin de tester le modèle SLEXS (Discussion).
  - ∩ Robert Espesser (Ingénieur d'études au CNRS, LPL), pour sa contribution technique importante et les nombreuses heures de travail passées ensemble (Chapitre 1, Partie IV).
  - ∩ Michael Eid (P.O., FPSE), Paolo Ghisletta (M.A., FPSE & Centre interfacultaire de gérontologie), et Olivier Renaud (M.E.R., FPSE, Université

---

<sup>1</sup> Noms agencés alphabétiquement.



de Genève), pour les heures de consultation en statistique qu'ils m'ont accordé. Un grand merci à Olivier pour sa disponibilité inconditionnelle et son attention.

- ∩ Michel Grzeskowiak (FPSE, Université de Genève) pour sa disponibilité et son aide avec l'utilisation du logiciel d'expérimentation E-prime.
  - ∩ La Locutrice qui a prêté sa voix et donné son temps pour les enregistrements utilisés dans les études du Chapitre 1, Partie IV.
  - ∩ Min et Edmond, du « Centre for Multilevel Modelling, Londres (UK) », pour leur aide avec MLwiN.
- ▷ Tous les intervenants auprès desquels j'ai sollicité un avis ou des discussions sur des points techniques ou théoriques :
- ∩ mes collègues Doctorants ainsi que les Chercheurs et Enseignant-chercheurs d'Aix-en-Provence.
  - ∩ mes collègues Assistants, Maître-Assistants, Maîtres d'Enseignement et de Recherche et Professeurs Ordinaires de la FPSE à Genève, et du Graduate Program Cognitive Psychology (École doctorale Suisse).
  - ∩ les Chercheurs suivants, extérieurs aux institutions qui m'ont accueillies :
    - Alain Content (Université Libre de Bruxelles, Belgique).
    - Edgar Erfelder (Universität Mannheim, Deutschland).
    - Aditti Lahiri (Universität Konstanz, Deutschland).
    - Sven Mattys (University of Bristol, United Kingdom).
    - Jay McClelland (Carnegie Mellon University, United States of America).
    - James McQueen (Max Planck Instituut voor Psycholinguïstiek, Nederland).
    - Boris New (Université Paris V, France).
    - Henning Reetz (Universität Konstanz, Deutschland).
    - Jan Peter de Ruiter (Max Planck Instituut voor Psycholinguïstiek, Nederland).
- ▷ Tous les participants aux expériences effectuées dans le cadre de ce doctorat.
- ▷ Mes Parents, Famille, Ami(e)s, et Proches. Un très grand merci pour votre soutien non professionnel mais essentiel.

J'adresse également des remerciements « institutionnels » :

- ▷ Au Fonds National Suisse de la Recherche Scientifique, pour le financement de formations en statistiques (Swiss Summer School 2001-2002 et 2002-2003, Lugano, Suisse).
- ▷ À l'École Doctorale de Psychologie Cognitive et Expérimentale (Genève), pour le financement d'une formation sur les Réseaux de neurones (2005, ESPCI, Paris, France).

## Conventions de notation

### Notation de la significativité des résultats

- ▷ En rouge gras : effets significatifs ( $p \leq .05$ )
- ▷ En bleu, normal : tendances ( $.05 < p < .1$ )
- ▷ En italique : non significatif ( $.1 \leq p$ )

### Notation des mots

- ▷ Notation orthographique : caractères de l'alphabet latin, en italique, entre guillemets ou parenthèses (exemple : « *chemise* » ou (*chemise*))
- ▷ Notation phonétique : segments transcrits en API entre crochets (exemple : [ʃømiz])
- ▷ Notation phonologique : segments transcrits en API entre barres de fraction (exemple : /ʃømiz/)

## Remarque à l'attention du lecteur concernant la structure du manuscrit

Le texte est structuré en niveaux de détail, le niveau maximal correspondant aux notes de bas de page, qui fournissent des exemples et des justifications aux idées ou aux faits avancés dans le corps du texte. Pour ne pas briser la fluidité de la lecture, leur parcours est facultatif.

# Table des matières

<b>1</b>	<b>INTRODUCTION .....</b>	<b>9</b>
1.1	GÉNÉRALITÉS .....	9
1.2	LOCALISER LES FRONTIÈRES LEXICALES DANS LE FLUX DE PAROLE : UN ÉCUEIL THÉORIQUE ET PRATIQUE.....	10
1.3	QUELLES SOLUTIONS ? INTRODUCTION AU TRAVAIL EFFECTUÉ.....	12
1.4	PRÉSENTATION DE L'ORGANISATION DU MANUSCRIT .....	13
<b>2</b>	<b>PARTIE II – THÉORIE.....</b>	<b>16</b>
2.1	CHAPITRE 1 – MODÈLES DE LA RECONNAISSANCE DES MOTS PARLÉS .....	17
2.1.1	<i>Généralités sur les modèles psycholinguistiques de la reconnaissance des mots parlés .....</i>	<i>17</i>
2.1.2	<i>Fonctions du système de reconnaissance des mots parlés au niveau calculatoire.....</i>	<i>20</i>
2.1.3	<i>Présentation d'un modèle implémenté : le modèle TRACE.....</i>	<i>32</i>
2.1.4	<i>Récapitulation - Partie II, Chapitre 1 .....</i>	<i>41</i>
2.2	CHAPITRE 2 – PROSODIE ET DÉMARCATIION DES UNITÉS LEXICALES .....	42
2.2.1	<i>Quelques rappels sur la prosodie.....</i>	<i>42</i>
2.2.2	<i>Le rôle de l'information prosodique dans la reconnaissance des mots parlés.....</i>	<i>44</i>
2.2.3	<i>Les travaux en reconnaissance automatique de la parole.....</i>	<i>56</i>
2.2.4	<i>Récapitulation - Partie II, Chapitre 2 .....</i>	<i>60</i>
2.3	CHAPITRE 3 - L'ORGANISATION SUPERPOSANTE DU LEXIQUE : QUEL RÔLE SUR LA RECONNAISSANCE DES MOTS PARLÉS ?.....	61
2.3.1	<i>L'organisation superposante du lexique .....</i>	<i>61</i>
2.3.2	<i>L'étude du traitement des superpositions lexicales dans la littérature : principaux courants théoriques et approches expérimentales.....</i>	<i>72</i>
2.3.3	<i>Les prédictions des modèles psycholinguistiques de la reconnaissance des mots parlés face aux superpositions lexicales réelles .....</i>	<i>74</i>
2.3.4	<i>Récapitulation - Partie II, Chapitre 3 .....</i>	<i>76</i>
2.4	CHAPITRE 4 – RÉCAPITULATION DE LA PARTIE II : IDÉES CLÉ ET DES HYPOTHÈSES FONDANT CE TRAVAIL DE RECHERCHE .....	77
<b>3</b>	<b>PARTIE III – MÉTHODOLOGIE.....</b>	<b>78</b>
3.1	CHAPITRE 1 - PRÉSENTATION DES PROBLÈMES MÉTHODOLOGIQUES ET TECHNIQUES RENCONTRÉS....	79
3.1.1	<i>Description des facteurs influençant la production des données et des résultats .....</i>	<i>80</i>
3.1.2	<i>Étapes de l'étude où des procédés peuvent être appliqués pour réduire ou contrôler la variance indésirable .....</i>	<i>86</i>
3.2	CHAPITRE 2 - L'ORIGINE DE LA MESURE DES TEMPS DE RÉPONSE LORS DE LA RECONNAISSANCE D'UN MOT PARLÉ ISOLÉ .....	93
3.2.1	<i>Introduction : .....</i>	<i>93</i>
3.2.2	<i>Origine de la mesure et conséquences sur la théorie.....</i>	<i>94</i>
3.2.3	<i>Solution technique proposée au problème de l'origine de la mesure dans l'étude de la reconnaissance des mots parlés.....</i>	<i>106</i>
3.2.4	<i>Conclusions .....</i>	<i>106</i>

3.3	CHAPITRE 3 - LE TRAITEMENT STATISTIQUE DES TEMPS DE RÉACTION POUR RECONNAÎTRE UN MOT PARLÉ ISOLÉ.....	107
3.3.1	<i>Introduction</i> : .....	107
3.3.2	<i>La préparation des données pour leur analyse statistique</i> .....	108
3.3.3	<i>L'analyse statistique des données</i> .....	117
3.3.4	<i>Présentation de la technique de régression multiniveaux</i> .....	123
3.3.5	<i>Conclusions</i> .....	131
3.3.6	<i>Résumé récapitulatif</i> .....	132
3.4	CHAPITRE 4 - MÉTHODOLOGIE EXPÉRIMENTALE .....	134
3.4.1	<i>Création du matériel expérimental</i> .....	134
3.4.2	<i>Tâches utilisées</i> .....	147
3.4.3	<i>Population testée et obtention des participants</i> .....	150
3.4.4	<i>Recueil des données</i> .....	150
3.4.5	<i>Analyse des données</i> .....	154
3.5	CHAPITRE 5 - PLAN SYNTHÉTIQUE RÉCAPITULANT LES ASPECTS TECHNIQUES ET PRÉSENTANT LES ASPECTS THÉORIQUES POUR CHAQUE EXPÉRIENCE .....	155
<b>4</b>	<b>PARTIE IV – EXPÉRIMENTATION .....</b>	<b>159</b>
4.1	CHAPITRE 1 - EXPLORATION DU RÔLE DES PARAMÈTRES SUPRASEGMENTAUX DANS LE GUIDAGE SENSORIEL DE L'ACTIVATION LEXICALE.....	161
4.1.1	<i>Présentation du contenu et de l'organisation du Chapitre 1</i> .....	161
4.1.2	<i>Étude préalable</i> .....	163
4.1.3	<i>Expérience 1 : Test perceptif de discrimination – Comparaison d'items désémanés par synthèse</i> <i>196</i>	
4.1.4	<i>Expérience 2 : Test perceptif de détection et de localisation consciente de frontières lexicales</i>	210
4.1.5	<i>Expérience 3 : Effets du débit d'élocution et de la hauteur de la voix sur la perception</i> .....	220
4.1.6	<i>Conclusions relatives aux études présentées dans le Chapitre 1, Partie IV</i> .....	228
4.2	CHAPITRE 2 - LINÉARISATION D'ÉLÉMENTS SUPERPOSÉS : QUELLE ORGANISATION POUR LE SYSTÈME DE RECONNAISSANCE DES MOTS PARLÉS ? .....	231
4.2.1	<i>Introduction</i> .....	231
4.2.2	<i>Expérience 4 : Effet d'une superposition lexicale initiale réelle et/ou virtuelle sur la reconnaissance auditive d'un mot monosyllabique en modalité auditive</i> .....	234
4.2.3	<i>Expérience 5 : Effet d'une superposition lexicale initiale réelle sur la reconnaissance d'un mot monosyllabique en modalité auditive – contrôle phonétique</i> .....	242
4.2.4	<i>Expérience 6 : Effet d'une superposition lexicale réelle, initiale et/ou finale, sur la reconnaissance d'un mot bisyllabique, en modalité auditive</i> .....	248
4.2.5	<i>Expérience 7 : Réplication et contrôle</i> .....	256
4.2.6	<i>Expérience 8 : Effet d'une superposition lexicale initiale réelle sur la reconnaissance d'un mot monosyllabique, en modalité visuelle</i> .....	266
4.2.7	<i>Expérience 9 : Sondage de la charge de calcul dans le système, aux différentes étapes de la reconnaissance d'un mot monosyllabique impliqué dans des superpositions lexicales initiales réelles et virtuelles, en modalité auditive</i> .....	272
4.2.8	<i>Expérience 10 : Sondage de la charge de calcul dans le système de traitement, à différentes étapes de la reconnaissance d'un mot bisyllabique renfermant une superposition lexicale réelle, initiale et/ou finale, en modalité auditive</i> .....	288
4.2.9	<i>Expérience 11 : Détection de syllabes non lexicales du Français I – Exploration du rôle de la fréquence positionnelle</i> .....	300
4.2.10	<i>Expérience 12 : Détection de syllabes non lexicales du Français II – Poursuite de l'exploration du rôle de la fréquence positionnelle</i> .....	311
4.2.11	<i>Expérience 13 : Détection de syllabes lexicales I – Exploration du rôle de la fréquence positionnelle en initiale de mots</i> .....	318
4.2.12	<i>Expérience 14 : Détection de syllabes lexicales II – Exploration du rôle de la fréquence positionnelle en non initiale de mots</i> .....	324
4.2.13	<i>Analyse exploratoire</i> .....	330
4.2.14	<i>Conclusions des Expériences 11 à 14, et des Analyses supplémentaires et exploratoire</i> .....	336

4.2.15	<i>Conclusions relatives aux études présentées dans le Chapitre 2.....</i>	338
4.3	CHAPITRE 3 - EXPLORATION DE L'INFLUENCE COMBINÉE DES ASPECTS SUPRA-SEGMENTAUX DU SIGNAL ET ARCHITECTURAUX DU LEXIQUE MENTAL SUR LE TRAITEMENT LEXICAL .....	341
4.3.1	<i>Expérience 15 : Exploration de l'influence de la cohésion acoustique et psychologique d'une chaîne parlée sur sa perception I .....</i>	342
4.3.2	<i>Expérience 16 : Exploration de l'influence de la cohésion acoustique et psychologique d'une chaîne parlée sur sa perception II.....</i>	350
4.3.3	<i>Conclusions relatives aux études présentées dans le Chapitre 3, Partie IV.....</i>	354
4.4	CHAPITRE 4 – EXPLORATION DES DONNÉES ISSUES DES CHAPITRES 2 ET 3 : RELATION ENTRE CARACTÉRISTIQUES LINGUISTIQUES ET PROSODIQUES DES ITEMS, ET ENTRE PROSODIE ET TEMPS DE RÉPONSE.....	355
4.4.1	<i>Présentation du contenu et de l'organisation du Chapitre 4 .....</i>	355
4.4.2	<i>Présentation des effets.....</i>	357
4.4.3	<i>Discussion .....</i>	367
4.4.4	<i>Conclusions relatives aux analyses présentées dans le Chapitre 4, Partie IV .....</i>	371
4.5	CHAPITRE 5 – RÉCAPITULATION SYNTHÉTIQUE DES RÉSULTATS PRÉSENTÉS DANS LA PARTIE EXPÉRIMENTALE .....	372
<b>5</b>	<b>PARTIE V - DISCUSSION .....</b>	<b>374</b>
5.1.1	<i>L'exploration du rôle de la prosodie dans la reconnaissance des mots parlés.....</i>	375
5.1.2	<i>L'exploration du rôle des superpositions lexicales dans la reconnaissance des mots parlés .....</i>	377
5.1.3	<i>Description du modèle de reconnaissance des mots SLEXSI .....</i>	382
5.1.4	<i>Test du modèle SLEXSI.....</i>	395
<b>6</b>	<b>CONCLUSION GÉNÉRALE - LES APPORTS DE CE TRAVAIL.....</b>	<b>407</b>
	<b>Annexes</b>	423
	<b>Index des Équations</b>	570
	<b>Index des Figures</b>	573
	<b>Index des Tableaux</b>	577
	<b>Index des Annexes</b>	583

# Introduction

## 1 Introduction

### 1.1 Généralités

Bien que la perception du langage parlé soit un phénomène commun et involontaire, les processus de traitement de l'information qui le sous-tendent sont extrêmement complexes. Les connaissances que nous en avons sont parcellaires, souvent controversées, et manifestement insuffisantes pour permettre la création d'un dispositif technique reproduisant artificiellement les capacités du système naturel avec un niveau de performance comparable. Pourtant, des enjeux scientifiques, technologiques, économiques et sociaux très importants en découlent, la parole étant envisagée comme une interface naturelle entre l'Homme et la machine.

Percevoir le langage parlé revient à :

- ▷ identifier chaque seconde 10 à 15 phonèmes, 3 à 4 syllabes et 2 à 3 mots (soit, à la minute, environ 750 phonèmes, 200 syllabes, 150 mots). Cela implique donc d'être capable de retrouver immédiatement et avec un taux d'erreur quasi nul, les informations formelles (phonologie, morphosyntaxe) et conceptuelles (syntaxe, sémantique, pragmatique) rattachées à chaque mot, dans un stock moyen de 50 000 à 100 000 mots mémorisés (lexique mental)
- ▷ analyser le message produit par le locuteur au fur et à mesure qu'il est énoncé pour en comprendre le sens (intégration syntaxique, sémantique et pragmatique).

Il s'agit donc d'un processus en deux étapes successives :

- ▷ reconnaître des mots présents dans le flux sonore (accès au lexique mental, ou activation des informations lexicales représentées en mémoire à long terme pour les rendre disponibles dans la mémoire de travail)
- ▷ construire le sens du message à partir des informations activées depuis le lexique mental à l'étape précédente (élaboration en mémoire de travail d'une information non représentée en mémoire à long terme, et propre à la situation de communication).

Le travail de recherche présenté dans ce manuscrit s'est focalisé sur la première phase du processus, la reconnaissance des mots parlés. Elle se produit sous l'influence de 2 sources d'information :

- ▷ La première source est extérieure au système de traitement. Il s'agit de l'information acoustique, structurée de façon complexe sur plusieurs niveaux (information linguistique segmentale et suprasegmentale).

- ▷ La deuxième source d'information est interne au système. Il s'agit de l'information représentationnelle, qui correspond aux connaissances linguistiques stockées en mémoire à long terme, et aux règles de fonctionnement du système qui influencent sa dynamique de fonctionnement.

Le travail de recherche présenté dans ce manuscrit s'est plus particulièrement focalisé sur les aspects du traitement relatifs à la démarcation des unités lexicales : guidage ascendant de la reconnaissance des mots par l'information prosodique, et guidage descendant de la reconnaissance des mots par la structure superposante du lexique, séparément et en interaction.

Le but de cette double focalisation est de mieux comprendre l'un des points essentiels rattachés à la reconnaissance des mots parlés : la segmentation du flux de parole en mots. Segmenter le flux de parole en mots implique en effet de localiser les frontières lexicales dans le flux de parole. Or, ce phénomène pose un problème de compréhension.

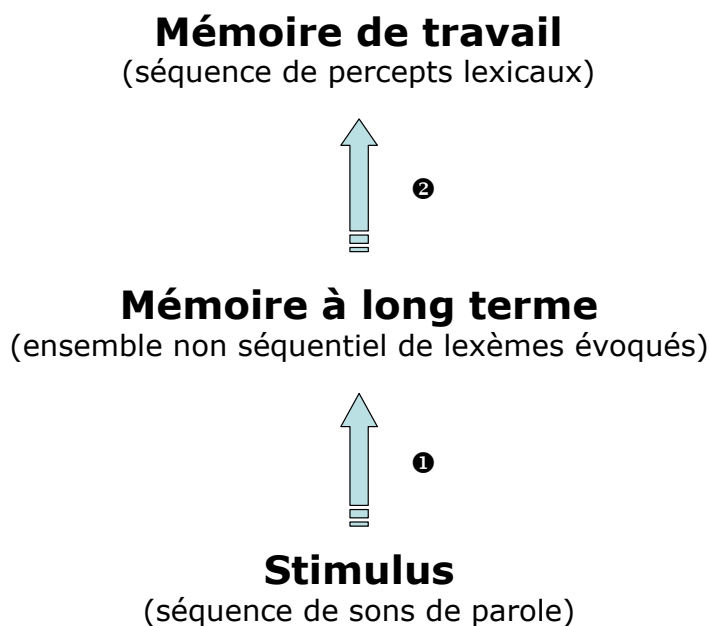
## **1.2 Localiser les frontières lexicales dans le flux de parole : un écueil théorique et pratique**

Le problème de la reconnaissance des mots parlés tel qu'il est souvent abordé, en psycholinguistique comme en reconnaissance automatique de la parole, gravite autour du problème de la localisation des frontières lexicales dans un flux de parole continu. Contrairement à l'écrit où les frontières entre mots sont systématiquement marquées par un élément caractéristique (espace vide), l'oral ne présente pas de corrélats acoustiques spécifiques et systématiques de frontière lexicale. L'onde sonore peut par exemple ne pas être interrompue entre deux mots, ou l'être au sein d'un mot (phonèmes « plosifs »), sans que la perception discrète des unités lexicales ne soit altérée.

Pour générer un percept lexical, le système naturel procède probablement à partir d'une séquence de représentations phonémiques (issue de l'analyse perceptive du stimulus) et des connaissances lexicales dont il dispose (lexique mental).

Le système effectuerait une série d'opérations visant à apparier les lexèmes contenus dans le lexique mental avec une séquence infra-lexicale de représentations phonémiques, procédure qui implique une transition d'états logiques de type OU-ET-OU (passage d'une organisation séquentielle (sons de parole) à une organisation séquentielle (percepts lexicaux) via une organisation non séquentielle (lexique mental)), comme représenté dans la Figure 1.

Figure 1 : Déroulement schématisé de la reconnaissance des mots



À cause de l'utilisation combinatoire du code phonémique pour générer les mots (deuxième articulation), les lexiques des langues humaines sont des structures très fortement superposantes. Cela signifie que les mots qu'ils renferment peuvent partager des caractéristiques formelles communes (un même phonème ou une même séquence de phonèmes). Deux catégories de superpositions formelles peuvent être recensées :

- ▷ Les superpositions non lexicales. Les mots partagent une même séquence de phonèmes, qui n'est pas lexicale (exemple : abc et bcd ont en commun la séquence bc qui n'est pas un mot).
- ▷ Les superpositions lexicales. Les mots entrent dans un rapport d'inclusion mutuelle (exemple : le mot abcd renferme le mot abc).

Les superpositions non lexicales concernent tous les mots d'une langue, tandis que les superpositions lexicales en concernent la majorité. On estime par exemple que moins de 2% des mots de l'Anglais ne sont pas impliqués dans une superposition lexicale (Cutler, McQueen, Jansonius, & Bayerl, 2002).

Or, l'existence de superpositions formelles complique le processus de reconnaissance des mots. Les superpositions non lexicales augmentent le degré de confusion entre mots, tandis que les superpositions lexicales entraînent un rapport équivoque entre séquence de sons de parole et séquence de mots, ambiguïté dont résulte le problème de la localisation des frontières de mots dans le flux de parole. Par exemple, une séquence de sons de parole a-b-c-d (stimulus) peut évoquer les lexèmes



abcd, a, ab, cd, et bcd en mémoire à long terme. Le système doit donc « linéariser » le lexique, c'est à dire faire un choix parmi ces alternatives, pour générer une séquence de percepts lexicaux s'excluant mutuellement (exemple : ab + cd), et étant en adéquation avec la séquence de mots produite.

L'écueil théorique et pratique que pose ce traitement est relatif aux procédés par lesquels le système naturel parvient, avec un taux de performance maximal jamais égalé par les systèmes artificiels, à localiser les frontières des mots : quels sont ces procédés ?

### **1.3 Quelles solutions ? Introduction au travail effectué**

Le système peut appairer directement le lexique mental sur l'information acoustique sensorialisée, ou appairer le lexique mental sur une information représentationnelle infra-lexicale, phonétique ou phonologique.

Dans le premier cas, le système dispose d'une précision maximale de l'entrée (richesse maximale de l'information), ce qui lui permet de ne pas activer les lexèmes qui ne correspondent pas aux mots produits.

Dans le deuxième cas, il ne dispose que d'une information filtrée, donc appauvrie<sup>2</sup>. Il doit alors résoudre la structure superposante du lexique sur la base de procédures algorithmiques, qui visent à faire émerger les lexèmes les plus probables de l'ensemble des lexèmes en compétition pour la reconnaissance. Ces procédures sont cependant coûteuses en énergie, et parfois en temps.

La psycholinguistique a longtemps privilégié les systèmes appariant le lexique sur une information représentationnelle. La conclusion à laquelle elle parvient avec ce type de modèle est qu'un mot impliqué dans des superpositions lexicales est plus difficile à traiter qu'un mot qui n'est pas impliqué dans des superpositions lexicales (exemple : TRACE (McClelland & Elman, 1986)). Aucun modèle ne prédit qu'une telle configuration soit plus facile à traiter.

Étant donné l'universalité des superpositions lexicales dans les langues humaines, il semble étrange qu'un mot impliqué dans ce type de configuration soit plus coûteux à traiter qu'un mot qui ne l'est pas. On peut en effet s'interroger sur la pertinence d'un

---

<sup>2</sup> L'intérêt de travailler à partir d'une information représentationnelle est d'évacuer le problème de la variabilité de la parole et de la coarticulation des phonèmes.

système qui serait en difficulté à chaque mot à reconnaître, ou qui produirait des percepts lexicaux sans rapport avec les mots produits par le locuteur. Comment un tel système se serait-il seulement développé au cours de l'évolution ? De plus le choix d'évacuer certaines informations présentes dans le signal de parole par filtrage phonétique ou phonologique, pour des raisons techniques plutôt que théoriques, signifie que les modèles de la reconnaissance des mots parlés sont peut-être incomplets. En réalité, les superpositions lexicales sont peut-être une source de facilitation de la reconnaissance des mots parlés ?

Partant de ces observations et interrogations, ce travail se propose :

- ▷ D'explorer la dimension suprasegmentale de la source externe d'information, et de tester les effets qu'elle produit sur l'organisation du traitement perceptif et lexical. Dans ce but, le travail a d'abord consisté à rechercher dans le signal acoustique la présence d'un code suprasegmental susceptible de démarquer les unités lexicales (Analyses acoustiques, Simulations 1 & 2). Ensuite, il a fallu montrer que l'information ainsi isolée est utilisable (Expérience 1) et utilisée (Expérience 2) pour la localisation des frontières de mots. Enfin, quelques pistes ont été proposées pour comprendre comment la prosodie guide le décodage de la parole (Expérience 3).
- ▷ De tester l'incidence réelle des superpositions lexicales sur la reconnaissance des mots parlés, afin de déterminer si elles ont effectivement un impact défavorable (Expériences 4 à 7). Dans un deuxième temps, il a fallu chercher à comprendre quels mécanismes sont en jeu lors du traitement de ce type de configuration (Expériences 8 à 14, Analyse exploratoire).
- ▷ De tester l'influence combinée de la prosodie et des superpositions lexicales sur la prégnance perceptive du mot porteur (Expérience 15) et du mot enchâssé (Expérience 16), afin de savoir si la prosodie contraint les traitements lexicaux, ou si elle remplit seulement un rôle de balisage sensoriel en vue de la segmentation du flux sonore en mots. Par ailleurs, le matériel utilisé dans les expériences relatives aux superpositions lexicales a aussi fait l'objet d'analyses acoustiques afin de tenter de préciser le lien entre prosodie et lexique.

Ce travail s'achève avec la proposition d'un modèle mathématique de prégnance perceptive des mots parlés, SLEXS. Ce modèle laisse entrevoir une réponse possible à l'écueil de la segmentation lexicale. Cette dernière dépendrait directement de la prégnance perceptive des mots, liée tout à la fois et entre autres, à la ligne suprasegmentale du signal de parole et à l'organisation superposante du lexique.

## **1.4 Présentation de l'organisation du manuscrit**

Le manuscrit est organisé en 6 parties. La première partie correspond à cette introduction, la deuxième partie expose les aspects théoriques fondant ce travail, la troisième partie expose la méthodologie suivie, la quatrième partie expose les

développements expérimentaux réalisés, la cinquième partie développe une discussion sur les résultats obtenus, et la sixième partie ou conclusion clos le travail.

De manière plus détaillée :

1. La partie théorique (Partie II) est subdivisée en 4 chapitres.
  - ▷ Le premier chapitre présente de manière synthétique les connaissances dont on dispose actuellement sur le fonctionnement du système naturel de reconnaissance des mots parlés (architecture logique du système), et présente également un modèle implémenté, le modèle TRACE, utilisé comme pôle de référence dans ce travail.
  - ▷ Le rôle démarcatif de l'information prosodique dans la reconnaissance des mots parlés est exploré dans le deuxième chapitre, les questions portant sur le niveau du traitement où la prosodie est susceptible d'être utilisée dans le processus, et les bénéfices qu'elle est susceptible d'apporter à la reconnaissance des mots. L'idée d'un code prosodique porté par la syllabe, ou code prososyllabique, et visant à guider le traitement selon un mode ascendant est proposée.
  - ▷ Le troisième chapitre de la partie théorique aborde le rôle des superpositions lexicales dans la reconnaissance des mots parlés. Y est présenté un ensemble de statistiques lexicales décrivant la structure superposante du lexique français et la répartition des superpositions lexicales dans les mots, une revue de question sur les recherches en psycholinguistique concernant l'effet des superpositions lexicales sur la reconnaissance des mots, et les résultats de simulations de TRACE sur ce type de configuration lexicale.
  - ▷ Le quatrième chapitre récapitule les idées clé.
2. La partie méthodologique (Partie III) est subdivisée en 5 chapitres.
  - ▷ Le premier chapitre détaille les problèmes méthodologiques rencontrés dans l'étude.
  - ▷ Le deuxième chapitre traite du problème de l'origine de la mesure des temps de réponse lors de la reconnaissance d'un mot parlé isolé.
  - ▷ Le troisième chapitre est relatif à la méthodologie statistique mise en place pour l'analyse des temps de réaction.
  - ▷ Le quatrième chapitre détaille la méthodologie expérimentale elle-même.
  - ▷ Le cinquième chapitre expose un récapitulatif des aspects théoriques et techniques mis en œuvre à travers l'ensemble des expériences et analyses effectuées dans la thèse.
3. La partie expérimentale (Partie IV) est subdivisée en 5 chapitres.
  - ▷ Le premier chapitre présente un ensemble d'analyses et d'expériences relatives à l'exploration du rôle des paramètres suprasegmentaux dans le guidage sensoriel de l'activation lexicale (Étude préalable, Expériences 1 à 3).
  - ▷ Le deuxième chapitre présente un ensemble d'analyses et d'expériences relatives à la détection d'éléments impliqués dans des superpositions lexicales (Expériences 4 à 14, Analyse exploratoire).
  - ▷ Le troisième chapitre explore de manière combinée l'impact des aspects suprasegmentaux du signal et architecturaux du lexique sur le traitement lexical (Expériences 15 et 16).

- ▷ Le quatrième chapitre explore les données acoustiques issues d'une analyse des items expérimentaux utilisés dans les Expériences 4 à 7 et 11 à 14, afin d'évaluer la relation entre prosodie et structure lexicale dans des configurations spécifiques de superpositions lexicales.
  - ▷ Le chapitre 5 récapitule les résultats obtenus dans la partie expérimentale.
4. La discussion (Partie V) synthétise les effets obtenus dans ce travail, en discute les implications théoriques, et fournit quelques propositions pour les expliquer. Un modèle de prégnance perceptive est présenté dans cette cinquième partie, pour rendre compte des phénomènes paradoxaux obtenus expérimentalement dans ce travail de recherche. Enfin, les points à retenir de ce travail sont récapitulés, et les développements ultérieurs à lui donner sont introduits.

## 2 Partie II – Théorie

La partie théorique présente les bases à partir desquelles ce travail de recherche a été élaboré. Elle est organisée en trois chapitres :

- ▷ Chapitre 1 – Modèles de la reconnaissance des mots parlés
- ▷ Chapitre 2 – Prosodie et démarcation des unités lexicales
- ▷ Chapitre 3 – L'organisation superposante du lexique : quel rôle sur la reconnaissance des mots parlés ?

Le premier chapitre récapitule de façon synthétique les connaissances dont on dispose actuellement en psycholinguistique sur la reconnaissance des mots parlés. Il présente ensuite un modèle implémenté de reconnaissance des mots parlés, le modèle TRACE, ainsi que les résultats d'une simulation de ce modèle lors du traitement d'un mot impliqué dans des superpositions lexicales.

Le deuxième chapitre introduit succinctement la prosodie (aspects acoustiques, linguistiques, et fonctionnels). La prosodie étant un objet vaste et complexe, la problématique est ensuite précisée (unité porteuse du code prosodique, rôle de ce code dans la démarcation lexicale). Les différentes hypothèses sur la manière dont la démarcation lexicale pourrait s'effectuer en Français sont passées en revue et illustrées par quelques travaux expérimentaux. Enfin, les travaux de recherche en reconnaissance automatique de la parole du CLIPS-IMAG et de l'IPC, à partir desquels la partie de ce travail relative à l'existence d'un code prosodique a été élaborée, sont rapidement présentés.

Le troisième chapitre présente des données sur l'organisation superposante du lexique du Français, ainsi qu'un comparatif avec d'autres langues européennes, afin de montrer l'importance du phénomène et son universalité dans les langues humaines. Des travaux de recherche sur l'impact qu'une telle structure a sur la reconnaissance des mots parlés sont ensuite présentés. Les prédictions de quelques modèles centraux en psycholinguistique de la reconnaissance des mots parlés, relatives au comportement du système face à des superpositions de mots, sont ensuite proposées.

La partie théorique s'achève par un récapitulatif des idées fondamentales à la base de ce travail de recherche, et pose les deux hypothèses autour desquelles s'est articulée la partie expérimentale.

## 2.1 Chapitre 1 – Modèles de la reconnaissance des mots parlés

Ce premier chapitre présente synthétiquement les connaissances actuelles dont on dispose en psycholinguistique sur le phénomène de la reconnaissance des mots parlés (Sections 2.1.1 et 2.1.2). Il présente ensuite le modèle TRACE (Section 2.1.3 page 32), utilisé dans ce travail comme référence (architecture et principes de fonctionnement), ainsi que les prédictions de ce modèle face à une configuration de type superposition lexicale. Le Chapitre 1 se termine par un bref récapitulatif des idées clés qu'il contient (Section 2.1.4 page 41).

### 2.1.1 Généralités sur les modèles psycholinguistiques de la reconnaissance des mots parlés

La modélisation d'un phénomène cognitif nécessite 3 niveaux de conceptualisation (Marr, 1982; Rumelhart & McClelland, 1985) :

- ▷ Le niveau implémentational (quel dispositif physique permet d'effectuer le calcul déterminé au niveau représentationnel ?)
- ▷ Le niveau représentationnel (comment sont calculées les fonctions définies au niveau calculatoire ?)
- ▷ Le niveau calculatoire (quelles fonctions sont réalisées par le système ?)

Au niveau implémentational, la plupart des modèles psycholinguistiques de la reconnaissance des mots parlés sont des dispositifs connexionnistes<sup>3</sup> (TRACE (McClelland & Elman, 1986), SHORTLIST (Norris, 1994), DCM (Gaskell & Marslen-Wilson, 1997), PARSYN (Luce, Goldinger, Auer, & Vitevitch, 2000), FUL (Lahiri & Marslen-Wilson, 1991; Reetz, 2000), RAW (Dashorst, 1997; van Kuijk, Wittenburg, & Dijkstra, 1996)), les autres modèles n'étant pas implémentés (modèles « verbaux » - COHORT (Marslen-Wilson, 1987; Marslen-Wilson & Welsh, 1978), modèle de Forster (Forster, 1976) ;

---

<sup>3</sup> Réseaux de neurones artificiels (analogie neurobiologique). Les réseaux neuromimétiques sont des approximateurs universels (toute fonction bornée peut être approchée uniformément par un réseau de neurones comportant une couche cachée) parcimonieux (obtention de l'approximation la plus satisfaisante avec un nombre minimum de paramètres (Dreyfus et al., 2002)). La connaissance est contenue de 2 manières dans ces dispositifs : le niveau d'activation des unités (mémoire à court terme – chaque unité ou groupe d'unités correspond à une proposition dont la véracité est proportionnelle au degré d'activation des unités), et la propension des connexions entre unités à transmettre l'activation (mémoire à long terme). Les propriétés fondamentales des réseaux connexionnistes sont essentiellement perceptives, mémorielles et associatives (exemples : mise en correspondance linéaire ou non linéaire de 2 espaces, adaptation, intégration contextuelle, robustesse, tolérance au bruit et généralisation (reconnaissance de formes dégradées, prédiction), compression de données, classification et décision, etc.), mais il est possible d'obtenir des traitements symboliques avec une architecture appropriée (Haykin, 1999).

modèles « mathématiques » - modèle des « Logogens » (Morton, 1969), FLMP (Massaro, 1989; Massaro & Oden, 1980), NAM (Luce & Pisoni, 1998). NAM a été ultérieurement implémenté sous la forme du modèle PARSYN), ou l'étant sous une autre forme (exemple du modèle de Klatt, LAFS (Klatt, 1979), proche de l'ingénierie, implémenté sous la forme d'un programme non connexionniste de recherche de mots dans un lexique, à partir de l'analyse de formes spectrales).

Au niveau représentationnel, les systèmes connexionnistes posent un problème d'intelligibilité à cause de leur opacité (relation obscure entre le niveau physique et le niveau représentationnel<sup>4</sup>). Cette difficulté est généralement contournée par l'adoption d'un mode localiste de représentation<sup>5</sup> de l'information dans le dispositif, qui permet de spécifier un rapport transparent entre entités physiques et linguistiques, et d'intégrer des règles transparentes de fonctionnement dans le dispositif<sup>6</sup>. Le système tend ainsi à devenir symbolique et modulaire<sup>7</sup>, en dépit de sa structure en réseau. Au niveau

---

<sup>4</sup> Un système de ce type peut disposer de représentations dynamiques, émergeant grâce à des phénomènes d'auto-organisation et de réverbération (Amit, 1995). Dans ce cas, les représentations ne sont pas des entités abstraites stockées physiquement, mais des propriétés émergentes générées par le fonctionnement du système. De plus, un système connexionniste ne présente pas forcément une association stricte entre niveaux de représentation et étapes de traitement, du fait que les niveaux de représentation peuvent s'intégrer les uns dans les autres à un même niveau physique.

<sup>5</sup> Une unité physique du réseau correspond à une unité abstraite représentée. Les unités physiques sont structurées en niveaux correspondant à des niveaux abstraits.

<sup>6</sup> Cependant, le mode localiste de représentation peut aboutir à un effet pervers sur la conception psychologique du traitement lorsque le mot est considéré comme une unité et non plus comme un élément composite. Une limitation des propriétés dynamiques du modèle obtenu peut en découler, un modèle connexionniste à représentation non localiste disposant d'une plasticité comportementale plus importante qu'un modèle connexionniste localiste (variation du traitement selon les situations – exemple : reconnaissance directe de l'unité lexicale pour un débit rapide, analyse infra-lexicale préalable pour un débit lent - Les observations de (Radeau, Morais, Mousty, & Bertelson, 2000) concernant les effets de point d'unicité en fonction du débit d'élocution laissent penser que le système dispose de plusieurs manières de traiter l'information). Cette plasticité comportementale serait liée à la présence d'attracteurs phonologiques et lexicaux dans le système (McLeod, Plaut, & Shallice, 2001; McLeod, Shallice, & Plaut, 2000). Les niveaux de représentation déduits d'observations comportementales correspondraient, dans ce cas, à une granularité statistique dans la mémoire et dans le stimulus. Un modèle connexionniste à représentation localiste ne peut réagir que d'une seule manière (niveaux successifs d'analyse imputables à la présence de niveaux de représentation physiquement distincts dans la structure du dispositif - (Sendlmeier, 1995)).

Il faut aussi remarquer que ce mode de représentation correspond à un codage de l'information particulièrement peu économique et mal adapté à des structures complexes devant renfermer un nombre élevé d'éléments (redondance élevée).

<sup>7</sup> Un système connexionniste dont le mode de représentation confonds niveau physique et niveau représentationnel peut être qualifié de modulaire. Un système connexionniste peut avoir un mode de représentation modulaire parce que chaque unité de traitement code pour une unité abstraite, ou parce que les unités de traitement sont hiérarchisées en niveaux distincts, chaque niveau physique codant pour un niveau représentationnel.

Un système connexionniste dont le mode de représentation n'est pas défini par un rapport univoque entre niveau physique et niveau représentationnel est qualifié de distribué.

représentationnel, il existe donc deux <sup>8</sup> types de modèles à implémentation connexionniste :

- ▷ les modèles à représentations localistes (exemples : TRACE (traits, phonèmes, mots), PARSYN (phonèmes, phonèmes, mots), FUL (traits, mots), RAW (phonèmes, mots))
- ▷ les modèles à représentations distribuées <sup>9</sup> (pas d'exemple en psycholinguistique – certains auteurs font cependant appel au distributionnalisme des représentations lexicales pour justifier des effets liés aux variations de microstructure dans le signal de parole (Marslen-Wilson & Warren, 1994))

Seuls les modèles de premier type sont utilisés en psycholinguistique, car ils permettent d'ouvrir la « boîte noire ».

Au niveau calculatoire, enfin, les modèles s'accordent globalement sur les fonctions logiques qui doivent être remplies par le système.

Une synthèse de ces fonctions est donnée à la section suivante. Pour un descriptif exhaustif des modèles de la reconnaissance des mots parlés, se reporter aux manuscrits sur le sujet (livres, articles, thèses)<sup>10</sup>.

<sup>8</sup> Une troisième catégorie peut être proposée : les modèles hybrides (SHORTLIST (traits, couche cachée, mots), DCM (traits, couche cachée, représentation lexicale distribuée (sémantique distribuée et phonologique distribuée (séquence de phonèmes organisée en syllabes – attaque, noyau, coda))). Modèles connexionnistes qui adoptent un mode localiste de représentation dans certaines de leurs couches d'unités seulement (entrée et sortie).

<sup>9</sup> Le distributionnalisme est un mode de représentation qui permet d'encoder très efficacement la redondance de l'information (le nombre de représentations qui peuvent être mémorisées dans un réseau est supérieur au nombre d'unités qui composent celui-ci). Les connaissances sont mémorisées de manière économique dans le système et sont protégées en cas de lésion du dispositif (dégradation progressive vs. disparition avec un système localiste). La contrepartie fonctionnelle de ce mode de représentation, du fait qu'il rend continu l'espace de représentation, est la survenue possible de phénomènes de confusion (illusions). Enfin, il rend difficile le codage de représentations abstraites dans la structure, ainsi que le décodage des cartes d'activation à travers la structure.

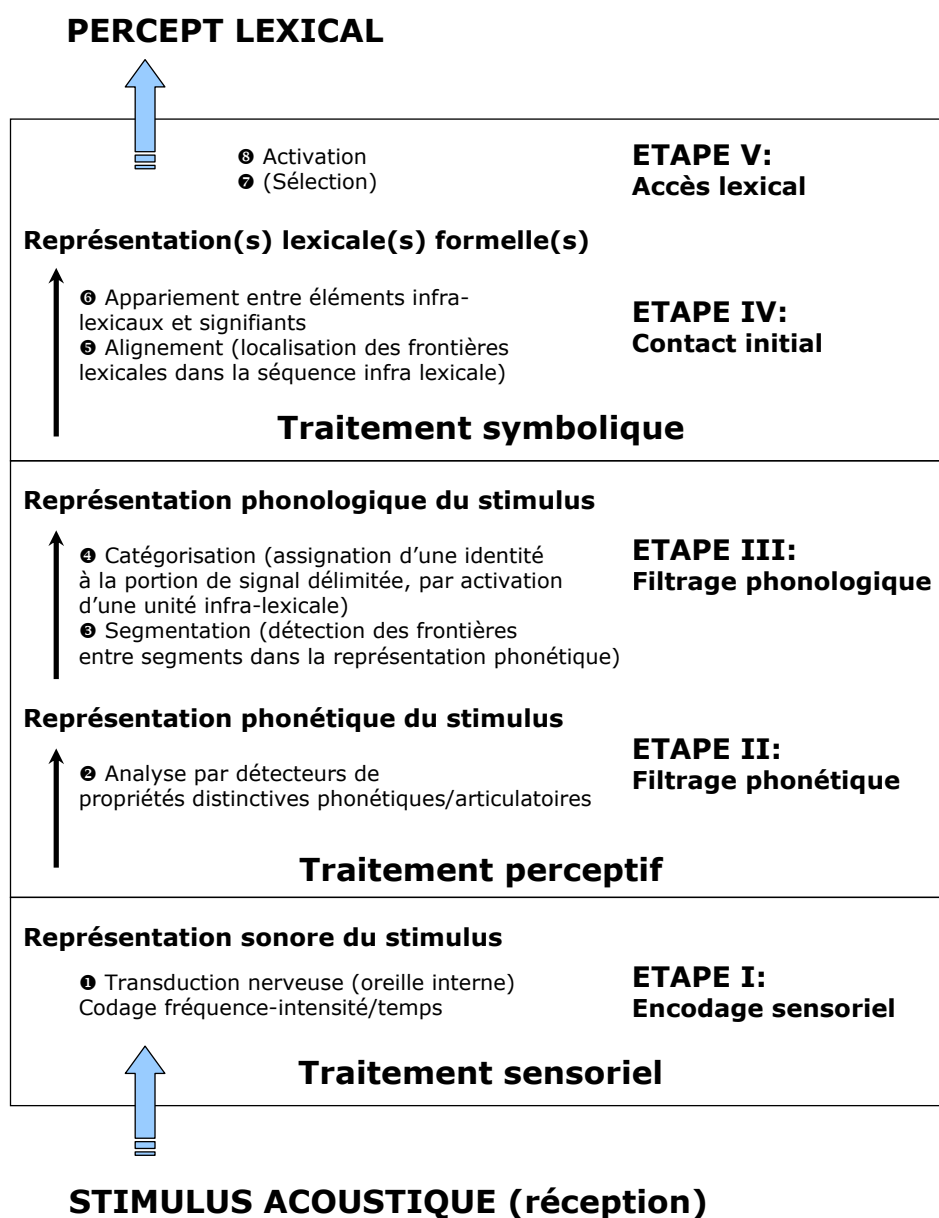
<sup>10</sup> Remarque : les divers modèles correspondent souvent à une alternative au niveau représentationnel, voire implémentationnel, d'un même niveau calculatoire.



## 2.1.2 Fonctions du système de reconnaissance des mots parlés au niveau calculatoire

La Figure 2 ci-dessous fournit un plan d'ensemble sur la logique supposée du déroulement de la reconnaissance d'un mot à partir d'une stimulation sonore. Ce plan est organisé selon 3 types de traitement (sensoriel, perceptif, symbolique), en 5 étapes (macro structure), dans chacune desquelles des fonctions logiques ou processus s'opèrent (micro structure). Le texte qui suit ce schéma décrit chacune des étapes en détail.

Figure 2 : Déroulement détaillé du processus d'accès ascendant au lexique (modèle calculatoire)



### 2.1.2.1 Traitement sensoriel

L'information acoustique reçue (captée par l'oreille externe et transmise par l'oreille moyenne) est traduite dans un format nerveux par l'oreille interne (transduction nerveuse (❶) - Étape I), en même temps qu'elle est analysée spectralement en termes d'amplitude de fréquence dans le temps (codage tonotopique), donnant lieu à une représentation sonore de nature sensorielle (carte sensorielle).

### 2.1.2.2 Traitement perceptif

Le traitement perceptif peut être envisagé comme l'action d'un filtre basé sur des détecteurs, constitué de 2 sous-filtres (phonétique et phonologique) servant à normaliser l'information<sup>11</sup>. La représentation sonore est d'abord analysée par des détecteurs de traits distinctifs (neurones – sous-filtre phonétique)<sup>12</sup>, pour être décomposée en constituants segmentaux élémentaires (traits distinctifs correspondant aux contrastes phonologiques minimaux – Étape II / détection de propriétés distinctives (❷)). Le but de ce type de traitement est d'extraire une information abstraite (représentation phonétique) afin de résoudre le problème de la variabilité de la parole (prosodie, timbre, accent régional, particularités de la prononciation du locuteur, etc.). Ce premier sous-filtre est responsable de la perte de l'information supra-segmentale, qui aurait pu être utile pour guider le processus d'alignement lors de l'Étape IV. Une procédure de reconnaissance de phonèmes, voire de syllabes<sup>13</sup> (sous-filtre phonologique), a ensuite lieu à partir de ce niveau des traits (Étape III / segmentation (❸) et catégorisation (❹))<sup>14</sup>. Le traitement perceptif de l'Étape III correspond à la résolution d'un problème de

<sup>11</sup> La normalisation de l'information s'effectue par la détection des propriétés idéales dont un prototype représentationnel est doté. L'activation des représentations, selon ce principe, s'effectue de manière graduelle, proportionnellement au degré de correspondance du stimulus avec le prototype (inactivé = complètement divergeant ; totalement activé = correspondance complète).

<sup>12</sup> Mise en évidence :

- au niveau comportemental par la procédure d'adaptation sélective (par exemple : (Eimas, Cooper, & Corbit, 1973; Tartter & Eimas, 1975))
- au niveau électrophysiologique (par exemple : (Miller, Escabi, & Schreiner, 2001)).

<sup>13</sup> Les travaux de Kolinsky et collègues sur les illusions lexicales provoquées par le phénomène de fusion phonologique appuient l'idée d'un codage infra-lexical syllabique de l'information au cours du traitement auditif : la présentation dichotique de 2 pseudo-mots produit un percept lexical illusoire si la première syllabe de l'un est la première syllabe du mot perçu, et si la deuxième syllabe de l'autre est la deuxième syllabe du mot perçu (exemple : /bitɔ̃/ + /kɔʒu/ = /biʒu/). L'illusion est moins fréquente lorsque les syllabes des pseudo-mots ne correspondent pas aux syllabes du mot perçu (exemple : /kitɔ̃/ + /bɔʒu/, ou /kiʒɔ̃/ + /bɔtu/ - (Kolinsky, Morais, & Cluytens, 1995) ; (Kolinsky & Morais, 1996)).

<sup>14</sup> Il semble nécessaire d'envisager une analyse du stimulus en composantes infra-lexicales abstraites pour rendre compte de la perception de pseudo-mots et de mots nouveaux. La distinction entre niveaux de traitement

reconnaissance de formes<sup>15</sup>. Son but est d'extraire d'un signal coarticulé, des invariants linguistiques représentationnels élémentaires (représentations infra-lexicales), qui permettront d'asseoir l'étape suivante du traitement (traitement symbolique) sur une base fiable<sup>16</sup>. Toutefois, en cas d'erreur à ce niveau, les étapes ultérieures de l'accès lexical conduisent à des résultats aberrants<sup>17</sup> ou à l'échec pur et simple de la reconnaissance<sup>18</sup>, et la communication s'en trouve altérée (exploration d'hypothèses alternatives et compréhension d'un autre message que celui émis par le locuteur). Puisqu'il est souvent impossible d'identifier des phonèmes isolés de leur contexte (à cause de la coarticulation), on s'attendrait à ce que le système ne soit pas fiable. La rareté des erreurs en situation de communication naturelle laisse au contraire penser que le système dispose de procédures intégratives prenant en compte les caractéristiques du contexte gauche et droit pour procéder à l'identification des segments (interférences

---

par leur influence sur la reconnaissance des mots (infra-lexicaux facilitateurs, lexicaux inhibiteurs) semble appuyer l'idée d'une analyse infra-lexicale (par exemple : (Vitevitch & Luce, 1999; Vitevitch, Luce, Pisoni, & Auer, 1999)). Certains auteurs considèrent cependant qu'il n'est pas nécessaire de procéder à une analyse infra-lexicale d'éléments linguistiques comme les phonèmes au moment de la reconnaissance, ce type d'analyse n'étant nécessaire qu'au moment de la structuration du système lors de l'apprentissage de mots nouveaux (Klatt, 1979, 1981). Le traitement lexical se ferait directement à partir du niveau des traits distinctifs. Ceci semble accrédité par l'influence directe des modifications coarticulatoires du spectre d'une voyelle (vers la consonne suivante) sur le choix lexical d'auditeurs humains (visibles 25 à 50 ms. avant la fin de la voyelle - (Marslen-Wilson & Warren, 1994)).

<sup>15</sup> Il existe un débat quant à la nature du traitement des sons de parole (perceptif vs. auditif – voir aussi la note 14 page 21). Les sons de parole seraient traités différemment des sons non langagiers car :

- ils paraissent être spécifiquement traités par l'hémisphère gauche chez l'être humain (par exemple : (Shtyrov, Kujala, Palva, Ilmoniemi, & Naatanen, 2000))
- la capacité des individus à discriminer les sons de parole de leur langue s'effectue selon un filtre phonologique (« perceptual magnet effect » - par exemple : (Walley & Sloane, 2001))
- le fonctionnement de ce filtre phonologique est affecté par la quantité d'éléments qu'il renferme (par exemple : (Costa, Cutler, & Sebastian-Galles, 1998))
- ce filtre s'acquiert et est donc représentationnel (par exemple : (Anderson, Morgan, & White, 2003; McClelland, Fiez, & McCandliss, 2002)).

Cependant, des données montrent que la perception catégorielle des sons de parole existe chez les animaux, qui ne disposent pas d'un système de représentations phonologiques. Ce fait accréditerait l'idée d'un traitement auditif catégoriel des sons de parole, ou du fondement acoustique des contrastes phonologiques (par exemple (Kluender, 2000) et (Kluender, Lotto, & Holt, sous presse)).

<sup>16</sup> La coarticulation représente cependant une source de continuité et de cohésion acoustique du signal (composition spectrale), et est nécessaire à la perception des unités lexicales (la concaténation de segments non coarticulés donne un signal inintelligible).

<sup>17</sup> Activation du mauvais mot (exemple : identification du phonème /k/ à la place du phonème /g/, reconnaissance du mot /kʁɛ/ à la place du mot /gʁɛ/).

<sup>18</sup> Exemple : /kʁa/ à la place de /gʁa/.

latérales (considération de la coarticulation) et descendantes (syllabe, lexique)<sup>19</sup>, ou que la perception est réalisée directement à partir d'unités lexicales<sup>20</sup>.

### 2.1.2.3 Traitement symbolique

Le traitement symbolique s'effectue en 2 étapes :

- ▷ contact initial (activation d'un sous-ensemble de signifiants ou lexèmes dans le lexique)
- ▷ accès lexical (sélection d'un signifiant de ce sous-ensemble et activation du mot correspondant)

#### 2.1.2.3.1 Contact initial

Avec les constituants infra-lexicaux abstraits, le système dispose d'un code qui est supposé permettre la localisation d'éléments formels (signifiants ou lexèmes) dans le lexique (Étape IV - contact initial). Le contact initial peut être ciblé (c'est à dire ne concerner qu'un lexème) ou diffus (c'est à dire concerner une zone du lexique renfermant le lexème du mot à reconnaître et les lexèmes ayant des points communs avec), selon la richesse de l'information utilisée par le système (carte sensorielle vs. information représentationnelle phonétique ou phonologique) et le contenu du lexique. Le contact initial est réalisé par la mise en correspondance de la séquence d'unités représentationnelles infra-lexicales (clé) avec les unités représentationnelles du niveau

---

<sup>19</sup> La modélisation localiste entraîne un débat sur la manière de concevoir l'intégration des unités élémentaires dans les unités supérieures :

- Selon le courant interactiviste, des connexions descendantes (comme dans TRACE), depuis le niveau lexical sur le niveau phonémique, permettent une rétroaction de l'information, dans le but d'installer une résonance dynamique devant favoriser l'identification des éléments du niveau inférieur dans les cas de situation ambiguë (par exemple : (Elman & McClelland, 1988)).
- Selon le courant autonomiste, un flux descendant d'information depuis le niveau lexical sur les niveaux infra-lexicaux n'apporte rien de plus au processus de reconnaissance des mots. Il est introduit dans les modèles pour des raisons techniques visant à augmenter la performance du modèle pour reconnaître les phonèmes (Norris, McQueen, & Cutler, 2000a, 2000b).

Cette opposition n'est pas nécessaire dans un modèle basé sur un mode distribué de représentation, où les unités physiques du réseau qui sont impliquées dans une représentation de haut niveau, le sont aussi dans des unités de niveau inférieur. Cette optique sur l'intégration des niveaux de représentation les uns dans les autres est défendue par des auteurs comme Massaro (selon le modèle FLMP, les unités supérieures sont constituées de sous-unités, elles-mêmes constituées de sous-unités, jusqu'aux primitives sensorielles).

<sup>20</sup> Les représentations lexicales seraient composées de traits distinctifs, ce qui permettrait de rendre compte de l'effet perceptif de variations dans la microstructure du signal de parole (Marslen-Wilson & Warren, 1994). Selon les mêmes auteurs les effets phonologiques seraient liés à l'acquisition d'une conscience méta phonologique lors de l'apprentissage de la lecture (ils disparaissent chez les illettrés ou dans les langues qui ne présentent pas une correspondance transparente entre graphèmes et phonèmes (exemple du Chinois)), et la capacité du système à percevoir des pseudo-mots ou des mots nouveaux serait liée à une médiation phonologique par le lexique.

supérieur (serrure), à la manière de l'introduction d'une clé dans une serrure. Deux processus logiques permettent d'effectuer l'opération :

- ▷ l'alignement (⑤) : localisation, dans la séquence d'éléments du niveau immédiatement infra-lexical, d'un point d'entrée dans le lexique mental (ajustement de la clé en face de la serrure).
- ▷ l'appariement (⑥) : mise en contact du niveau immédiatement infra-lexical avec un signifiant du lexique mental (introduction de la clé dans la serrure).

### 2.1.2.3.1.1 **Alignement**

Deux types d'alignement sont envisagés :

- ▷ l'alignement initial
- ▷ l'alignement exhaustif

C'est au cours du processus d'alignement que la correspondance équivoque entre séquence infra-lexicale et éléments du lexique pose un problème en cas de superpositions lexicales (1 clé, plusieurs serrures). Le but des procédures d'alignement est donc d'éviter au système, autant que possible, de s'engager dans des « garden path » lexicaux<sup>21</sup>.

- ▷ l'alignement initial est préférablement ou uniquement effectué sur les phonèmes ayant une probabilité élevée de correspondre à une initiale de mot (le système sait que la clé ne peut aller que dans certaines des serrures qui sont à sa disposition et recherche celles qui ont la plus grande chance de correspondre). L'alignement initial implique que le système dispose de procédures de localisation des initiales effectives de mot (car tout phonème est une initiale de mot potentielle) ou de certaines initiales de mot seulement (toutes les initiales probables ne sont pas des initiales effectives – exemple : /p/ du mot « *charpente* »). On peut citer les procédures suivantes : MSS<sup>22</sup>

---

<sup>21</sup> Par exemple, la séquence de phonèmes /bɔʁdyʁ/ peut former un mot (une bordure) ou deux mots (un bord dur). Si le système ne dispose pas d'une autre source d'information (redondance phonématique et morphématique, prosodie, etc.) que la séquence d'éléments infra-lexicaux abstraits, son choix se fera au hasard et le conduira dans 50% des cas à une erreur d'alignement.

Cependant, on remarque que certaines séquences acoustiques présentées sans contexte conduisent quasi automatiquement à un « garden path » lexical – par exemple la séquence de phonèmes /maʃʃãdaʁ/ active le mot « marchandage » très fortement parce que la séquence de phonèmes, dernier non compris, concorde avec cette hypothèse lexicale. Le système doit réviser sa solution à l'arrivée du dernier phonème pour effectuer une segmentation lexicale différente (marchand d'art).

<sup>22</sup> MSS pour « Metrical Segmentation Strategy » : règle de segmentation lexicale de l'Anglais basée sur l'utilisation de la métrique (rythmicité), avec contact initial par les syllabes fortes (75% des syllabes fortes sont des initiales de mot en Anglais - (Cutler & Carter, 1987)). La partie de ce travail relative à la prosodie révèle qu'il pourrait exister un code prosodique au niveau de la syllabe en Français, plus complexe qu'une simple opposition de type fort/faible. Mattys et collègues (Integration of multiple speech segmentation cues : a hierarchical framework – en préparation) et Mattys (2004) remarquent que pour l'Anglais, l'influence de cette stratégie de segmentation serait limitée aux cas extrêmes où l'audibilité est fortement réduite. Dans les circonstances habituelles de communication, son influence serait donc secondaire par rapport à d'autres indices (traits phonétiques, phonotactique, lexicalité de la séquence de phonèmes, sémantique).

(Cutler & Norris, 1988; Norris, 1994; Norris, McQueen, & Cutler, 1995), SOSH<sup>23</sup> (Content, Kearns, & Frauenfelder, 2001; Dumay, Frauenfelder, & Content, 2002), PLSS<sup>24</sup> (L. A. Taft, 1984), PWC<sup>25</sup> (Cutler et al., 2002; Norris, McQueen, Cutler, & Butterfield, 1997), exploitation de la redondance du code linguistique<sup>26</sup>, utilisation des silences<sup>27</sup>. Ces procédures, utilisées en

<sup>23</sup> SOSH pour « Syllable Onset Segmentation Heuristic » : règle de segmentation basée sur la détection des initiales de syllabes (utilisation des variations de coarticulation entre segments (alternance faible-forte), supérieure entre 2 segments consécutifs d'une même syllabe qu'entre 2 segments consécutifs de 2 syllabes, et plus faible sur l'attaque pour ce qui est de la coarticulation entre le dernier phonème de la syllabe qui précède et l'attaque de la syllabe considérée, que sur le dernier phonème de la syllabe qui précède). Cette règle présuppose que les initiales de mots correspondent majoritairement aux initiales de syllabes (les cas de resyllabation ont une fréquence approximative de 5% en Français (estimation selon Fougeron, C. (2005). Communication personnelle). Ils concernent un faible nombre de consonnes finales (/n/, /z/, /t/, /p/, /k/) et de mots). Elles peuvent donc être traitées comme des exceptions, voir être représentées spécifiquement.

<sup>24</sup> PLSS pour « Post Lexical Segmentation Strategy » : règle de segmentation implicite, basée sur la reconnaissance du mot qui précède (fournit une information sur la localisation de la frontière lexicale initiale), l'identification des mots parlés procédant selon une temporalité « gauche-droite », et le début du premier mot correspondant au début de la phrase (selon Cole et Jakimik (1980)).

<sup>25</sup> PWC pour « Possible Word Constraint » : règle de segmentation basée sur la plausibilité lexicale du découpage de la séquence phonémique. Pour produire cet effet de manière ascendante, les hypothèses qui conduisent à un découpage renfermant une consonne isolée ou un groupe de consonnes sont éliminées (se reporter à la Section 2.1.2.3.1.2 page 27 relative à l'appariement : la règle PWC est une autre manière de concevoir un appariement syllabique). Cette règle permettrait d'éliminer 73% (Anglais) et 61% (Néerlandais) des mots enchâssés (exemple : « bell » dans « belt » - (Cutler et al., 2002)).

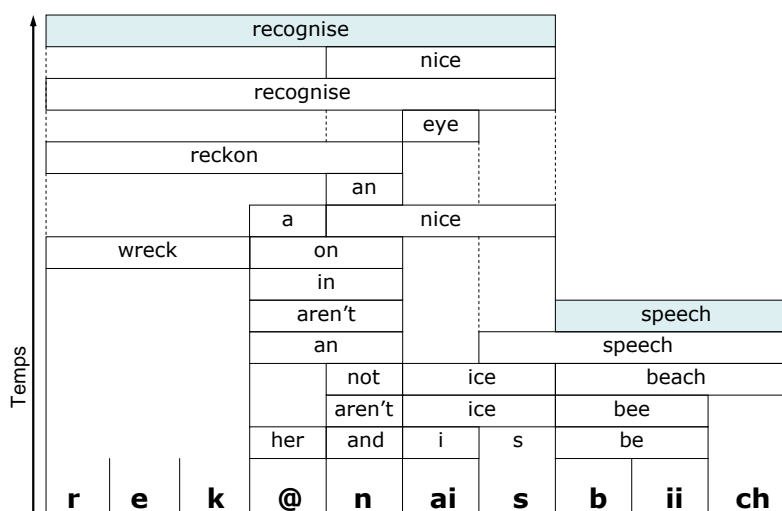
<sup>26</sup> Plus un système est redondant, plus il est organisé, plus il contient d'information utilisable, et plus il est stable. La redondance réduit donc le degré d'incertitude dans le code linguistique en organisant l'information selon des probabilités de cooccurrence dès les plus bas niveaux, déterminant ainsi des contraintes sur l'organisation séquentielle du signal (redondance phonématique (organisation des phonèmes dans les syllabes et des syllabes dans les mots), redondance morphématique (organisation syntaxique et sémantique des mots dans les phrases)).

- la redondance phonématique permet au système de prédire la survenue d'un segment à partir des 5 phonèmes précédents (mémoire d'ordre 5 dans les langues humaines naturelles), et ce faisant, influence la phase du contact initial. Cette prédictibilité a une incidence sur le traitement : elle facilite la segmentation du flux sonore en mots (une séquence de phonèmes fortement prédictible présentera une cohésion perceptive élevée, et la probabilité d'être en présence d'une frontière lexicale sera faible (exemple : [jō] dans [ɛfɔkmasjō] – exemple : (McQueen, 1998)). Une séquence de phonèmes faiblement prédictible présentera une cohésion perceptive faible ou nulle, et la probabilité d'être en présence d'une frontière lexicale sera élevée ou maximale (exemple : [ɜf] dans [bɑɜɜfɪltɜɑ]). La redondance phonématique génère donc une hétérogénéité perceptive de la séquence de phonèmes, utile pour contraindre la localisation des frontières lexicales (40.8% des frontières lexicales peuvent être correctement détectées (Xanthos, 2004) par la procédure proposée par Harris (« predictability strategy » - (Harris, 1955))).
- la redondance morphématique permet au système de prédire la survenue d'un mot à partir des mots qui précèdent (la mémoire porte sur des classes abstraites d'éléments qui peuvent prendre certaines positions dans la chaîne à l'exclusion de toute autre (exemple pour les classes formelles : après un article, le mot ne peut pas être un verbe. Une séquence de phonèmes comme /deʁɑʒe/ ne pourra donc pas être segmentée en /de/ et /ʁɑʒe/ bien que ces formes phonologiques correspondent séparément aux mots « des » et « ranger »). La redondance morphématique influence donc la sélection probabiliste d'un mot parmi un ensemble.

combinaison, permettraient de guider le système. Dans le cas d'un alignement initial, la linéarisation du lexique peut donc s'effectuer directement, sans une étape intermédiaire de sélection (accès direct).

- ▷ l'alignement exhaustif est effectué en aveugle sur chaque élément infra-lexical (le système essaie la clé dans toutes les serrures qu'il a à disposition<sup>28</sup>). Un alignement exhaustif signifie que le système n'est pas capable de localiser le début des mots et compense cette lacune par une exploration systématique des possibilités, comme illustré dans la Figure 3 (le temps nécessaire pour reconnaître « recognize » est supérieur à celui nécessaire pour reconnaître « speech », indépendamment de la durée des 2 mots, car la séquence de phonèmes correspondant au premier renvoie à un nombre de mots supérieur à la séquence de phonèmes du deuxième). Dans ce cas, il est supposé que la linéarisation du lexique nécessitera une procédure supplémentaire de sélection (⊙) des candidats à la reconnaissance au cours de l'Étape V (accès par sélection - détails à la Section 2.1.2.3.2 page 30).

Figure 3 : Exploration exhaustive du lexique selon un alignement sur chaque segment, lors de la reconnaissance automatique de la séquence « recognize speech » par une machine (reproduit à partir de (Shillcock, 1990))



<sup>27</sup> Les silences dont il s'agit ici correspondent plutôt à des marques de frontières entre syntagmes. En effet, il n'existe pas de correspondance stricte entre les vides acoustiques et les vides perceptifs : l'onde sonore peut présenter des interruptions au sein des mots à cause des caractéristiques de certains phonèmes (exemple des plosives non voisées) sans que l'unité perçue du mot ne soit altérée. Inversement, l'onde acoustique peut être continue entre deux mots sans que cela n'empêche la perception d'éléments distincts.

<sup>28</sup> L'exploration du lexique à partir d'un alignement exhaustif peut se faire séquentiellement (exemple présenté dans la Figure 3) ou parallèlement (exemple du modèle TRACE présenté à la Section 2.1.3 page 32). Quel qu'en soit le mode, elle prend d'autant plus de temps que la chaîne de phonèmes renvoie à un nombre élevé de mots, entraînant un allongement de la durée nécessaire pour reconnaître un mot :

- si le système procède par exploration séquentielle, la durée totale du traitement résulte de la sommation des durées de traitement de chaque hypothèse considérée dans le processus (relation de type  $F(t) = k.N + r$ , avec  $k$  le temps de traitement d'une hypothèse,  $N$  le nombre d'hypothèses considérées,  $r$  un résidu (R. D. Luce, 1986)).
- si le système procède par exploration parallèle, la durée totale du traitement n'est plus linéairement proportionnelle au nombre de cas examinés, mais à la difficulté du choix, qui dépend du nombre de possibilités et de leur degré de similitude.

L'alignement phonémique exhaustif apparaît comme une procédure excessivement coûteuse en termes de ressources, de calculs et de temps. Au contraire, un alignement exhaustif syllabique représente un compromis avantageux sur le plan énergétique pour le système, puisque la majorité des mots commence par l'initiale d'une syllabe (sauf cas de resyllabations – mais voir note 23 page 25).

### **2.1.2.3.1.2 Appariement**

Deux conceptions de l'appariement sont envisageables, relativement à la manière dont le système gère la temporalité du stimulus. Elles dépendent du niveau de représentation à partir duquel s'effectue le contact initial (P. Warren & Marslen-Wilson, 1987) :

- ▷ Appariement continu (clé de type « code numérique ») :

l'appariement s'effectue à partir d'un niveau de représentation sensoriel (exemple : LAFS (Klatt, 1979)) ou phonétique (exemple : FUL (Lahiri & Marslen-Wilson, 1991; Lahiri & Reetz, 2002; Reetz, 2000)), au fur et à mesure que le signal acoustique est reçu et traité. Au niveau lexical, cela se traduit par une variation progressive du degré d'activation des représentations concernées (fonction de leur degré de correspondance avec les caractéristiques du stimulus). L'intérêt de ce type d'appariement est de préserver un bon degré de richesse de l'information, autrement fortement dégradé par un filtrage phonologique (Klatt, 1989). Les phénomènes de granularité perceptive expérimentalement mis en évidence (effets de séquençage, de perception catégorielle, et effets syllabiques) trouveraient une explication dans des facteurs acoustiques et/ou dans l'organisation des représentations lexicales en mémoire (structure des représentations, structure du lexique).

- ▷ Appariement discret (clé de type « objet ») :

l'appariement s'effectue à partir d'un niveau phonologique de représentation (phonémique ou syllabique), c'est-à-dire par empans temporels correspondant à des empans mémoriels<sup>29</sup>. Ces empans mémoriels sont responsables de la granularité perceptive (discrétude) qui peut être observée expérimentalement. Si un gabarit phonémique est utilisé, le système peut présenter des effets de séquençage phonémique comparables à ceux obtenus dans le cas d'un appariement continu (exemple : effets de point d'unicité)<sup>30</sup>. Si un gabarit syllabique est utilisé, la directionnalité gauche-droite du flux d'information est

---

<sup>29</sup> Les représentations infra-lexicales peuvent donc être considérées comme une mémoire tampon, qui stocke des empans d'information acoustique.

<sup>30</sup> Remarque : la séquentialité temporelle d'un stimulus n'équivaut pas strictement à sa séquentialité phonémique, dans la mesure où les phonèmes possèdent des durées moyennes intrinsèques propres (variant d'un phonème à l'autre dans une proportion de 1 à 3 pour un débit donné), et qu'il peut s'agir d'une propriété pertinente pour le traitement. Un modèle à appariement phonémique ne correspond donc pas totalement à un modèle à appariement continu.



partiellement évacuée (empans de signal de plusieurs phonèmes)<sup>31</sup>, et les effets de séquençage phonémique ne sont plus possibles<sup>32</sup>.

Un appariement continu n'entraîne pas forcément une granularité perceptive. Si un modèle à appariement continu possède des représentations lexicales composites (phonème(s), syllabe(s), mot(s) représentés dans une même couche d'unités – il s'agit donc d'un système à représentations lexicales distribuées), le niveau de précision de l'analyse est fonction du degré de résolution de l'information acoustique, donc lié à la prosodie du locuteur. Le système est ainsi structurellement libre d'effectuer une reconnaissance d'unités infra-lexicales, en plus de la reconnaissance d'unités lexicales, ou non, en fonction des situations auxquelles il est confronté (voir note 6 page 18). Ce type de modèle peut alors présenter des effets de point d'unicité et de syllabe simultanément.

Par contre, un modèle à appariement discret ne peut procéder de la sorte : il est structurellement contraint à traiter les unités infra-lexicales selon un ordre hiérarchique, avant de traiter les unités lexicales. Il devrait donc présenter une granularité perceptive identique dans toutes les situations.

Si la manière dont s'effectue la projection d'information d'un niveau à l'autre (appariement discret ou continu) est relativement aisée à spécifier dans un modèle, ceci est plus délicat pour la dynamique du système<sup>33</sup>. En effet, la projection de l'information d'un niveau à l'autre doit absolument permettre de :

- ▷ conserver l'information infra-lexicale et lexicale jusqu'à ce que le mot soit effectivement reconnu, tout en libérant les détecteurs pour qu'ils soient de nouveau immédiatement opérationnels ( $aab \neq ab$ <sup>34</sup>). Une solution est la transmission de l'activation au niveau supérieur, avec réinitialisation

<sup>31</sup> Cependant, la reconnaissance des mots parlés s'effectue toujours selon un traitement gauche-droite. Nooteboom montre que la reconnaissance de mots parlés à partir de fragments initiaux conduit à un meilleur taux de réussite (.89) qu'une reconnaissance à partir de fragments finaux (.61), ces fragments renfermant suffisamment d'information pour distinguer le mot de tous les autres mots du lexique (Nooteboom, 1981).

<sup>32</sup> Par contre, un appariement syllabique rends compte à la fois de la règle de filtrage PWC (dans laquelle une consonne ou un groupe de consonnes ne représentent jamais une syllabe), de la prise en compte de la coarticulation pour la catégorisation des phonèmes, et d'un alignement syllabique de type SOSH.

<sup>33</sup> Aspects pour partie liés aux caractéristiques temporelles du stimulus, pour partie à l'évolution du traitement dans le temps.

<sup>34</sup> Exemples : coauteurs, nihilisme, féérique, maharaja. Toutefois, selon LEXIQUE (New, Pallier, Ferrand, & Matos, 2001), ce phénomène ne concerne que les voyelles /e/, /i/, /u/, /o/, /a/ (295 mots), et les consonnes /m/, /t/ et /s/ (39 mots - deuxièmement, honnêteté, transsibérien) – soit au total moins de 0.26% du contenu de la base de données. Le phénomène pourrait toutefois se rencontrer plus fréquemment lors du traitement de la parole en situation naturelle, puisqu'il implique aussi les frontières lexicales (exemple : attributs unificateurs), bien que plusieurs phonèmes conduisent à la production de liaisons empêchant une répétition immédiate (un ingrat, son ombre, ses états, etc.). Le rôle des liaisons et des resyllabations serait peut-être à chercher dans cette direction.

immédiate de l'unité émettrice. La rémanence des unités représentationnelles doit donc varier selon le niveau de représentation, en fonction de la durée (variable) de l'élément représenté. Elle doit ainsi être très brève au niveau des traits distinctifs et des phonèmes<sup>35</sup>, et varier au niveau des syllabes et des mots en fonction du nombre de phonèmes qui les constituent. Cependant, pour faire face aux cas de « garden path » lexicaux, le système doit disposer en mémoire, pendant suffisamment de temps, d'une trace des alternatives lexicales. La durée de rétention de l'information lexicale (et éventuellement syllabique ?) doit donc être supérieure à celle de l'unité lexicale, comme le souligne Mattys (1997)<sup>36</sup> (exemple de la séquence /maʁʃādaʁ/ - voir note 21 en page 23). Il faut donc, soit que la rémanence des unités lexicales soit supérieure à la durée du mot qui leur correspond, soit qu'elle soit conservée en mémoire de travail pour permettre de libérer le système perceptif sans perdre l'information nécessaire au traitement du message reçu. Cette dernière solution offre l'avantage supplémentaire de permettre un transfert économique d'information de la ou des représentations erronées à la ou aux représentations correctes.

- ▷ coder la position d'un élément infra-lexical dans la séquence (aab ≠ aba<sup>37</sup>). Ce point est souvent laissé de côté car implicitement lié à la séquentialité du signal, qui imprime sa propre dynamique sur la progression du traitement, et doit donc absolument être prise en considération par le système. Par exemple, dans le modèle TRACE, tous les mots qui commencent par un phonème donné sont activés, puis entrent en compétition mutuelle. Les mots qui renferment le deuxième phonème de la séquence présentée gagnent un meilleur score d'activation, et l'emportent donc sur le conflit qui les oppose avec les autres mots, etc. Les mots « *bâton* » et « *tomba* » ne sont donc pas confondus grâce à la dynamique du traitement, bien que tous deux reçoivent la même quantité d'activation depuis le niveau phonémique. Deux autres solutions pourraient toutefois être considérées pour résoudre le codage positionnel des éléments infra-lexicaux dans les éléments lexicaux :

- ∩ l'utilisation de représentations lexicales phonologiquement composites<sup>38</sup> (en remplacement ou non des niveaux infra-lexicaux de représentation), sur lesquelles le résultat du traitement sensoriel est directement projeté. La diffusion de l'activation serait alors ciblée (seuls sont activés les mots qui renferment un phonème à une position donnée). Ce type de modèle implique cependant qu'une procédure infaillible de localisation des

---

<sup>35</sup> Pour les phonèmes, la variation est fonction du phonème, du débit du locuteur, et de la position du phonème dans la chaîne parlée.

<sup>36</sup> Les observations de Grosjean semblent appuyer l'idée que le traitement lexical se poursuit au-delà de la fin acoustique du mot. Beaucoup de mots monosyllabiques devenant uniques après leur fin ne reçoivent un score de certitude maximal pour leur reconnaissance que durant le traitement du mot ou des mots suivants (Grosjean, 1985). Ce phénomène avec des mots monosyllabiques se justifie par leur propension à être enchâssés dans d'autres mots, donc à être impliqués dans des « garden path » lexicaux.

<sup>37</sup> Exemple : /asidyle/ - /elysida/.

<sup>38</sup> Il s'agit donc, du point de vue des unités infra-lexicales dans la représentation lexicale composite, d'une mémoire à exemplaires. C'est-à-dire que dans le mot /kaʁakas/, les 2 /k/ et les 3 /a/ sont des représentations distinctes. Chaque phonème peut, par ailleurs, être représenté sous une forme isolée au niveau lexical, s'il correspond à un mot (la reconnaissance du phonème [a] pour le mot « à » implique alors une autre représentation phonologique que celles contenues dans /kaʁakas/).

frontières ait préalablement lieu, dont nous n'avons pas connaissance à l'heure actuelle (influence possible de la prosodie combinée à d'autres facteurs + dynamique).

- ∩ l'existence d'une mémoire tampon entre niveaux distincts de représentation (phonèmes / syllabes / mots). Il s'agirait d'une mémoire de type échoïque permettant de stocker des éléments abstraits de durée variable (Mattys, 1997)<sup>39</sup>. L'écho renfermerait une information de rang, car cet espace de mémoire se remplirait au fur et à mesure de la reconnaissance des unités dont il est chargé de conserver la trace. Il se viderait dès qu'une unité supérieure correspond à son contenu (activation de l'unité supérieure). L'écho lexical pourrait renfermer plusieurs alternatives lexicales pour permettre une correction en cas de « garden path » lexical (voir plus haut le point relatif au maintiens de l'information en mémoire de travail).

L'organisation du modèle présentée dans la Figure 2 page 20 serait donc trop simple. Des modules / processus annexes pourraient venir s'y ajouter, ou modifier sa structure.

### 2.1.2.3.2 Accès lexical

L'accès lexical correspond à la dernière étape du traitement. Il aboutit à l'émergence du percept à la conscience du sujet, produit par l'activation de l'unité lexicale (⊗) via l'intégration du signifiant cible et des informations abstraites qui lui sont liées (signifié – substance (sémantique) et forme (syntaxe)). Dans l'analogie de la clé et de la serrure, la reconnaissance du mot peut être imagée par l'accès au contenu du coffre (signifié)<sup>40</sup>. Le percept lexical correspond à l'aboutissement du traitement, réussi lorsque le mot perçu est le mot produit par le locuteur.

Le système conscient ne pouvant gérer les mots que séquentiellement, un à la fois<sup>41</sup>, il y a obligation d'une phase de sélection avant l'élaboration d'un percept lexical si

---

<sup>39</sup> Selon Mattys (1997), les travaux des années 70 sur la perception auditive révèlent l'existence d'une mémoire sensorielle stockant une durée d'information de 250 ms. environ. La limite inférieure de résolution de cette mémoire sensorielle auditive serait comprise entre 120 et 170 ms. selon (Efron, 1970) (en-deca de cette résolution temporelle limite, les caractéristiques des stimuli fusionneraient et apparaîtraient indistinctes). Cet empan sensoriel correspondrait approximativement à l'ordre de grandeur de la durée syllabique. Toutefois, la durée d'une syllabe étant très variable (environ de 100 à 500 ms.), l'information pourrait plutôt être maintenue par un autre système de mémoire. Une mémoire de type échoïque est envisagée. Elle se baserait sur l'utilisation de circuits réverbérants (maintiens de l'activité neuronale dans une assemblée cellulaire après disparition du stimulus - (Amit, 1995)), et permettrait de conserver une trace de l'information phonologique une fois décodée, avec éventuellement la possibilité d'effectuer dynamiquement l'intégration de l'information.

<sup>40</sup> À ce niveau, des traitements supplémentaires doivent parfois encore avoir lieu pour résoudre le rapport équivoque entre signifiant et signifié (cas d'homophonie).

<sup>41</sup> Selon l'analogie de la clé utilisée pour illustrer le processus, la logique exclusive du système conscient serait imputable à l'incapacité de l'individu à utiliser plusieurs clés en même temps.

le contact initial est diffus (1 clé – plusieurs serrures). Si le contact est ciblé sur un seul signifiant (1 clé – 1 serrure), l'activation de cet élément peut avoir lieu sans autre (accès direct).

Le processus de sélection varie selon les modèles. Par exemple, il peut correspondre à un ciblage par réduction de la cohorte (Marslen-Wilson & Welsh, 1978) ou à une différenciation par niveau d'activation (McClelland & Elman, 1986). Quelle que soit la méthode utilisée, sélectionner un élément parmi d'autres implique de résoudre un problème de détection d'une cible au sein d'un groupe d'éléments similaires, desquels elle se démarque peu ou tardivement, afin de la faire émerger. La détection peut être envisagée comme découlant d'un processus :

- ▷ de facilitation ascendante portant sur la cible (séquençage phonémique gauche-droite ; l'information nouvelle renforce un des candidats de la cohorte).
- ▷ d'inhibition de l'environnement mental de la cible (filtrage plus ou moins actif des alternatives) :
  - ∞ inhibition ascendante : utilisation de la dynamique temporelle du signal pour éliminer les candidats qui deviennent inadéquats au fur et à mesure que l'information arrive (séquençage phonémique gauche-droite ; l'information nouvelle discrédite peu à peu les membres de la cohorte) – jusqu'à ce qu'il n'en reste plus qu'un (effet de point d'unicité<sup>42</sup> - modèles de type COHORT (Marslen-Wilson, 1987; Marslen-Wilson & Welsh, 1978))
  - ∞ inhibition latérale : utilisation de la dynamique interne du système de reconnaissance afin d'éviter qu'un élément de niveau immédiatement infra-lexical ne soit simultanément utilisé dans plus d'un élément du niveau lexical (les signifiants s'inhibent mutuellement et proportionnellement à leur degré d'activation, le moins inhibé et/ou le plus inhibiteur étant celui qui sera reconnu - modèles de type TRACE (McClelland & Elman, 1986))

Les procédures de sélection sont consommatrices d'énergie et de temps : il faudra d'autant plus de temps pour reconnaître un élément que l'ensemble duquel il doit être extrait, par facilitation ou par inhibition<sup>43</sup>, est important. Théoriquement, il apparaît donc intéressant de devoir éviter la phase de sélection ou de limiter l'étendue du contact

<sup>42</sup> Le rôle du point d'unicité serait cependant limité pour la sélection. Des statistiques obtenues à partir d'un lexique anglais de 20000 mots montrent que seulement 39% des mots ne sont pas uniques avant leur dernier phonème (probabilité pondérée par la fréquence d'occurrence des mots), 23% divergent au dernier phonème et 38% après leur fin (P. A. Luce, 1986).

<sup>43</sup> Facilitation ascendante : plus la cohorte de voisins est importante, plus il faut d'information pour caractériser la cible, donc plus il faut de temps. Inhibition ascendante : plus la cohorte est importante, plus il faut d'information pour démarquer la cible, donc plus il faut de temps ; inhibition latérale : plus la cohorte est importante, plus il y a de compétition, donc plus il faut de temps pour résorber le conflit.

initial. En d'autres termes, plus le contact initial est ciblé, mieux cela vaut. Pour y parvenir, le système peut :

- ▷ utiliser tous les indices acoustiques à sa disposition, en particulier les indices prosodiques et coarticulatoires, au lieu de les éliminer par filtrage lors du traitement perceptif.
- ▷ posséder une organisation qui facilite la reconnaissance des structures superposantes, vu leur universalité et leur fréquence, plutôt qu'une organisation qui génère un traitement complexe et des conflits avec ce type de configuration.

Ces deux idées composent le cœur théorique de ce travail de recherche. Chacune d'elles est reprise et explorée plus loin, dans un chapitre qui lui est dédié (respectivement Chapitres 2 et Chapitre 3, Partie II). Avant cela, un modèle implémenté est présenté dans la section suivante. Il s'agit du modèle TRACE, un classique de la reconnaissance des mots parlés en psycholinguistique.

### **2.1.3 Présentation d'un modèle implémenté : le modèle TRACE**

Le modèle TRACE (McClelland & Elman, 1986) est un système de reconnaissance des mots parlés. Il est issu de l'adaptation pour l'oral d'un dispositif connexionniste localiste, créé pour la reconnaissance des mots écrits (McClelland & Rumelhart, 1981). Il s'inspire des systèmes de perception de la parole de l'ingénierie (HEARSAY (Erman & Lesser, 1980; Reddy, Erman, Fennell, & Neely, 1973) et « Neural spectrogram » (Crowder, 1978, 1981)). Ce modèle a fortement influencé le champ théorique de la psycholinguistique au cours des deux dernières décennies, grâce aux prédictions claires et fortes qu'il permet de fournir. Étant donné ses succès lors de validations expérimentales chez l'humain, il a été choisi comme référence théorique dans la partie du travail relative aux superpositions lexicales, et une section de présentation lui est donc dédiée.

#### **2.1.3.1 Organisation du système et principes de fonctionnement**

TRACE est implémenté sous la forme d'un réseau connexionniste comportant un très grand nombre d'unités, organisées en 3 niveaux de traitement (entrée, traitement, sortie, représentant respectivement les traits acoustiques, les phonèmes et les lexèmes), dans lesquels chaque unité se consacre à l'évaluation d'une hypothèse sur l'existence d'un objet particulier survenant à un moment donné dans le stimulus. TRACE peut donc être utilisé pour la reconnaissance de phonèmes ou de lexèmes.

Le système fonctionne selon le principe de l'activation interactive (activation d'unités interagissant entre elles), à partir d'une information acoustique déjà analysée et transcrite dans 7 dimensions. Ces 7 dimensions ont été choisies par les auteurs à partir des travaux de Jakobson et collègues en linguistique (1952), et correspondent à une partie des traits distinctifs de nature acoustique définis par les linguistes. Ce sont : la puissance (distinction voyelle-consonne), le caractère vocalique (présence de formants), la diffusion (dispersion de l'énergie sur l'ensemble des fréquences du spectre), la finesse (concentration d'énergie dans les fréquences élevées du spectre), le caractère consonantique (faible énergie totale et obstruction dans le chenal phonatoire), le voisement, et l'explosion (distinction des consonnes occlusives).

Ces 7 traits sont utilisés par la première couche d'unités comme information d'entrée, chacun y étant codé sur une échelle de 1 à 9, selon une norme établie par les auteurs. Structurellement, cela entraîne l'usage de 9 unités ou détecteurs par dimension (une unité par valeur de codage), soient 7 groupes de 9 unités<sup>44</sup>.

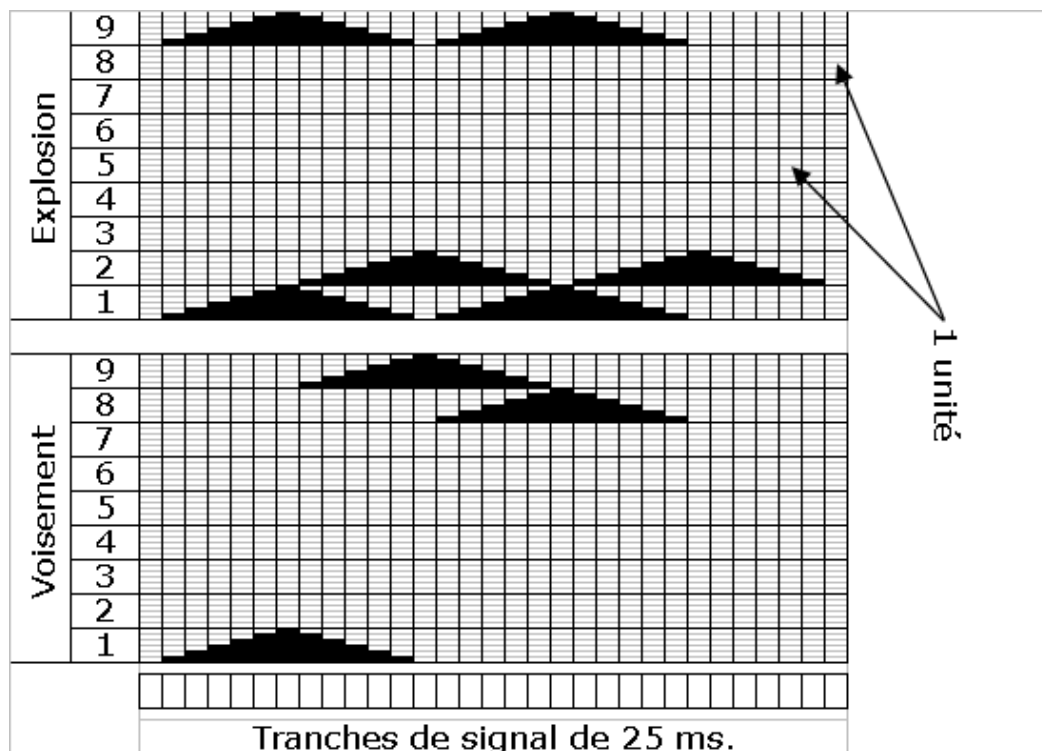
La dimension temporelle est codée dans le système par la réplication des 7 groupes de 9 unités. Chaque groupe traite une information issue d'un empan de 25 ms. de signal acoustique, sa réplication ayant lieu autant de fois que nécessaire pour traiter la totalité du stimulus. Le codage de 25 ms. de signal s'effectue donc sur 63 unités de la première couche (Figure 4).

Un trait phonétique est codé sur 11 périodes successives (soit une durée de signal de 275 ms.). La probabilité du trait ou valeur d'activation de l'unité, symbolisée par la hauteur du noircissement d'une case dans la Figure 4, croît linéairement depuis la période initiale (valeur de présence minimale) jusqu'à la période n°6 (valeur de présence maximale), puis décroît linéairement jusqu'à la période n°11. Les pics d'activation de deux traits successifs sont distants de 7 périodes, entraînant un chevauchement des spécifications de ces deux traits sur 5 périodes.

---

<sup>44</sup> Exemple : le phonème /r/ est respectivement codé par les valeurs [7, 7, 1, 2, 3, 8, -] pour les traits puissance, voyelle, diffusion, finesse, consonne, voisement, explosion. Cela signifie que le phonème /r/ active l'unité n°7 dans le groupe d'unités chargées de la détection du trait "puissance", l'unité n°2 dans le groupe de chargé de la détection du trait "finesse", etc.

Figure 4 : Codage du système de traits dans la couche d'entrée – exemple pour deux des sept traits



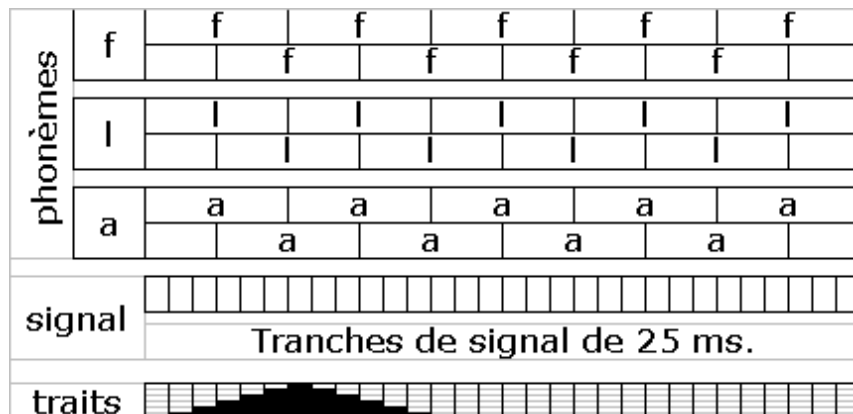
Les niveaux de 1 à 9 correspondent aux niveaux de codage pour l'activation, du plus faible au plus élevé, et chaque case correspond à une unité de traitement, dupliquée horizontalement pour effectuer le traitement sur une nouvelle tranche de 25 ms. de signal. Les triangles noirs correspondent à la probabilité d'occurrence d'un trait dans le temps (dans l'exemple, le trait de voisement a la probabilité d'occurrence la plus élevée (9) sur la 13<sup>ème</sup> période, mais sa probabilité reste encore élevée (8) à la 19<sup>ème</sup> période).

Lorsque 2 unités parmi les 9 d'une même dimension et traitant une même période sont simultanément activées, un phénomène d'inhibition latérale se produit (compétition). L'inhibition que les unités exercent les unes sur les autres est proportionnelle à leur degré d'activation.

Au niveau de la couche des phonèmes, le système dispose d'un détecteur pour chacun des phonèmes de la langue considérée. Chaque unité est chargée d'évaluer la probabilité d'occurrence d'un phonème à un moment particulier à partir de l'état de la couche des traits. Une unité phonémique traite une durée de 150 ms., et reçoit donc l'excitation en provenance de 6 périodes depuis le niveau des traits. Chaque phonème est excitable de façon identique par le motif de traits qui lui correspond. Une même unité phonémique est dupliquée toutes les 3 périodes selon le même principe détaillé pour les

traits. Deux unités phonémiques adjacentes se chevauchent donc partiellement sur 3 périodes (Figure 5).

Figure 5 : Organisation de la couche des phonèmes

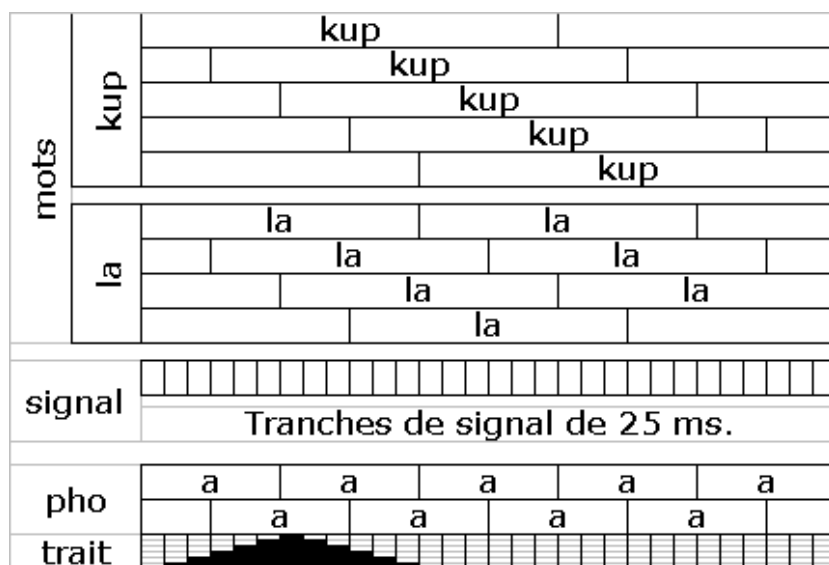


Comme pour le niveau des traits, un phénomène d’inhibition latérale existe entre les unités de la couche des phonèmes. Chaque unité phonémique inhibe les autres proportionnellement à l’intensité de son activation à un même instant. Deux unités phonémiques centrées sur une même période s’inhibent donc 2 fois plus que deux unités successives centrées à une distance de 3 périodes, aucune inhibition n’ayant lieu entre 2 unités centrées à 6 périodes de distance.

L’activation des unités phonémiques est ensuite transmise aux unités de la couche des lexèmes. Un phonème excite un lexème indépendamment du nombre de phonèmes qui composent celui-ci. Il existe un détecteur pour chacun des mots reconnus par le système. Chaque unité possède une étendue de traitement correspondant à la totalité des périodes nécessaires à la description du mot pour lequel elle code, relativement au nombre de phonèmes qui composent ce mot (Figure 6). Le réseau comporte une copie de chaque unité, décalée toutes les 3 périodes, comme dans le cas des phonèmes. Il n’y a pas de différence dans l’excitabilité des lexèmes en fonction de leur fréquence d’occurrence dans la langue.



Figure 6 : Organisation de la couche des lexèmes



Comme pour le niveau des traits et celui des phonèmes, des phénomènes d'inhibition ont lieu entre unités lexicales selon le principe déjà exposé, dépendant directement du nombre de périodes de chevauchement. Un détecteur de mot inhibe un autre détecteur de mot proportionnellement à l'importance du chevauchement des 2 mots dans le temps, 2 mots qui ne se chevauchent pas ne s'inhibant pas mutuellement.

La couche des lexèmes entretient une boucle de rétroaction avec la couche des phonèmes dans le but d'augmenter la performance du système dans les tâches d'identification de phonèmes grâce aux connaissances lexicales. Le Tableau 1 récapitule l'organisation des connexions entre unités, tant entre niveaux qu'au sein d'un même niveau.

Tableau 1 : Organisation des connexions entre unités

Lien / Unité	Type de connexion	Valeur
Trait-phonème	+	.02
Phonème-trait	+	0
Phonème-lexème	+	.05
Lexème-phonème	+	.03
Trait-trait	-	.04
Phonème-phonème	-	.04
Lexème-lexème	-	.03

Avant le traitement de chaque stimulus, le niveau d'activation des unités est initialisé. L'information entrante est appliquée période par période, 1 période correspondant à 1 cycle de calcul dans le réseau. Ceci permet de respecter la directionnalité temporelle de l'information acoustique, et d'imprimer une dynamique au

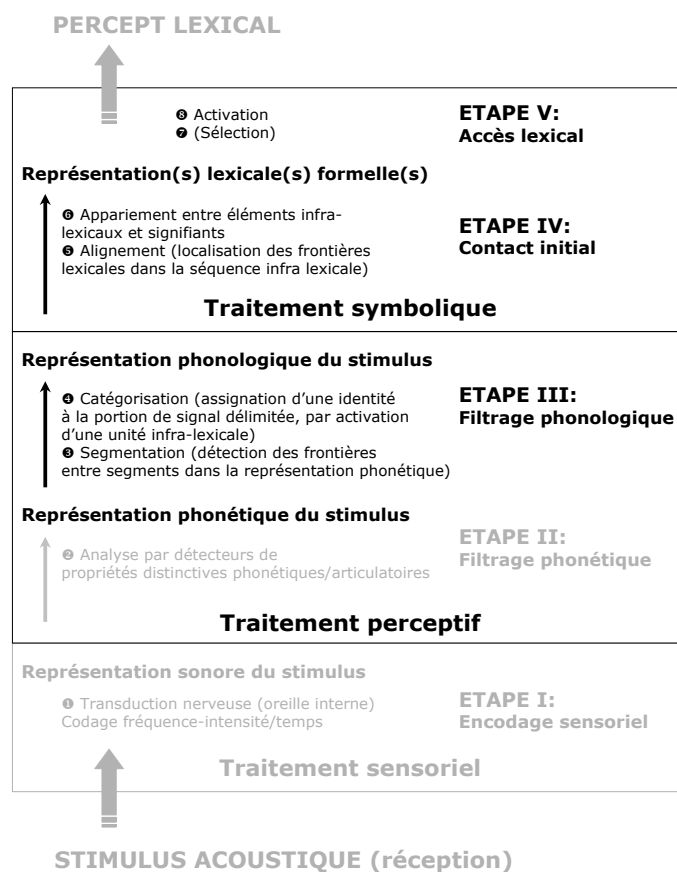
traitement en rapport avec la réalité. Pour chaque période, le calcul correspond à une intégration de l'information nouvelle avec l'ancienne contenue dans le réseau sous forme d'activation. Le calcul peut se poursuivre après la fin de la présentation du signal d'entrée, jusqu'à ce que le système se stabilise dans un état donné sous l'effet de l'activation interactive. Le mécanisme de décision se base sur l'utilisation de seuils d'activation relatifs ou absolus, la lecture de la solution s'effectuant depuis l'extérieur du système.

Le mécanisme de fonctionnement de TRACE ne fait aucune distinction entre mécanismes perceptifs et mécanismes de la mémoire de travail, puisque la trace d'activation résultant du traitement est à la fois processus perceptif et mémoire de travail.

### 2.1.3.2 Le niveau calculatoire du modèle TRACE

Le système honore les phases perceptives et symboliques du modèle calculatoire initialement présenté dans la Figure 2 page 20, et repris dans la Figure 7 ci-après, le texte en gris y figurant les étapes qui ne sont pas considérées dans le modèle TRACE.

Figure 7 : Niveau calculatoire dans le modèle TRACE



Il faut cependant souligner qu'à cause du processus d'activation interactive, l'ordre dans lequel est effectué chaque processus ne correspond plus à celui initialement proposé. Dans TRACE, la segmentation et la catégorisation s'effectuent en même temps lors de la projection de l'information depuis le niveau des traits sur le niveau des phonèmes. Par ailleurs, l'alignement, l'appariement, l'activation et la sélection ne peuvent plus être distingués comme une séquence de processus, et ont lieu simultanément lors de la projection de l'information depuis le niveau des phonèmes sur le niveau des lexèmes.

### 2.1.3.3 Comportement général du système

TRACE exhibe un ensemble de comportements comparables à ceux du système naturel : capacité à prendre en considération le contexte phonétique antérieur et postérieur pour compenser la coarticulation et favoriser l'identification correcte d'un phonème, effets phonotactiques, perception catégorielle des consonnes<sup>45</sup>, restauration phonémique, difficulté plus importante à identifier un mot sans avoir procédé à l'analyse du mot suivant, etc. Concernant la segmentation de la chaîne parlée en phonèmes et en mots, TRACE opère sans l'aide d'un processus de marquage des frontières ou d'une influence supra-lexicale (syntaxe, sémantique), mais par le seul procédé de l'activation interactive et grâce aux connaissances phonologiques et lexicales du système.

TRACE a cependant été critiqué pour des déficiences de nature :

- ▷ Structurale, qui entraînent la nécessité d'une normalisation de l'information d'entrée, aboutissant à une perte de la richesse phonétique<sup>46</sup> et prosodique<sup>47</sup>, et rendent impossible la prise en compte du phénomène d'amorçage par répétition observé chez l'humain<sup>48</sup>. De plus, les dimensions utilisées en entrée du système ne sont peut-être pas pertinentes pour le traitement perceptif, bien qu'elles le soient du point de vue de la théorie linguistique.
- ▷ Fonctionnelle, qui se traduisent essentiellement par une difficulté à identifier les mots courts en dépit d'un biais privilégiant les mots courts en début de traitement, un échec fréquent face aux segmentations multiples (les configurations dans lesquelles le premier mot est le plus long possible sont privilégiées), et l'impossibilité de reconnaître le mot stimulus en cas de compétition avec un gang de mots, phénomène probablement fréquent.

---

<sup>45</sup> Et hélas aussi des voyelles : la perception des voyelles est continue dans le système naturel.

<sup>46</sup> Les traits ont une saillance et une utilité identique, leur chevauchement entre 2 phonèmes est identique, etc.

<sup>47</sup> Tous les phonèmes sont de taille et d'intensité identique, les modulations mélodiques n'existent plus, les variations de débit sont annulées, etc.

<sup>48</sup> Deux présentations d'un même stimulus excitent deux copies distinctes d'une même représentation.

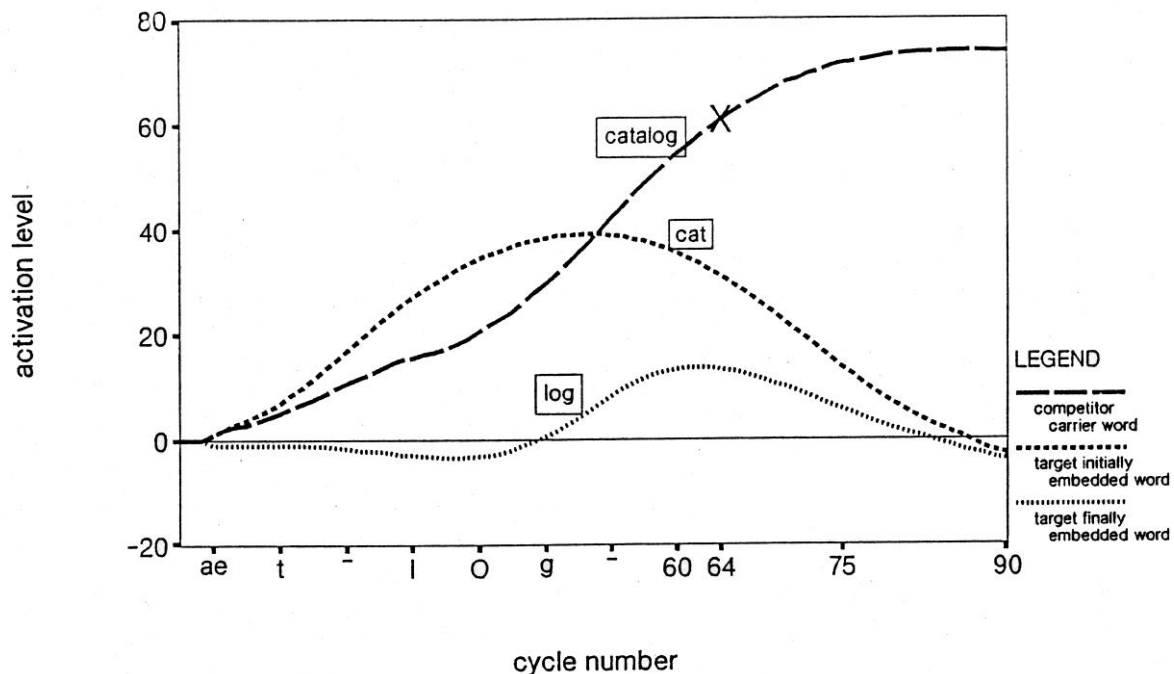
### 2.1.3.4 Comportement spécifique du modèle face aux superpositions lexicales

Le modèle fait des prédictions comportementales très précises en cas de superpositions de mots, et ce à travers 2 mécanismes :

- ▷ le mécanisme d'inhibition latérale (deux unités d'un même niveau activées par une même portion de parole s'inhibent mutuellement et proportionnellement à leur degré d'activation) entraîne des temps de reconnaissance des mots impliqués dans des superpositions lexicales plus importants que ceux de mots qui ne sont pas impliqués dans des superpositions lexicales. Cette durée supplémentaire est due au temps nécessaire pour que les représentations parviennent à un niveau d'activation suffisant pour se départager des compétitrices, qui tendent à les empêcher d'y parvenir.
- ▷ la dynamique du traitement liée à la temporalité du stimulus entraîne une modulation de l'effet de compétition, caractérisée par une diminution de sa force entre le début et la fin du mot. Cette modulation est due au fait qu'une compétition en début de mot a un impact sur le niveau d'activation du mot à reconnaître, car le système ne dispose pas de suffisamment d'information contextuelle pour départager les candidats. Au contraire, une compétition sur la fin du mot est généralement empêchée par la représentation du mot porteur, puisque celle-ci est tellement activée qu'elle inhibe très fortement la représentation du mot enchâssé, l'empêchant de parvenir à un degré d'activation pouvant avoir une incidence mesurable sur la représentation du mot porteur.

Il s'ensuit qu'une superposition lexicale initiale produit selon TRACE un effet négatif sur la reconnaissance d'un mot (augmentation de la durée du processus), tandis qu'une superposition lexicale finale est quasiment sans effet sur les temps de reconnaissance. Ceci est illustré dans la Figure 8, issue d'une simulation réalisée par Frauenfelder et Peeters (1990) du traitement d'une configuration de superpositions lexicales réelles (« *catalog* », « *cat* », « *log* »). Cette configuration implique à la fois une superposition lexicale initiale (« *cat* », responsable d'une inflexion de la courbe d'activation du lexème « *catalog* » (retard de la reconnaissance)) et une superposition lexicale finale (« *log* », sans effet sur le déroulement temporel de l'activation du lexème « *catalog* »).

Figure 8 : Simulation de la reconnaissance du mot « *catalog* » par le modèle TRACE. Évolution des courbes d'activation du lexème stimulus (« *catalog* ») et lexèmes des 2 mots enchâssés qu'il renferme (« *cat* » & « *log* »)



Le modèle fait de plus une distinction entre superpositions réelles (exemple précédent) et superpositions virtuelles. Ces dernières ont une influence moins marquée, puisqu'il existe une différence d'excitation reçue depuis le niveau des phonèmes, le mot porteur virtuel n'étant que partiellement activé par rapport au mot enchâssé réel. Toutefois, l'origine et le sens des effets demeurent identiques aux situations de superpositions lexicales réelles.

### **2.1.4 Récapitulation - Partie II, Chapitre 1**

Lors du traitement d'un mot impliqué dans des superpositions lexicales, le système active vraisemblablement plusieurs lexèmes simultanément. S'il fait cela, des procédures de sélection doivent ensuite être mises en œuvre pour ne conserver qu'une seule alternative. Or, ces procédures de sélection sont coûteuses en temps et/ou en énergie. Elles doivent donc avoir un impact mesurable sur le processus de reconnaissance d'un mot.

Le modèle TRACE traduit ce coût supplémentaire par une augmentation des temps de reconnaissance des mots impliqués dans des superpositions lexicales initiales, les superpositions lexicales non-initiales ayant un impact minimal, qui peut passer inaperçu.

L'examen détaillé de la situation révèle qu'il y aurait un avantage certain à ce que le système prenne en considération la prosodie et la coarticulation, information à même de guider la segmentation lexicale et de réduire la quantité de possibilités erronées, au lieu d'évacuer ces informations par un filtrage phonétique et phonologique lors de l'étape perceptive.

## 2.2 Chapitre 2 – Prosodie et démarcation des unités lexicales

Ce deuxième chapitre introduit succinctement la prosodie (aspects acoustiques, linguistiques, et fonctionnels – Section 2.2.1 page 42). Le rôle de la prosodie dans la reconnaissance des mots parlés est ensuite développé (Section 2.2.2 page 44 - unité porteuse du code prosodique, rôle de ce code dans la démarcation lexicale, hypothèses sur la manière dont la démarcation lexicale pourrait s'effectuer en Français). Les travaux de recherche en reconnaissance automatique de la parole, qui ont été utilisés pour fonder la partie de ce travail relative à l'existence d'un code prosodique, sont ensuite présentés (Section 2.2.3 page 56). Le Chapitre 2 se termine par une récapitulation des idées clé.

### 2.2.1 Quelques rappels sur la prosodie

#### 2.2.1.1 Niveaux d'organisation

La prosodie correspond à l'ensemble des phénomènes accentuels, mélodiques et rythmiques dans la parole. Elle peut être abordée selon 3 niveaux d'organisation, quelle que soit la langue considérée (Lacheret-Dujour & Beaugendre, 1999) :

- ▷ Le niveau phonético-acoustique (fréquence fondamentale, intensité, durée)
- ▷ Le niveau phonologique (représentations abstraites, telles que les groupes intonatifs et accentuels)
- ▷ Le niveau fonctionnel. La prosodie remplit 3 fonctions (Caelen-Haumont, 1996; Caelen-Haumont & Keller, 1997)<sup>49</sup> :
  - ∩ La fonction démarcative (processus qui « aident à entendre », c'est à dire à donner à l'auditoire des informations sur la composition et la structuration de l'énoncé en démarquant les unités linguistiques des différents niveaux qui le composent (unité lexicale, groupe syntaxique, phrase))
  - ∩ La fonction discriminative (processus qui aident à comprendre le sens du message en hiérarchisant les unités linguistiques en fonction du principe de l'information prioritaire)
  - ∩ La fonction élective (moyens de partager connaissances et croyances, articulés sur les principes de la fonction discriminative)

---

<sup>49</sup> Caelen-Haumont propose de subdiviser les rôles de la prosodie en « faire-savoir » (comprenant le « faire-entendre » rattaché à la fonction démarcative, et le « faire-comprendre » rattaché à la fonction discriminative) et « faire-croire » (rattaché à la fonction élective - (Caelen-Haumont & Keller, 1997)).

### 2.2.1.2 Organisation générale de la prosodie du Français

Les linguistes décrivent l'organisation prosodique générale du Français selon le schéma présenté dans le Tableau 2 (Lacheret-Dujour & Beaugendre, 1999).

Tableau 2 : Les différentes composantes de la prosodie du Français

Prosodie lexicale	Prosodie supra-lexicale
Accent morphologique	Groupements syntactico-sémantiques (morphèmes intonatifs terminaux) Opérations énonciatives (intonation, accent de focalisation) Rythme (intonation, accents rythmiques primaire et secondaire)

Cet ensemble de structures prosodiques relativement élaborées est décrit à partir de la perception de l'auditeur (prosodie perçue), ainsi que des corrélats acoustiques de la prosodie perçue (prosodie instrumentale).

### 2.2.1.3 Corrélats prosodiques entre signal acoustique, production et perception

Les relations qui existent entre les aspects acoustiques, la production et la perception, sont présentées pour rappel dans le Tableau 3.

Tableau 3 : Paramètres prosodiques : relations entre acoustique, production, et perception

Acoustique	Production	Perception		
		Locale	Globale	
$f_0^{50}$	Tension des cordes vocales	Hauteur	Rythme	Accentuation
	Pression sous-glottique		Mélodie	
Intensité <sup>51</sup>	Tension des cordes vocales	Sonie	Rythme	Accentuation
	Pression sous-glottique			
Durée	Débit	Étendue		

<sup>50</sup>  $f_0 = 3P_s + 50Q + k$

Avec  $P_s$  pour pression sous-glottique,  $Q$  pour rapport masse/tension,  $k$  constante.

<sup>51</sup>  $I = P_s - 2Q + k'$

$k'$  constante.



#### **2.2.1.4 Focalisation du travail effectué**

Les quelques éléments de rappel jetés au lecteur en guise de préambule permettent de dresser un premier bilan de la situation : l'étude du rôle de la prosodie sur la reconnaissance des mots parlés est l'étude du rôle d'un objet complexe, donc une étude a priori difficile.

Toutefois, ce problème de complexité a pu être contourné d'emblée en se focalisant sur les corrélats de la démarcation lexicale au niveau acoustique plutôt qu'au niveau phonologique. Ce positionnement a permis de prendre de la distance par rapport à la théorie linguistique, pour faciliter l'abord d'une hypothèse avant tout psychophysique de la démarcation lexicale.

L'approche suivie dans ce travail ne considère donc pas la grille de lecture donnée par les linguistes, dont certains aspects correspondent à un niveau d'élaboration déjà élevé de l'organisation prosodique (la plupart des éléments prosodiques décrits pour le Français sont déjà supra-lexicaux (Tableau 2)). Elle porte au contraire sur un niveau plus élémentaire de type code prosodique infra-lexical, de nature phonético-acoustique (voire phonologique), qui supporterait toutes ces structures, et fonctionnerait selon des principes universels indépendants de la langue considérée. Un tel code servirait à baliser le signal de parole afin de permettre au système de démarquer les éléments lexicaux en dépit d'un tramage prosodique supra-lexical accentuel et intonational non concordant avec les frontières des mots, et ce dès les plus bas niveaux du processus perceptif.

### **2.2.2 Le rôle de l'information prosodique dans la reconnaissance des mots parlés**

La présentation donnée à la section 2.1.2 page 20 du fonctionnement du système de reconnaissance des mots parlés permet de comprendre pourquoi un modèle qui n'utilise que l'information segmentale se trouve la plupart du temps face à une multiplicité de solutions lexicales pour segmenter le flux de parole en mots. Des procédures de sélection sont donc nécessaires pour parvenir à une solution linéaire, mais elles sont complexes et coûteuses, ou simples et insuffisantes, et leur performance ainsi que leur fiabilité n'est pas toujours estimée ou connue.

L'information suprasegmentale est théoriquement susceptible de guider la segmentation du flux de parole en mots, en complétant ou remplaçant les procédures de sélection lexicale. En effet, la fonction démarcative donne à l'auditoire des informations sur la composition et la structuration de l'énoncé à travers les trois paramètres de la prosodie, utilisés séparément ou en combinaison. Cependant, la manière dont la prosodie

prends corps dans le signal acoustique et dont elle influence la perception, en particulier pour la démarcation des unités lexicales, reste difficile à expliciter. En Français, les corrélats prosodiques absolus de frontières lexicales ne sont pas spécifiques et systématiques, bien qu'il existe des indices de frontière généralement dynamiques, en particulier sur les fins de mots polysyllabiques de contenu<sup>52</sup>. Les initiales de mots de contenu sont plus difficiles à déceler car prosodiquement moins contrastées. Les marques de frontières lexicales seraient donc probablement de nature contrastive (Caelen-Haumont, 1996), et si tel est le cas, un code prosodique porté par une unité infra-lexicale devrait exister.

### **2.2.2.1 Quelle unité de la ligne segmentale pour supporter quel code prosodique ?**

#### **2.2.2.1.1 L'unité plausible**

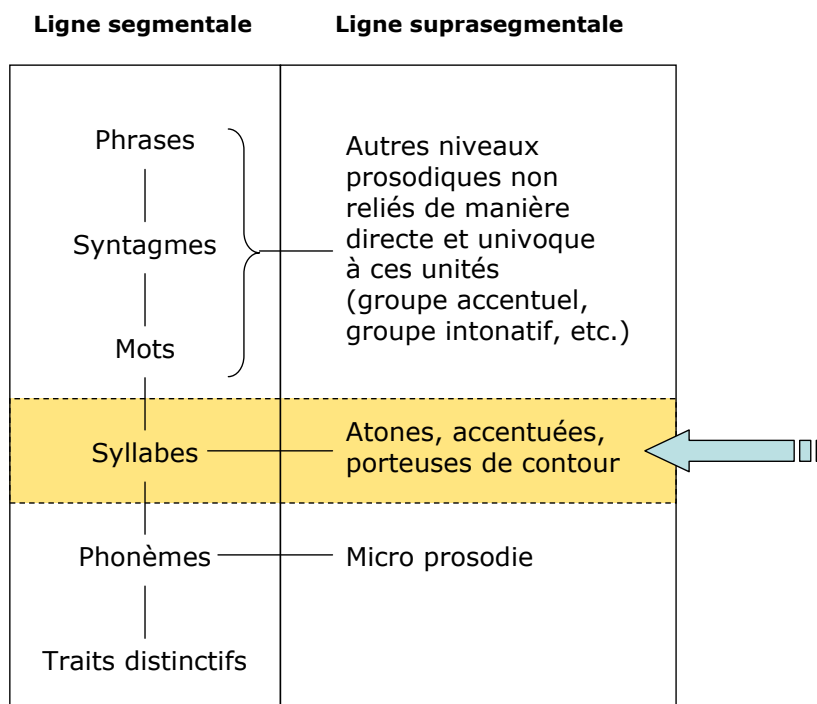
La Figure 9 met en parallèle les niveaux de hiérarchisation linguistique dans la ligne segmentale avec les niveaux hiérarchiques de la ligne supra-segmentale décrits dans la littérature par les modèles phonologiques des structures prosodiques (par exemple (Booij, Nespors, & Vogel, 1988)<sup>53</sup>).

---

<sup>52</sup> Contraste important d'énergie et de f<sub>0</sub> entre la syllabe qui précède et la dernière syllabe du mot, allongement des segments de la dernière syllabe. Toutefois, ce phénomène est fortement lié à la coïncidence entre la fin du mot et la fin du groupe accentuel.

<sup>53</sup> On trouve par exemple, par ordre de hiérarchisation du plus élémentaire au plus élevé, les unités suprasegmentales suivantes : le pied (plusieurs syllabes dont l'une est accentuée), le mot phonologique ou unité accentuelle (radical et affixes linéairement adjacents, mots composés, avec adjonction éventuelle de l'article), le syntagme phonologique ou groupe accentuel (déterminisme syntaxique fort mais non absolu), le syntagme intonatif ou groupe intonatif (dépendant de contraintes syntaxiques, sémantiques et pragmatiques faisant tendre la taille de ces syntagmes à un équilibre dans une même phrase), l'énoncé phonologique (correspondant souvent à la phrase mais pouvant la dépasser).

Figure 9 : Hiérarchisation de l'information linguistique



Hormi les traits distinctifs, trop élémentaires pour autoriser des phénomènes suprasegmentaux, et les mots, les syntagmes et les phrases, trop élaborés et variables en longueur pour permettre une correspondance univoque avec des unités de la ligne suprasegmentale, seuls les phonèmes et les syllabes pourraient permettre de supporter des phénomènes suprasegmentaux élémentaires.

La syllabe apparaît cependant comme le support le plus adéquat pour un code fondé sur la modulation des 3 paramètres prosodiques, le phonème étant une unité trop élémentaire pour porter les 3 paramètres simultanément (exemple des phonèmes non voisés pour la  $f_0$ ).

La syllabe est de plus une unité de base en production de la parole, puisqu'elle correspond à une « bouchée de parole » plus ou moins pleine (exemple : « a » ou « schtroumpf »). Par ailleurs, elle correspond au gabarit minimal du mot. Le locuteur peut la produire rapidement ou lentement, avec une voix forte ou faible, et en ciblant une fréquence haute ou basse, voire en créant des modulations de durée, de mélodie, et d'intensité à l'intérieur de l'unité, et de façon relativement indépendante de la syllabe qui précède et de celle qui suit. La syllabe permet donc des variations suffisantes de durée, de mélodie, et d'intensité, pour supporter des phénomènes contrastifs élémentaires, qui se traduisent au niveau perceptif par le rythme, le contour mélodique, et l'accentuation (Tableau 3).

Les caractéristiques prosodiques d'une syllabe étant elles-mêmes impliquées dans la formation d'unités suprasegmentales plus élaborées, le contrôle suprasegmental des unités supérieures à la syllabe passe aussi par le contrôle suprasegmental des syllabes produites. Les syllabes peuvent donc servir d'unité pour supporter les niveaux suprasegmentaux plus élaborés à travers un système de code.

Parallèlement au statut de la syllabe comme unité de production, la syllabe serait aussi une unité fondamentale pour la perception de la parole et la reconnaissance des mots parlés (par exemple (Content et al., 2001; Cutler, McQueen, Norris, & Somejuan, 2001; Dumay et al., 2002; Mehler, Dommergues, Frauenfelder, & Segui, 1981; Tabossi, Collina, Mazzetti, & Zoppello, 2000; R. M. Warren, Healy, & Chalikia, 1996)). Cette idée a été suggérée pour plusieurs raisons :

- ▷ La syllabe correspond au gabarit minimal de l'unité lexicale
- ▷ Elle permet de limiter les erreurs de segmentation lexicale (utilisation d'un phonème dans plus d'un mot, ou omission d'un phonème de la séquence – exemple : PWC).
- ▷ Elle permet de localiser les frontières lexicales (exemple : SOSH). La syllabe étant une unité à forte cohésion acoustique, la coarticulation des segments *y* est plus importante qu'entre 2 segments successifs de 2 syllabes. Ceci introduit une hétérogénéité acoustique utile dans la séquence de phonèmes.
- ▷ La syllabe véhicule des contrastes prosodiques pouvant être exploités pour localiser l'initiale des mots dans certaines langues métriques comme l'Anglais (exemple : MSS).
- ▷ Elle est la seule unité infra-lexicale suffisamment élaborée pour renfermer toujours un noyau vocalique, à la base des phénomènes mélodiques dans la voix

Ces quelques observations appuient la candidature de la syllabe comme l'unité la plus adéquate pour supporter un code prosodique élémentaire, désigné dans la suite du texte par la terminologie de « code prososyllabique ».

### **2.2.2.1.2 Le code prososyllabique**

Le code prososyllabique peut être défini a priori comme une organisation des paramètres  $f_0$ , intensité et durée au niveau de la syllabe, en fonction de facteurs morphosyntaxiques (lexicalité ou grammaticalité d'un mot), topologiques (position de la syllabe dans un mot polysyllabique), et autres (non déterminés. Ce peuvent être des facteurs syntaxiques, sémantiques ou d'une autre nature, qui viennent pondérer le code prososyllabique de base (morphosyntaxique et topologique)), de manière exploitable par le système perceptif.

Deux types de code prososyllabique peuvent être envisagés pour démarquer les unités lexicales selon un traitement ascendant. Le premier est un code absolu ou statique basé sur des caractéristiques prosodiques fixes de la syllabe. Le second est un code relatif ou dynamique basé sur des phénomènes contrastifs entre deux syllabes successives.

### 2.2.2.1.2.1 Code prososyllabique absolu

Les linguistes catégorisent les syllabes du Français d'un point de vue perceptif pour ce qui est de leurs caractéristiques suprasegmentales, grâce à la dimension de l'accentuation (syllabes atones ou accentuées) et de la mélodie (syllabes porteuses de contour (déclaration, interrogation, etc.) ou non). Il s'agit cependant d'une catégorisation grossière. Puisque 3 paramètres acoustiques permettent de déterminer la prosodie dans la substance, 8 types de syllabes peuvent théoriquement être spécifiés dans le cas imaginaire où ces 3 paramètres ont des modalités binaires (Tableau 4).

Tableau 4 : Types de syllabes pouvant être théoriquement définis sur le plan acoustique

Durée		Brève (B)		Longue (L)	
Intensité		Faible (F)	Élevée (E)	Faible (F)	Élevée (E)
f0	Grave (G)	BFG	BEG	LFG	LEG
	Aigue (A)	BFA	BEA	LFA	LEA

Cependant, une telle catégorisation est caricaturale, étant donné d'une part que les paramètres prosodiques n'ont pas des modalités binaires mais varient sur un continuum (la caractérisation prosodique des syllabes serait donc plus subtile et le nombre de catégories prososyllabiques pourrait être supérieur à 8<sup>54</sup>), et d'autre part, que ces paramètres sont parfois fortement corrélés entre eux (f0 et intensité, ou intensité et durée), et ne permettent donc pas toujours des combinaisons orthogonales telles que celles présentées dans le Tableau 4. De plus, toutes les formes prososyllabiques possibles ne sont peut-être pas « signifiantes » pour le système perceptif, ou exploitables, ce dernier ne réagissant peut-être qu'à certaines configurations et non à d'autres pour des raisons psychophysiques. Ainsi, certaines catégories prososyllabiques pourraient être perçues comme appartenant à une même catégorie. Par ailleurs, le système locuteur peut n'utiliser que, ou préférablement, certaines configurations prosodiques particulièrement « signifiantes » ou particulièrement faciles à produire, et donc privilégier ces catégories, tandis que d'autres sont peu ou jamais produites.

<sup>54</sup> Par exemple, pour la f0, certains linguistes utilisent une catégorisation basée sur 4 niveaux (suraigu, haut, bas, infra-grave) - (Lacheret-Dujour & Beaugendre, 1999)

Pour toutes ces raisons, il est difficile de savoir a priori combien de catégories utiles de syllabes la prosodie permet de distinguer dans un système de code absolu, ni si un tel ciblage est effectivement produit étant donné la variabilité prosodique.

#### **2.2.2.1.2.2 Code prososyllabique relatif**

Puisque l'idée de catégories absolues ne semble pas être la plus adéquate pour définir un code prososyllabique impliqué dans la démarcation d'unités lexicales, le code prososyllabique pourrait être préférablement basé sur des catégories relatives, c'est-à-dire utilisant des phénomènes contrastifs et dynamiques à travers le déploiement temporel du flux de parole. Avec un tel code, la différence entre types de syllabes selon la grille de classification proposée dans le Tableau 4 reste valable avec des paramètres à modalités non binaires. On peut par exemple avoir une syllabe BFG parce que plus grave qu'une syllabe BFA qui la précède ou la suit, même si d'un point de vue absolu, la syllabe BFG n'est pas produite avec une  $f_0$  basse.

L'idée d'un code prososyllabique dynamique est donc probablement plus réaliste que celle d'un code absolu. Elle permet de justifier la difficulté pour trouver dans la prosodie des marques de frontière de mots, sans exclure que de telles marques ne soient présentes dans la parole. Cette idée a été privilégiée et explorée dans le travail de recherche présenté dans ce rapport.

### **2.2.2.2 L'influence de la prosodie sur la segmentation du flux de parole en mots**

En psycholinguistique, le rôle de la prosodie dans la démarcation lexicale est étudié essentiellement chez l'adulte normal (Cutler, Dahan, & van Donselaar, 1997). La prosodie est alors considérée comme une information guidant le traitement lexical de manière ascendante (démarcation lexicale ascendante), les questions gravitant autour du type d'indice prosodique utilisé pour effectuer la segmentation en mots, en particulier pour résoudre les ambiguïtés lexicales (Gaskell & Marslen-Wilson, 2001). Les psycholinguistes se sont également interrogés sur l'existence de représentations mentales prosodiques guidant le traitement de manière descendante (Cutler et al., 1997) (démarcation lexicale descendante), hypothèse plus délicate à démontrer que la précédente.

#### **2.2.2.2.1.1 Démarcation lexicale ascendante**

En Français, la démarcation ascendante des unités lexicales par la prosodie peut être envisagée, dans le cadre des théories linguistiques, comme appuyée par l'accent

primaire, localisé sur les fins des groupes accentuels (qui correspondent à des fins de mots de contenu<sup>55</sup>), et par l'accent secondaire, localisé facultativement sur le début des mots de contenu. Elle peut aussi être envisagée hors de ce cadre, comme supportée par des phénomènes prososyllabiques accentuels et non accentuels, qui exploitent des propriétés psychophysiques élémentaires. C'est cette dernière position qui a été adoptée dans ce travail. En dépit de ce choix théorique, l'hypothèse d'une démarcation lexicale basée sur les accents primaire et secondaire est rapidement considérée afin de montrer d'une part que l'accent peut être utile à la démarcation lexicale parce qu'il se fonde sur ce code prososyllabique, mais d'autre part qu'il est insuffisant pour tenir un rôle systématique compte tenu de sa distribution.

#### **2.2.2.2.1.1 Démarcation lexicale ascendante basée sur l'accent primaire**

L'accent primaire est impliqué dans la réalisation de la trame métrique en Français, et est considéré comme régulier, logique, normal et attendu par les linguistes (Lacheret-Dujour & Beaugendre, 1999). L'accent primaire se caractérise essentiellement par un allongement de la durée des segments composant la dernière syllabe du groupe accentuel, et éventuellement par une montée de  $f_0$  (Lacheret-Dujour & Beaugendre, 1999).

Contrairement à une langue comme l'Anglais, qui présente une concordance élevée entre la trame métrique et l'initiale des mots de contenu (75% des syllabes fortes ou à voyelle pleine sont localisées sur les initiales de mots de contenu, et 90% des initiales de mots de contenu sont des syllabes fortes - (Cutler & Carter, 1987))<sup>56</sup>, l'accent primaire du Français ne représente pas a priori le candidat le plus adapté pour guider l'accès lexical, étant donné qu'il est distribué sur la fin des mots de contenu et non sur leur début, et qu'il l'est de façon facultative.

Dans le cadre d'une recherche portant sur le rôle de l'accent primaire dans la segmentation du flux de parole en mots, Bagou et collègues (2002) ont toutefois montré, à l'aide du paradigme d'acquisition d'un « mini-langage artificiel », que les auditeurs

---

<sup>55</sup> L'accent primaire apparaît comme une marque non systématique de fins de mots de contenu. Il est parfois qualifié d'accent lexical.

<sup>56</sup> En Anglais, il est supposé que la présence d'une syllabe forte serve de point d'entrée privilégiée dans le lexique (Cutler & Norris, 1988; McQueen, Norris, & Cutler, 1994), et permette d'appuyer l'utilisation de stratégies de segmentation telles que la MSS proposée par (Cutler & Norris, 1988). En effet, la régularité métrique permet au système de générer des attentes perceptives qui peuvent le conduire à se focaliser sur des zones particulières de la séquence d'événements (Jones, 1976; Jones, Kidd, & Wetzel, 1981). Une syllabe forte correspondrait de plus à une zone du message ayant une définition acoustique meilleure.

francophones effectuent la segmentation du flux de parole en unités constitutives (« mots » de la langue artificielle) avec de meilleures performances lorsque des marques de frontière « lexicale » finale sont spécifiées par l’allongement de la durée des segments en plus d’une spécification par des indices phonotactiques (82% de réponses correctes), que lorsqu’elles ne sont spécifiées que par des indices phonotactiques (64.5% de réponses correctes). La durée finale serait donc un indice important en Français pour démarquer les unités lexicales par leur fin.

Il en va de même du rôle de la  $f_0$  dans une organisation prosodique de type accent primaire du Français (Bagou et al., 2002). La performance des auditeurs francophones pour segmenter le flux sonore en unités constitutives (« mots » de la langue artificielle) est en effet meilleure lorsque des marques de frontière « lexicale » finale sont spécifiées par une modulation de la  $f_0$  en plus d’une spécification par des indices phonotactiques (89.9% de réponses correctes), ou par une modulation combinée de  $f_0$  et de durée (en plus de la phonotactique - 85.5%), que lorsqu’elles ne sont spécifiées que par des indices phonotactiques (64.5%).

Ces quelques résultats révèlent l’effet bénéfique de la présence de marques prosodiques de type accent primaire pour la démarcation ascendante des unités lexicales en Français.

Toutefois, les effets de type accent primaire sur la démarcation lexicale mis en évidence par Bagou et collègues peuvent tout aussi bien être la manifestation psychologique de l’utilisation par le système de phénomènes psychoacoustiques plus universels, non limités à ce seul type d’organisation accentuelle. En effet, quelle que soit la langue considérée, les modulations de durée et de  $f_0$  permettent des modulations locales dans la résolution acoustique des éléments constituant le message parlé, ainsi que des modulations dans la dynamique du traitement, avec une alternance entre périodes d’affluence importante d’information abstraite à traiter et périodes d’affluence plus faible. Rietveld (Rietveld, 1980) montre par exemple que la durée des segments correspond à une information essentielle pour orienter la segmentation en mots et lever les ambiguïtés lexicales en Français, dans les cas d’homophrases du type « *le couplet complet* » vs. « *le couple est complet* ».

#### **2.2.2.2.1.1.2 Démarcation lexicale ascendante basée sur l’accent secondaire**

Théoriquement, l’accent secondaire représenterait un candidat plus adéquat que l’accent primaire pour la démarcation ascendante des unités lexicales en Français, puisqu’il se localise sur l’initiale des mots et non sur leur fin. L’accent secondaire



présente cependant une distribution variable (il peut être réalisé aussi bien sur les mots de classe fermée que sur les mots de classe ouverte) et facultative (son apparition dépendrait de phénomènes de régulation rythmique, principalement lorsque la composante linguistique de l'énoncé ne permet pas de produire un nombre suffisant d'accents primaires (Lacheret-Dujour & Beaugendre, 1999)), ce qui a conduit les linguistes à proposer l'existence de 3 accents secondaires distincts (Lacheret-Dujour & Beaugendre, 1999) :

- ▷ L'accent rythmique ou ictus mélodique (relief mélodique accompagné éventuellement d'une augmentation de l'intensité), associé à la syllabe initiale des mots.
- ▷ L'accent énonciatif ou de focalisation (relief mélodique et intensitaire plus important que dans l'ictus mélodique). Il est produit sur la première syllabe accentuable du groupe, correspondant à la syllabe initiale du premier mot de contenu s'il commence par une consonne (***m**aison, **l'**enfant, **les** enfants*), ou à la deuxième syllabe du premier mot de contenu s'il ne commence pas par une consonne (*ad**m**irablement*).
- ▷ L'accent emphatique, véhiculé par l'intonation, qui réalise une fonction de type discriminatif plutôt que démarcatif.

À cause de sa variabilité et de son caractère facultatif, l'accent secondaire n'offrirait pas un tramage de la chaîne parlée suffisant pour démarquer les unités lexicales de manière fiable. Comme l'accent primaire, l'accent secondaire représenterait toutefois une aide possible pour la segmentation lexicale, probablement grâce à son incidence sur le système perceptif à travers des phénomènes psychophysiques élémentaires. Aucune étude psycholinguistique n'est disponible sur le rôle de l'accent secondaire en Français dans la démarcation lexicale ascendante.

#### **2.2.2.2.1.1.3 Démarcation lexicale ascendante basée sur un guidage psychophysique du traitement**

La trame accentuelle représentant une structuration prosodique relativement élaborée du flux de parole, d'une part, et d'autre part étant basée sur une organisation particulière des paramètres prosodiques responsables de l'orientation du traitement perceptif, il est possible que la prosodie permette une démarcation des unités lexicales sur la base d'une organisation plus élémentaire que l'organisation accentuelle décrite précédemment (celle-ci étant par ailleurs insuffisante pour démarquer tous les mots). Une telle organisation fondamentale de la prosodie correspond au code prososyllabique qui a déjà été introduit. Ce code serait porteur de tous les phénomènes déjà énumérés, ainsi que d'autres plus élaborés ou plus simples. Ces phénomènes plus simples interagiraient avec le système perceptif sur la base de phénomènes psychophysiques élémentaires sans pour autant former nécessairement des structures prosodiques remarquables et inventoriées telles que les accents, pour produire un effet démarcatif au

niveau lexical. Ce code permettrait de distinguer les mots monosyllabiques de contenu des mots outils monosyllabiques (Caillaud, Antoine, Caelen, & Caelen-Haumont, 1993; Caillaud, Munteanu, Serignat, & Caelen, 1998; Lachaud, Caelen-Haumont, Pynte, & Espesser, 1999; Munteanu, Caillaud, Serignat, & Caelen-Haumont, 1997), ou encore les mots monosyllabiques d'initiale ou fin de mot polysyllabique en cas d'enchâssement lexical, comme l'a montré Davis (Davis, 2000) pour l'Anglais (les auditeurs sont capables de discerner précocement (avant la fin de la syllabe expérimentale, donc avant la divergence phonémique) entre le mot monosyllabique superposé à l'initiale d'un mot polysyllabique (« *his captain* ») et la forme libre de ce mot monosyllabique (« *his cap tucked* »), grâce à une information de durée, des segments (écart moyen de -48 ms. entre conditions) et du délai de voisement (-19 ms.)).

Comme suggéré plus haut et par l'étude de Davis, le débit permettrait de moduler la résolution acoustique des segments, des syllabes et des mots (un phonème produit avec une durée plus importante est décrit par une quantité d'information plus importante), et servirait à adapter la quantité de calculs à réaliser par unité de temps en fonction de la difficulté cognitive du traitement. Ainsi, la présence en Français d'un allongement sur les fins de mots de contenu porté par l'accent primaire pourrait correspondre à une temporisation du flux d'information entrant pour laisser au système le temps de terminer ses traitements lexicaux.

La  $f_0$  et l'intensité agiraient différemment en ayant un rôle de guidage attentionnel (modulation du niveau vigile de l'auditeur en fonction des difficultés locales du traitement). Cette idée est suggérée par l'existence d'une série restreinte mais universelle (indépendante de la culture considérée) de contours dans les productions maternelles adressées au nourrisson, liés à des contextes spécifiques de soin (augmentation de l'information de contour aux dépens d'une articulation claire et d'un contenu informatif), et qui induisent un état vigile particulier chez l'enfant<sup>57</sup>. De telles observations laissent penser que les contours mélodiques seraient en lien avec une base biologique de la communication (H. Papousek & Papousek, 1987), ces contours pouvant

---

<sup>57</sup> Les contours montants ou en forme de cloche, par exemple, sont utilisés pour attirer et maintenir l'attention des nourrissons (M. Papousek & Papousek, 1981), alors que les contours descendants avec une gamme étroite de hauteurs sont appliqués pour calmer et endormir le nourrisson (M. Papousek & Papousek, 1981). Les contours plus variables servent, quant à eux, à renforcer un aspect positif (Stern, Spieker, & MacKain, 1982). En outre, le tempo et le rythme des énoncés maternels s'accordent aux besoins supposés du nourrisson. Ainsi, les énoncés rythmiques lents sont utilisés chez les nourrissons attentifs, les énoncés avec un tempo accéléré chez les nourrissons inattentifs, ceux avec des rythmes et des tempos variables chez les nourrissons agités, et ceux avec un tempo progressivement décroissant chez les nourrissons en train de s'endormir (M. Papousek & Papousek, 1981).

correspondre à des formes prosodiques élémentaires utilisées plus tard dans la parole adulte comme unités de traitement, linguistiques ou attentionnelles.

Le code prososyllabique pourrait donc servir à baliser la séquence et permettre au système de s’y situer dès les plus bas niveaux de traitement grâce à des phénomènes simples de modulation de la résolution de la ligne segmentale. Il permettrait également de capaciter le système de traitement en modulant le niveau vigile de l’auditeur.

#### **2.2.2.2.1.2 Démarcation lexicale descendante**

Concevoir un phénomène de démarcation lexicale descendante par la prosodie implique d’envisager l’existence de représentations mentales de l’information suprasegmentale, associées aux représentations phonologiques des mots ou non :

- ▷ Si les représentations phonologiques des mots intègrent l’information suprasegmentale, cela implique que le système puisse disposer de lexèmes différents en fonction de la forme prosodique donnée à la ligne segmentale (langues à accent lexical).
- ▷ Si les représentations phonologiques des mots n’intègrent pas l’information suprasegmentale, le système peut néanmoins disposer de représentations suprasegmentales distinctes des lexèmes, et dotées d’une signification particulière en lien avec la démarcation lexicale.

L’hypothèse d’une représentation mentale de la prosodie n’est pas une idée farfelue. Il faut bien que certains éléments prosodiques soient représentés dans le système pour être produits et utilisés dans la communication.

##### **2.2.2.2.1.2.1 Démarcation lexicale descendante basée sur l’accent lexical**

Contrairement à d’autres langues romanes comme l’Espagnol ou l’Italien, le Français n’a pas conservé l’accent tonique du latin, qui était un accent distinctif (par exemple, en Italien, « *fòrmica* » désigne le matériel (*formica*), et « *formica* » désigne l’animal (*fourmi*)). Une démarcation lexicale descendante basée sur une représentation lexicale de la prosodie ne serait donc pas applicable au Français.

##### **2.2.2.2.1.2.2 Démarcation lexicale descendante basée sur l’utilisation de représentations prosodiques indépendantes du lexique**

L’hypothèse d’un guidage de la segmentation lexicale à partir de représentations prosodiques indépendantes du lexique peut impliquer l’existence de détecteurs de motifs prosodiques particuliers. Ces détecteurs réagiraient par exemple à des motifs prosodiques fréquemment rencontrés aux frontières lexicales, et assureraient une robustesse du processus de démarcation des mots. Ils pourraient aussi réagir à des motifs prosodiques définissant l’organisation générale d’un mot, et servir à activer un

sous-ensemble du lexique mental (exemple : tous les mots de contenu de x syllabes), et ainsi limiter le nombre de solutions lexicales alternatives à considérer<sup>58</sup>.

Par exemple, les résultats de Lindfield et collègues (1999a; 1999b), pour une langue sans accent lexical comme l'Anglais (le traitement mental de FORbear équivaut à celui de forBEAR selon Cutler (1986)), suggèrent que certaines composantes de la prosodie d'un mot sont représentées en mémoire, ce qui pourrait donc aussi être le cas en Français. Ces chercheurs obtiennent en effet, avec une tâche de « gating », un meilleur taux de reconnaissance et des réponses plus précoces lorsque l'information segmentale de début de mot donnée à l'auditeur est complétée par une information prosodique relative à la durée du mot, à l'accentuation, et à son schéma syllabique (nombre de syllabes).

L'hypothèse des détecteurs de motifs prosodiques est techniquement difficile à décorrélérer de l'hypothèse relative à l'influence ascendante de la prosodie, bien que les observations de Lindfield et collègues indiquent l'existence d'une connaissance de ce type et sa mise en œuvre dans la reconnaissance des mots parlés. Le travail présenté dans ce rapport n'a donc pas considéré l'hypothèse des détecteurs de motifs prosodiques, au profit de l'hypothèse d'un guidage ascendant exploitant des mécanismes perceptifs fondamentaux pour orienter la dynamique des traitements ultérieurs.

#### **2.2.2.2.2 Conclusion**

L'étude classique du rôle de la prosodie sur la démarcation lexicale est dans une impasse lorsqu'elle cherche à attribuer ce rôle à une organisation qui varie d'une langue l'autre, et qui n'est pas systématiquement présente dans les productions (accents, etc.). Une autre approche développée en reconnaissance automatique de la parole est maintenant présentée.

---

<sup>58</sup>

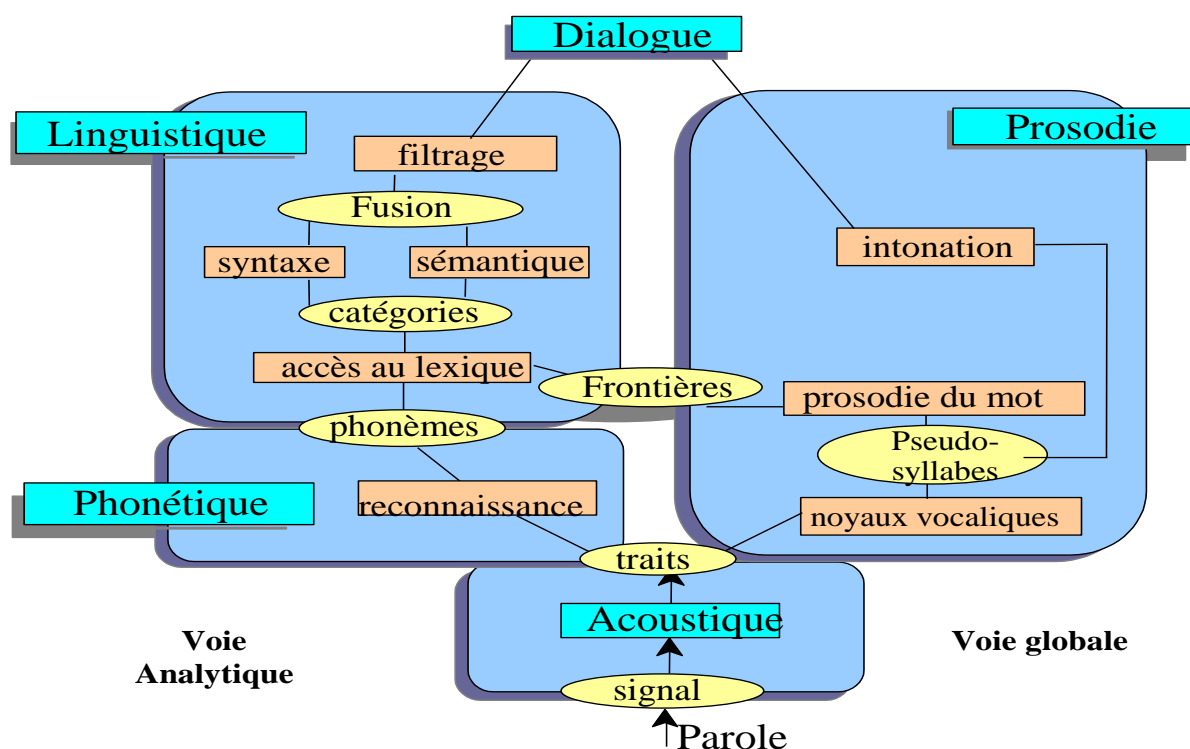
Comme semblent le suggérer les résultats de Wingfield et collègues (1997).

### 2.2.3 Les travaux en reconnaissance automatique de la parole

L'approche technique qui a été développée dans ce travail s'inspire directement des recherches en reconnaissance automatique de la parole développées au CLIPS-IMAG<sup>59</sup> et à l'ICP<sup>60</sup> dans les années 90 (Caillaud, 1996; Caillaud et al., 1993; Caillaud et al., 1998; Munteanu, 1996; Munteanu et al., 1997). Elles sont succinctement résumées dans cette section.

MICRO (Modélisation Informatique de la Cognition en Reconnaissance de l'Oral) est un projet de recherche en compréhension automatique de la parole inspiré des théories cognitives, duquel est née une architecture logicielle permettant d'aboutir à un système robuste de traitement de la parole. L'architecture de ce système, modulaire et interactive, est donnée dans la Figure 10 (Munteanu, 1996).

Figure 10 : Architecture du système MICRO



Avec : Rectangles bleus (sombre): modules; Rectangles oranges (clair): agents du module; Ellipses jaunes: éléments traités.

<sup>59</sup> Communication Langagière: Interaction Personne-Système, Institut de Mathématiques Appliquées de Grenoble.

<sup>60</sup> Institut de la Communication Parlée - Institut National Polytechnique de Grenoble.

MICRO dispose d'un module dédié au traitement de la prosodie (module « Prosodie »), chargé de détecter les noyaux vocaliques, de localiser les frontières lexicales et syntagmatiques probables, et de modéliser l'intonation. L'hypothèse sous-jacente à l'existence de ce module est que la prosodie reflète des phénomènes lexicaux, syntaxiques, sémantiques et pragmatiques du discours, et que l'information qu'elle véhicule est suffisante pour offrir des points d'ancrage, nécessaires à la création d'hypothèses de segmentation lexicale lors du traitement linguistique. Dans MICRO, l'agent « prosodie du mot » assure cette fonctionnalité en fournissant des informations sur la localisation la plus probable des frontières de mots, ainsi que sur leur catégorie morphosyntaxique (mots outils ou mots de contenu). Cette tâche est définie comme un problème de classification de syllabes, utilisant en entrée la durée des syllabes et leur  $f_0$  et intensité moyennes, mesurées par l'agent « noyaux vocaliques » du module prosodique. L'estimation est effectuée à partir de portions de signal de 3 syllabes et porte sur la syllabe centrale de l'échantillon. Elle se base donc sur l'utilisation de 12 valeurs (avec un supplément éventuel d'information si une pause discursive est présente entre 2 de ces 3 syllabes). L'agent « noyaux vocaliques » attribue au noyau vocalique de la syllabe centrale l'une des 8 catégories issues du croisement complet entre type de mot (2 modalités : mot grammatical (G) vs. mot de contenu (L)) et localisation de la syllabe dans le mot (4 modalités : début (D), non-début non-fin (M), fin de mot plurisyllabique (F), mot monosyllabique ()), soient DG, MG, LG, G, DL, ML, FL, L.

L'agent « prosodie du mot » met en œuvre l'algorithme « PmBc » (Caillaud, 1996; Munteanu, 1996). Il s'agit d'une heuristique de recherche produisant et corrigeant ses propres règles de classification (généralisations successives à partir d'exemples)<sup>61</sup>. Ses performances de classification<sup>62</sup>, testées sur un corpus de parole lue<sup>63</sup> sont données dans le Tableau 5.

---

61 Les règles sont de type "IF  $f_0(n) \leq 0.35$  AND  $E(n+1) \geq 0.59$  AND  $d(n-1) \geq 0.13$  AND  $d(n) \leq 0.18$  THEN G". Pour plus de détails sur le fonctionnement du système, se reporter à Caillaud, Munteanu et col., "Prosodic knowledge acquisition for lexical access improvement", *Communication and cognition. Artificial Intelligence. The Journal for the integrated study of AI cognitive science and applied epistemology*, 1998, 15-3, pp 255-278, ainsi qu'aux thèses de doctorat de Caillaud (Bertrand Caillaud, *Apprentissage de connaissances prosodiques pour la reconnaissance automatique de la parole*, Institut national polytechnique de Grenoble, 1996) et de Munteanu (Paul Munteanu, *Extraction de connaissances dans les bases de données Parole: apport de l'apprentissage symbolique*, Institut national polytechnique de Grenoble, 1996).

62 Taux de détection: [Nombre d'exemples de la catégorie correctement classés] / [Nombre total d'exemples présentés de cette catégorie]

Taux de fiabilité de la détection: [Nombre d'exemples de la catégorie correctement classés] / [(Nombre d'exemples de la catégorie correctement classés) + (Nombre d'éléments classés par erreur dans la catégorie considérée)]

Taux de certitude sur la détection: [Taux de détection] \* [Taux de fiabilité] / 100

Tableau 5 : Performances de classification de l'algorithme PmBc (taux en %)<sup>64</sup>

	Détection	Fiabilité	Certitude
<b>DL</b>	16,56	60,63	10,04
<b>ML</b>	9,29	65,83	6,12
<b>FL</b>	45,50	73,58	33,48
<b>L</b>	11,40	60,50	6,90
<b>G</b>	51,59	73,58	37,96
<b>Moy</b>	<b>26,87</b>	<b>66,82</b>	<b>18,90</b>

Remarque :

DG et FG renfermaient trop peu d'items pour autoriser un apprentissage de règles, tandis que MG ne renfermait aucun item. Ces 3 catégories ne sont donc pas renseignées dans le Tableau 5.

Le Tableau 6 présente les moyennes de détection, de fiabilité et de certitude de 3 algorithmes classificateurs (PmBc, ID3, NOPT - repris du travail de thèse de (Munteanu, 1996)) testés sur le corpus précédemment présenté, selon la même procédure.

Tableau 6 : Performance moyenne des 3 algorithmes dans la catégorisation des syllabes

Performance moyenne des 3 algorithmes par catégorie de syllabe (%)	Détection	Fiabilité	Certitude
<b>DL</b>	24.08	56.29	13.64
<b>ML</b>	11.18	45.41	5.90
<b>FL</b>	53.70	66.56	35.45
<b>L</b>	16.85	53.85	8.95
<b>G</b>	57.73	69.18	39.86
<b>Moyenne</b>	<b>32.71</b>	<b>58.26</b>	<b>20.76</b>

Les performances sont du même ordre de grandeur que celles de l'algorithme PmBc. Avec un taux moyen de 33% de détection correcte et une fiabilité moyenne de 58%, l'information prosodique a permis de catégoriser en moyenne 1 syllabe sur 5 avec certitude.

63 1591 items extraits de la base de données LABIS de BDSONS (Descout, Serignat, Cervantes, & Carré, 1986) Les items utilisés dans le test sont distribués de la manière suivante dans les catégories :

DG	MG	FG	G	DL	ML	FL	L	Total
27	0	27	556	345	141	345	150	1591

64 Deux catégories supplémentaires (début d'un mot grammatical, fin d'un mot grammatical) sont distinguées dans l'étude, mais les résultats obtenus ne sont pas présentés (lacunaires compte tenu de la faible quantité d'items, la présente étude n'ayant par ailleurs pas porté sur ces catégories).

Dans le cadre de ce corpus, les résultats montrent que les monosyllabes grammaticaux correspondent à la catégorie la mieux identifiée et présentent le meilleur taux de fiabilité, immédiatement suivie des fins de plurisyllabes lexicaux. Ce sont ensuite les débuts de plurisyllabes lexicaux, tandis que monosyllabes lexicaux et milieux de plurisyllabes lexicaux présentent des taux de reconnaissance et de fiabilité faibles. On aurait donc la meilleure détection pour les mots outils et pour les frontières de mots de contenu plurisyllabiques.

En dépit de performances de détection ne dépassant pas 60% (G), le procédé s'avère suffisant selon les auteurs pour améliorer de façon très significative les taux de reconnaissance des mots par le système, grâce à la réduction qu'il permet du nombre d'hypothèses de segmentations lexicales erronées. Les auteurs précisent qu'aucune amélioration des taux de classification prosodique des syllabes n'est remarquée avec une augmentation de la précision de l'analyse (nombre de sous-catégories supérieur).

Une seconde étude (Caillaud, 1996; Munteanu, 1996) a porté sur l'exploration des performances de classification de syllabes à partir de l'information prosodique par un linguiste expert, dans le but de découvrir des règles de classification nouvelles ou plus performantes et d'améliorer les algorithmes existants. La procédure de classification est similaire à celle utilisée avec les algorithmes (excepté le format de présentation des données (graphique), et la normalisation des données (par rapport à la valeur maximale atteinte pour chaque paramètre dans la production du locuteur considéré, ces valeurs étant donc exprimées en %)). Le test porte sur une quarantaine d'exemples, et les résultats sont présentés dans le Tableau 7.

Tableau 7 : Présentation des performances de classification des syllabes par un linguiste expert

<b>Décidé / Présenté</b>	<b>DL</b>	<b>FL</b>	<b>L</b>	<b>G</b>	<b>Total présentation</b>
<b>DL</b>	5	1	2	4	12
<b>FL</b>	4	9	4	0	17
<b>L</b>	0	3	0	2	5
<b>G</b>	1	1	0	8	10
<b>Total décision</b>	10	14	6	14	
	<b>Détection</b>		<b>Fiabilité</b>		<b>Certitude</b>
<b>DL</b>	41,67		50,00		20,83
<b>FL</b>	52,94		64,29		34,03
<b>L</b>	0,00		0,00		0,00
<b>G</b>	80,00		57,14		45,71
<b>Moyenne</b>	43,65		42,86		25,15



En dépit du faible nombre d'items utilisés dans cette évaluation, on retrouve une performance comparable à celle obtenue avec un algorithme : les monosyllabes grammaticaux sont les mieux identifiés, suivis des fins de polysyllabes lexicaux puis des débuts de polysyllabes lexicaux. Les monosyllabes lexicaux seraient mal reconnus (peu d'items présentés, toutefois).

Ces travaux montrent :

- ▷ La plausibilité de l'hypothèse d'un code prososyllabique, en rapport avec une information morphosyntaxique et topologique.
- ▷ L'utilité de ce code prososyllabique pour guider la démarcation lexicale (amélioration des performances du système).

#### **2.2.4 Récapitulation - Partie II, Chapitre 2**

La syllabe représente un bon candidat pour supporter un code prosodique à la base de tous les phénomènes prosodiques qui existent dans une langue. Les travaux de recherche effectués en reconnaissance automatique de la parole accèdent par ailleurs l'hypothèse d'un code prososyllabique, proposée comme une des bases à ce travail de doctorat.

## 2.3 Chapitre 3 - L'organisation superposante du lexique : quel rôle sur la reconnaissance des mots parlés ?

Le troisième chapitre présente des données sur l'organisation superposante du lexique français, issues d'analyses réalisées par l'auteur (Section 2.3.1). Ces données, mises en regard avec celles provenant d'études sur d'autres langues européennes, démontrent l'importance quantitative des superpositions lexicales formelles et son universalité dans les langues humaines, ce qui conduit à s'interroger sur son utilité. Un bilan des travaux de recherche disponibles dans la littérature sur le sujet est proposé dans la Section 2.3.2 page 72. Enfin, la Section 2.3.3 page 74 liste le comportement des principaux modèles psycholinguistiques de la reconnaissance des mots parlés face aux superpositions lexicales. Le Chapitre 3 se termine par une récapitulation des idées principales qu'il présente.

### 2.3.1 L'organisation superposante du lexique

Deux catégories de superpositions formelles existent dans le lexique : les superpositions non lexicales et les superpositions lexicales. La première catégorie de superpositions formelles ne sera pas abordée dans ce travail. Elle renvoie à des superpositions de diverse nature (traits, phonèmes, syllabes, morphèmes non lexicaux). La deuxième catégorie de superpositions est par contre au cœur de cette étude.

Deux types fondamentaux de superpositions lexicales formelles peuvent être définis : les superpositions lexicales réelles et les superpositions lexicales virtuelles.

Les superpositions lexicales réelles correspondent à la présence dans un mot d'au moins une forme phonologique enchâssée renvoyant à un ou plusieurs mots (mots fantômes - (Cutler, 2000)). Par exemple, le mot « *charpente* » renferme 11 formes phonologiques (/ʃa/, /ʃak/, /ak/, /akp/, /akpã/, /akpãt/, /pã/, /pãt/, /ãt/, /a/, /ã/) renvoyant à 21 mots d'après BRULEX (Content, Mousty, & Radeau, 1990). Cette configuration existe dans le lexique mental aussi bien que dans le signal acoustique. Elle est donc qualifiée de réelle puisqu'elle a une existence physique. Plus un mot renferme de phonèmes, plus il est probable qu'il soit impliqué dans une superposition lexicale réelle.

Les superpositions lexicales virtuelles correspondent à la possibilité pour un mot d'être enchâssé dans un autre (« enchâssabilité » d'un mot). Par exemple, la forme phonologique /pla/ (correspondant au mot « *plat* ») est enchâssable dans 144 mots

d'après BRULEX (/plas/, /plastik/, /kōtāplatif/, etc.). Cette configuration existe donc seulement dans le lexique mental, mais pas dans le signal acoustique. N'ayant pas d'existence physique, elle est donc qualifiée de virtuelle. Plus un mot renferme un nombre faible de phonèmes, plus il est probable qu'il soit impliqué dans une superposition lexicale virtuelle.

#### Quelques chiffres sur le phénomène des superpositions lexicales formelles

Les analyses statistiques effectuées sur de larges corpus en Anglais en Néerlandais révèlent que le phénomène des superpositions lexicales correspond à une règle dans l'organisation du lexique, et non à une exception. Plus de 98%<sup>65</sup> des mots anglais renferment au moins un mot superposé, moins de 2% n'en renferment aucun (Cutler et al., 2002). En Néerlandais, ces proportions sont respectivement de 99.4% et 0.6% (Cutler et al., 2002). Un mode de comptabilisation conservateur (concordance entre superposition lexicale et structure syllabique du mot porteur) ne réduit pas de façon notable cette proportion : 83.8% des mots plurisyllabiques anglais renferment toujours au moins un mot superposé (McQueen, Culter, Briscoe, & Norris, 1995).

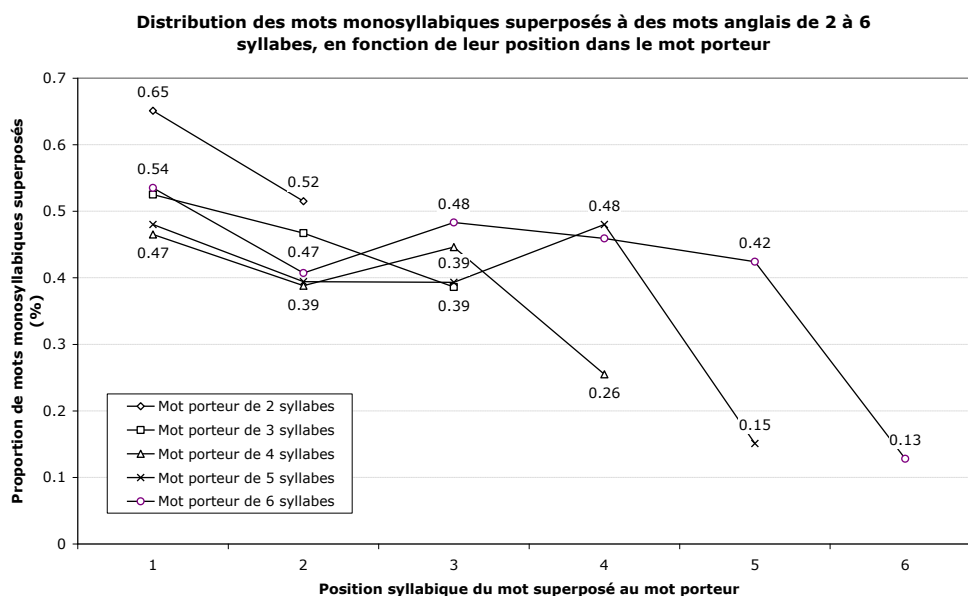
La répartition des superpositions lexicales dans le lexique anglais ne semble pas être aléatoire. Les superpositions lexicales qui impliquent des mots monosyllabiques composant les syllabes du mot porteur se distribuent selon un gradient opposant le début (quantité élevée de superpositions (53.12% en moyenne)) et la fin des mots porteurs (quantité moindre (28.7% en moyenne, soit 1.85 fois moins))<sup>66</sup>.

---

<sup>65</sup> Analyse d'un corpus de 71187 mots.

<sup>66</sup> Analyse sur 24279 mots Anglais (McQueen et al., 1995).

Figure 11 : Distribution positionnelle des mots monosyllabiques dans des mots anglais de 2 à 6 syllabes



Cela implique que plus d'un mot anglais polysyllabique sur 2 aurait pour première syllabe un mot, et moins d'un mot anglais polysyllabique sur 3 aurait pour dernière syllabe un mot. Pour ce qui concerne la répartition des mots polysyllabiques superposés, ce gradient disparaît (11.9% de mots polysyllabiques superposés en initiale, 11.5% en position finale). La quantité de mots anglais commençant par un mot polysyllabique (plus d'un mot sur 10) serait donc approximativement identique à la quantité de mots anglais finissant par un mot polysyllabique, et serait fortement réduite par rapport à la quantité de mots commençant ou finissant par un mot monosyllabique. L'existence d'un gradient de distribution des superpositions lexicales opposant le début et la fin des mots porteurs serait donc relatif à un phénomène de combinatoire lexico-syllabique.

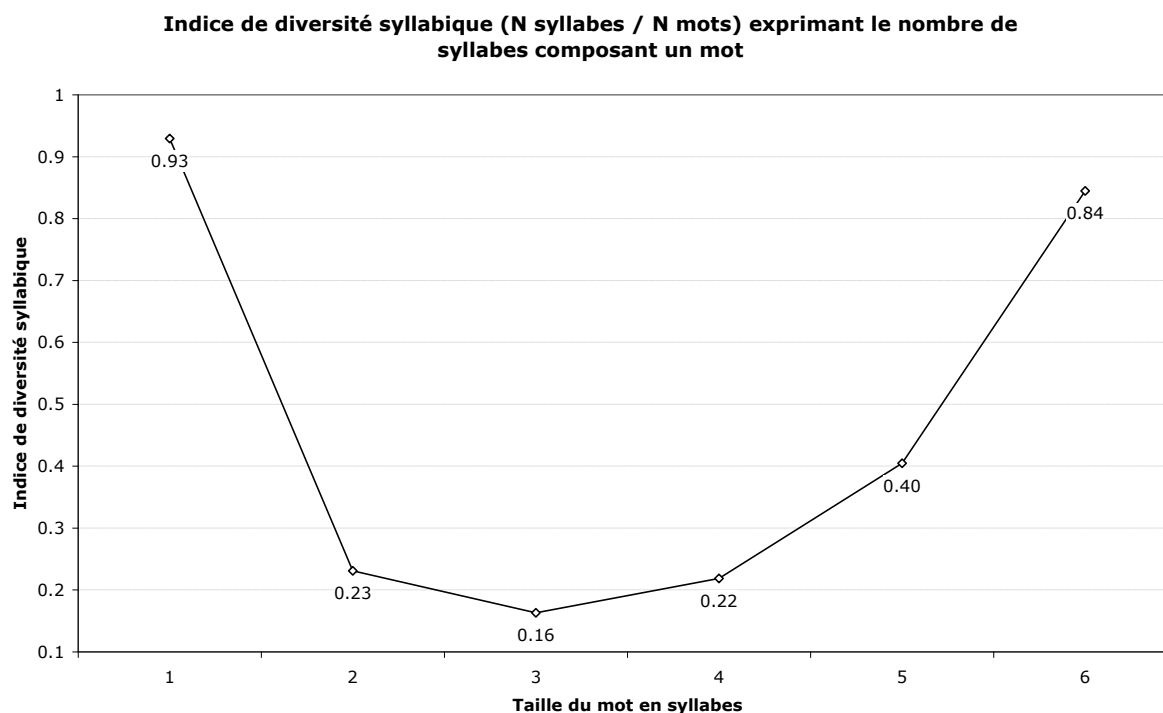
Partant de ces quelques observations, une analyse syllabique d'un lexique français issu de la base de données lexicales BRULEX a été effectuée à l'aide de scripts (Awk / Linux - C. Lachaud)<sup>67</sup>. Les résultats présentés plus bas portent sur 32375 entrées de 1 à 6 syllabes, constituées de 3972 syllabes. Toutes les catégories grammaticales sont représentées excepté les interjections. Les mots d'origine étrangère et les mots composés ont été filtrés. Le but de cette analyse était :

<sup>67</sup> La base de données BRULEX a été préalablement syllabée par l'auteur. Les frontières syllabiques ont été localisées sur la base du modèle de syllabation du français de Laporte (Laporte, 1993), opérationnalisé par la notion d'effort articulatoire (la production de la séquence de syllabes requiert un effort d'articulation minimal pour produire chaque syllabe - exemple : /kak/+/tys/, et non /ka/+/ktys/, la séquence /ktys/ nécessitant un effort articulatoire supérieur à la séquence /tys/, tandis que la séquence /kak/ n'est pas plus difficile, bien que plus complexe, à produire que la séquence /ka/).

- ▷ de disposer d'un outil de travail pour l'étude du traitement de la langue utilisée dans cette recherche.
- ▷ de disposer d'informations concernant l'organisation du lexique français.
- ▷ de confirmer les connaissances issues de l'analyse de lexiques anglais et néerlandais, avec le lexique français.

La Figure 12 présente un indice de diversité syllabique par mot (nombre de syllabes différentes / nombre de mots), en fonction de la taille des mots.

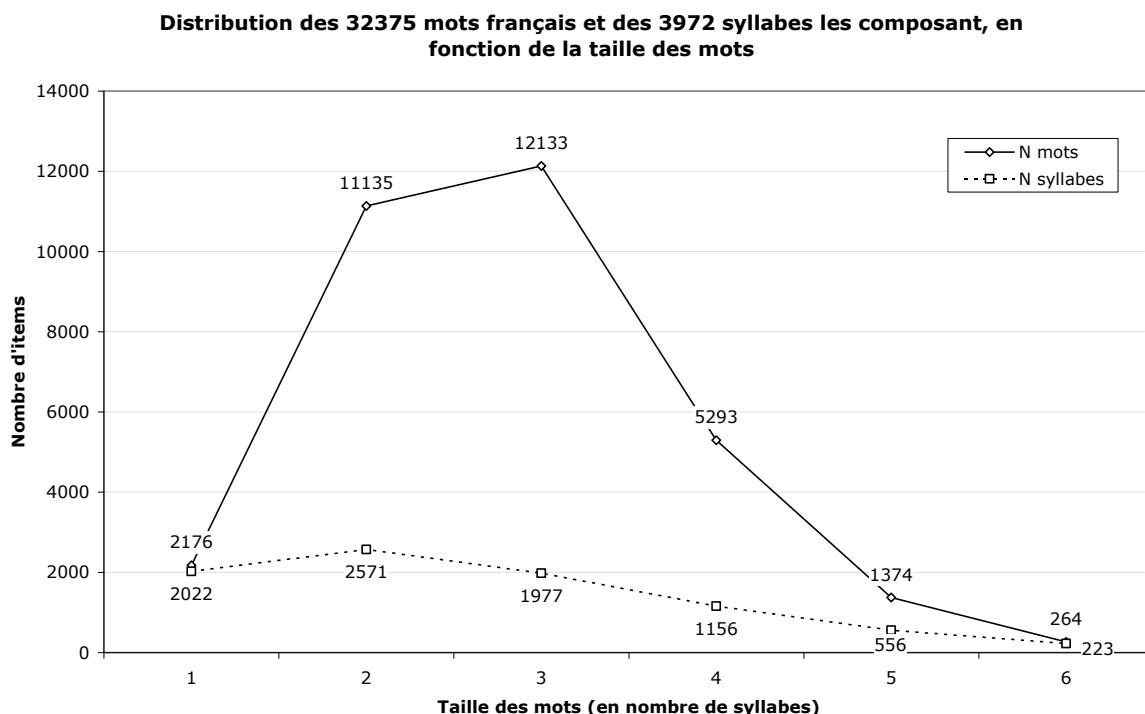
Figure 12 : Diversité syllabique des mots en fonction de leur taille



Une explosion de la combinatoire syllabique a lieu pour les mots de 2, 3 et 4 syllabes (N syllabes / N mots suit une courbe asymptotique, avec un minimum pour les mots de 3 syllabes). Ce sont donc les mots de 2 à 4 syllabes qui présentent la plus grande probabilité d'être impliqués dans des superpositions lexicales. L'indice n'est pas égal à 1 pour les mots monosyllabiques à cause de l'homophonie.

La Figure 13 montre la distribution des mots et des syllabes de ce lexique, en fonction de la taille des mots et de la position syllabique dans le mot.

Figure 13 : Distribution des mots du lexique analysé et des syllabes les composant, respectivement en fonction du nombre de syllabes qu'ils renferment / de la position syllabique dans le mot



La majorité des mots du lexique analysé sont bi, tri et quadrisyllabiques (au total 88.2% - respectivement 34.4%, 37.5% et 16.3%), les mots monosyllabiques ne représentant que 6.7% du corpus, et les mots de 5 et 6 syllabes totalisant le pourcentage restant (5.1%)<sup>68</sup>. La majorité des superpositions lexicales concernerait donc des mots bi ou trisyllabiques (71.9%), qui peuvent être porteurs de mots mono ou bisyllabiques.

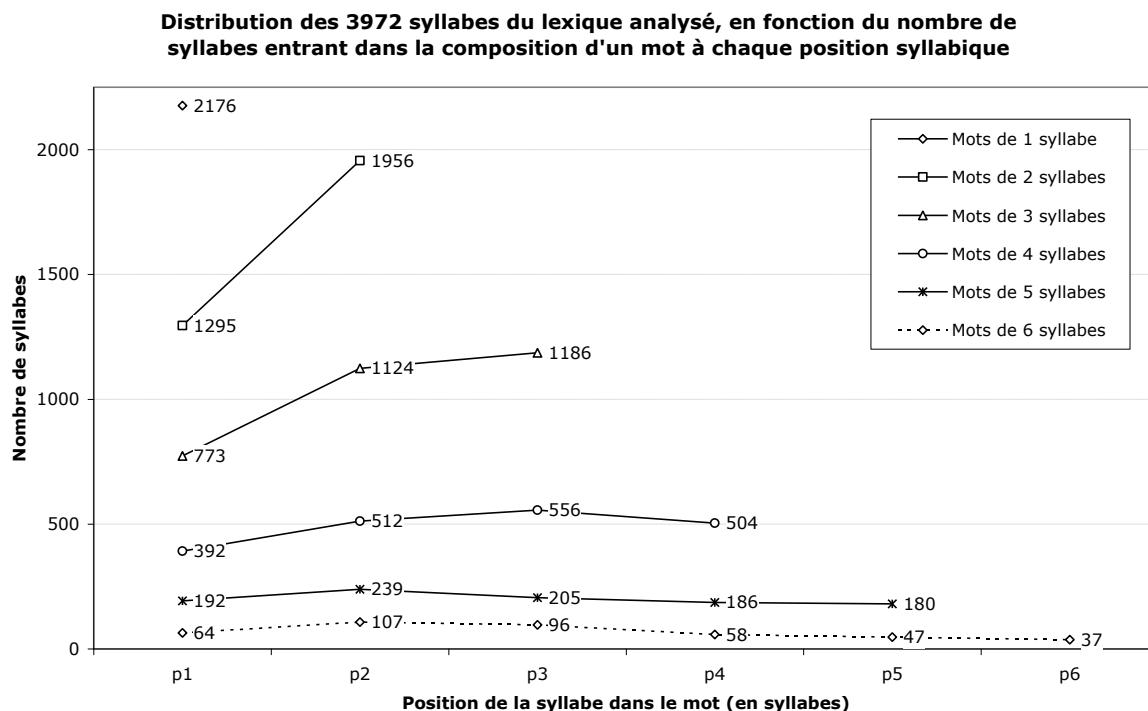
La diversité syllabique est plus importante sur les 2 premières positions syllabiques. Elle décroît ensuite presque linéairement (le calcul des pourcentages n'a pas été fait car une même syllabe peut être utilisée à plusieurs positions).

La Figure 14 présente l'organisation syllabique du lexique français analysé.

<sup>68</sup>

Le nombre de mots de plus de 6 syllabes est très faible.

Figure 14 : Organisation syllabique du lexique français analysé

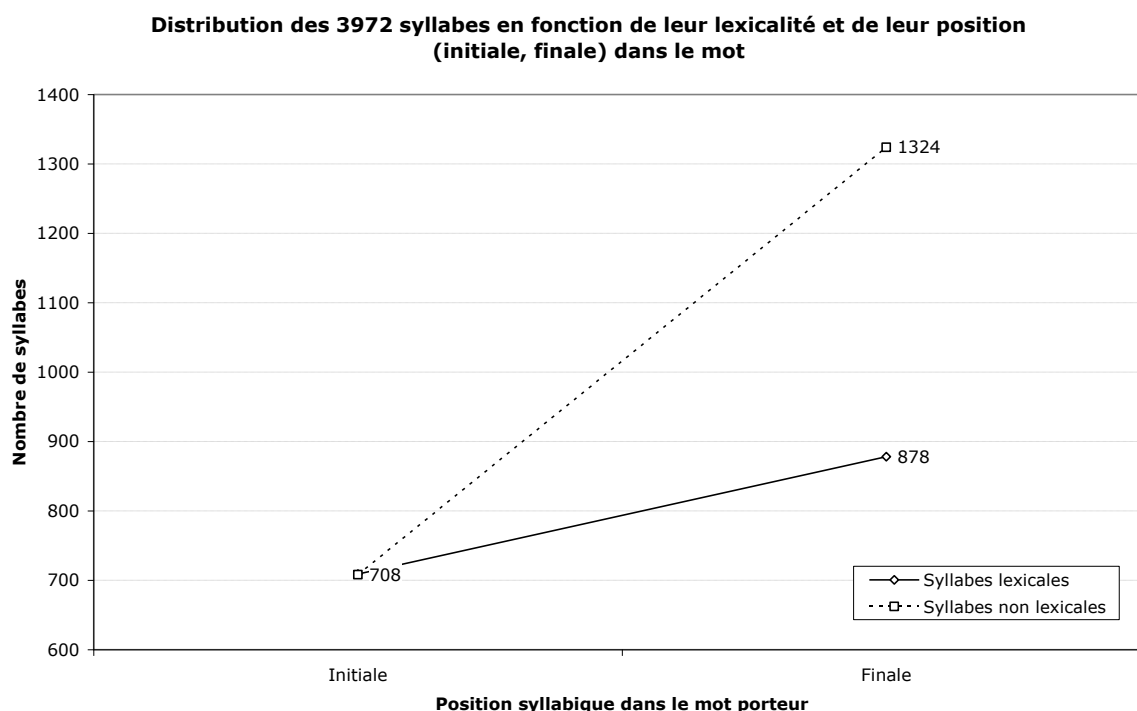


Elle montre que la diversité syllabique du lexique est toujours plus réduite sur la première syllabe d'un mot polysyllabique que sur la syllabe suivante. Ceci est particulièrement vrai pour les mots de 2 à 4 syllabes, qui composent la majorité des mots du lexique analysé (88.2%), et pour lesquels la diversité syllabique par mot est minimale. Grâce à cette organisation, le système pourrait disposer d'une facilité plus importante pour trouver des points d'entrée dans le lexique, du fait qu'ils sont moins nombreux (une faible combinatoire syllabique sur l'initiale des mots favoriserait le contact initial), et éventuellement, du fait qu'ils sont suivis d'un contraste combinatoire (la brusque augmentation des possibilités après la 1<sup>ère</sup> syllabe aurait pour conséquence de ralentir le traitement exploratoire et éviterait une segmentation lexicale en cours de mot). Cette structuration syllabique du système influencerait donc sa dynamique.

Une analyse plus poussée de la répartition des syllabes dans le lexique, prenant en considération leur lexicalité et leur localisation sur les zones-frontière (initiale vs. finale)<sup>69</sup>, est présentée dans la Figure 15.

<sup>69</sup> Le rang « milieu » de mot n'est pas considéré car il n'implique pas une frontière lexicale, et implique un nombre variable de positions selon la taille des mots considérés.

Figure 15 : Distribution positionnelle des syllabes en début et en fin de mots plurisyllabiques, en fonction de leur lexicalité



On y observe :

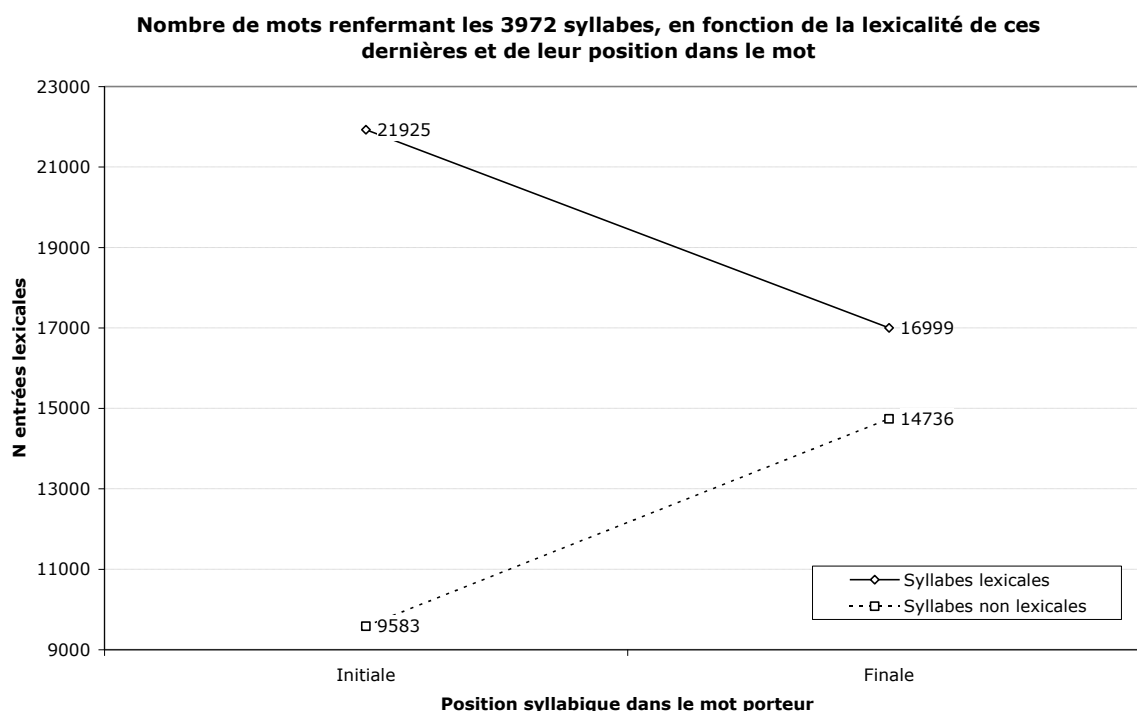
- ▷ une diversité syllabique identique entre syllabes lexicales et non lexicales sur l'initiale des mots du lexique analysé
- ▷ une diversité syllabique 1.5 fois plus importante des syllabes non lexicales par rapport aux syllabes lexicales pour la fin des mots
- ▷ une diversité syllabique 1.6 fois supérieure en moyenne sur la fin des mots, par rapport à la position initiale (1.9 fois plus importante pour les syllabes non lexicales, 1.2 fois plus importante pour les syllabes lexicales)

Cette analyse confirme que la diversité syllabique initiale est plus faible que la diversité syllabique finale. Ceci est vrai pour les syllabes lexicales comme pour les syllabes non lexicales.

Une diversité syllabique moindre sur l'initiale des mots peut être liée à un taux de superposition lexicale plus important à cette position. C'est ce qui a été confirmé par l'analyse, dont le résultat est présenté dans la Figure 16.



Figure 16 : Nombre d'entrées lexicales commençant / finissant par une syllabe lexicale / non lexicale

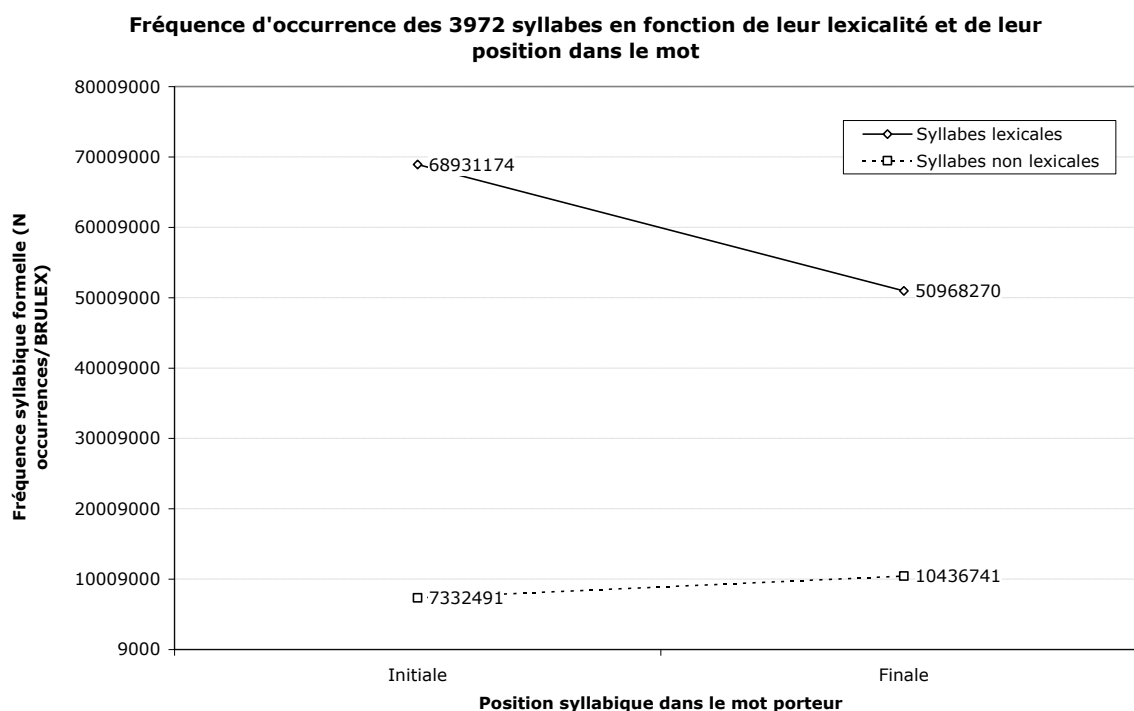


Le nombre de mots commençant par une syllabe lexicale est 2.3 fois plus important que le nombre de mots commençant par une syllabe non lexicale. Comparativement, cet écart est quasi insignifiant sur les fins de mots (1.2 fois plus de mots finissant par une syllabe lexicale que par une syllabe non lexicale)<sup>70</sup>.

La distribution des syllabes en fonction de la fréquence d'occurrence fournit approximativement un schéma identique (Figure 17).

<sup>70</sup> Remarque : la somme du nombre de mots pour la position initiale ou finale devrait correspondre au nombre de mots de la base de données analysée (N = 32375). Mais ce n'est pas le cas (N = 31508 pour la position initiale, N = 31735 pour la position finale). Cet écart (respectivement de 867 et 640) peut provenir d'une erreur dans un script ou d'une erreur de segmentation syllabique de la base de données. Le problème n'a pas pu être localisé et corrigé. Les résultats sont tout de même présentées en l'état, l'imprécision ayant un ordre de grandeur faible (script identique pour les 2 positions, écart moyen < 3 %). Cette erreur s'est aussi probablement répercutée sur l'estimation de la fréquence d'occurrence des syllabes (plus bas).

Figure 17 : Fréquence syllabique formelle en fonction de la lexicalité des syllabes et de leur position dans les mots



La fréquence avec laquelle les syllabes sont rencontrées est incomparablement plus élevée (6.7 fois) pour les syllabes lexicales que pour les syllabes non lexicales, aussi bien en position initiale (9.4 fois) qu'en position finale (4.9 fois). Il semblerait donc que le système préfère employer des mots qui renferment des unités à la fois infra-lexicales et lexicales.

Pour récapituler ces observations :

- ▷ On retrouve pour le Français ce qui était observé pour l'Anglais et le Néerlandais
- ▷ le lexique présente une diversité syllabique 1.6 fois moindre sur l'initiale des mots que sur leur fin
- ▷ il contient 2.3 fois plus de mots commençant par une syllabe lexicale que par une syllabe non lexicale (alors qu'en position finale, la différence s'estompe)
- ▷ dans l'usage de la langue française, on rencontre :
  - ∩ 9.4 fois plus souvent une syllabe lexicale à l'initiale, qu'une syllabe initiale non lexicale (cet écart diminue en position finale)
  - ∩ 6.7 fois plus souvent une syllabe lexicale en frontière de mot, qu'une syllabe non lexicale

Toutefois, ces proportions devraient être modulées si on considère les éléments apportés par le travail de thèse de Davis (2000). Reprenant et vérifiant le travail de

McQueen et collègues (1995), Davis effectue un comptage syllabique des superpositions lexicales initiales sur un lexique anglais (CELEX (Baayen, Pipenbrook, & Guilikers, 1995)) et remarque qu'une grande quantité de superpositions lexicales a une origine morphologique. Si les bases de données utilisées par McQueen et collègues (McQueen et al., 1995) et par Lachaud (BRULEX) ne renferment pas de formes fléchies (morphologiquement transparentes), elles contiennent en revanche des formes dérivées (estimées être en majorité sémantiquement transparentes (82.7%), contre 4.7% désignées comme opaques selon Davis (2000)) et des mots composés (ces derniers ont cependant été éliminés de l'analyse effectuée par Lachaud – celle-ci portait donc aussi sur des mots morphologiquement dérivés). Dans les cas où la superposition lexicale initiale implique la racine du mot porteur (exemple : « *blanc* » et « *blancheur* »), la structure de l'item serait formellement (et sémantiquement) transparente, et la reconnaissance du mot porteur serait favorisée (le système ciblerait d'abord la famille morphologique du mot dans le lexique (« *blanc* » et dérivés), puis préciserait sa sélection au sein du sous-ensemble). Par conséquent, l'influence des superpositions lexicales initiales morphologiques sur la reconnaissance des mots pourrait être contraire à celle des superpositions lexicales initiales non morphologiques. Aussi, il serait nécessaire de distinguer les cas de figure dans une estimation de la distribution des superpositions lexicales dans le lexique.

Davis fournit les résultats d'une analyse distinguant les superpositions lexicales initiales de nature morphologique des superpositions lexicales initiales non morphologiques, repris dans le Tableau 8.

Tableau 8 : Proportion de mots enchâssés à l'initiale, en fonction du caractère morphologique et non morphologique des superpositions lexicales

Caractéristiques du comptage	Proportion (%)
Tous types de superpositions initiales syllabiques	50.5
Sans superpositions initiales morphologiques de type dérivationnel	39
Sans superpositions initiales morphologiques (mots dérivés + mots composés)	26.9

Un peu plus du quart des mots anglais renfermerait une superposition lexicale monosyllabique initiale non morphologique, ce qui représente une proportion encore importante, et de surcroît sous-estimée par le comptage syllabique. On peut donc supposer qu'il en est de même pour le Français, vu que les analyses présentées précédemment montrent les mêmes résultats que celles de Cutler et collègues (2002) pour l'anglais. Ainsi, en Français, la reconnaissance d'au moins 25% des mots serait susceptible de poser un problème (la quantité réellement rencontrée par les auditeurs, prenant en considération la fréquence d'usage des mots, ne peut cependant pas être

estimée ici). Ce chiffre estimé restant élevé, il faut en déduire que l'organisation superposante du lexique a probablement une utilité particulière pour le processus de reconnaissance des mots, en tous cas pour ce qui concerne les superposition lexicales initiales. On peut proposer que la présence de superpositions lexicales initiales facilite :

▷ la reconnaissance des mots

les superpositions monosyllabiques initiales favoriseraient la reconnaissance des mots porteurs parce que les lexèmes sont organisées selon un schéma superposant qui leur permet de se potentialiser mutuellement. Les superpositions lexicales finales seraient au contraire perturbatrices (morcelante) ou sans influence (reconnaissance déjà effectuée ou suffisamment avancée).

▷ la segmentation du flux sonore en mots

∩ Soit parce que les superpositions monosyllabiques initiales ralentissent la reconnaissance du mot porteur, permettant ainsi au système de traiter plus d'information pour localiser les frontières des mots et avoir le temps de lever les ambiguïtés, qui représentent véritablement un problème dans le traitement (les superpositions monosyllabiques finales seraient moins perturbatrices ou seraient sans influence (reconnaissance déjà effectuée ou suffisamment avancée du mot porteur)).

∩ Soit parce que les superpositions monosyllabiques initiales sont sans influence sur la reconnaissance du mot porteur<sup>71</sup>, mais elles modulent le poids du début des mots dans l'espace perceptif, favorisant ainsi la localisation dans la chaîne parlée des points d'entrée pour l'accès lexical (les superpositions lexicales finales doivent donc être moins fréquentes pour autoriser un codage de ce type).

Bien que l'influence des mots enchâssés soit probablement réduite lors du traitement de la parole grâce à l'intégration de sources multiples d'information formelle et conceptuelle <sup>72</sup>, l'évaluation de l'impact des superpositions lexicales sur la reconnaissance des mots parlés isolés nécessite une investigation particulière, étant donné l'ampleur de ce type d'organisation et le caractère paradoxal d'une telle structure,

---

<sup>71</sup> Au début du traitement, à cause du traitement séquentiel de la parole, le système n'a reconnu qu'un mot et se trouve face à des superpositions lexicales virtuelles. L'influence perturbatrice des superpositions lexicales initiales pourrait être plus tardive au cours du traitement (lorsque la superposition devient réelle), mais elle serait alors limitée (le système laisse tomber la première hypothèse pour une autre sans conséquence).

<sup>72</sup> Par exemple, information formelle de nature acoustique (prosodique et articulatoire, qui permet de moduler la prégnance sensorielle des mots enchâssés à l'initiale, donc de diminuer leur impact perceptif), syllabique (la règle PWC permettrait d'éliminer jusqu'à 73% des mots enchâssés en Anglais (Cutler et al., 2002)). Toutefois, cette règle est inutile dans le cas d'un comptage syllabique comme ceux dont les statistiques ont été présentées et discutées dans cette section), grammatical (la redondance morphématique (voire note 26 page 25) pourrait restreindre les possibilités lexicales en créant des attentes du système vis à vis de catégories grammaticales, et en excluant d'autres. Cependant, l'influence de ce filtre grammatical n'est pas claire : 33.8% des mots d'une catégorie grammaticale donnée renferment à leur initiale des mots enchâssés de la même catégorie (McQueen et al., 1995)), information sémantique (le contexte sémantique permettrait de focaliser le système sur un jeu restreint de mots sémantiquement congruents).

si elle s'avère être, comme le prétendent certains modèles de la psycholinguistique, d'un traitement plus difficile.

### 2.3.2 L'étude du traitement des superpositions lexicales dans la littérature : principaux courants théoriques et approches expérimentales

La majorité des études qui abordent le phénomène des superpositions lexicales :

- ▷ ont cherché à montrer que les mots enchâssés (superpositions lexicales réelles) étaient activés lors du traitement du mot porteur, en mesurant l'effet résiduel de leur activation dans une tâche postperceptive. À cette fin, le paradigme d'amorçage (formel ou sémantique) a généralement été utilisé dans une tâche de décision lexicale (par exemple (Davis, Marslen-Wilson, & Gaskell, 2000, 2002; Isel & Bacri, 1999; Shillcock, 1990; Spinelli, Segui, & Radeau, 2001; Vroomen & van Gelder, 1997)), avec les notions de densité de voisinage et de fréquence lexicale en toile de fond théorique. Selon ces travaux, aussi bien les mots enchâssés à l'initiale (effets de cohorte : /vɛɪk/ amorce /vɛɪtɪz/ (Spinelli et al., 2001)) que les mots enchâssés en fin de porteur (« trombone » amorce « rib » par le mot enchâssé « bone » (Shillcock, 1990)) seraient activés lors du traitement du mot porteur. Cependant, les observations de Vroomen et van Gelder (1997) (un effet mesurable d'amorçage est obtenu lorsque le mot est enchâssé à l'initiale d'un pseudo-mot, et non lorsqu'il est enchâssé à l'initiale d'un mot) et d'Isel et Bacri (Isel & Bacri, 1999) (il n'y a pas d'effet résiduel mesurable de la présence d'un mot enchâssé à l'initiale d'un mot porteur), indiquent que l'activation d'un mot monosyllabique enchâssé initial est très éphémère, et qu'elle est masquée dans le temps par la reconnaissance du mot porteur. L'étude de Bowers et collègues (2005), apporte également des éléments en faveur de l'activation des mots enchâssés grâce à une procédure originale (tâche de « compétition sémantique », qui mesure également un résidu du traitement dans le système, et non l'effet direct des superpositions sur le traitement du mot stimulus (la variable dépendante mesurée est relative à une question posée après le traitement visuel d'un mot)). Les mots superposés aux items (superpositions réelles et superpositions virtuelles) sont sémantiquement activés quelle que soit la position de la superposition (initiale, médiale ou finale).
- ▷ ont étudié les implications de la présence d'un enchâssement (superposition lexicale réelle) pour la segmentation du flux de parole en mots (par exemple (Cutler & van Donselaar, 2001; Davis, Gaskell, & Marslen-Wilson, 1997a, 1997b; McQueen et al., 1994)).
- ▷ ont parfois inversé le problème en étudiant la reconnaissance des mots enchâssés (superposition lexicale réelle - par exemple (Davis et al., 1997a; Davis et al., 2000)), au lieu de la reconnaissance des mots porteurs de mots enchâssés.
- ▷ ont abordé l'influence des superpositions lexicales virtuelles comme un effet de cohorte et de voisinage (travaux de Marslen-Wilson, Luce, Goldinger et Vitevitch), éventuellement comme un phénomène morphologique (par exemple (Andrews, 1986)).
- ▷ se sont intéressés au traitement des superpositions morphologiques (superpositions lexicales réelles), essentiellement dans l'optique de tester

l'hypothèse décompositionnelle des mots morphologiquement complexes (par exemple (Andrews, 1986; M. Taft & Forster, 1976)).

On peut donc dire que l'organisation superposante du lexique, et plus particulièrement les superpositions lexicales, représente un problème central de la psycholinguistique, mais qu'il n'a pas, jusqu'ici, été formellement identifié comme tel par la discipline, puisque abordé de manière atomiste à travers des sous-catégories spécifiques de phénomènes (morphologie, segmentation lexicale, voisinage, point d'unicité). Toutes ces sous-catégories, reliées à un problème plus générique (organisation superposante du lexique mental), se confondent en effet partiellement avec la supercatégorie « superposition lexicale ».

Laisant de côté les études qui ont abordé l'impact des superpositions lexicales virtuelles sur la reconnaissance des mots stimuli (mots pouvant être l'initiale d'un mot plus long - problèmes de réduction de cohorte et de traitement morphologique), seulement 2 études essayant de comprendre l'effet direct de superpositions lexicales réelles sur la reconnaissance de mots porteurs ont pu être trouvées dans la littérature.

La première de ces 2 études a été réalisée par Taft et Forster (1976). Elle traitait de la reconnaissance de mots plurisyllabiques, mono ou pluri-morphémiques, en modalité visuelle. La seconde, réalisée par Luce et Lyons (1999), traitait de la reconnaissance de mots bisyllabiques monomorphémiques, en modalité auditive.

### **2.3.2.1 L'étude de Taft et Forster (1976)**

Focalisée autour du problème du stockage des mots plurisyllabiques et de leur accès, cette étude montre que la reconnaissance visuelle des mots s'effectue probablement sur la base de la première syllabe, et non des syllabes non initiales. La 5<sup>ème</sup> expérience décrite dans l'article (tâche de décision lexicale) révèle que la reconnaissance du mot porteur est affectée par la fréquence du mot superposé qui compose sa première syllabe. Comparant 20 paires de mots porteurs (exemple : « *headstand* » et « *loincloth* »), appariés en fréquence, les auteurs trouvent que celui qui renferme initialement le mot le plus fréquent (« *head* ») est reconnu le plus rapidement (598 vs. 629 ms.). L'interprétation qu'ils en donnent, dans le contexte théorique d'il y a 30 ans (modèle de Forster), est que les mots sont accédés sur la base de leur première syllabe lors de leur traitement visuel.

### 2.3.2.2 L'étude de Luce et Lyons (1999)

Dans 2 expériences utilisant toutes deux une tâche de décision lexicale et une tâche de répétition (« shadowing »), les auteurs explorent l'incidence sur la reconnaissance auditive d'un mot porteur bisyllabique, de la présence d'un mot enchâssé en position initiale et formant la première syllabe du porteur (Expérience 1 – N = 18 paires (« *cherish* » (« *chair* ») vs. « *flourish* »)), ou en position finale et formant la deuxième syllabe du porteur (Expérience 2 – N = 26 paires (« *chloride* » (« *ride* ») vs. « *chlorine* ») – superpositions non morphologiques). Les items sont fortement contrôlés de manière à être appariés sur des critères de fréquence, de durée, de position du point d'unicité et de familiarité. Les données sont analysées classiquement, par ANOVAs séparées sur les sujets et sur les items.

Les résultats montrent que les mots renfermant une première syllabe lexicale sont classés et répétés plus rapidement que ceux dont la première syllabe n'est pas lexicale (tâche de décision lexicale : 939 vs. 991 ms. ; shadowing : 819 vs. 858 ms.). Par contre, la mesure ne permet pas de montrer que le traitement des mots porteurs ait été modifié par la présence d'un enchâssement final (tâche de décision lexicale : 1069 vs. 1071 ms. ; « shadowing » : 975 vs. 965 ms.).

Les auteurs concluent de ces résultats que :

- ▷ La présence d'un mot enchâssé initial favorise la reconnaissance du mot porteur, phénomène contradictoire avec les prédictions théoriques qui prévoient une interférence lexicale (exemple : TRACE).
- ▷ Bien que le mot enchâssé en position finale soit probablement activé (étude de Vroomen et de Gelder (1997)), son impact sur la reconnaissance du porteur est réduit et non mesurable.

En résumé, la littérature indique que les mots superposés sont transitoirement activés lors du traitement des mots stimuli. Elle fournit cependant très peu d'observations relatives à l'impact effectif des superpositions réelles non morphologiques sur la reconnaissance du mot stimulus ou porteur. Lorsqu'elle le fait, des résultats paradoxaux sont observés.

### 2.3.3 Les prédictions des modèles psycholinguistiques de la reconnaissance des mots parlés face aux superpositions lexicales réelles

Le Tableau 9 présente les prédictions de quelques modèles psycholinguistiques de la reconnaissance des mots parlés, spécifiquement lors du traitement de superpositions lexicales réelles.

Tableau 9 : Temps de reconnaissance prédits par quelques modèles psycholinguistiques de la reconnaissance des mots parlés, lors du traitement de superpositions lexicales réelles initiale ou finale

Prédictions	Superposition initiale	Superposition finale
TRACE, SHORTLIST et PARSYN	Allongement des TR. E et P entrent en compétition au début du traitement (inhibition latérale). Ensuite, P est trop activé par les données et emporte le conflit.	Pas d'effet ou faible allongement des TR. P est trop activé pour être perturbé par une compétition tardive, et empêche E de s'activer (inhibition latérale).
COHORT, FUL	Allongement des TR seulement en cas de superposition lexicale initiale virtuelle (décalage du PU après la fin du mot)	Pas d'effet. Le système continue à apparier l'information tant qu'elle correspond à l'un des candidats de la cohorte.
DCM	Allongement des TR seulement en cas de superposition lexicale initiale virtuelle. Les effets de compétition existent mais disparaissent dès qu'il y a suffisamment d'information ascendante pour distinguer E de P.	

Avec : E = mot enchâssé ; P = mot porteur

Aucun des modèles théoriques actuels de la psycholinguistique n'envisage de mécanismes susceptibles de produire des effets facilitateurs liés aux superpositions lexicales formelles, tels que ceux qui sont rapportés dans les deux études citées à la section précédente. Pour cette raison, les effets obtenus par les 2 études citées à la Section 2.3.2 page 72 correspondent à une aberration théorique.



### **2.3.4 Récapitulation - Partie II, Chapitre 3**

Le lexique est une structure massivement superposante, dans laquelle les superpositions lexicales impliquant des mots enchâssés monosyllabiques ne sont pas réparties aléatoirement. Elles se répartissent massivement à l'initiale des mots porteurs plurisyllabiques, ce qui laisse penser que cette organisation joue un rôle pour la recherche et/ou l'activation lexicale.

Quelques travaux sur la reconnaissance des mots parlés rapportent un phénomène de facilitation dû à la présence de telles superpositions lexicales à l'initiale des mots stimuli. Ces effets ne s'expliquent pas dans le cadre théorique actuel de la psycholinguistique. Celui-ci prévoit, en cas de superposition lexicale initiale, soit une difficulté accrue du traitement, soit une absence d'effet, mais jamais de facilitation.

Les observations sur l'organisation du lexique corroborées aux effets de facilitation dus aux superpositions lexicales initiales, suggèrent l'existence possible d'une controverse théorique.

## 2.4 Chapitre 4 – Récapitulation de la Partie II : idées clé et des hypothèses fondant ce travail de recherche

Au terme de la présentation du cadre théorique dessiné pour ce travail de recherche, on dispose des éléments suivants :

- ▷ Le lexique présente une structure massivement superposante (superpositions lexicales). L'accès ascendant au lexique serait donc équivoque durant au moins l'une des étapes du processus.
- ▷ Les modèles psycholinguistiques de la reconnaissance des mots parlés se basent uniquement sur le traitement de la ligne segmentale. À cause d'un filtrage phonétique et phonologique, et à cause de la structure superposante du lexique, ces systèmes doivent faire face à une explosion combinatoire des solutions lexicales. Ils résolvent ce problème par la mise en œuvre de procédures de sélection.
- ▷ Quelles que soient les procédures de sélection utilisées, elles impliquent un traitement plus complexe, plus difficile et parfois plus long. Le modèle TRACE de la reconnaissance des mots parlés, pris comme cadre de référence pour la partie de ce travail relative au rôle des superpositions lexicales sur la reconnaissance des mots (Partie IV, Chapitre 2), propose un mécanisme d'inhibition réciproque des candidats qui se traduit par un allongement des temps de reconnaissance du mot stimulus.
- ▷ Les travaux en psycholinguistique et en reconnaissance automatique de la parole montrent que l'information de la ligne suprasegmentale est importante pour guider la démarcation lexicale. La prosodie permettrait d'éliminer les alternatives inutiles, soit en évitant l'activation de candidats multiples, soit en appuyant les processus de sélection. Par conséquent, les modèles psycholinguistiques qui ne prennent pas en considération cette dimension de la parole sont forcément incomplets, voire faux.

À partir de ces 4 points fondamentaux, les 2 hypothèses de travail suivantes ont été proposées :

- ▷ Les syllabes véhiculent des formes prosodiques en lien étroit avec des facteurs topologiques et morphosyntaxiques. Ces formes prosodiques décrivent un code « prososyllabique », qui est exploité par le système pour procéder à la démarcation lexicale.
- ▷ La structure superposante du lexique (superpositions lexicales), qui est prédite comme un handicap pour la reconnaissance des mots parlés selon le modèle TRACE, serait au contraire utile au processus. En effet, dans le cas contraire, cette structure se rencontrerait exceptionnellement, ou le niveau de performance avec lequel nous reconnaissons les mots parlés serait moindre.

### 3 Partie III – Méthodologie

Le but de la partie méthodologique est de présenter un ensemble de problèmes techniques, relatifs à l'étude de la reconnaissance des mots parlés à l'aide du paradigme de la chronométrie mentale (tâche de décision lexicale, tâche de détection de cibles), ainsi que les solutions qui ont été adoptées pour les résoudre. Cette partie vise également à présenter en détail et une fois pour toutes la méthodologie expérimentale commune aux expériences. Lors de la présentation de chaque expérience, le lecteur y est donc renvoyé.

La partie méthodologique comporte 5 chapitres :

- ▷ Chapitre 1 : Présentation des problèmes méthodologiques et techniques rencontrés
- ▷ Chapitre 2 : L'origine de la mesure des temps de réponse lors de la reconnaissance d'un mot parlé isolé
- ▷ Chapitre 3 : Le traitement statistique des temps de réaction pour reconnaître un mot parlé isolé
- ▷ Chapitre 4 : Méthodologie expérimentale
- ▷ Chapitre 5 : Récapitulation technique et théorique pour chaque expérience

Le premier chapitre donne une vue d'ensemble des principaux problèmes méthodologiques et techniques qui ont été rencontrés dans l'étude, ainsi que les divers moyens qui ont été mis en œuvre pour les résoudre aux différentes étapes du travail.

Le deuxième chapitre est spécifiquement consacré au problème de l'origine de la mesure des temps de réaction des auditeurs, qui est absolument crucial avec l'utilisation de stimuli temporels. Ce chapitre présente également la solution qui a finalement été adoptée pour éviter tout biais dû à la temporalité des stimuli.

Le troisième chapitre est dédié à la méthodologie statistique, avec une justification des choix qui ont conduit à la mise en place de la procédure adoptée.

Le quatrième chapitre présente la méthodologie expérimentale elle-même.

Le cinquième chapitre fournit deux tableaux synoptiques des aspects techniques et théoriques mis en œuvre dans les expériences (plan expérimental de l'étude).

### **3.1 Chapitre 1 - Présentation des problèmes méthodologiques et techniques rencontrés**

Ce premier chapitre est relativement général, et doit donc être considéré comme une introduction à la troisième partie de ce rapport. Il présente les principaux problèmes méthodologiques et techniques qui ont été rencontrés dans l'étude, et introduit les points techniques qui ont permis de les résoudre aux différentes étapes du travail. Ces aspects techniques visent pour l'essentiel à l'obtention d'une procédure de recueil des données précise et fiable, et de résultats aussi exacts que possible. Le chapitre est subdivisé en deux sections :

- ▷ La première section récapitule les éléments influençant d'une part la production des données comportementales par les participants aux expériences, et d'autre part l'obtention des résultats, afin de décrire la complexité de la situation au lecteur.
- ▷ La deuxième section présente les étapes du travail où des solutions ont été appliquées pour contrôler la variance indésirable.

En théorie, l'expérimentation permet de disposer de mesures d'une pureté élevée, c'est à dire relatives et seulement relatives à la relation qui est étudiée entre variable(s) indépendante(s) et variable(s) dépendante(s).

Pour de multiples raisons, il peut cependant être techniquement difficile :

- ▷ de manipuler une VI, de mesurer une VD, de contrôler une VP
- ▷ de ne pas biaiser les résultats avec la procédure d'analyse des données (préparation des données, méthode d'analyse).

Afin de comprendre pourquoi ces difficultés existent, les facteurs qui influencent la production des données et des résultats dans une étude psycholinguistique sur la reconnaissance des mots parlés basée sur le paradigme de la chronométrie mentale, sont d'abord présentés.

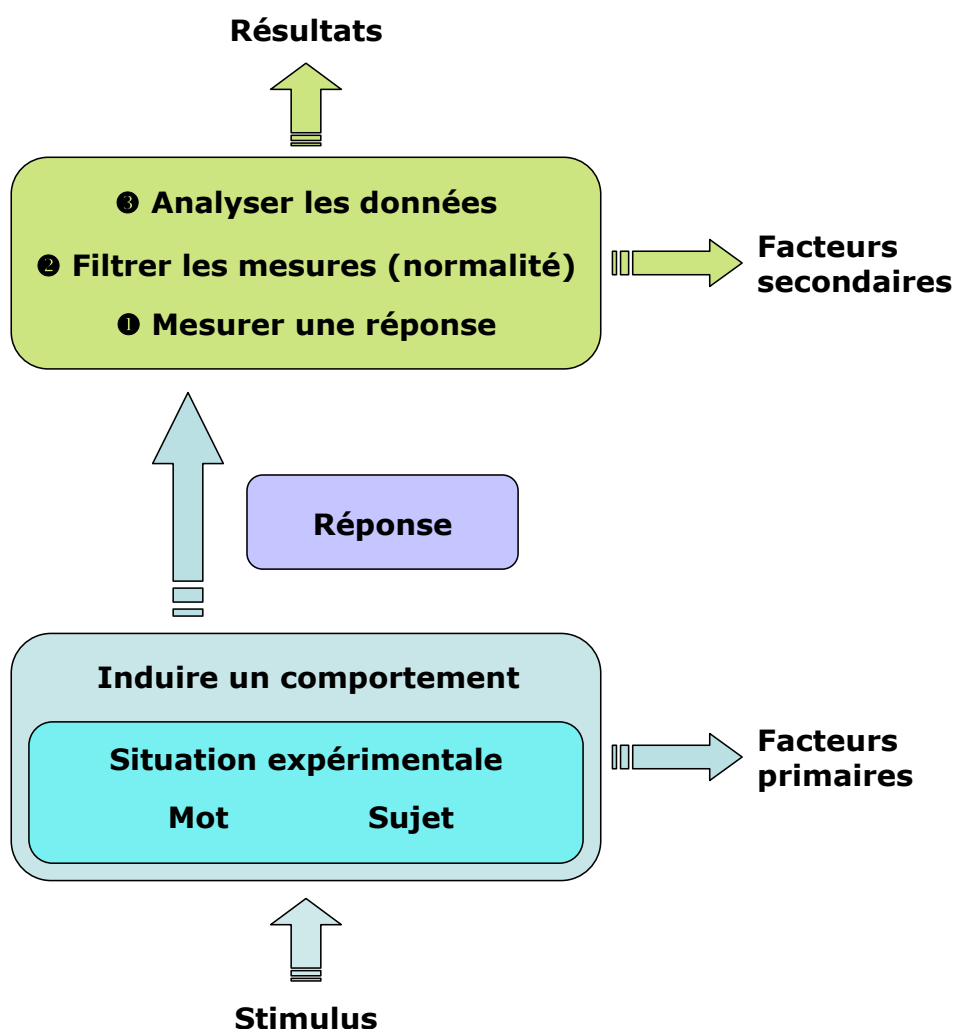
Dans la deuxième section de ce chapitre, un plan présentant les étapes d'une expérimentation, auxquelles une procédure peut être appliquée pour résoudre un problème particulier, est présenté au lecteur.

### 3.1.1 Description des facteurs influençant la production des données et des résultats

En psycholinguistique, les difficultés pour manipuler, contrôler et analyser une situation expérimentale sont multiples. Elles sont répertoriées ci-dessous comme étant relatives aux :

- ▷ Facteurs influençant la production des réponses mesurées (facteurs primaires, liés aux mots, aux sujets et à la situation expérimentale)
- ▷ Facteurs influençant la production de résultats interprétables (facteurs secondaires, liés à la mesure des données, au protocole de préparation des données pour les analyses, et aux méthodes d'analyse des données).

Figure 18 : Facteurs influençant les résultats dans une étude expérimentale sur la reconnaissance des mots parlés



### 3.1.1.1 Facteurs primaires influençant la production des réponses mesurées

Il s'agit de facteurs ou dimensions qui déterminent la réaction d'un individu à un mot donné. Ils sont relatifs aussi bien aux mots et à leurs représentations mentales dans les sujets (multidimensionnalité des mots) qu'aux sujets eux-mêmes (multidimensionnalité des sujets). Ils font aussi référence à des éléments de la situation expérimentale (autres dimensions de la situation expérimentale influençant la production des réponses).

La maîtrise des situations étudiées expérimentalement en psycholinguistique est d'une grande complexité, à cause de l'existence de nombreuses variables parasites, qui peuvent être corrélées aux VI et/ou aux VD<sup>73</sup>. Les paragraphes qui suivent détaillent les dimensions impliquées dans le comportement d'un individu face à un mot donné.

#### 3.1.1.1.1 Multidimensionnalité des mots

Tout mot parlé se décrit sur un ensemble de dimensions répertoriées selon 3 catégories :

- ▷ Dimensions physiques : elles correspondent aux aspects du monde physique qui permettent de décrire formellement les aspects linguistiques d'un mot selon un code propre au substrat physique considéré (mot physique).
- ▷ Dimensions linguistiques : elles correspondent aux aspects formels et abstraits des mots, déterminés par consensus social, qui permettent de décrire chaque item d'une langue de façon théorique (mot théorique).
- ▷ Dimensions psychologiques : elles correspondent aux aspects représentationnels des mots, qui relèvent de l'incarnation dans l'individu des dimensions linguistiques formelles et abstraites, et éventuellement, des dimensions physiques d'un mot. Les dimensions psychologiques sont supposées être partiellement spécifiques à l'individu (« mots dans les sujets »), et partiellement partagées par l'ensemble des locuteurs d'une langue (aspect supra-individuel relatif au niveau des items).

##### 3.1.1.1.1.1 Dimensions physiques des mots parlés (mot physique)

Les dimensions physiques qui caractérisent un mot dans l'onde acoustique sont au nombre de 3 :

- ▷ Durée
- ▷ Intensité (amplitude de l'onde acoustique)
- ▷ Fréquence (vitesse de l'oscillation)

---

<sup>73</sup> Par exemple, ces corrélations peuvent avoir lieu entre variables liées aux items (exemple : corrélation entre le nombre de phonèmes (mot) et le nombre de voisins phonologiques (mot)), ou entre variables liées aux items et variables liées aux sujets (exemple : corrélation entre le nombre de phonèmes (mot) et la familiarité conceptuelle (sujet)).

**3.1.1.1.1.2 Dimensions linguistiques des mots parlés (mot théorique)**

Sur le plan linguistique, un mot parlé se décrit dans 5 dimensions, elles-mêmes réparties en forme (aspects relatifs au signifiant) et contenu (aspects relatifs au signifié) :

- ▷ Dimension phonologique (forme) : Description des mots parlés en termes de traits distinctifs, de phonèmes, d'agencement de ces phonèmes selon des règles phonotactiques, de structure syllabique, de structure prosodique.
- ▷ Dimension morphosyntaxique (forme) : Description de la manière dont les mots véhiculent les contrastes grammaticaux (morphologie inflexionnelle), et de la manière dont de nouveaux mots peuvent être créés (morphologie dérivationnelle).
- ▷ Dimension syntaxique (contenu) : Description de la manière dont les mots, répartis en catégories syntaxiques, peuvent être agencés en phrases, selon des règles permettant d'exprimer les rapports de sens et les relations entre les signes.
- ▷ Dimension sémantique (contenu) : Description des relations entre les signes (connaissances linguistiques) et le monde (connaissances extralinguistiques).
- ▷ Dimension pragmatique (contenu) : Description du rapport entre les signes et les usages de ces signes (rôle du contexte et des connaissances partagées dans les choix d'utilisation des mots).

On peut cependant réorganiser cette structure théorique selon une perspective plus factuelle, puisqu'elle renseigne en fin de compte sur :

- ▷ La composition des mots
- ▷ Les relations entre mots
- ▷ Les relations entre les mots et le monde

**3.1.1.1.1.2.1 La composition des mots**

- ▷ Éléments constitutifs
  - ∩ Forme : traits phonétiques, phonèmes, syllabes, morphèmes
  - ∩ Contenu : traits sémantiques, morphèmes
- ▷ Organisation des éléments constitutifs
  - ∩ règles phonotactiques d'enchaînement des segments et structure des syllabes
  - ∩ agencement des morphèmes
  - ∩ rapport entre signifiant et signifié.

**3.1.1.1.1.2.2 Les relations entre mots**

- ▷ Forme :
  - ∩ Relations de superposition dans le lexique (phonétiques (traits partagés), phonémiques (phonèmes en commun), morphologiques (morphèmes en commun), lexicales (enchâssement, enchâssabilité))
  - ∩ Relations syntaxiques :

- relations catégorielles d'appartenance (catégorie grammaticale, genre, etc.)
  - relations catégorielles d'interdépendance (règles d'association des mots entre eux, etc.)
- ▷ Contenu : liens sémantiques (nombre de traits sémantiques partagés)

### **3.1.1.1.2.3 Les relations entre les mots et le monde**

- ▷ Sémantique
- ▷ Pragmatique

### **3.1.1.1.3 Dimensions psychologiques des mots parlés (mot dans sujet)**

Les dimensions physiques et linguistiques d'un mot parlé sont intégrées dans le mental d'un individu selon un format qui fait toujours l'objet de recherches. Cette intégration conduit à devoir décrire un ensemble de facteurs ou dimensions psychologiques des mots. Les dimensions psychologiques d'un mot comportent deux aspects indissociables : structural et fonctionnel<sup>74</sup>.

#### **3.1.1.1.3.1 Dimensions structurales**

Il s'agit de composantes représentées comme des entités ayant une existence permanente dans le mental d'un individu (mémoire à long terme), et constituant le lexique mental. Ce dernier peut être défini par :

- ▷ Son contenu (phonologie, syntaxe, morphologie, sémantique)
- ▷ Son volume (nombre de mots connus)
- ▷ Son organisation :
  - ∟ stratification temporelle (âge d'acquisition des mots)
  - ∟ relations entre éléments au sein du lexique mental (liens formels et conceptuels entre mots)

#### **3.1.1.1.3.2 Dimensions fonctionnelles**

Il s'agit des états spécifiques du système et des propriétés émergentes associées au fonctionnement de la structure (mémoire de travail), et qui dépendent :

- ▷ De l'excitabilité spécifique des représentations (familiarité conceptuelle, fréquence d'usage, charge affective)
- ▷ De l'excitabilité contextuelle des représentations (point d'unicité d'un mot, pragmatique et utilisation dans le rapport au monde, charge affective)

---

<sup>74</sup> La fonction est indissociable de la structure dans une perspective systémique, même si des structures différentes peuvent assurer une fonction identique.



### 3.1.1.1.2 Multidimensionnalité des sujets (sujet)

Comme dans le cas des mots, les sujets présentent une complexité constitutive qui est impliquée dans la production des données recueillies lors d'une expérience, et qui peuvent être répertoriées selon deux catégories de facteurs :

- ▷ Dimensions psychologiques du sujet
  - ∩ Expérience et aptitudes (connaissances extralinguistiques, compétences linguistiques, compétences particulières ayant une incidence sur la réalisation de la tâche expérimentale (exemple : manipulation de consoles de jeux), etc.)
  - ∩ Caractère (émotivité, vivacité intellectuelle, qualités psychologiques susceptibles d'avoir une influence sur la stratégie de réponse des sujets (exemple : audace, peur, désir de plaire en faisant bien, etc.), etc.)
- ▷ Dimensions biologiques du sujet (sexe, âge, niveau vigile, latéralisation manuelle, dextérité motrice, etc.)

### 3.1.1.1.3 Autres dimensions de la situation expérimentale influençant la production des réponses

Le contexte expérimental peut agir à différents niveaux et être source de biais expérimentaux. Par exemple :

- ▷ organisation du plan d'expérience (des groupes de mots / sujets)
- ▷ ordre de présentation des items (apprentissage, fatigue, amorçage)
- ▷ effet du contexte (exemple : distracteurs pouvant contribuer à un phénomène d'amorçage formel / sémantique)
- ▷ type de tâche demandée (exemple : certains sujets sont incapables d'effectuer correctement une détection de cibles sonores. Ils ne perçoivent pas la différence de sonorité entre la cible et le son distracteur. Ou encore : mesure-t-elle bien le niveau de traitement souhaité ?)
- ▷ etc.

### 3.1.1.1.4 Récapitulation concernant les facteurs primaires :

La réponse fournie par un individu lors d'un test chronométrique sur la reconnaissance des mots parlés renferme à la fois :

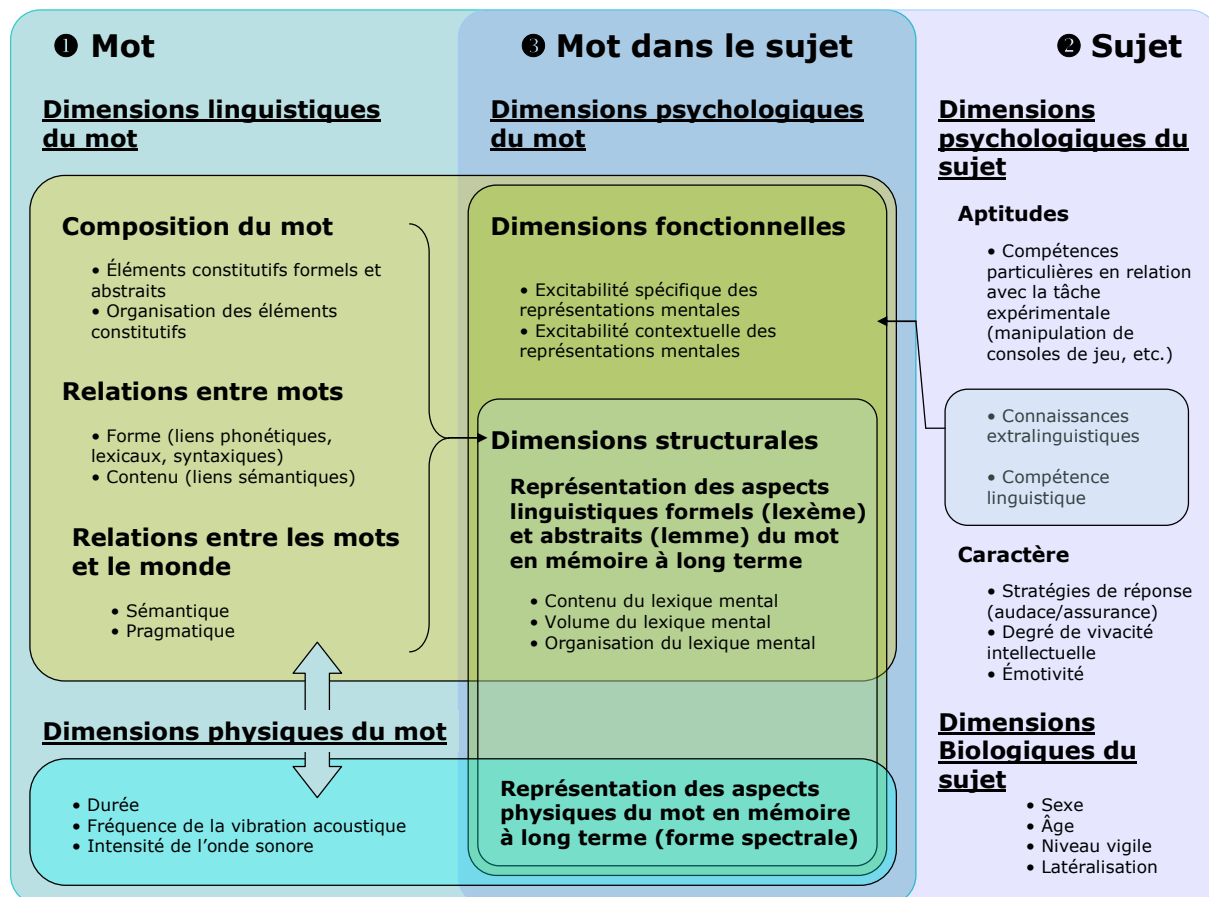
- ▷ Un effet lié aux caractéristiques physiques et linguistiques du mot présenté (❶)
- ▷ Un effet lié aux caractéristiques du sujet qui répond (❷)
- ▷ Un effet lié à l'intégration du mot dans le sujet (aspects représentationnels structuraux et fonctionnels - ❸)
- ▷ Un effet lié aux caractéristiques de la situation expérimentale (❹)

La mesure dépend donc de plusieurs dimensions qu'il est nécessaire de décorrélérer pour pouvoir conclure théoriquement sur ce qui est étudié : un effet des aspects

physiques, linguistiques, ou psychologiques du mot, sur la manière dont il est représenté et accédé en mémoire à long terme pour générer un percept lexical.

Le schéma ci-dessous récapitule la situation en présentant les facteurs primaires qui influencent la production des données dans les expériences sur la reconnaissance des mots parlés, et la manière dont ils sont organisés en niveaux et entrent en relation.

Figure 19 : Facteurs primaires influençant la production des données



La recherche présentée dans ce manuscrit s'est focalisée sur l'impact des facteurs primaires relatifs à la dimension physique des mots (prosodie) et à la relation formelle entre mots (superpositions lexicales), sur le comportement d'auditeurs humains. Toutes les autres dimensions relatives aux mots (1) et aux sujets (2) représentent des sources de variation indésirable qui devaient être contrôlées.

Le point central de ce travail de doctorat est donc de parvenir à comprendre comment les mots intégrés dans le sujet (3) sont accédés, en manipulant (1) et (2).

### **3.1.1.2 Facteurs secondaires influençant le recueil des données et la production de résultats interprétables**

Il s'agit d'aspects de l'expérimentation susceptibles de biaiser l'information produite par les sujets. Ces biais peuvent survenir au moment de la mesure, ou ultérieurement, lors de l'analyse des données.

#### **3.1.1.2.1 Facteurs secondaires influençant le recueil des données**

Il s'agit de facteurs susceptibles de fausser la mesure de la réponse produite par les sujets. Deux exemples ont pu être répertoriés à travers les expériences présentées dans ce rapport :

- ▷ Choix de l'origine de la mesure (problèmes liés à la mesure des temps de réponse pour le traitement de stimuli temporels).
- ▷ Précision des outils de mesure.

#### **3.1.1.2.2 Facteurs secondaires influençant la production de résultats interprétables**

Il s'agit de facteurs susceptibles d'induire des modifications des données mesurées expérimentalement, lors des phases d'analyse statistique (facteurs liés à la préparation des données pour l'analyse, et à leur analyse statistique)<sup>75</sup>. Il s'agit donc d'éléments pouvant induire des erreurs dans les résultats en dépit d'un contrôle très strict de l'expérience et d'une mesure très précise.

### **3.1.2 Étapes de l'étude où des procédés peuvent être appliqués pour réduire ou contrôler la variance indésirable**

Certains des problèmes exposés dans la section 3.1.1 peuvent être résolus, partiellement ou totalement, à l'étape du travail où ils surviennent. Les problèmes partiellement résolus font l'objet d'un complément de contrôle statistique a posteriori.

Pour d'autres problèmes, une solution directe n'est pas possible. L'application a posteriori de procédés statistiques permettant de corriger les biais qui n'ont pu être évités est alors utilisée.

---

<sup>75</sup> L'interprétation n'est pas considérée comme faisant partie de l'aspect purement technique de l'étude expérimentale (elle survient à un niveau ultérieur, à partir des résultats produits par les analyses statistiques). Impliquant à la fois les connaissances/compétences du chercheur et son imaginaire, l'interprétation peut aussi être source de biais.

Enfin, certains problèmes sont sans solution technique, à ce jour et en l'état actuel de mes connaissances personnelles.

Cette section propose une énumération rapide des problèmes en fonction du moment auquel une solution peut leur être appliquée durant l'étude. Ils sont organisés selon la structure en 3 points qui vient d'être décrite (problèmes directement résolus, problèmes résolus par correction statistique des données, problèmes insolubles).

### **3.1.2.1 Problèmes directement résolus**

Ils sont relatifs à des éléments manipulables et contrôlables, et sont traités lors de la création de la situation expérimentale, lors de la mesure, et lors des analyses.

#### **3.1.2.1.1 Contrôle lors de l'élaboration de l'expérience**

▷ Contrôle des items

Il est effectué lors de la construction du matériel expérimental, par la manipulation de deux groupes de facteurs :

- ∨ Facteurs linguistiques et psychologiques, contrôlés à travers le choix des mots via des bases de données lexicales.
- ∨ Facteurs phonétiques et acoustiques, contrôlés lors de l'enregistrement des items (organisation des listes à lire pour répartir les biais acoustiques de façon homogène sur toutes les conditions, production contrôlée par le locuteur (exemple : prosodie), sélection des items parmi plusieurs versions enregistrées).

Le contrôle des items est complété a posteriori, lorsque nécessaire, par des procédés statistiques correctifs (se reporter à la Section 3.1.2.2).

▷ Contrôle des sujets :

Il s'effectue à travers la sélection de la population testée, en fonction de caractéristiques :

- ∨ biologiques : sujets adultes normaux, sans atteinte des systèmes auditif, nerveux central, et moteur
- ∨ linguistiques : sujets francophones de langue maternelle française et non bilingues ; contrôle des connaissances linguistiques pour les autres langues, et du niveau de compétence dans ces autres langues (questionnaire)

De la même façon que pour le contrôle des items, un complément de contrôle est effectué a posteriori lors des analyses statistiques.

▷ Contrôle de la passation :

Il est effectué à travers l'organisation du plan d'expérience, l'ordre aléatoire de présentation des items, le contexte lexical neutre (pas d'amorçage, etc.), et l'isolement du sujet lors de la passation (absence de stimulations autres que celle délivrées pour le but de l'expérience : isolement physique par un caisson insonorisé et un casque de studio). Tous les aspects de la passation qui ne sont pas totalement contrôlés à cette étape font l'objet d'un complément de contrôle a posteriori, lors des analyses statistiques.

### **3.1.2.1.2 Contrôle lors de la mesure**

Il s'agit de s'assurer de la précision des outils de mesure (précision du système testée par simulation d'une expérience, détails fournis dans le chapitre « Méthodologie expérimentale », à la Section 3.4.4).

### **3.1.2.1.3 Contrôle lors de la préparation des données pour l'analyse**

Il s'agit de s'assurer que la procédure de tri des données contrôle effectivement la normalité des réponses, des sujets et des items, et qu'elle ne filtre que les éléments qui s'écartent trop du centre de la distribution normale considérée. La procédure adoptée est détaillée dans la Section 3.3.3 page 117.

### **3.1.2.1.4 Contrôle lors de l'analyse statistique des données**

Ce contrôle est essentiellement réalisé par le choix d'une technique autorisant l'analyse de données hiérarchiquement organisées, permettant de travailler avec des plans incomplets et des valeurs manquantes, et permettant d'intégrer des covariables. La présentation de cette technique et des justifications de son choix est détaillée dans les Sections 3.3.3 page 117 et 3.3.4 page 123).

### **3.1.2.2 Problèmes résolus a posteriori par correction statistique**

Les éléments partiellement ou non manipulables / contrôlables à l'étape du travail où ils devraient être maîtrisés ont biaisé les réponses recueillies. Celles-ci font donc l'objet d'une procédure de correction a posteriori lors des analyses statistiques. Le principe consiste à intégrer les variables non maîtrisées ainsi que les interactions dans lesquelles elles sont impliquées au modèle statistique, pour en distinguer l'effet de celui des facteurs expérimentaux. La méthode appliquée permet ainsi de partitionner la variance présente dans les données de manière précise, de connaître la quantité de variance qui est due aux facteurs étudiés, et de quantifier leur effet dans la dimension de la variable dépendante. Plus de détails sont disponibles à la Section 3.3.4 page 123.

### 3.1.2.2.1 Problèmes partiellement résolus ou non résolus lors de l'élaboration de l'expérience

#### 3.1.2.2.1.1 Limite des possibilités de contrôle des items

##### 3.1.2.2.1.1.1 Contraintes lexicales de niveau linguistique et psychologique

En fonction des critères de sélection liés aux facteurs manipulés, la quantité de mots présentant les caractéristiques voulues peut être excessivement réduite<sup>76</sup>.

Dans ces conditions, il est souvent impossible de contrôler des facteurs parasites, ce qui aboutirait à une restriction supplémentaire de la quantité d'items, et interdirait toute expérimentation.

Une solution consiste à tester avec un matériel parfois moins bien contrôlé, puis à inclure dans l'équation du modèle statistique, lors des analyses, les facteurs parasites linguistiques et psychologiques pour lesquels on dispose de renseignements, ainsi que les interactions théoriquement pertinentes qu'ils font entre eux et avec les autres facteurs de la situation.

##### 3.1.2.2.1.1.2 Contraintes lexicales de niveau phonétique et acoustique

Il n'est pas possible qu'un locuteur contrôle précisément et à volonté la durée, la fréquence et l'intensité de ses productions, lorsqu'il traduit l'information linguistique en format acoustique<sup>77</sup>. Ceci entraîne deux problèmes majeurs :

- ▷ La perfection du contrôle du matériel pour des facteurs lexicaux comme le nombre de syllabes ou de phonèmes au niveau du « mot théorique » peut être anéantie au niveau du « mot physique » (durée moyenne des segments variable, variations de débit du locuteur, etc.)
- ▷ la position du point de reconnaissance peut varier en fonction de la durée des phonèmes qui précèdent ce point, du débit du locuteur, de la quantité de signal nécessaire pour identifier le phonème correspondant au PU, etc.

---

<sup>76</sup> Il n'est pas rare de disposer d'un corpus d'une dizaine d'items par condition, ce qui pose ensuite le problème de la généralisation des effets obtenus, puisque ceux-ci sont relatifs au traitement d'une catégorie très restreinte des mots de la langue (néanmoins, l'étude d'un cas particulier peut aussi renseigner sur le fonctionnement du système).

<sup>77</sup> L'utilisation de la synthèse vocale à des fins de contrôle est extrêmement complexe et soulève d'autres problèmes liés au manque de naturel et d'audibilité de l'information acoustique. De plus, imposer artificiellement des contraintes aux caractéristiques physiques des segments (comme leur durée) peut s'avérer avoir des répercussions inattendues sur le processus de perception, si bien qu'en voulant contrôler la situation, on introduit des biais. De même avec l'utilisation de procédures de normalisation acoustique appliquées a posteriori (exemple : normalisation de la durée ou de la f0).

Il faut cependant contrôler les caractéristiques physiques des items, puisqu'elles ont des conséquences importantes dans le traitement perceptif. Un moyen de le faire est d'intégrer les facteurs susceptibles de produire une variation dans la réponse des sujets (durée des items, position du début du PU, position du PR,  $f_0$ , intensité) dans les modèles d'analyse statistique en tant que covariables, pour en distinguer les effets des effets des variables indépendantes.

#### **3.1.2.2.1.2 Limite des possibilités de contrôle des sujets**

La latéralité et le sexe des sujets sont un exemple de contrainte biologique qui peut théoriquement avoir une incidence sur les réponses fournies<sup>78</sup>, et qui peut être difficile à contrôler lors de la passation de l'expérience (difficulté à obtenir la même quantité d'individus dans chaque condition expérimentale).

Intégrer ces deux facteurs comme covariables dans les modèles d'analyse permet de vérifier qu'ils sont sans effet sur les résultats, ou le cas échéant, de prendre en considération leur incidence sur la variance des réponses.

#### **3.1.2.2.1.3 Limite des possibilités de contrôle de la passation**

L'ordre de présentation des items est partiellement contrôlé lors de la passation, puisqu'il est rendu aléatoire de manière à répartir les effets de fatigue et d'apprentissage de façon homogène à travers les conditions de l'expérience. Cependant, ces effets existent dans les données, et peuvent varier en fonction des conditions expérimentales. Ils doivent donc être contrôlés ou corrigés a posteriori, lors de l'analyse statistique des données (ordre de présentation utilisé comme covariable dans le modèle d'analyse).

### **3.1.2.2.2 Problèmes partiellement résolus ou non résolus lors de la mesure**

L'effet de l'origine de la mesure n'est absolument pas contrôlable, dans le cas de la mesure de temps de réponse par tâche de décision lexicale, lors d'expériences utilisant des mots parlés phonétiquement non appariés. La seule manière de le résoudre est d'utiliser une méthode d'analyse statistique permettant de décorrélérer les facteurs déterminant les temps de réponse. Les temps de réaction sont alors mesurés depuis le

---

78 Sexe : la latéralisation inter hémisphérique n'est pas la même en fonction du sexe pour ce qui concerne les fonctions du langage (asymétrie plus marquée chez l'homme). Cette différence peut avoir une influence comportementale mesurable.

Latéralité manuelle : peut avoir une influence comportementale mesurable selon que l'ordre moteur est donné dans le même hémisphère que celui traitant le langage ou dans l'hémisphère controlatéral.

début du mot, et le modèle d'analyse statistique intègre comme covariables, la durée du mot, la position du point de reconnaissance, et la position du début du point d'unicité phonologique, ainsi que les interactions deux à deux et entre les trois facteurs (Cf. chapitre 2 de cette partie).

### 3.1.2.3 Problèmes insolubles

Des facteurs non manipulables et non contrôlables existent dans les situations expérimentales impliquant des êtres humains et des mots. Il s'agit aussi bien de facteurs psychologiques (exemple : émotivité, résonance affective d'un mot spécifique pour un individu, etc.), que de facteurs inconnus (qui peuvent être relatifs aux sujets, aux items, à la situation expérimentale, etc.). Dans les modèles d'analyse, leur influence apparaît sous forme de variance résiduelle (non modélisée), provenant des sujets, des items, ou d'autre chose (résidus de l'équation générale). Si ces facteurs n'influencent pas remarquablement la distribution des données, ils sont ignorés. Leur présence peut donc toujours causer un biais dans les résultats (exemple d'un facteur partiellement confondu avec une des VI, qui rend l'effet de celle-ci significatif alors qu'il n'y a pas d'effet). Si au contraire ces facteurs influencent remarquablement la distribution, un sujet ou un item peut apparaître comme déviant, et être éliminé des analyses lors du tri des données.

### 3.1.2.4 Conclusions concernant la résolution des problèmes de contrôle de l'étude

Cinq difficultés majeures à résoudre pour la faisabilité de l'étude ont été identifiées :

- ▷ 4 parmi l'ensemble des éléments présentés comme nécessaires à maîtriser dans ce chapitre introductif :
  - ∩ décorrélérer les facteurs présents dans la situation expérimentale
  - ∩ mesurer la durée du processus d'accès au lexique, sans mesurer aussi les caractéristiques physiques du stimulus
  - ∩ préparer les données pour l'analyse sans modifier l'information qu'elles renferment
  - ∩ disposer de la technique statistique adéquate pour analyser les données
- ▷ une 5<sup>ème</sup> difficulté, non mentionnée jusqu'ici, est implicite au contrôle de variables psychologiques : obtenir des informations non disponibles, relatives à deux facteurs psychologiques à contrôler (âge d'acquisition des mots utilisés dans les corpus, familiarité conceptuelle des sujets avec ces mots).

Les solutions aux 4 premiers problèmes seront développées en détail dans les 3 chapitres suivants de la partie méthodologique :



- ▷ Chapitre 2 – L'origine de la mesure des temps de réponse lors de la reconnaissance d'un mot parlé isolé
- ▷ Chapitre 3 – Le traitement statistique des temps de réaction pour reconnaître un mot parlé isolé
- ▷ Chapitre 4 – Méthodologie expérimentale

La solution adoptée face au 5<sup>ème</sup> problème est présentée dans l'Annexe 1 page 423.

Le Tableau 10 fournit une liste récapitulative des facteurs manipulés dans les expériences de cette étude et la manière dont ils ont été contrôlés.

Tableau 10 : Facteurs manipulés dans les expériences

	Variable	Type de variable	Type de contrôle adopté
Facteurs primaires	Catégorie grammaticale	VP	Sélection des mots
	Liens lexicaux d'enchaînement	VI	Sélection des mots
	Prosodie	VI	Production
	Liens sémantiques inter-mots	VP	Sélection des mots
	Liens sémantiques enchaînements/porteurs	VP	Sélection des mots
	Liens phonémiques	VP	Sélection des mots + analyses statistiques
	Morphologie	VP	Sélection des mots
	Durée	VP	Analyses statistiques
	Intensité	VP	Analyses statistiques
	f0	VP	Analyses statistiques
	Âge d'acquisition	VP	Analyses statistiques
	Familiarité conceptuelle	VP	Analyses statistiques
	Position phonémique du point d'unicité	VP	Sélection des mots + analyses statistiques
	Position acoustique du début du PU (en ms.)	VP	Analyses statistiques
	Position acoustique du PR (en ms.)	VP	Analyses statistiques
	Fréquence d'occurrence	VP	Sélection des mots + analyses statistiques
	Sexe	VP	Analyses statistiques
	Latéralité manuelle	VP	Analyses statistiques
	Compétence linguistique	VP	Sélection des sujets
Facteurs secondaires	Origine de la mesure	VP	Analyses statistiques
	Précision des outils de mesure	VP	Contrôle du matériel par une simulation
	Tri des données	VP	Méthode de tri
	Technique d'analyse statistique et aspects associés	VP	Réflexion avec des statisticiens, formation technique

## **3.2 Chapitre 2 - L'origine de la mesure des temps de réponse lors de la reconnaissance d'un mot parlé isolé**

Ce chapitre décrit un problème méthodologique relatif à la mesure des temps de réaction dans les tâches portant sur la reconnaissance de mots parlés isolés, comme la tâche de décision lexicale. Le problème, lié au choix d'une origine pour la mesure, peut être généralisé à tout type de stimulus parvenant progressivement à un système récepteur dans le temps, dont le traitement s'effectue au fur et à mesure de l'arrivée de l'information, et pour lequel on souhaite mesurer la durée des traitements psychologiques qui lui sont appliqués.

### **3.2.1 Introduction :**

L'information décrivant physiquement un mot parlé, véhiculée par une onde acoustique, parvient au système perceptif graduellement dans le temps. À cause de cela, le traitement auditif d'un stimulus lexical se fait au fur et à mesure que le signal devient disponible. Il arrive un moment où une quantité suffisante d'information a été analysée : il est alors possible de reconnaître le mot présenté.

Ce moment, qui peut survenir avant la fin du mot, est variable :

- ▷ d'un mot à l'autre, en fonction de la structure phonétique de l'item et des caractéristiques acoustiques de la forme produite.
- ▷ d'un individu à l'autre pour un même mot, en fonction du degré de familiarité de l'auditeur avec l'item et du lexique de l'individu.

Les effets liés aux traitements psychologiques des stimuli lexicaux étant extrêmement fins (modulations dans la vitesse des participants à fournir leur réponse, de l'ordre de quelques dizaines de millisecondes au mieux), il est important de s'assurer que les variations mesurées entre conditions expérimentales ne sont pas dues aux variations des caractéristiques physiques des items (durée, position temporelle du moment de la reconnaissance ou de position du point d'unicité phonologique, etc.).

Considérer avec précaution le choix de l'origine de la mesure des temps de réponse revient à considérer ces paramètres plus attentivement, car du choix d'une origine dépendra aussi le contenu de la mesure. D'où ces deux questions :

- ▷ Que souhaite-t-on mesurer ?

- ▷ Qu'est-ce qui a réellement été mesuré ?

Classiquement, ce problème n'est pas discuté dans les travaux employant la tâche de décision lexicale avec des mots parlés, par inconscience ou par consensus. La mesure est en général faite à partir du début du mot, ou plus rarement de sa fin, ces deux points étant les plus naturels à considérer et les plus aisés à utiliser. Pourtant, 4 autres origines pourraient être employées : le point de reconnaissance (PR)<sup>79</sup> et 3 origines relatives au point d'unicité phonologique (PUP)<sup>80</sup>. La première partie du développement s'attache donc à une réflexion sur le choix d'une origine aux temps de réponse lors de la reconnaissance de mots parlés, afin de permettre une meilleure compréhension du problème et de l'étendue de ses conséquences sur la théorie.

Lors de l'étude expérimentale de la reconnaissance des mots parlés à l'aide du paradigme de la chronométrie mentale, les propriétés physiques d'un item devraient être décorrélées de ses propriétés psychologiques afin d'autoriser des conclusions ne portant que sur les traitements psychologiques du langage. En effet, à cause des caractéristiques physiques du stimulus, le traitement sensoriel, qui s'effectue au fur et à mesure que l'information acoustique parvient au système, est aussi reflété dans le traitement perceptif. Toutefois, à cause des caractéristiques psychologiques du stimulus, le traitement perceptif ne peut se réduire à l'aspect sériel du décours de l'information, des phénomènes intégratifs ayant lieu lors de l'accès au lexique et lors du traitement de la phrase. Techniquement, cette décorrélation est difficile à réaliser, et ne peut en tous cas pas être obtenue par le seul choix d'une origine, comme cela est montré dans la première partie du développement qui suit. La seconde partie du développement s'attache donc à proposer une solution basée une méthode d'analyse statistique.

### **3.2.2 Origine de la mesure et conséquences sur la théorie**

#### **3.2.2.1 Présentation des diverses origines possibles**

Six origines sont utilisables pour la mesure des temps de reconnaissance d'un mot parlé :

- ▷ Le début du mot (DM)
- ▷ La fin du mot (FM)
- ▷ Le point de reconnaissance (PR)

---

<sup>79</sup> Quantité minimale de signal acoustique nécessaire à la reconnaissance effective d'un mot.

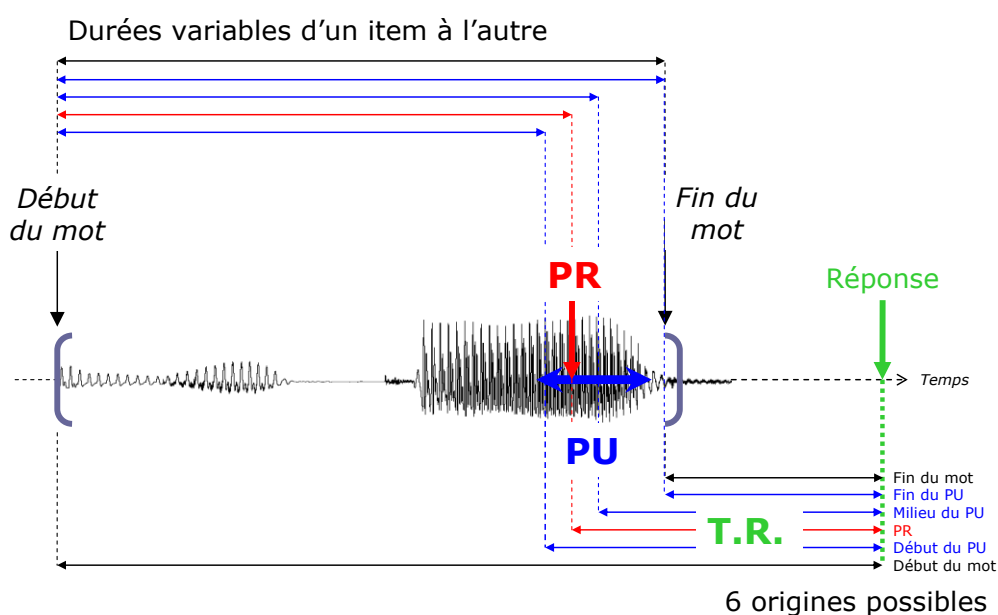
<sup>80</sup> Phonème rendant théoriquement unique un mot d'une langue donné.

Le point d'unicité phonologique étant un phonème, soit une durée (qui peut atteindre plus de 250 ms.), plusieurs origines relatives au PUP sont possibles :

- ▷ Le début du point d'unicité (DPU), frontière gauche du phonème<sup>81</sup>
- ▷ Le milieu du point d'unicité (MPU)
- ▷ La fin du point d'unicité (FPU), frontière droite du phonème

La première origine, le début du mot, présente l'intérêt d'être fixe et valable pour tous les items. Les 5 suivantes varient d'un mot à l'autre. Les 3 origines relatives au PUP (DPU, MPU et FPU) peuvent, en plus, varier d'un individu à l'autre, la position réelle du PU d'un mot, qui dépend du lexique de chacun, ne correspondant pas forcément à sa position théorique, liée au lexique de la langue. Le schéma ci-dessous présente la localisation de ces 6 origines dans un mot à PUP pré-final (exemple fictif).

Figure 20 : Localisation des différentes origines possibles dans un mot



Dans un mot à PUP post-final, les 3 origines relatives au PU sont confondues avec la fin du mot. L'utilité de chacune de ces origines se justifie par des arguments théoriques (relatifs aux buts de la recherche) et techniques (relatifs à la facilité d'utilisation) :

- ▷ Début du mot : moment à partir duquel le traitement commence. Facile à localiser et à utiliser, universel, classique.

<sup>81</sup> La frontière gauche du PU ne correspond pas forcément au PR, puisqu'il est nécessaire d'entendre une quantité suffisante et variable de signal acoustique correspondant au PU pour identifier ce phonème et reconnaître le mot.

- ▷ Fin du mot : moment auquel l'auditeur dispose de toute l'information décrivant le stimulus pour être en mesure de fournir une réponse relative à la stimulation. Facile à localiser, facile d'emploi, mais plutôt inutilisé.
- ▷ Point de reconnaissance : moment à partir duquel l'auditeur dispose de l'information acoustique minimale décrivant le mot stimulus pour être certain qu'il n'a pas fourni de réponse pour un autre mot. Difficile à localiser : cette position est localisée perceptivement par une analyse auditive du stimulus et visuelle de l'oscillogramme correspondant. La valeur moyenne des relevés fournis par plusieurs personnes est utilisée. Peu ordinaire, et inutilisée à cause de la lourdeur du procédé de localisation.
- ▷ Point d'unicité : position théorique définie par le dictionnaire, correspondant au segment à partir duquel un mot n'est plus confondu avec un autre mot du lexique de la langue. Classique mais peu utilisé car long à localiser dans un corpus de mots à PUP pré-final (avant la fin du mot), et difficile à localiser avec certains phonèmes (coarticulation). Pose de plus un problème d'hétérogénéité entre mots à PUP pré-final et mots à PUP post-final (après la fin du mot - voir ci-dessous).

PR et début du PUP correspondent la plupart du temps à 2 points différents dans le signal acoustique. Deux situations sont rencontrées :

- ▷ Dans le cas des mots à PUP pré-final, le début du PUP se situe avant le PR. L'auditeur doit en effet entendre suffisamment de signal pour identifier le phonème correspondant au PU et ainsi reconnaître le mot. Il n'a pas cette possibilité au moment du début du PU, la quantité de signal manquante dépendant de la coarticulation du PU avec le phonème précédent, et du phonème correspondant au PU.
- ▷ Dans le cas des mots à PUP post-final, le début du PUP est après le PR. En effet, le début du PUP correspond alors à l'arrêt du signal (pour un mot isolé), et donc, est confondu avec la fin du mot. Par contre, le PR se trouve avant la fin du mot, puisqu'il n'est généralement pas nécessaire d'entendre la totalité du dernier phonème (dont la durée est souvent allongée par rapport à la durée moyenne) pour qu'il soit identifié. Il n'est pas non plus nécessaire d'attendre l'arrêt du signal pour prévoir que le mot touche à sa fin (indices prosodiques et coarticulatoires), ce qui permet d'éliminer les alternatives lexicales avant le PU (exemple : le mot /ʃømiz/ possède un PUP post-final puisque des mots comme « *chemiserie* » et « *chemisier* » existent. Grâce à l'information prosodique, qui permet de distinguer un item bisyllabique d'un item trisyllabique, /ʃømizɛi/ et /ʃømizje/ seront éliminés de la cohorte, aboutissant à une anticipation du moment de la reconnaissance).

### 3.2.2.2 Importance de l'origine

Un examen approfondi des conséquences liées au choix de l'une des 6 origines présentées révèle qu'aucune d'elles ne garantit une mesure exempte de biais et d'incertitudes. Deux types de biais peuvent survenir :

- ▷ Les biais sur le sens (direction) des effets
- ▷ Les biais sur l'existence des effets

### 3.2.2.2.1 Biais sur le sens des effets

En fonction de l'origine choisie, le sens (direction) des effets peut s'inverser. Les trois graphiques ci-dessous illustrent ce phénomène à partir des données de l'Expérience 4 (Section 4.2.2 page 234), qui porte sur la reconnaissance de mots parlés monosyllabiques, et dans laquelle sont manipulés 2 facteurs A et B à 2 modalités chacun. Ces graphiques montrent comment une même mesure peut trouver une interprétation différente selon que l'origine est choisie au début ou à la fin du mot, ou au début du PUP. Les TR, enregistrés à partir du début du fichier son, ont été transformés par simple soustraction de la valeur correspondant à la position de l'origine dans le fichier son, pour obtenir un TR mesuré depuis ce point. Les graphiques figurent simplement les temps de réponse moyens par condition. Ces moyennes varient entre 79 et 629 ms., l'échelle des graphiques n'est volontairement pas homogène, de manière à laisser l'accent sur les écarts entre conditions, pour mieux représenter les variations du sens des effets en fonction de l'origine choisie.

Figure 21 : Exemple de temps de réponse mesurés depuis le début du mot

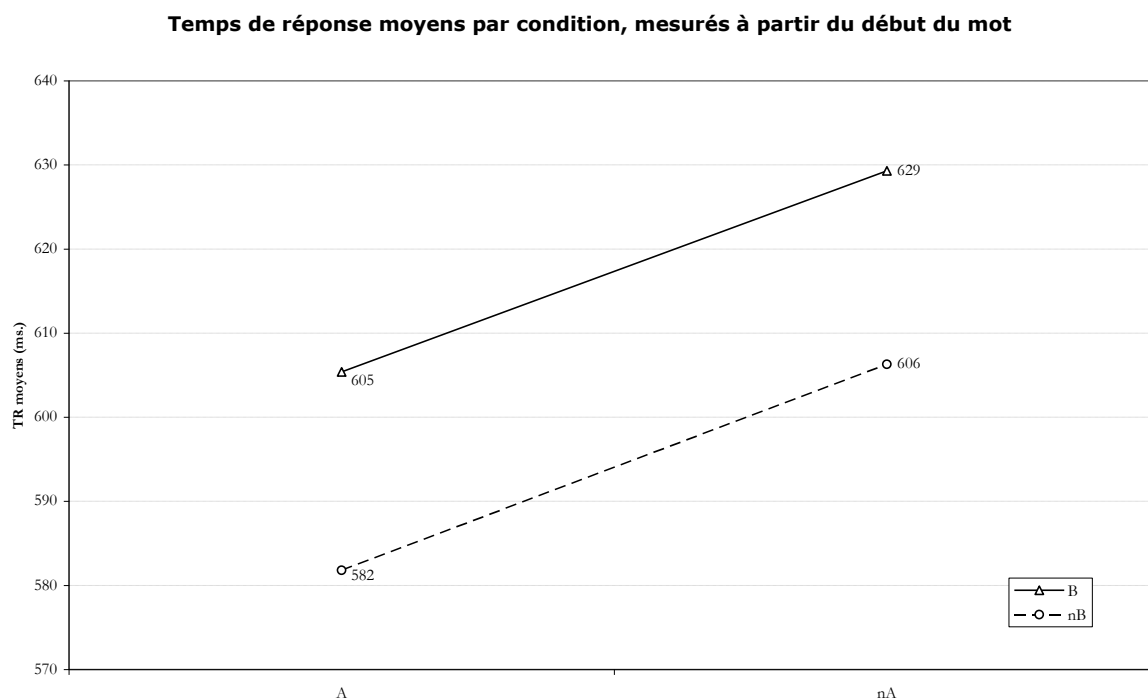


Figure 22 : Exemple de temps de réponse mesurés depuis la fin du mot

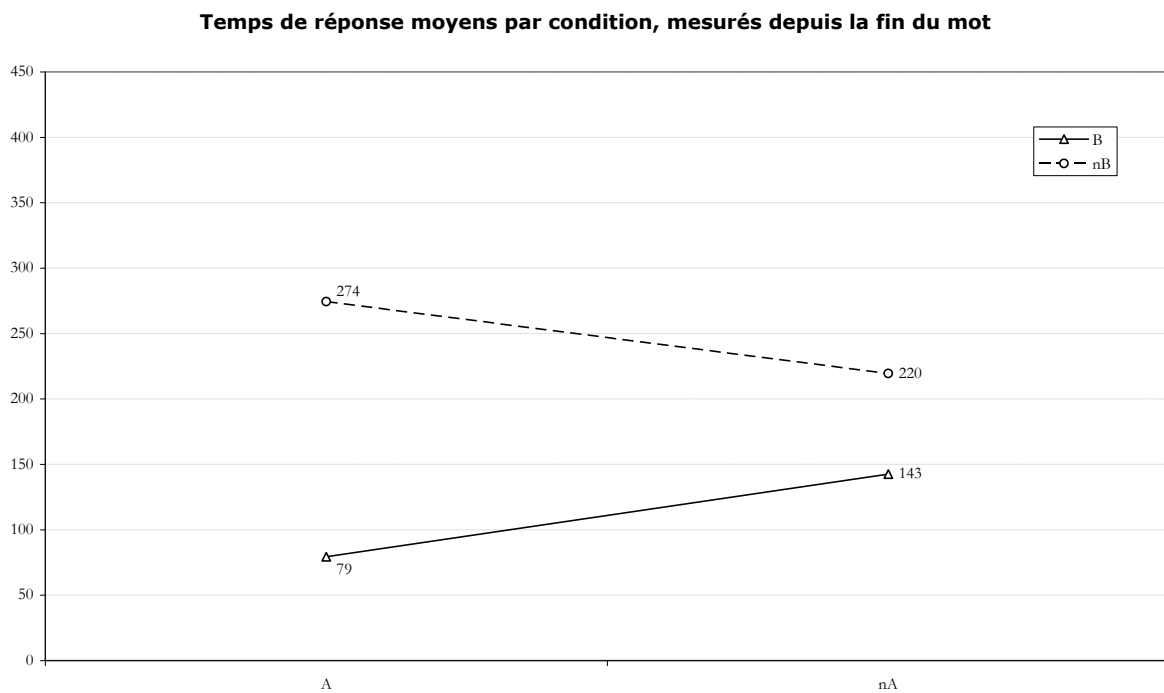
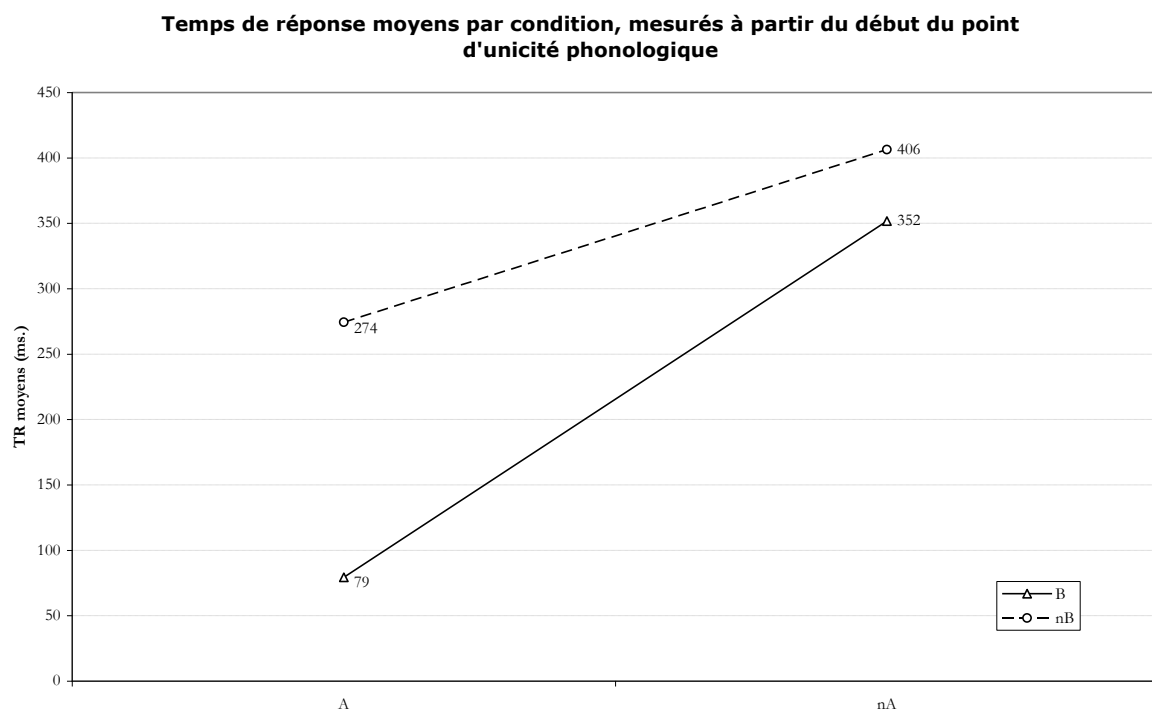


Figure 23 : Exemple de temps de réponse mesurés depuis le début du PUP



Le Tableau 11 récapitule les effets obtenus en fonction de l'origine adoptée à la mesure des temps de réponse.

Tableau 11 : Récapitulation du sens des effets de chaque facteur, en fonction de l'origine

TR mesuré depuis	A	B	A * B
Début du mot	nA < A	B < nB	-
Fin du mot	A < nA	-	Oui
Début du PU	A < nA	B < nB	Oui

### 3.2.2.2 Biais sur l'existence des effets

Des effets peuvent être induits ou annulés à cause de la variabilité inter items des durées séparant le début du mot des 5 autres origines (FM, PR, DPU, MPU, FPU), au gré des combinaisons entre variance des temps de réaction et variance de la position de l'origine par rapport au moment réel de la reconnaissance, d'une condition à l'autre.

Cette variabilité peut aussi induire quelques modifications dans l'élimination des valeurs non normales, lors du tri des données (une réponse donnée par un sujet à un item est conservée avec une origine, mais pas avec une autre).

Les conséquences sont multiples :

- ▷ des effets réels peuvent ne pas apparaître à cause du choix de l'origine
- ▷ des effets qui n'existent pas peuvent apparaître pour la même raison

#### Résultats des analyses de variance obtenus pour les trois exemples précédents

Ci-dessous, les résultats d'analyses de variance (par sujets et par items) sont présentés pour les trois exemples précédents (mesure des TR depuis le début du mot, depuis la fin du mot, depuis le début du PUP). Le Tableau 12 présente les effets principaux des deux facteurs expérimentaux. Le Tableau 13 récapitule l'existence d'un effet, pour chaque facteur et pour leur interaction, ainsi que dans sa partie droite, le sens de l'effet, en fonction du point origine.

Tableau 12 : Effets principaux des 2 facteurs expérimentaux

TR	ANOVA par sujets				ANOVA par items			
	A		B		A		B	
	F <sub>(1,18)</sub>	p<	F <sub>(1,18)</sub>	p<	F <sub>(1,56)</sub>	p<	F <sub>(1,56)</sub>	p<
Début mot	12.8	<b>0.003</b>	19.8	<b>.0004</b>	2.0	0.17	2.7	0.11
Fin du mot	0.0	0.84	505.3	<b>2<sup>E-14</sup></b>	0.0	0.95	55.8	<b>6<sup>E-10</sup></b>
Début du PU	1109.6	<b>2<sup>E-17</sup></b>	385.6	<b>2<sup>E-13</sup></b>	141.5	<b>6<sup>E-17</sup></b>	52.8	<b>2<sup>E-09</sup></b>

Tableau 13 : Synthèse des effets

TR	A	B	A * B		
Début du mot	ns.	ns.	ns.	-	-
Fin du mot	ns.	s.	ns.	-	B < nB
Début du PU	s.	s.	s.	A < nA	B < nB



### 3.2.2.2.3 Conclusion

Il est possible, avec l'utilisation du paradigme de la chronométrie mentale appliqué à l'étude de la reconnaissance des mots parlés, de valider ou d'invalider une hypothèse en fonction de l'origine des temps de réponse mesurés que choisit le chercheur.

### 3.2.2.3 Intérêt de chaque origine et biais encourus

Cette section présente :

- ▷ l'intérêt relatif à l'utilisation de chaque origine possible aux TR
- ▷ les biais à contrôler avec chaque origine.
- ▷ les biais non contrôlables et donc les erreurs introduites dans l'étude.

#### 3.2.2.3.1 Mesure depuis le début du mot

C'est une mesure facilement réalisable par tous (ne nécessitant pas d'outils ni de connaissances techniques avancés), universellement utilisée, ce qui permet une comparaison entre études.

Cependant, avec l'emploi de cette origine, le chercheur doit faire face à deux inconvénients au moins :

- ▷ Les temps de réponse mesurés depuis le début du mot seront plus longs pour des mots de durée importante que pour des mots de durée brève. Or, les facteurs modifiant la durée d'un mot sont multiples :
  - ∩ nombre de phonèmes
  - ∩ durée de chacun des phonèmes constitutifs
  - ∩ débit d'élocution.
 Ils ne peuvent être aisément contrôlés. Le contrôle de la durée des mots à partir des deux premiers de ces paramètres ne garantit pas le résultat final :
  - ∩ contrôle du nombre de phonèmes seul : durée variable des segments, en fonction des phonèmes.
  - ∩ contrôle de la durée des segments : difficulté technique et problème possible de naturel à la perception.
  - ∩ contrôle du nombre de phonèmes et de la durée des phonèmes à partir de leur moyenne théorique : débit variable à la production. Un traitement artificiel du fichier son avec un algorithme de compression doit encore avoir lieu, mais une hétérogénéité acoustique peut être créée dans le matériel (variations de débit entre items : certains sont compressés, d'autres non). La méthode n'est donc pas recommandable dans la mesure où on ignore ce qu'elle induit.

Le contrôle simultané de tous ces paramètres (nombre de phonèmes, durée théorique moyenne des phonèmes constitutifs, débit d'élocution) est techniquement irréaliste.

- ▷ Dans de nombreux cas, on reconnaît un mot isolé avant sa fin. Ce moment varie en fonction de la prosodie, de la coarticulation, et de la position du point d'unicité. Un contrôle très strict de la durée doit donc aussi s'accompagner du contrôle de ces autres facteurs, ce qui est techniquement infaisable.

### 3.2.2.3.2 Mesure depuis la fin du mot

Comme pour une mesure depuis le début du mot, il s'agit d'une origine techniquement facile à utiliser, mais qui est rarement employée. Elle offre l'avantage de résoudre le premier problème précédemment exposé (durée des mots), mais laisse le chercheur dépourvu face au second :

- ▷ un mot peut être reconnu avant sa fin s'il devient phonétiquement unique (point d'unicité)
- ▷ s'il est acoustiquement reconnaissable avant sa fin, à cause de la :
  - ∟ coarticulation renseignant sur le dernier phonème
  - ∟ prosodie indiquant le nombre de syllabes
  - ∟ quantité de signal nécessaire à l'identification du dernier phonème.

Le contrôle de la position du point d'unicité, après le dernier phonème du mot afin que la fin du mot et le début du PU soient confondus :

- ▷ n'annule pas l'influence des paramètres acoustiques précédents.
- ▷ certaines études peuvent nécessiter des items dont le point d'unicité se trouve avant la fin du mot.

#### Exemples :

Les mots français /bɛs/ (*baisse*) et /bak/ (*bac*) sont tous deux monosyllabiques, composés de 3 phonèmes, et présentent un PUP post-final (/bɛs/ peut commencer un mot comme /bɛstjɔl/ (*bestiole*) et /bak/ un mot comme /bakteʁi/ (*bactérie*)).

Le mot /bɛs/ sera reconnu aux alentours de 45% de sa durée dans le cas de l'enregistrement présenté ci-dessous (Figure 24, page 102 - marque rouge - 300 ms. après le début du mot et 355 ms. avant sa fin) : il n'est pas nécessaire d'entendre la totalité du dernier phonème ([s]), qui dure ici environ 395 ms., pour pouvoir identifier l'item sur la base de paramètres acoustiques et prosodiques. Le mot /bak/ sera reconnu 290 ms. après son début (Figure 25, page 102), environ 20 ms. après l'explosion du [k], soit après l'écoute de 77% du stimulus, et 110 ms. avant la fin du mot.

Figure 24 : Oscillogramme et spectrogramme du mot « *baisse* »

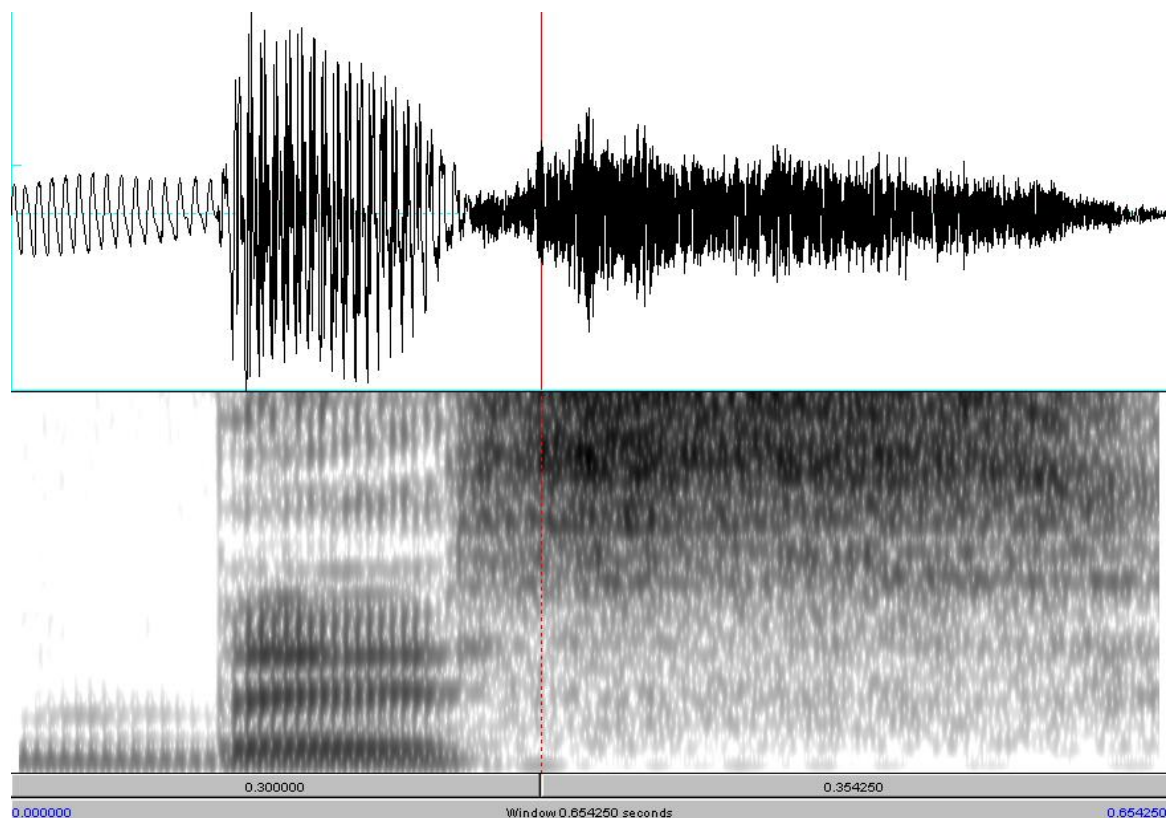
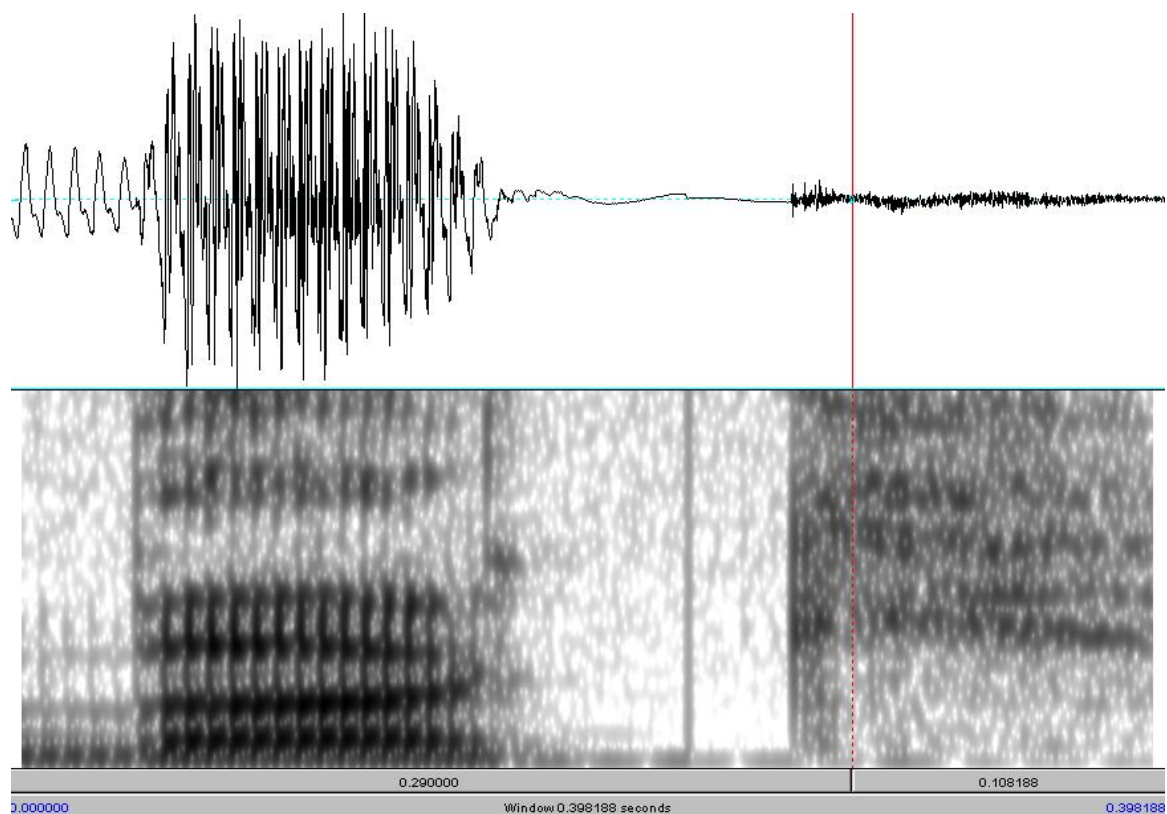


Figure 25 : Oscillogramme et spectrogramme du mot « *bac* »



Une mesure aveugle des temps de réponse depuis la fin des mots serait entachée d'une erreur absolue de 355 ms. dans le premier cas, et de 110 ms. dans le second. La comparaison des temps de réponse entre les deux mots, à supposer que la réponse des auditeurs soit initiée à ces points, révélerait une différence de 245 ms., qui n'est imputable qu'aux caractéristiques physiques des stimuli et au moment où le mot est reconnaissable.

Il faut ici remarquer qu'une mesure des temps de réponse depuis la fin des mots peut conduire à des valeurs négatives, ou trop faibles pour être reconnues comme des temps de réponse valides. Ces données peuvent être éliminées de la distribution par filtrage, lors du tri des réponses, malgré leur validité psychologique.

Comme le contrôle des paramètres acoustiques est techniquement irréaliste, le choix de l'origine doit préférablement porter sur un autre point : le point de reconnaissance peut permettre de contourner plus judicieusement ces difficultés.

### 3.2.2.3.3 Mesure depuis le point de reconnaissance

Le PR est peu ou pas utilisé. Il nécessite des connaissances phonétiques et des procédés de validation assez compliqués pour sa localisation (« gating » précis, validation de la position par consensus de plusieurs auditeurs). Son intérêt est de prendre en considération à la fois les aspects acoustiques et psychologiques de la reconnaissance d'un mot, puisque sa localisation se base sur la perception statique d'un auditeur humain<sup>82</sup>, et non sur un arbitraire théorique.

L'idée derrière le choix de cette origine est que :

- ▷ si la reconnaissance du mot a lieu au PR, le processus cognitif mesuré a une temporalité qui se confond avec celle du traitement perceptif.
- ▷ si la reconnaissance du mot s'effectue après le PR, l'activation lexicale nécessite probablement des calculs post-perceptifs supplémentaires.

Cependant, en cas de reconnaissance après le PR, il n'y a aucune possibilité d'être certain que ce retard n'est pas dû à des activités sans rapport avec l'activation lexicale (exemple : l'auditeur tarde à fournir sa réponse à cause d'un problème ponctuel d'inattention).

---

<sup>82</sup> C'est-à-dire dans des conditions d'écoute du stimulus desquelles la dimension évanescence du signal de parole est annulée par la visualisation de l'onde acoustique et la possibilité de réécouter le signal à volonté.

On remarque de plus que les TR mesurés depuis cette origine sont parfois inférieurs à 100 ms. et mêmes négatifs, surtout avec les mots plurisyllabiques, indiquant que la reconnaissance peut être anticipée avant ce point (reconnaissance au PUP).

### 3.2.2.3.4 Mesure depuis le point d'unicité (début, milieu ou fin)

Le PU représente un point théorique dont la validité est largement admise en psycholinguistique (modèle de référence : COHORT (Marslen-Wilson & Welsh, 1978)). Comme pour le PR cependant, des connaissances techniques sont nécessaires à sa localisation dans le signal acoustique, mais même en disposant de ces connaissances, il peut être difficile de localiser les frontières de certains phonèmes à cause de la coarticulation.

- ▷ Un phonème n'est pas un point sur l'échelle du temps : c'est une durée (de quelques dizaines à plusieurs centaines de millisecondes) qui varie selon les segments et le contexte dans lequel ils se trouvent.
- ▷ Tous les phonèmes ne sont pas traités de manière identique : la quantité de signal nécessaire à l'identification d'un phonème varie selon :
  - ∟ les segments.
  - ∟ le contexte phonétique.
  - ∟ la position du segment dans la syllabe et dans le mot.

Puisque le point d'unicité est une durée, l'origine de la mesure peut techniquement se faire à partir du début ou de la fin de cette durée, ou encore d'un compromis entre les deux. Théoriquement, quel que soit le point choisi, il y a cependant des problèmes :

- ▷ si le matériel renferme des mots à point d'unicité pré-final et des mots à point d'unicité post-final, seul le début du point d'unicité permet une mesure homogène dans ces deux cas de figure. Toutefois, la mesure impliquera alors deux origines différentes :
  - ∟ le début du point d'unicité (mots à PUP pré-final).
  - ∟ la fin du mot (mots à PUP post-final).
 Il existe donc un risque méthodologique important de biais de la mesure.
- ▷ Si le matériel ne renferme que des mots à point d'unicité post-final, la question ne se pose pas, puisqu'elle renvoie à une mesure à partir de la fin du mot.
- ▷ si le matériel ne renferme que des mots à point d'unicité pré-final, 3 choix sont possibles (début, milieu ou fin du phonème), mais :
  - ∟ avec le début du PUP :
    - on ignore quelle quantité de signal est nécessaire pour identifier le phonème correspondant au PU
    - en cas de segments fortement coarticulés, d'une part il est difficile de localiser la frontière, et d'autre part, l'auditeur utilise

l'information phonétique déjà à disposition dans le phonème précédent pour le traitement (il peut être avantagé par rapport à un mot où le segment correspondant au PUP est peu coarticulé avec le segment qui précède – la zone de coarticulation pouvant s'étendre sur plusieurs dizaines de millisecondes).

- ∩ avec la fin du PUP, le problème de la durée variable des phonèmes et de la quantité de signal nécessaire à l'identification d'un segment donné se repose.
- ∩ avec le milieu du PUP, on n'évacue pas complètement les questions de la durée variable des segments et de la quantité minimale de signal nécessaire à l'identification d'un phonème donné, mais on peut réduire leur influence, et se rapprocher d'une position voisine du PR.

Par ailleurs, le choix du PU comme origine implique un contrôle du matériel pour que, entre conditions comparées, la position moyenne du point d'unicité soit identique. Si la position est pré-finale :

- ▷ la durée du phonème correspondant au point d'unicité doit être identique.
- ▷ la durée de signal précédent le point d'unicité doit être identique.
- ▷ la quantité de signal acoustique nécessaire à l'identification du phonème correspondant au point d'unicité doit être identique.

Le contrôle de tous ces éléments est techniquement irréaliste.

Le point d'unicité d'un mot pour un individu donné n'est de plus pas forcément le point d'unicité théorique. Aucun contrôle n'est techniquement possible pour cet aspect de la réalité.

### 3.2.2.3.5 Récapitulatif

Tableau 14 : Avantages et inconvénients de chaque origine

	<b>Avantages</b>	<b>Inconvénients</b>
<b>Début du mot</b>	facile d'emploi, universel (permet une comparaison entre études)	biais à cause de la durée variable des mots
<b>PU</b>	validé par la théorie	statut différent entre mots à PU pré-final et mots à PU post-final, difficile à localiser objectivement
<b>PR</b>	référence = la perception de l'auditeur, et non l'arbitraire ou la théorie	ne permet pas d'évacuer de la mesure les caractéristiques physiques de l'item, difficile à localiser objectivement
<b>Fin du mot</b>	facile d'emploi	non-utilisé, un mot peut être reconnu avant sa fin, confondu avec la position du PU dans le cas de mots à PU post-final

### 3.2.2.3.6 Conclusions

- ▷ Aucune origine ne permet de garantir une mesure non biaisée au chercheur. Chacune présente ses intérêts et ses inconvénients.
- ▷ Le problème posé ne se résout pas avec le choix de l'origine la moins pire : on mesure toujours aussi autre chose que le phénomène cognitif étudié.

### 3.2.3 Solution technique proposée au problème de l'origine de la mesure dans l'étude de la reconnaissance des mots parlés

La solution technique qui a été adoptée est de nature statistique. Elle se base sur l'utilisation de la méthode de régression multiniveaux, qui permet de prendre en considération dans un même modèle plus d'un facteur aléatoire (sujets et items), les facteurs expérimentaux, et des covariables liées aux sujets et/ou aux items. Cette méthode est présentée à la section 3.3.4 page 123.

La technique possède toutefois ses limites, puisqu'il s'agit d'une exploration des relations linéaires entre variations d'un facteur et d'un autre, des relations non linéaires existant peut-être.

### 3.2.4 Conclusions

Face au problème de l'origine de la mesure des temps de réponse qui a été abordé dans ce chapitre, il ressort qu'aucune origine ne convient. D'une part, toutes autorisent des biais dans les données. D'autre part, les résultats peuvent radicalement varier en fonction de l'origine choisie. Il n'existe donc pas de solution simple, comme le choix d'une origine.

La solution qui a été adoptée dans ce travail consiste en un contrôle a posteriori de la situation, au moment des analyses statistiques. Le procédé utilisé consiste à modéliser tous les facteurs identifiés comme influençant les TR, afin de distinguer la part de variance qui est due à chacun d'eux dans les données, pour éviter de confondre ces influences parasites avec celle des facteurs expérimentaux. Si cette méthode peut avoir des limites (facteur non modélisé ou mal modélisé, influence non linéaire d'un facteur, etc.), elle offre deux avantages majeurs :

- ▷ Elle donne suffisamment de liberté au chercheur pour utiliser un matériel moins contrôlé, ou une situation présentant des aspects non contrôlables, mais modélisables.
- ▷ Dans tous les cas, elle permet de produire des données propres par la prise en compte d'une multitude de sources de variance, jusqu'alors ignorées dans les travaux de recherche pour des raisons techniques.

### **3.3 Chapitre 3 - Le traitement statistique des temps de réaction pour reconnaître un mot parlé isolé**

Ce chapitre décrit un ensemble de problèmes méthodologiques relatifs au traitement statistique des temps de réponse obtenus lors de la reconnaissance de mots parlés isolés et mesurés avec une tâche de décision lexicale (TDL). Les aspects abordés concernent le tri des données, la gestion des valeurs manquantes, et la méthode d'analyse statistique. Compte tenu des précautions nécessaires pour se prémunir de conclusions théoriques erronées dans une étude, l'ANOVA paraît être une technique difficile d'emploi, et inadéquate dans certaines situations. Une autre méthode d'analyse, l'analyse de régression multiniveaux, est proposée en remplacement, et justifiée par la résolution des problèmes qui seront exposés dans le développement (relatifs aux points qui précèdent). C'est cette technique qui a été utilisée pour l'analyse des expériences présentées dans ce manuscrit.

#### **3.3.1 Introduction :**

En psycholinguistique, les mesures comportementales recueillies à l'aide d'une tâche de décision lexicale sont couramment analysées par ANOVA, selon une procédure de double validation des effets fixes (analyses séparées par items et par sujets), après élimination des réponses aberrantes (« outliers ») et remplacement des valeurs manquantes par une moyenne.

Ces pratiques peuvent parfois introduire des biais dans les résultats, biais qui ont été regroupés en deux catégories :

- ▷ Les biais liés aux procédures de tri des données et de gestion des valeurs manquantes.
- ▷ Les biais liés à une utilisation inadéquate de l'ANOVA.

Le chapitre est organisé en 6 sections, dont cette introduction représente la première.

- ▷ La deuxième section (Section 3.3.2) est consacrée à un examen détaillé des pratiques de préparation des données en vue de leur analyse avec une ANOVA (tri des données et gestion des valeurs manquantes), afin de montrer quels sont les risques encourus sur les résultats. Quelques pistes sont proposées pour éviter certains biais.
- ▷ La section suivante (Section 3.3.3) est consacrée à l'utilisation des ANOVA en psycholinguistique. Elle vise à montrer que :
  - ∩ l'ANOVA n'est généralement pas la meilleure technique à utiliser avec des expériences à mesures répétées, essentiellement parce que les données



qui en sont issues renferment une structure hiérarchique et plus d'un facteur aléatoire.

- ∩ la routine de double validation des effets fixes par le cumul des tests  $F_1$  et  $F_2$ , qui s'est généralisée et institutionnalisée en psycholinguistique, ne garantit pas au chercheur la possibilité de conclure sur l'existence d'un effet fixe.
- ▷ La Section 3.3.4 présente la méthode d'analyse de régression multiniveaux, en mettant l'accent sur les avantages qu'elle offre par rapport à l'ANOVA pour traiter des données recueillies dans une expérience à mesures répétées.
- ▷ La Section 3.3.5 présente les conclusions des points abordés dans le Chapitre 3.
- ▷ Une dernière section (Section 3.3.6) clos le chapitre avec un récapitulatif des problèmes et solutions pour le traitement des données, la gestion des valeurs manquantes, et l'analyse statistique des données.

### 3.3.2 La préparation des données pour leur analyse statistique

#### 3.3.2.1 Tri des données

Le tri des données a pour but d'éliminer de la distribution les réponses sans rapport avec le traitement demandé, c'est à dire trop éloignées de l'ensemble des observations<sup>83</sup>, qui représentent du bruit dans la distribution. De plus, les techniques d'analyse statistique généralement employées en psychologie (analyse de régression, ANOVA) sont théoriquement conçues pour traiter des distributions normales<sup>84</sup>.

##### 3.3.2.1.1 Réflexion à partir de la procédure communément suivie en psycholinguistique

La procédure de tri couramment employée en psycholinguistique consiste à éliminer en aveugle les valeurs de la distribution générale situées au-delà de  $\pm 2$  écart-types autour de la moyenne (95,44% des valeurs sont conservées).

Hormis l'élimination de 4.66% d'observations normales (qui devraient être conservées), ce procédé, par un filtrage mal ciblé de l'information, peut conduire à transformer la distribution réellement produite et créer des biais dans les résultats :

---

<sup>83</sup> On considère que les phénomènes psychologiques se manifestent généralement par un comportement gaussien. Une observation qui s'écarte trop du centre de la distribution normale peut être considérée comme étant sans rapport avec le phénomène étudié, ou révéler une perturbation provenant de l'intervention d'un autre phénomène.

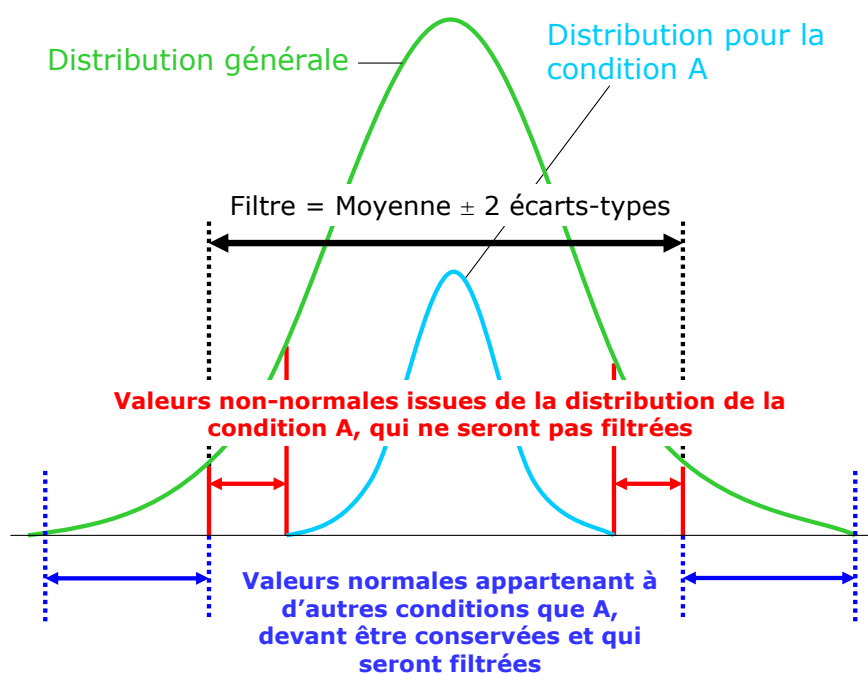
<sup>84</sup> Remarque : les réponses qui seront éliminées sur un critère de normalité peuvent être différentes en fonction du choix de l'origine choisie pour mesurer les temps de réaction (se reporter au chapitre 2 de cette partie).

- ▷ lorsque des valeurs non normales d'une condition ont été conservées alors qu'elles auraient dû être éliminées.
- ▷ lorsque des valeurs normales d'une condition ont été éliminées alors qu'elles auraient dû être conservées.

Pour mieux se représenter les conséquences de ce problème de mauvais ciblage du filtre, la Figure 26 illustre un premier cas de figure où l'une des conditions (A), sous-ensemble de la distribution, est centrée sur la moyenne générale, mais présente une étendue plus restreinte que la distribution de la totalité des données. L'élimination des valeurs situées au-delà de  $\pm 2$  écart-types autour de la moyenne générale conduit à :

- ▷ une perte d'information de la distribution générale.
- ▷ la conservation de valeurs non normales de la distribution A.

Figure 26 : Premier exemple d'effets indésirables liés à un filtre mal adapté

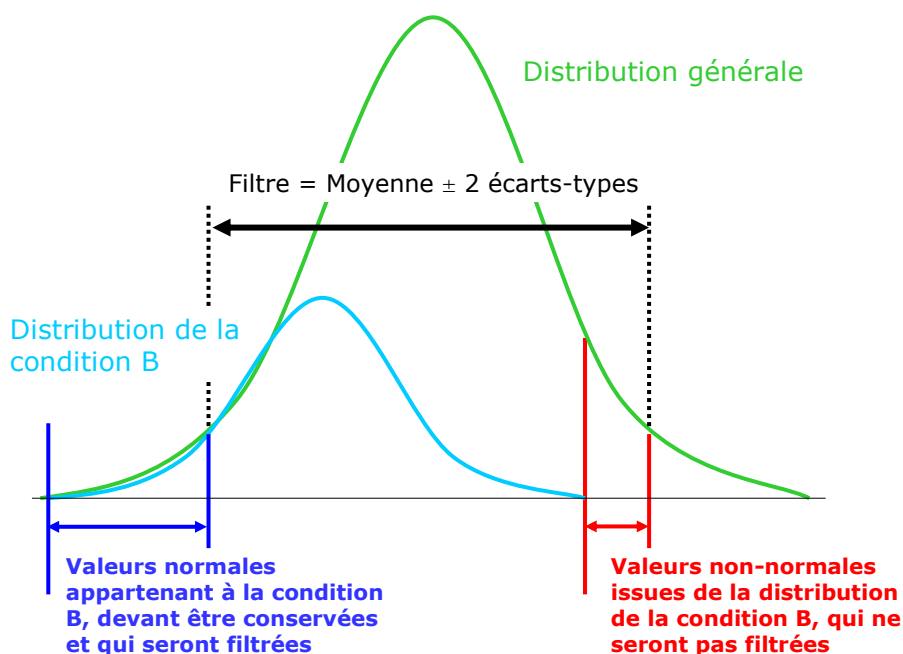


Ce procédé de tri des données ne modifie pas la moyenne de A, qui correspond à la moyenne de la distribution générale.

Le schéma suivant (Figure 27) illustre un second cas de figure, dans lequel une condition (B), sous-ensemble de la distribution, possède une moyenne différente de la moyenne générale. Dans cette situation, un filtrage à  $\pm 2$  écart-types conduit à :

- ▷ éliminer asymétriquement des observations de B distribuées normalement
- ▷ conserver asymétriquement des observations de B non normales

Figure 27 : Deuxième exemple d'effets indésirables liés à un filtre mal adapté



Les conséquences sont beaucoup plus graves que dans l'exemple de la condition A :

- ▷ le procédé rend la courbe non gaussienne
- ▷ il augmente ou diminue la moyenne de la condition B, en plus d'une modification de l'étendue de la distribution de cette condition

Dans la mesure où l'exemple de la condition B est le plus probablement rencontré relativement à l'exemple de la condition A, ce procédé de tri des données ne devrait pas être utilisé.

### 3.3.2.1.2 Quelles solutions pour le tri des données ?

A priori, trois solutions existent. Elles se distinguent par le niveau de ciblage du filtre (précision) :

- ▷ filtrage à  $\pm 3$  écarts-type sur la distribution totale
- ▷ filtrage à  $\pm 3$  écarts-type sur la distribution par condition
- ▷ filtrage à  $\pm 3$  écarts-type sur la distribution des observations pour le niveau unitaire (distribution des réponses d'un sujet ou pour un item)

Si les 2 premières solutions apportent une amélioration à la méthode usuelle de tri des données présentées en Section 3.3.2.1.1 page 108 (le filtrage à  $\pm 3$  écarts-type de la distribution totale conserve 99.73% des valeurs et résout le problème de l'élimination

possible des valeurs normales d'une condition ; le filtrage à  $\pm 3$  écarts-type de la distribution de chaque condition résout en plus le problème de la conservation des valeurs non normales d'une condition), elles ne remplissent pas entièrement le but pour lequel une procédure de tri est utilisée (élimination du bruit). Ces deux premières solutions sont donc mal adaptées, du fait qu'elles induisent encore une distorsion des distributions analysées, relativement aux distributions réellement produites (le filtrage à  $\pm 3$  écarts-type de la distribution totale ne permet pas de résoudre les problèmes qui découlent de la conservation des valeurs non normales d'une condition ; le filtrage à  $\pm 3$  écarts-type de la distribution de chaque condition conserve des réponses non normales dans les distributions des réponses d'un sujet ou pour un item).

La troisième solution (filtrage croisé à  $\pm 3$  écarts-type pour la distribution de chaque sujet et de chaque item<sup>85</sup>) est la plus difficile à mettre en œuvre, mais c'est aussi la plus précise. Elle permet une élimination ciblée des réponses non normales, sans déformation des distributions recueillies, et réduit donc fortement le risque d'un mauvais filtrage. C'est cette solution qui a été utilisée pour la préparation des données de toutes les expériences présentées dans ce manuscrit.

Il faut toutefois noter que :

- ▷ lorsque le nombre de réponses par item ou par sujet est petit, l'estimation de la moyenne et de l'étendue peut être problématique, et la décision aux frontières de la distribution, être entachée d'incertitude.
- ▷ étant donné que plusieurs niveaux existent dans les réponses obtenues (observation, item, sujet, condition expérimentale, distribution générale), ce filtrage ne dispense pas d'un tri préalable des items et des sujets « anormaux », sur la base de la distribution de leur moyenne.

### 3.3.2.2 Gestion des valeurs manquantes

Effectuer une ANOVA implique de disposer d'un jeu complet de données, dans lequel chaque sujet a fourni une réponse à chaque item. En cas de valeur manquante (non-réponse ou réponse non normale éliminée par filtrage), la technique n'est plus directement applicable. Avant de solutionner le problème, il faut déterminer si la cause des valeurs manquantes est en lien avec un traitement (« non-missing ») ou résulte d'un phénomène aléatoire (« missing at random ») :

- ▷ Dans le cas de valeurs manquantes à cause d'un traitement, aucun remplacement ne devrait être effectué :

---

<sup>85</sup> Son principe : la distribution des données est examinée pour chaque item et pour chaque sujet. Seuls les « outliers » de chaque distribution sont éliminés.

∩ généralement, la cause de leur absence est liée à un item ou un sujet, qui doivent être éliminés des analyses. Ceci implique ensuite d'équilibrer les autres conditions en retirant autant d'items et de sujets de manière à en avoir la même quantité dans toutes les conditions, sous peine de disposer d'un jeu de données incomplet. Cependant :

- le choix des items et des sujets à retirer pour équilibrer les conditions risque d'être arbitraire : le procédé peut être éthiquement douteux.
- éliminer des items peut aboutir à déséquilibrer le matériel à travers les conditions expérimentales, en fonction des items retirés.

Aussi serait-il préférable de travailler avec une technique d'analyse permettant de traiter des plans d'expérience incomplets, comme la méthode d'analyse multiniveaux.

∩ l'autre cause possible à ces valeurs manquantes est liée au traitement d'un des facteurs expérimentaux (exemple : fréquence d'usage d'un mot. Les mots fréquents ont statistiquement plus de chances d'être connus par tous les sujets, et donc statistiquement moins de chances de conduire à des valeurs manquantes (non-réponses, réponses éliminées car fournies avec un temps de latence beaucoup trop long par rapport à la moyenne). Ce n'est pas le cas des mots rares, qui risquent de paraître inhabituels voire inconnus, et pour lesquels il y aura plus de valeurs manquantes). Ceci soulève un problème théorique plus fondamental, et plus grave pour la faisabilité de l'analyse. L'expérimentateur doit en effet comprendre pourquoi il a obtenu un tel phénomène et y remédier. La solution qui consisterait à remplacer ces valeurs manquantes par une moyenne serait une erreur méthodologique.

▷ Dans le cas de valeurs manquant à cause d'un phénomène aléatoire, le remplacement peut se faire :

∩ par une valeur moyenne :

- générale
- par condition
- par item et sujet ( $\Sigma \text{valeurs}_{\text{item}} + \Sigma \text{valeurs}_{\text{sujet}} / N \text{ valeurs}_{\text{item et sujet}}$ )

D'autres techniques plus élaborées existent (exemple : procédure de Yates (Yates, 1933)). Elles sont à choisir en fonction de la structure expérimentale analysée.

∩ Par des valeurs estimées, générées à l'aide d'algorithmes, à partir des paramètres qui décrivent la distribution considérée (moyenne, étendue, kurtosis, etc.). Il s'agit donc de valeurs appartenant à la distribution, qui auraient pu être produites par les sujets dans l'expérience.

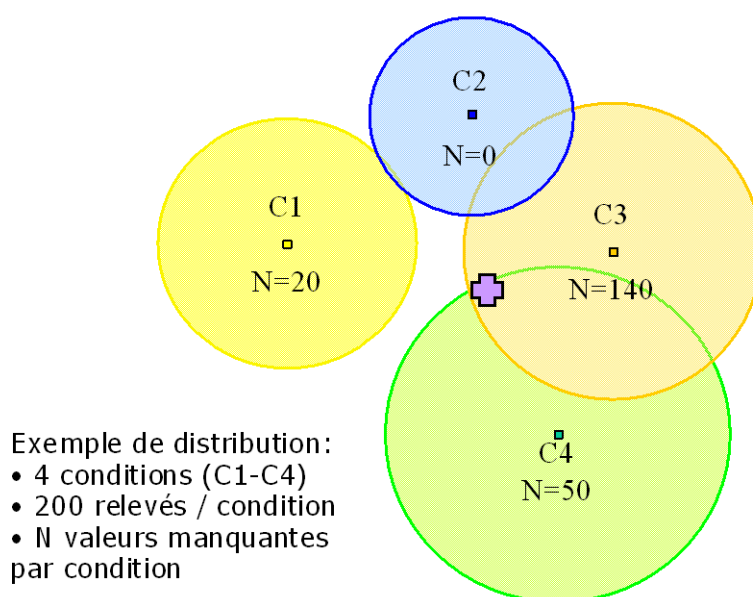
Il est préférable de procéder au remplacement des valeurs manquantes par des valeurs estimées, bien que le plus souvent et pour des raisons de facilité, la valeur moyenne de la distribution générale soit utilisée. Le problème qu'il y a à utiliser une valeur moyenne est illustré ci-dessous à l'aide d'un exemple fictif (distorsion de la variance, déplacement des moyennes, modification des effets).

### 3.3.2.2.1 Conséquences de la procédure de remplacement des valeurs manquantes sur la distribution

#### 3.3.2.2.1.1 Présentation de l'exemple fictif

Le diagramme suivant (Figure 28) montre le détail de la dispersion des valeurs (surfaces circulaires colorées) autour de la moyenne par condition (carrés colorés au centre des cercles), la moyenne générale de la distribution étant marquée par la croix violette. Le but de cette figure est d'aider à se représenter les modifications induites sur la moyenne (centre de gravité) et l'étendue de la distribution (dispersion autour de la moyenne) par les différents procédés de remplacement des valeurs manquantes. Il s'agit d'une schématisation des notions de centre de gravité et de dispersion autour de la moyenne, et non de la représentation bivariée d'une situation réelle ou imaginée.

Figure 28 : Schématisation des notions de centre de gravité et de dispersion autour de la moyenne pour un exemple fictif de distribution



Dans cet exemple, l'étendue des distributions et le nombre de valeurs manquantes par condition ont été choisis inégaux, comme c'est vraisemblablement le cas dans une situation réelle. À cause du nombre inégal de valeurs manquantes par condition, certaines conditions (comme C3) seront plus affectées par le remplacement des valeurs manquantes avec une moyenne que d'autres (comme C1), tandis qu'une condition dans cet exemple ne sera aucunement affectée puisqu'elle ne renferme pas de valeur manquante (C2).

Le remplacement des valeurs manquantes par une moyenne produira d'une part une distorsion des variances de chaque condition, et d'autre part une distorsion inégale de la variance d'une condition à l'autre.

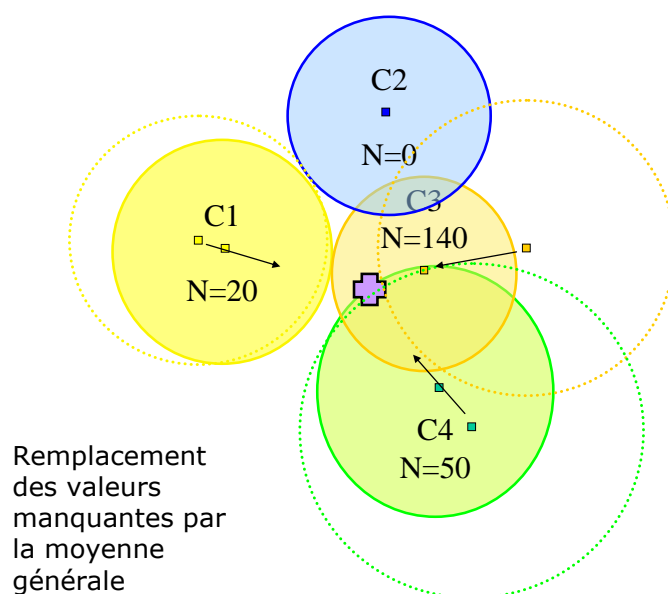
À la section précédente, 4 méthodes de remplacement des valeurs manquantes ont été présentées et discutées (par une moyenne (générale, par condition, par item et sujet), par des valeurs estimées). Dans les 4 sections suivantes, les conséquences de ces procédés de remplacement sur la distribution de l'exemple fictif qui vient d'être présenté sont discutées.

### **3.3.2.2.1.2 Conséquences d'un remplacement des valeurs manquantes par la moyenne générale**

Le diagramme ci-dessous (Figure 29) montre qu'avec cette méthode :

- ▷ La dispersion des valeurs autour de la moyenne de chaque condition (cercles pleins) est moins importante que dans la distribution d'origine (cercles pointillés - écrasement de la variance intragroupe)
- ▷ Le centre de gravité de chaque condition se décale vers la moyenne générale (écrasement de la variance intergroupe)
- ▷ Il se décale proportionnellement au nombre de valeurs manquantes (création d'une distorsion entre conditions)

Figure 29 : État de la distribution après remplacement des valeurs manquantes par la moyenne générale



Cette procédure entraîne :

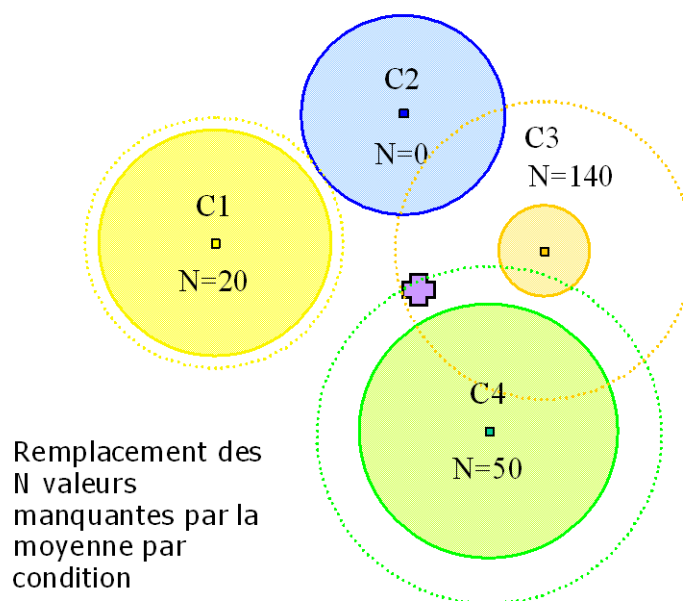
- ▷ le risque de faire disparaître des effets principaux existant (erreurs de type II).
- ▷ le risque de faire apparaître des interactions qui n'existent pas (erreurs de type I<sup>86</sup>).

On justifie généralement son usage par son avantage à ne révéler que des effets principaux robustes. Par contre, il ne faut pas l'employer si l'étude porte aussi sur des interactions. Cette méthode ne devrait pas être utilisée.

### 3.3.2.2.1.3 Conséquences d'un remplacement des valeurs manquantes par la moyenne par condition

Le diagramme ci-dessous (Figure 30) montre qu'avec ce procédé, la dispersion des valeurs autour de la moyenne de chaque condition (cercles pleins) est réduite par rapport à la distribution d'origine (cercles pointillés – écrasement de la variance intragroupe seulement).

Figure 30 : État de la distribution après remplacement des valeurs manquantes par la moyenne par condition



En conséquence, aucun effet principal ne peut disparaître. Au contraire, à cause de la diminution de la variance intragroupe :

- ▷ des effets principaux peuvent apparaître

<sup>86</sup>  $H_0$  est rejetée alors qu'elle est vraie, ou encore, un facteur ne produit pas d'effet mais on conclue l'inverse.



- ▷ des interactions peuvent apparaître

Cette procédure favorise donc les erreurs de type I, et serait pire que la précédente du point de vue des retombées de son usage sur la théorie. Aussi est-il préférable de l'éviter.

#### **3.3.2.2.1.4 Conséquences d'un remplacement des valeurs manquantes par la moyenne par sujet&item**

Une meilleure précision peut être obtenue en utilisant la dernière méthode de remplacement d'une valeur manquante par une valeur moyenne par sujet ET par item ( $\Sigma \text{valeurs}_{\text{item}} + \Sigma \text{valeurs}_{\text{sujet}} / N \text{ valeurs}_{\text{item et sujet}}$ ). Cependant, comme tout remplacement de valeur par une moyenne entraîne l'écrasement de la variance et une transformation de la distribution réelle des réponses, il serait préférable d'utiliser une méthode d'estimation des valeurs manquantes.

#### **3.3.2.2.1.5 Conséquences d'un remplacement des valeurs manquantes par des valeurs estimées avec un algorithme<sup>87</sup>**

Cette procédure est intéressante car elle n'écrase pas la variance et a peu d'incidence sur les caractéristiques de la distribution (à condition que la proportion « valeurs disponibles / valeurs manquantes » soit suffisamment élevée pour permettre de décrire la distribution de façon assez précise). C'est la méthode préférable parmi les 4 méthodes listées et précédemment discutées.

### **3.3.2.2.2 Quelles solutions au problème des valeurs manquantes ?**

- ▷ Si l'ANOVA doit être utilisée, le remplacement des valeurs manquantes doit se faire selon une procédure qui :
  - ∩ ne modifie pas ou modifie le moins possible la position de la moyenne (variance intergroupes)
  - ∩ ne modifie pas ou modifie le moins possible la dispersion des valeurs autour de cette moyenne (variance intragroupe)
  - ∩ porte sur le niveau le plus unitaire possible.

Il est donc préférable de recourir à des procédures d'estimation des valeurs manquantes qu'à leur remplacement par une valeur moyenne.
- ▷ S'il n'est pas impératif d'utiliser l'ANOVA, l'idéal est de faire appel à une procédure d'analyse capable de travailler sur un jeu de données incomplètes,

---

<sup>87</sup> Exemple d'algorithme: « Expectation Maximization » (Dempster, Laird, & Rubin, 1977). Procédure utilisée pour approximer une fonction de probabilité. Elle permet de générer, grâce à un processus itératif, des valeurs appartenant à la distribution, à partir de la moyenne, de la matrice de covariance, et des corrélations entre variables et valeurs manquantes.

comme l'analyse de régression multiniveaux. Cette méthode offre l'avantage supplémentaire de pouvoir analyser des données recueillies selon un plan incomplet (nombre différent mais du même ordre de grandeur de sujets et/ou d'items par groupe).

### 3.3.3 L'analyse statistique des données

Les expériences présentées dans ce rapport ont une structure factorielle : un même individu fournit des réponses à une série d'items, ces mêmes items étant présentés à plusieurs individus (expériences de type « mesures répétées »). Elles comportent un ou plusieurs facteurs fixes (variables étudiées) et au moins deux facteurs aléatoires (sujets, items<sup>88</sup>). Analyser des données ayant une telle structure avec une ANOVA implique deux problèmes liés :

- ▷ La difficulté à traiter une hiérarchie dans les données (les réponses d'un sujet peuvent être corrélées entre elles – problème de la sphéricité)
- ▷ La difficulté à traiter simultanément deux facteurs aléatoires

Ces deux problèmes sont détaillés dans les deux paragraphes suivants.

#### La présence d'une structure hiérarchique dans les données

Les réponses obtenues dans une expérience de type « mesures répétées » peuvent ne pas être indépendantes (corrélations intra classe ou par sujet). Ceci est la traduction de l'existence d'une structure hiérarchique dans les données (niveau des sujets, niveau des items considérés comme les représentations mentales des sujets). Ignorer ce fait revient à courir le risque de sous-estimer une erreur de type I. Aussi, la technique d'ANOVA choisie doit être adaptée pour traiter correctement cette structure hiérarchique.

#### La présence de deux facteurs aléatoires

Lorsque deux facteurs aléatoires sont présents dans la situation (sujets et items), ils doivent être simultanément considérés dans l'analyse sous peine de courir le risque d'une erreur de type I. S'ils sont considérés séparément, il n'est pas possible de conclure quant à la généralisation des effets des facteurs étudiés à tous les sujets et tous les items, ni même conclure quant à la généralisation des effets des facteurs étudiés à tous

---

<sup>88</sup> En réalité, la situation est plus complexe : items et sujets sont aussi croisés. Le croisement fait référence au fait que chaque sujet dispose d'une organisation de son lexique mental qui lui est propre (mots connus, âge d'acquisition d'un mot, familiarité, vécu par rapport au concept ou à l'objet désigné par un mot (charge affective), etc.). La réactivité du système dépend alors à la fois du mot et du sujet. Dans ce travail, cette notion de croisement est modélisée par la structure hiérarchique du modèle, où les mots sont aussi considérés comme étant emboîtés dans les sujets.

les sujets ou à tous les items. En effet, agréger les deux niveaux de variance (item et sujet) pour faire comme s'il n'y en avait qu'un n'est pas suffisant pour retirer des données la variance liée au niveau qui n'est pas testé. Celle-ci s'ajoute à la variance de l'autre facteur aléatoire, mais aussi à la variance du ou des facteurs expérimentaux : elle peut donc biaiser les effets des facteurs étudiés et conduire à des conclusions erronées. La technique d'ANOVA choisie doit donc être adaptée pour traiter simultanément les 2 facteurs aléatoires.

Les paragraphes suivants expliquent les raisons pour lesquelles il est nécessaire de prendre des précautions en présence d'une structure hiérarchique dans les données, et en quoi la présence de 2 facteurs aléatoires entraîne l'obligation d'adapter ses méthodes d'analyse.

### 3.3.3.1 L'utilisation de l'ANOVA en psycholinguistique

Les décisions sur la significativité d'un effet sont souvent fondées, en psycholinguistique, sur une utilisation mal adaptée de l'ANOVA. Cette inadéquation provient du fait d'un décalage entre les besoins de l'analyse et les possibilités de la méthode choisie. Trois solutions existent pour remédier à cela :

- ▷ la première consiste à choisir correctement la technique d'ANOVA, à savoir celle qui correspond à la structure de l'expérience analysée.
- ▷ cependant, pour certains types d'expériences factorielles, il n'est pas toujours possible de construire le quotient F adéquat à partir de l'application directe des règles d'estimation des carrés moyens (Winer, Brown, & Michels, 1991), et donc de disposer du test statistique exact pour évaluer les effets du modèle. La deuxième solution consiste donc à calculer un quotient « quasi-F » (Clark, 1973; Kirk, 1968; Winer et al., 1991), en ajoutant ou retranchant certains termes des carrés moyens de manière à disposer d'un rapport permettant de tester l'hypothèse nulle sur le facteur souhaité (voir Clark (1973) pour les détails). Cette méthode peut être assez compliquée mathématiquement et nécessite de bonnes connaissances en statistiques.
- ▷ la troisième solution consiste à avoir recours à une technique plus complexe que l'ANOVA. L'analyse de régression multiniveaux a été choisie pour ce travail.

Au lieu de choisir l'une de ces 3 solutions face à des données présentant une structure hiérarchique et contenant 2 facteurs aléatoires, une procédure de double validation des effets étudiés est la plupart du temps employée. Elle consiste à tester conjointement par  $F_1$  (on suppose que le même comportement serait observé si les mêmes items étaient soumis à un nouveau lot de sujets) puis par  $F_2$  (on suppose que le même comportement serait observé si un nouveau lot d'items était proposé aux mêmes sujets) les mêmes données. Pour mieux comprendre en quoi cette pratique ne solutionne rien et peut avoir des conséquences dangereuses sur la théorie, un retour aux bases

statistiques est proposé à travers les deux sections suivantes. Le problème n'est pas nouveau et a déjà été signalé au début des années 70<sup>89</sup>.

### 3.3.3.1.1 Les conditions d'utilisation de l'ANOVA

Utiliser la technique de l'ANOVA nécessite que les 3 présupposés suivants soient suffisamment vrais :

- ▷ Normalité de la distribution des observations.
- ▷ Indépendance des erreurs (pas de corrélation liée à l'échantillonnage des sujets).
- ▷ Homogénéité de la variance intragroupe.

Pour les ANOVA traitant des facteurs à mesures répétées sur plus de 2 niveaux, la condition de sphéricité doit aussi être respectée<sup>90</sup>.

Selon Cochran (Cochran, 1947), il est difficile d'être dans une situation où tous ces présupposés sont satisfaits : une ANOVA fournit donc toujours des résultats relatifs. Cependant, si ces présupposés ne sont pas réunis, l'analyse, bien que théoriquement inapplicable, reste techniquement faisable. Elle peut alors conduire à des résultats erronés.

### 3.3.3.1.2 Ce que testent $F_1$ et $F_2$

Les bases mathématiques du test F sont présentées ci-dessous pour mieux comprendre ce qui est évalué avec la procédure de double validation d'un effet fixe (par  $F_1$  et par  $F_2$ ) et en quoi cette méthode peut conduire à des conclusions erronées. L'expérience qui est prise en exemple est de structure simple : il s'agit d'une passation à mesures répétées. Elle comporte des sujets, un traitement ou facteur fixe, et des items dans le traitement (mots différents d'une condition à l'autre).

---

<sup>89</sup> La méthode des quasi-F a été utilisée pour un temps en psycholinguistique, mais elle a graduellement disparu au profit d'un consensus nouveau (l'évolution, entre 1974 et 1997, du rapport « utilisation du  $F_{\min}^*$ /utilisation conjointe de  $F_1$  et  $F_2$  » dans les revues « Journal of verbal learning and verbal behavior » et « Journal of memory and language », est présentée dans Raaijmakers, Schrijnemakers et Gremmen (1999), p. 421).

<sup>90</sup> La sphéricité est une condition nécessaire et suffisante pour la validité du test F. Elle suppose que la variance des différences 2 à 2 entre groupes soient homogènes. Si le principe de sphéricité est violé, les degrés de liberté du test F doivent être corrigés, ou une autre méthode d'analyse doit être utilisée.

**3.3.3.1.2.1 Présentation de l'équation du modèle linéaire de base pour l'ANOVA de ce type d'expérience**

Équation 1 : Équation du modèle linéaire de base pour l'ANOVA d'une expérience à mesures répétées (reproduit à partir de (Raaijmakers et al., 1999))

$$Y_{ijk} = \mu + \alpha_k + \beta_{j(k)} + \pi_i + \alpha\pi_{jk} + \pi\beta_{ij(k)} + \varepsilon_{0(ijk)}$$

Avec :

*i* sujet, *j* item, *k* traitement

$\mu$  = moyenne générale

$\alpha_k$  = effet principal du traitement (T)

$\beta_{j(k)}$  = effet principal du mot dans le traitement (W(T))

$\pi_i$  = effet principal du sujet (S)

$\alpha\pi_{jk}$  = interaction traitement \* sujet (TS)

$\pi\beta_{ij(k)}$  = interaction sujet \* mot dans le traitement (W(T)S)

$\varepsilon_{0(ijk)}$  = erreur expérimentale

**3.3.3.1.2.2 Calcul des quotients  $F_1$  et  $F_2$** 

Tester la significativité d'une source de variation revient à calculer un quotient F de carrés moyens (CM), construit de manière à ce que la valeur attendue (espérance mathématique) du numérateur soit égale à l'espérance mathématique du dénominateur plus un terme qui reflète l'effet testé. Cela revient à montrer que la variance due au traitement ( $\sigma^2_T$ ) est plus grande que 0 (en théorie, si  $\sigma^2_T = 0$ , le quotient CMtraitement / CMerreur se distribue comme un F autour de la moyenne  $n / (n-2)$ , avec  $n$  le degré de liberté du carré moyen de l'erreur. Si ce quotient s'écarte suffisamment de cette moyenne, l'hypothèse nulle peut être rejetée).

Le calcul du F nécessite donc de connaître le carré moyen attendu pour chaque source de variation présente dans la situation expérimentale, et les degrés de liberté associés (pour la situation donnée en exemple (Clark, 1973; Raaijmakers et al., 1999), les carrés moyens attendus se calculent à partir des équations fournies dans le Tableau 15).

Tableau 15 : Calcul de l'espérance mathématique de chaque source de variation

Sources de variation	Label	dl	Carré moyen attendu (esp. math.)
Traitement	T	p-1	$\sigma^2_e + \sigma^2_{W(T)S} + q\sigma^2_{TS} + r\sigma^2_{W(T)} + rq\sigma^2_T$
Mot dans le traitement	W(T)	p(q-1)	$\sigma^2_e + \sigma^2_{W(T)S} + r\sigma^2_{W(T)}$
Sujet	S	r-1	$\sigma^2_e + \sigma^2_{W(T)S} + pq\sigma^2_S$
Traitement * Sujet	TS	(p-1)(r-1)	$\sigma^2_e + \sigma^2_{W(T)S} + q\sigma^2_{TS}$
Mot dans le traitement * Sujet	W(T)S	p(q-1)(r-1)	$\sigma^2_e + \sigma^2_{W(T)S}$

Avec  $p$  le nombre de niveaux pour la variable expérimentale,  $r$  le nombre de sujets,  $q$  le nombre d'items.

Les quotients  $F_1$  et  $F_2$  se calculent de la manière suivante :

Équation 2 : Calcul de  $F_1$

$$F_{1(p-1, (p-1)(r-1))} = \frac{CM_T}{CM_{TS}} = \frac{rq\sigma_T^2 + r\sigma_{W(T)}^2 + \sigma_e^2 + \sigma_{W(T)S}^2 + q\sigma_{TS}^2}{\sigma_e^2 + \sigma_{W(T)S}^2 + q\sigma_{TS}^2}$$

Équation 3 : Calcul de  $F_2$

$$F_{2(p-1, p(q-1))} = \frac{CM_T}{CM_{W(T)}} = \frac{rq\sigma_T^2 + r\sigma_{W(T)}^2 + \sigma_e^2 + \sigma_{W(T)S}^2 + q\sigma_{TS}^2}{\sigma_e^2 + \sigma_{W(T)S}^2 + r\sigma_{W(T)}^2}$$

Chacun de ces rapports présente 2 termes du numérateur qui ne se trouvent pas dans le dénominateur :

- ▷  $\sigma_T^2$ , le terme que l'on souhaite tester (variance liée au traitement), commun à  $F_1$  et  $F_2$
- ▷ un terme propre à  $F_1$  ( $\sigma_{W(T)}^2$ ) ou à  $F_2$  ( $\sigma_{TS}^2$ ).

Ainsi, une valeur significative de  $F_1$  peut être obtenue dans 3 cas de figure (Clark, 1973) :

- ▷ si  $\sigma_T^2 > 0$  et  $\sigma_{W(T)}^2 = 0$  (1)
- ▷ si  $\sigma_T^2 = 0$  et  $\sigma_{W(T)}^2 > 0$  (2)
- ▷ si  $\sigma_T^2 > 0$  et  $\sigma_{W(T)}^2 > 0$  (3)

Ce qui implique qu'un effet significatif obtenu par  $F_1$  peut renvoyer à :

- ▷ un effet du traitement (1)
- ▷ ou un effet des mots dans le traitement (2)
- ▷ ou un effet cumulé des deux (3)

Le seul cas qui permette une conclusion fiable et théoriquement correcte est le premier ( $\sigma_T^2 > 0$ , autre terme nul). Le 2<sup>ème</sup> cas conduit obligatoirement à une erreur d'interprétation puisque l'effet du traitement est nul, tandis que dans le 3<sup>ème</sup> cas, l'effet du traitement peut être non significatif, mais interprété comme significatif à cause du second terme ajouté. Comme le même résultat peut être obtenu au test  $F_1$  dans les trois cas sans qu'il ne soit possible de connaître lequel est considéré, il existe forcément une incertitude sur la cause de la significativité de l'effet. Le même raisonnement peut être tenu pour  $F_2$ .

### **3.3.3.1.2.3 Pourquoi la double validation d'un effet fixe, par $F_1$ et $F_2$ , ne permet pas de résoudre cette incertitude ?**

Utiliser  $F_1$  et  $F_2$  conjointement pour tester un effet ne solutionne pas l'incertitude sur la cause de la significativité du phénomène étudié. Le chercheur peut en effet jouer de malchance s'il est dans l'un des cas suivants :

- ▷  $\sigma^2_T = 0$  et  $\sigma^2_{W(T)} > 0$  pour  $F_1$  ET  $\sigma^2_T = 0$  et  $\sigma^2_{TS} > 0$  pour  $F_2$ .
- ▷  $\sigma^2_T > 0$  et  $\sigma^2_{W(T)} > 0$  pour  $F_1$  (effet non significatif du traitement mais effet significatif du cumul) ET  $\sigma^2_T > 0$  et  $\sigma^2_{TS} > 0$  pour  $F_2$  (idem).

Lors du test d'hypothèse ( $\sigma^2_T = 0$ ), il est ainsi possible d'obtenir une valeur significative de  $F_1$  à cause de la variance des mots dans le traitement ( $\sigma^2_{W(T)}$ ), et de  $F_2$  à cause de la variance de l'interaction sujets \* traitement ( $\sigma^2_{TS}$ ), c'est-à-dire pour des raisons qui ne sont pas liées à un effet principal du traitement. La conclusion sera pourtant que le facteur étudié a bien un effet significatif.

Selon Forster et Dickinson (1976), ces erreurs de type I surviendraient plus fréquemment qu'on ne le pense dans la littérature (« in extreme cases, the type I error rates for  $F_1$  and  $F_2$  can exceed the desired rate by a factor of at least 10 »), et comme le souligne Clark (1976), leur incidence sur la théorie ne serait ni anodine, ni aisément rectifiée (« (type I error is much more serious than the type II error) because highly significant results appear definitive and tend to discourage further investigation. What is more serious, to my mind, is that too many type I errors has served as foundation stones for highly influential theories or found their way into textbooks as 'definitive' findings »).

### **3.3.3.2 Quelles solutions techniques ?**

- ▷ Face au problème de la sphéricité :  
Si une structure hiérarchique existe dans les données, des techniques spécifiques d'ANOVA doivent être appliquées, ou les données doivent préférablement et plus simplement être traitées par analyse multiniveaux.
- ▷ Face au problème lié à la présence de 2 facteurs aléatoires dans la situation analysée :  
Le test F doit être effectué à partir d'un quotient calculé différemment (quasi-F), dont la formule est fonction de la situation expérimentale analysée. Une discussion technique et théorique autour de ce problème en psycholinguistique peut être trouvée dans les articles de Clark, de Forster et Dickinson, de Raaijmakers et col., et de Wike et Church (Clark, 1973, 1976; Forster & Dickinson, 1976; Raaijmakers et al., 1999; Wike & Church, 1976), ainsi que dans des ouvrages de statistique relatifs au sujet. L'utilisation d'une technique d'analyse de régression multiniveaux offre aussi la possibilité de résoudre ce problème.

### 3.3.3.3 Conclusions sur l'emploi de l'ANOVA

L'ANOVA apparaît comme une technique finalement difficile à utiliser correctement. Elle paraissait moins bien adaptée aux nécessités du travail que la technique d'analyse de régression multiniveaux, parce que :

- ▷ il était nécessaire de prendre en compte 2 facteurs aléatoires simultanément (sujets et items).
- ▷ le lot de réponses recueillies renfermait des valeurs manquantes ou le plan d'expérience n'était pas équilibré.
- ▷ des covariables liées aux items ou aux sujets devaient être incluses dans les analyses.

C'est donc l'analyse de régression multiniveaux qui a été employée à la place de l'ANOVA.

### 3.3.4 Présentation de la technique de régression multiniveaux

#### 3.3.4.1 Théorie

« L'analyse multiniveaux est une méthodologie pour l'analyse de données présentant des motifs complexes de variation, avec une focalisation particulière sur les sources emboîtées de variabilité : élèves dans les classes, employés dans les entreprises, etc. » (Snijders & Bosker, 1999). L'analyse de régression multiniveaux peut être considérée comme une extension de la technique de régression, dont l'ANOVA est un cas particulier. Il s'agit d'une technique basée sur la régression linéaire, dont le développement assez récent (années 1980 (Goldstein, 2003)) est lié à la recherche d'une solution technique pour :

- ▷ l'analyse des modèles à effets mixtes (présence de facteurs fixes et aléatoires)
- ▷ la prise en compte de l'influence du contexte dans les études en sciences sociales.

« Historiquement, les problèmes multiniveaux ont conduit à des approches d'analyse qui déplaçaient les variables sur un seul niveau d'intérêt, par agrégation ou désagrégation, avant de procéder à une régression multiple, analyse de variance ou quelque autre méthode d'analyse standard. Cependant, analyser des variables de niveaux différents à un seul niveau est inadéquat, parce que conduisant à deux problèmes distincts. » (Hox, 2002)

Ces deux problèmes sont de nature statistique et conceptuelle :

- ▷ d'un point de vue statistique, si les données sont agrégées, de l'information est perdue et l'analyse perd en puissance, alors que si les données sont



désagrégées, les valeurs des niveaux supérieurs sont démultipliées, produisant des effets significatifs qui n'existent pas.

- ▷ d'un point de vue conceptuel, deux types d'erreur d'interprétation peuvent survenir, par l'attribution d'un effet au niveau incorrect :
  - ∇ les erreurs écologiques surviennent lors de l'interprétation à un niveau individuel de données agrégées (on oublie qu'une partie des effets est liée au contexte en attribuant la totalité aux individus).
  - ∇ les erreurs atomistes correspondent à la situation inverse, lorsque des inférences sur le niveau supérieur sont faites à partir d'analyses portant sur le niveau inférieur.

Plus généralement, l'analyse de régression multiniveaux permet de modéliser très efficacement les sources de variance en rapport avec les facteurs présents dans une situation (fixes et aléatoires). Elle permet de plus de prendre en compte la covariance de facteurs non orthogonaux dans l'estimation d'un effet fixe. Elle offre ainsi au chercheur la possibilité de disposer d'un schéma précis de causalité dans la situation qu'il analyse (effets simples, interactions simples et complexes), au sein d'un même niveau et entre niveaux, entre les variables indépendantes et la ou les variables dépendantes. Face à des situations complexes, l'analyse de régression multiniveaux permet donc d'élaborer des modèles mathématiques plus vraisemblables qu'avec d'autres techniques d'analyse comme l'ANOVA ou les modèles de régression linéaire classiques (à un seul niveau), et donc de faire parler les données de manière beaucoup plus approfondie et exacte.

La méthode est fiable. L'évaluation d'un facteur porte sur l'ensemble des données, mais le calcul de la puissance ne porte que sur le nombre exact de valeurs concernées (diminution des risques d'erreur de type I et II), tandis que la quantité de données nécessaire, même pour analyser une situation déjà complexe, n'est pas exorbitante (contrairement à d'autres techniques comme la modélisation par équations structurales).

En dépit des avantages offerts par cette technique, l'exploration des rapports entre facteurs dans les données se limite aux relations linéaires entre variables. Or, les phénomènes psychologiques pourraient être basés sur des rapports non linéaires entre facteurs.

### 3.3.4.2 Quelques détails sur son utilisation

Le logiciel MLwiN (v1.1) a été utilisé pour élaborer les modèles d'analyse et tester la significativité des variables. Il s'agit d'un produit développé au Centre for Multilevel

Modelling©<sup>91</sup> (Institute of Education, London, UK), pour le système d'exploitation Microsoft® Windows®. Les données utilisées dans les modèles sont :

- ▷ Variable(s) dépendante(s) : données filtrées selon la procédure présentée à la section 3.3.2.1.2 (filtre à  $\pm 3$  écarts-type par item + par sujet) ; valeurs manquantes codées et non remplacées
- ▷ Prédicteurs : variables indépendantes, autres facteurs à contrôler (« covariables ») et interactions de prédicteurs ; prédicteurs catégoriels : codage numérique des catégories ; prédicteurs continus : données centrées sur la moyenne<sup>92</sup>.

Remarque :

Le codage adopté pour les variables indépendantes catégorielles à 2 modalités (modalité contrôle / modalité expérimentale pour le facteur) est au format 0 / 1, et non - 1 / 1.

La lecture des résultats se fait différemment selon le codage adopté, mais l'interprétation finale reste la même. Le premier format implique une lecture des résultats essentiellement basée sur un contraste de conditions, l'une d'elles étant considérée comme un contrôle. Cette procédure est préférée dans certaines disciplines (exemple : médecine). Dans une situation expérimentale classique en psycholinguistique (2 facteurs A et B à 2 modalités chacun (contrôle / test), croisement complet), chaque condition se traduirait en équation de la manière suivante :

- ▷  $Y_{00} = \mu + \varepsilon$  : condition contrôle – effet obtenu étant donné non-A et non-B
- ▷  $Y_{10} = \mu + A + \varepsilon$  : condition contrôle B, test A – effet de A étant donné non-B
- ▷  $Y_{01} = \mu + B + \varepsilon$  : condition contrôle A, test B – effet de B étant donné non-A
- ▷  $Y_{11} = \mu + A + B + A.B + \varepsilon$  : condition test A et B – effet de A, de B, et de leur interaction

*Avec  $\mu$  désignant la moyenne générale et  $\varepsilon$  le terme d'erreur.*

Les effets principaux des facteurs A et B et leur interaction seraient testés à partir des coefficients fournis dans l'équation du modèle. La procédure revient à comparer :

- ▷ la condition  $Y_{10}$  (coefficient du terme A) à la condition contrôle  $Y_{00}$  (effet de A).
- ▷ la condition  $Y_{01}$  (coefficient du terme B) à  $Y_{00}$  (effet de B).
- ▷ la somme  $Y_{00} + Y_{11}$  à la somme  $Y_{10} + Y_{01}$  (effet de l'interaction entre A et B).

---

<sup>91</sup> <http://multilevel.ioe.ac.uk/index.html>

<sup>92</sup> Centrer chaque prédicteur sur sa moyenne permet d'obtenir un effet estimé dans la dimension de la variable dépendante (facilite l'interprétation).

Le second codage implique une lecture des résultats basée sur un contraste de moyennes de conditions (procédure de type ANOVA). La situation expérimentale précédente s'écrirait comme suit en équations :

- ▷  $Y_{00} = \mu - A - B + A.B + \varepsilon$  : condition contrôle
- ▷  $Y_{10} = \mu + A - B - A.B + \varepsilon$  : condition contrôle B, test A
- ▷  $Y_{01} = \mu - A + B - A.B + \varepsilon$  : condition contrôle A, test B
- ▷  $Y_{11} = \mu + A + B + A.B + \varepsilon$  : condition test A et B,

et les effets seraient testés à partir de la comparaison des quotients :

- ▷  $(Y_{10} + Y_{11}) / 2$  et  $(Y_{01} + Y_{00}) / 2$  (effet de A).
- ▷  $(Y_{01} + Y_{11}) / 2$  et  $(Y_{10} + Y_{00}) / 2$  (effet de B).
- ▷  $(Y_{00} + Y_{11}) / 2$  et  $(Y_{10} + Y_{01}) / 2$  (effet de l'interaction entre A et B).

### 3.3.4.2.1 Le « modèle vide »

Les données sont d'abord testées par un modèle vide. Son but est d'estimer la nécessité d'une structure multiniveaux dans l'analyse, le modèle vide permettant de calculer un indice de corrélation intra classe (CIC ou  $\rho$ ), qui renseigne sur l'importance de la hiérarchisation des données. Dans ce travail, la structure comporte 2 niveaux (sujets, items).

Équation 4 : Calcul du degré de corrélation intra classe

$$\rho = \frac{\sigma_{u0}^2}{\sigma_{u0}^2 + \sigma_{e0}^2}$$

Avec  $\sigma_{u0}^2$  la variance due au niveau 2, et  $\sigma_{e0}^2$  la variance due au niveau 1

On considère qu'avec une corrélation intra classe supérieure à 10%, il est nécessaire d'avoir recours à un modèle multiniveaux pour éviter les erreurs décrites à la Section 3.3.4.1.

Équation 5 : Équation du modèle vide (exemple issu des données de l'Expérience 4)

$$TR_{It, Suj} \sim N(XB, \Omega)$$

$$TR_{It, Suj} = \beta_{0It, Suj} \text{Cons}$$

$$\beta_{0It, Suj} = 568.389(8.167) + u_{0Suj} + e_{0It, Suj}$$

$$[u_{0Suj}] \sim N(0, \Omega_u) : \Omega_u = [2448.324(596.498)]$$

$$[e_{0It, Suj}] \sim N(0, \Omega_e) : \Omega_e = [11476.330(357.591)]$$

$$-2 * \text{loglikelihood(IGLS Deviance)} = 25690.400(2100 \text{ of } 2400 \text{ cases in use})$$

« TR » correspond à la VD (temps de réponse), « It » et « Suj » correspondent respectivement aux niveau 1 et au niveau 2 de la structure. La distribution des TR est supposée normale (N), de moyenne  $X\beta$  et de variance  $\Omega$ .  $\beta_0$  est le coefficient correspondant à la moyenne générale de la distribution (« intercept ») estimée par le modèle (= 568.389), accompagnée de l'erreur standard de l'estimation donnée entre parenthèses (= 8.167 – de même pour chaque facteur qui sera inclus dans le modèle). « Cons » correspond à une constante nécessaire au calcul de la référence, «  $u_0$  » et «  $e_0$  » sont des termes d'erreur (variance aléatoire, respectivement au niveau des sujets et au niveau des items, supposés être distribués normalement (N), de moyenne 0 et de variance respective  $\Omega_u$  et  $\Omega_e$ ). « IGLS Deviance » est un score d'adéquation du modèle aux données, accompagné entre parenthèses du nombre d'items dans l'échantillon.

$$\rho = (2448.324) / (2448.324 + 11476.33) = 0.1758$$

Près d'un cinquième de la variance (17.58%) est liée au niveau 2 (sujets), et il est donc nécessaire de modéliser la structure multiniveaux des données dans l'analyse.

### 3.3.4.2.2 Construction du modèle d'analyse

Le modèle est élaboré :

- ▷ sur la base de la structure hiérarchique générale, déterminée à l'aide du modèle vide
- ▷ par l'adjonction de prédicteurs (facteurs, interaction de facteurs)
- ▷ par la spécification du niveau de variance à tester pour chaque prédicteur (effets fixes seuls (aucun niveau de variance spécifié), effets aléatoires de niveau 1, effets aléatoires de niveau 2, etc., interactions entre niveaux)

Le choix de la structure du modèle peut se faire selon deux procédures :

- ▷ guidée par la théorie. La structuration du modèle avance par l'adjonction d'éléments, justifiés par :
  - ∟ des bases théoriques
  - ∟ l'adéquation meilleure du modèle au jeu de données considéré (estimation mathématique)
  - ∟ la possibilité de parvenir à une estimation avec la nouvelle structure (convergence du modèle, qui dépend de la quantité de données et de dimensions dans les données, par rapport à la complexité du modèle).
- ▷ guidée par les données. On part d'un modèle complet (ou estimé tel, c'est à dire renfermant tous facteurs supposés avoir eu une influence sur les données mesurées), où toutes les relations entre facteurs, au sein d'un niveau et entre

niveaux, sont spécifiées. On restreint graduellement le modèle sur la base des 3 éléments précédents de justification. Cette deuxième méthode offre l'avantage de trouver plus aisément le modèle le mieux adapté aux données, sans risque d'oublier d'effets lors de la modélisation. Elle permet aussi et surtout de découvrir des effets insoupçonnés, qui auraient été omis avec la procédure précédente.

La construction d'un modèle nécessite donc un travail de recherche de la meilleure architecture, guidé par la pertinence théorique, la plausibilité, et la faisabilité.

Dans ce travail, la plupart des modèles ont été simplifiés en ne considérant que les effets fixes des prédicteurs (effets liés au traitement). Les effets aléatoires (liés aux sujets et aux items) sont exceptionnellement modélisés. Lorsqu'ils le sont, c'est uniquement dans un souci de précision du modèle (aucune justification théorique ; le seul intérêt est de les distinguer des effets fixes pour préciser d'avantage ces derniers – parfois cependant, le degré de précision supplémentaire est insignifiant), et à condition que le calcul puisse converger avec tous les effets aléatoires dans le modèle (en l'absence de justification théorique, comment choisir d'en modéliser certains et d'en rejeter d'autres ?).

### **3.3.4.2.3 Adéquation du modèle aux données**

L'adéquation du modèle est évaluée par une procédure qui permet de comparer deux architectures, afin de déterminer laquelle est préférable. La sélection d'une architecture se fait théoriquement selon un double critère :

- ▷ quantité de variance totale estimée par le modèle
- ▷ parcimonie (quantité de paramètres nécessaire pour expliquer une même quantité de variance)

Théoriquement, le meilleur modèle est celui qui explique le plus parcimonieusement le plus de variance dans les données. Pratiquement, des aspects théoriques modulent cette base.

### **3.3.4.2.4 Test des effets dans le modèle**

Le modèle, à sa convergence, fournit une estimation statistique des régularités linéaires qui existent entre les prédicteurs et la variable dépendante. C'est à partir de cette estimation que l'influence de chaque prédicteur sur les données est quantifiée dans la dimension de la variable dépendante, et que des valeurs moyennes par condition peuvent être calculées (somme des coefficients des facteurs). Il ne s'agit donc pas des

valeurs moyennes réelles ou mesurées par condition<sup>93</sup>, mais des valeurs moyennes estimées par le modèle sous l'effet de chaque prédicteur considéré (elles peuvent donc varier par rapport aux données réelles, et d'un modèle à un autre, en fonction des facteurs utilisés dans la modélisation).

La significativité des effets fixes et des contrastes (écarts entre conditions) est évaluée avec le logiciel MLwiN à l'aide d'un test de  $\chi^2$  (degré de liberté, N observations), comparant la distribution des valeurs estimées sous l'influence d'un facteur à la distribution des valeurs brutes.

Les résultats renseignés pour chaque expérience correspondent donc aux moyennes estimées par condition pour les effets fixes des facteurs étudiés, aux tests de significativité par  $\chi^2$  des effets fixes des prédicteurs et des contrastes entre conditions, et à la quantité de variance expliquée par chaque facteur expérimental (quantité totale de variance résiduelle dans un modèle sans le facteur moins quantité totale de variance résiduelle du modèle avec le facteur)<sup>94</sup>. Les moyennes brutes par condition sont également fournies au lecteur, à titre d'information.

### 3.3.4.2.5 Représentation graphique des résultats

Lorsqu'un graphique représente l'effet d'un facteur (1 condition = 1 modalité du facteur), 3 types d'information y figurent :

- ▷ La valeur moyenne estimée par condition (effets fixes)
- ▷ Les barres de significativité statistique (valeurs de l'erreur de l'estimation, directement fournies par le modèle pour le facteur)
- ▷ La valeur moyenne brute ou mesurée par condition

Lorsqu'un graphique représente l'effet simultané de plusieurs facteurs (1 condition résulte de la combinatoire de plus d'un facteur), 4 types d'information sont représentés :

- ▷ La valeur moyenne estimée par condition (effets fixes)

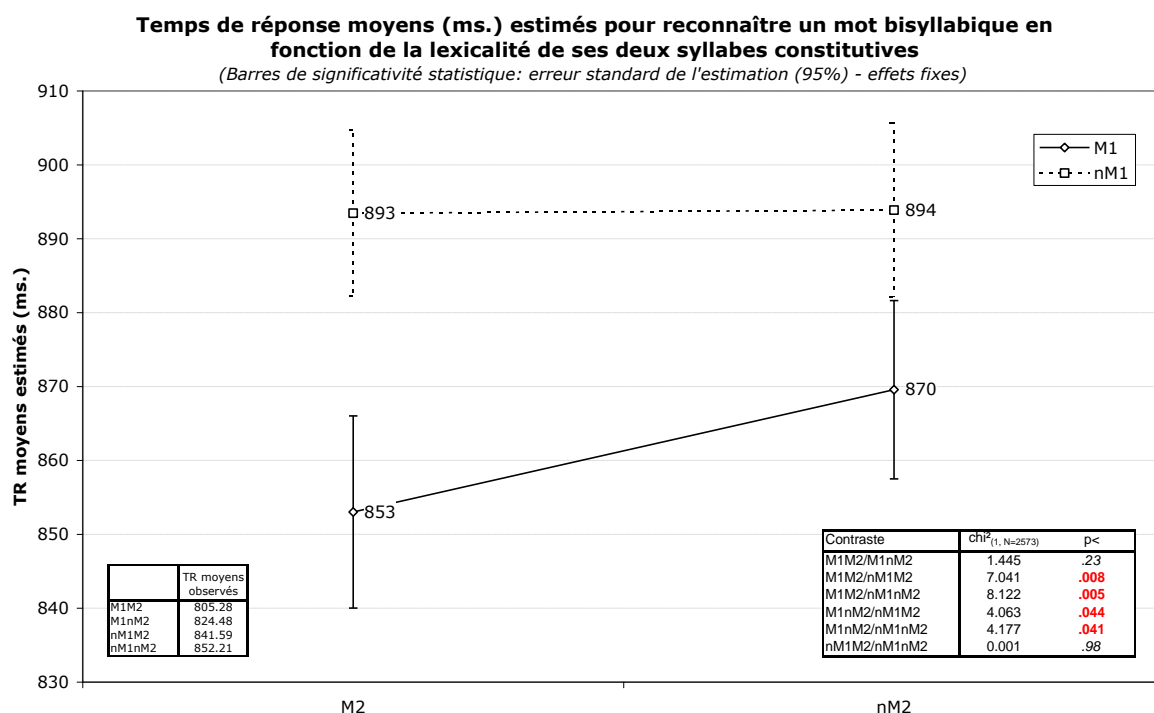
---

<sup>93</sup> Ces moyennes réelles ou mesurées par condition confondent les effets de tous les prédicteurs. L'utilisation des moyennes réelles ne serait recevable que dans une situation où aucune autre source de variation que la variable indépendante n'est présente. Dans le présent travail, ce n'est jamais le cas et la méthode évite donc des biais. Cependant, ces moyennes brutes sont fournies au lecteur à titre indicatif.

<sup>94</sup> Remarque : les % de variance expliquée par les facteurs sont souvent faibles. Ceci est vrai aussi avec des facteurs considérés comme influençant fortement le comportement (exemple : la fréquence d'occurrence d'un mot explique 0.008% de la variance totale dans le modèle étudié pour l'Expérience 7, et l'âge d'acquisition en explique 0.21% (même modèle)). Ce point est lié à la quantité élevée de sources de variance qui contribuent à la variance totale dans une mesure comportementale.

- ▷ Les barres de significativité statistique (calculées selon la procédure décrite plus bas)
- ▷ Le tableau des résultats statistiques présentant l'évaluation des contrastes entre conditions
- ▷ La valeur moyenne brute ou mesurée par condition (TR moyens observés)

Figure 31 : Exemple



Les barres de significativité statistique correspondent à un artifice cosmétique, statistiquement faux, dont le seul but est de fournir une aide pratique à la lecture des graphiques, basée sur l'erreur standard de l'estimation par le modèle de l'effet des facteurs étudiés. Elles permettent de visualiser une information directement interprétable en termes de significativité de la différence entre 2 conditions (si 2 barres d'erreur se touchent ou se chevauchent, l'écart des moyennes entre conditions comparées n'est pas significatif au seuil  $p = .05$ ). Il ne s'agit donc pas seulement d'une information relative à la dispersion des données autour de la moyenne (barres de déviation standard, barres d'erreur standard (déviation standard /  $\sqrt{N}$ )), mais d'une information prenant aussi en considération un seuil de significativité. Comme il est parfois difficile d'estimer visuellement si le seuil de significativité est dépassé, d'une part, et que l'information ainsi graphée est la traduction inexacte de résultats statistiques, d'autre part, le tableau des contrastes est fourni comme référence (valeurs statistiques exactes).

Procédure de calcul des barres de significativité statistique :

Pour chaque contraste de conditions, la valeur absolue de la différence estimée par le modèle entre les 2 moyennes ( $|C_1 - C_2|$ ) et un  $\chi^2$  pour cet écart sont calculés (procédure d'estimation du  $\chi^2$  sous MLwiN telle que décrite précédemment). Grâce à ces valeurs, l'erreur standard de l'estimation est calculée pour un intervalle de confiance de 95%, selon la formule :

$$E.S. = \frac{|C_1 - C_2|}{\sqrt{\chi^2}} * 1.96 / 2$$

où 1.96 est un coefficient qui détermine le seuil de l'intervalle de confiance, et 2 un coefficient correctif lié à la symétrie de la distribution normale.

Une valeur est ainsi obtenue pour chaque contraste. Étant donné qu'il est impossible de faire figurer graphiquement ces barres d'erreur dans une représentation simultanée de toutes les conditions (pour une même condition, on aurait une barre d'erreur différente pour chaque contraste), c'est la valeur moyenne des barres d'erreur par condition qui est représentée. À cause de cela, l'information de significativité qu'elles véhiculent est déformée et peut parfois contredire les données statistiques présentées dans le tableau. Un abaissement de la valeur du coefficient utilisé pour déterminer le seuil de l'intervalle de confiance dans l'équation du calcul de l'erreur standard peut permettre de résoudre cette incohérence (nouvelle valeur à utiliser dans tous les contrastes).

La méthode, bien qu'incorrecte statistiquement, est jugée comme efficace et adaptée au but visé, puisqu'elle a permis de représenter correctement dans les graphiques la significativité des écarts entre conditions pour toutes les expériences de l'étude.

### 3.3.5 Conclusions

La technique de régression multiniveaux est particulièrement intéressante pour ce travail, parce qu'elle permet de résoudre tous les points problématiques énumérés précédemment (structure hiérarchique des données, présence de plusieurs facteurs aléatoires, valeurs manquantes et plans non équilibrés, limitation avec les analyses de covariance). Elle a permis :

- ▷ De traiter des données présentant plusieurs niveaux de variance et de considérer simultanément plusieurs sources aléatoires de variance dans les données.



- ▷ De traiter directement des plans incomplets (valeurs manquantes, plans déséquilibrés) sans avoir recours à des procédés compliqués d'équilibrage.
- ▷ D'incorporer facilement au modèle (à sa création ou a posteriori) de nouveaux prédicteurs, indépendamment de leur rapport aux unités d'un niveau ou d'un autre (sujets ou items). Ceci est utile pour un contrôle statistique des facteurs parasites impossibles à éliminer (« covariables »), liés à la situation expérimentale, aux sujets ou aux items.
- ▷ De modéliser facilement des relations complexes entre niveaux et au sein d'un même niveau.
- ▷ De procéder a posteriori à des méta analyses (par exemple, portant sur les données de plusieurs expériences).
- ▷ D'utiliser aussi bien des données continues (modèles linéaires) que discrètes (modèles non linéaires de type « logit »).

Par conséquent, cette seule technique offre un grand spectre d'utilisation, qui permet d'apporter une souplesse certaine face aux contraintes multiples imposées par l'objet d'étude et les conditions de la recherche.

### 3.3.6 Résumé récapitulatif

Tableau 16 : Récapitulation des problèmes / solutions liés aux analyses statistiques

Opération	Problèmes	Solutions	Éviter
<b>Tri des données</b>	Peut induire des biais lorsque la méthode de filtrage choisie ne se limite pas à l'élimination des valeurs non normales	Éliminer les sujets anormaux, puis Éliminer les items anormaux, puis Éliminer les réponses anormales par item et/ou par sujet	Le filtrage en aveugle sur la distribution générale ou sur la distribution par condition
<b>Gestion des valeurs manquantes</b>	Peut induire des biais selon la méthode de remplacement choisie	Déterminer si la valeur manquante peut être remplacée (due au hasard) ou non (liée à un traitement) Si oui, remplacer par une valeur estimée ou utiliser une technique capable de travailler sur des données incomplètes (analyse multiniveaux)	Le remplacement des valeurs manquantes liées à un traitement et le remplacement des valeurs manquantes par une valeur moyenne

Opération	Problèmes	Solutions	Éviter
<b>Analyse statistique des données</b>	L'ANOVA est difficilement utilisable avec des données ayant une structure hiérarchique et présentant plus d'un facteur aléatoire L'utilisation de « covariables » liées aux sujets ou aux items n'est pas possible dans une ANOVA portant sur l'analyse d'une situation avec plus d'un facteur aléatoire	Choisir la technique d'ANOVA adaptée à la structure des données et utiliser des quotients « quasi-F » au lieu du test F classique, ou utiliser une autre technique que l'ANOVA (analyse multiniveaux)	L'utilisation routinière d'une technique d'analyse (s'assurer qu'elle est applicable dans les conditions de l'étude et qu'elle permet bien d'effectuer ce que l'on souhaite) L'utilisation de la recette de double validation des effets fixes, par $F_1$ et $F_2$

### 3.4 Chapitre 4 - Méthodologie expérimentale

Ce chapitre décrit tous les aspects de la procédure technique développée et utilisée dans l'expérimentation, et justifie les raisons des choix adoptés. Il expose les points communs aux différentes expériences afin d'éviter une redondance lors de la présentation de chaque expérience (seuls les points spécifiques seront alors détaillés).

Le Chapitre 4 est organisé en 5 sections, chacune relative :

- ▷ À la création du matériel expérimental
- ▷ Aux tâches utilisées
- ▷ À la population testée
- ▷ Au recueil des données
- ▷ Aux analyses effectuées sur les données

Il se termine par une sixième section récapitulant dans deux tableaux synoptiques tous les aspects techniques et théoriques développés dans l'expérimentation. Cette sixième section permet à la fois de synthétiser l'information présentée jusqu'alors et d'introduire la partie expérimentale (Partie IV).

#### 3.4.1 Création du matériel expérimental

Elle se déroule en 4 étapes : obtention des items linguistiques, enregistrement, création des stimuli, analyse acoustique des stimuli. Ces 4 étapes sont détaillées ci-après.

##### 3.4.1.1 Obtention des items linguistiques

La plupart des expériences nécessitent 2 types d'items :

- ▷ des items expérimentaux
- ▷ des items de remplissage

Selon les expériences, ces 2 catégories d'items peuvent être :

- ▷ des mots
- ▷ des pseudo-mots (séquence de phonèmes de la langue considérée, agencés de manière à respecter les contraintes phonotactiques de la langue sans pour autant former un mot existant)
- ▷ des non-mots (séquence de phonèmes de la langue considérée, agencés sans respecter les contraintes phonotactiques de la langue)

Leur obtention est décrite dans les 2 sous-sections suivantes.

### 3.4.1.1.1 Sélection des mots

La sélection des mots se fait en 3 étapes :

- ▷ extraction, à partir d'une base de données lexicale, d'un sous-ensemble de mots ayant les caractéristiques souhaitées, déterminées à partir des facteurs codés dans la base de données
- ▷ élimination manuelle de tous les mots de ce sous-ensemble qui possèdent des caractéristiques linguistiques et psychologiques indésirables, relativement à des facteurs qui ne sont pas codés dans la base de données (décrits plus bas)
- ▷ sélection manuelle des mots dans le sous-ensemble filtré, de manière à obtenir des conditions aussi homogènes que possible

Ces 3 étapes sont détaillées ci-dessous. Les deux premières visent à cibler le contenu de la base de données au plus près des besoins, afin de permettre, à la dernière étape, de se concentrer plus facilement sur l'homogénéité du matériel.

#### 3.4.1.1.1.1 Première étape : sélection d'un sous-ensemble dans une base de données lexicale

La base de données qui a été utilisée est BRULEX. La sélection dans cette base de données est faite selon les critères suivants :

- ▷ catégorie grammaticale « nom »
- ▷ nombre de syllabes « 1 ou 2 » selon l'expérience
- ▷ structure syllabique souhaitée (selon l'expérience. Toutes les syllabes utilisées commencent par une consonne)
- ▷ réduction des ambiguïtés (contrôle de l'homophonie ou contrôle de l'homogénéité de ce facteur à travers les conditions ou aux analyses statistiques)
- ▷ détermination de la position du point d'unicité phonologique du mot (pré ou post-final ; si pré-final, à une position particulière (exemple : dernier phonème))
- ▷ contrôle de l'absence de schwa et de gémination

#### 3.4.1.1.1.2 Deuxième étape : filtrage manuel des mots

Élimination des :

- ▷ mots pour lesquels l'information de fréquence n'est pas disponible dans BRULEX (exemple : « *pétanque* »)
- ▷ mots préfixés et pseudo-préfixés (exemple : « *préfixe* », « *préfet* »)
- ▷ mots composés (exemple : « *grand-père* »)
- ▷ mots argots ou populaires (exemple : « *clébard* »), insultants (exemple : « *crétin* ») ou pouvant l'être par leur ambiguïté (exemple : « *cochon* »), etc.

- ▷ mots d'origine étrangère (exemple : « *bowling* »)
- ▷ mots sexuellement connotés (exemple : « *capote* ») ou à enchâssements sexuellement connotés (exemple : « *confesse* »)
- ▷ mots malaisés à syllaber (exemple : « *tab/lée* » - « *ta/blée* »)
- ▷ mots ayant deux syllabes phonologiquement identiques (exemple : « *barbare* »)
- ▷ noms ressemblant à des adverbes ou à des adjectifs (exemple : « *petit* »)
- ▷ noms pouvant être des prénoms (exemple : « *clémence* »)

### **3.4.1.1.1.3 Troisième étape : sélection manuelle des items à partir du sous-ensemble filtré**

La sélection est faite de manière à ce que :

- ▷ les items soient suffisamment employés dans la population pour être connus de tous les participants testés
- ▷ les items satisfassent aux facteurs manipulés dans l'étude :
  - ∟ présence d'un enchâssement initial (E/nE – [plas] renferme le mot [pla])
  - ∟ enchâssabilité de l'item à l'initiale d'un mot plus long (I/nI – [plas] peut être le début de [plastik])
  - ∟ lexicalité d'une syllabe à une position précise dans un mot (M/nM – la première syllabe de [plastik] est un mot ([plas]))<sup>95</sup>
  - ∟ fréquence d'une syllabe dans le lexique (Rare / Fréquente (R/F))
  - ∟ fréquence positionnelle d'occurrence d'une syllabe dans les mots
    - à l'initiale (R/F)
    - en position non initiale (R/F)
- ▷ les items ne s'amorcent pas mutuellement :
  - ∟ relations sémantiques (exemple : « *vache* » et « *champ* ») et morphologiques (exemple : « *graine* » et « *grainetier* ») entre mots
  - ∟ relations sémantiques entre mots enchâssés et mots porteurs (exemple : « *lance* » dans « *balance* », en rapport sémantique avec « *bataille* »)
- ▷ qu'une même syllabe, lorsqu'elle existe dans un item expérimental, n'existe pas deux fois dans le matériel de l'expérience
- ▷ les items forment des conditions en moyenne appariées, pour autant de facteurs non expérimentaux que possible :
  - ∟ fréquence formelle phonologique d'occurrence
  - ∟ nombre de phonèmes
  - ∟ position du point d'unicité phonologique.

---

<sup>95</sup> La sélection des syllabes, ou des mots en fonction de leurs syllabes constitutives, de même que la création des pseudo-mots en fonction des caractéristiques des syllabes, se fait à l'aide de deux bases de données syllabiques issue de l'analyse de BRULEX, l'une produite par Jeremy Goslin, l'autre par l'auteur (non publiées).

D'autres facteurs spécifiques peuvent contraindre la sélection, en fonction des expériences (exemple : les mots bisyllabiques dont la deuxième syllabe est un mot présentent un point d'unicité identique à celui du mot enchâssé en deuxième syllabe). Ces spécificités sont détaillées lors de la présentation de chaque expérience.

### **3.4.1.1.2 Choix et création des pseudo-mots et des non-mots**

Les pseudo-mots utilisés dans les expériences sont de deux types :

- ▷ des syllabes du Français : items expérimentaux sélectionnés à partir des bases de données syllabiques obtenues par analyse de BRULEX (J. Goslin ; C. Lachaud)
- ▷ des syllabes et bisyllabes artificiels : items distracteurs créés pour permettre la réalisation des tâches de décision lexicale, de manière à respecter la phonotactique de la langue française, à partir de rien ou à partir de mots du Français selon les besoins de l'expérience.

Les non-mots utilisés sont des syllabes créées, qui n'existent pas en Français selon BRULEX. Une seule expérience (Expérience 12) utilise des non-mots.

#### **3.4.1.1.2.1 Sélection de syllabes pseudo-mot du Français**

Elle est faite en fonction des facteurs :

- ▷ nombre de phonèmes
- ▷ structure phonétique de la syllabe
- ▷ fréquence d'occurrence
- ▷ nombre de mots renfermant la syllabe
- ▷ position d'occurrence
- ▷ fréquence positionnelle d'occurrence
- ▷ nombre de mots renfermant la syllabe à une position donnée

#### **3.4.1.1.2.2 Création de pseudo-mots**

Les règles de création dépendent de ce qui est testé dans la situation expérimentale. D'une manière générale, les pseudo-mots ne doivent pas interférer avec le traitement des items expérimentaux. Il fallait donc être particulièrement vigilant à ce que les pseudo-mots :

- ▷ n'évoquent pas directement un mot expérimental (lien phonologique. Exemple : /teɪfəfɔm/ / « *téléphone* »)
- ▷ n'évoquent pas indirectement un mot expérimental (lien sémantique). Exemples :

- ∩ le pseudo-mot évoque par sa ressemblance phonologique un mot en lien sémantique avec un mot expérimental (/teʁefɔ̃m/ évoque « *téléphone* », en lien sémantique avec « *parole* »)
- ∩ le pseudo-mot renferme un mot enchâssé relié sémantiquement à un mot expérimental (le pseudo-mot /plasmuk/ renferme le mot « *place* » en lien sémantique avec le mot « *espace* »)

#### 3.4.1.1.2.3 Création des non-mots

Les non-mots sont créés de manière à :

- ▷ évoquer le moins possible de mots du Français (testé par enquête auprès de quelques étudiants). Il est cependant quasi-impossible d'avoir un non-mot qui n'évoque rien (le système réaliserait l'appariement le moins mauvais sur la base de la sonorité de l'item, et trouverait toujours quelque chose à appairer)
- ▷ être appariés aux syllabes pseudo-mot (syllabes du Français) utilisées dans les autres conditions pour respecter une structure phonétique commune de la syllabe

#### 3.4.1.1.3 **Remarques concernant la base de données lexicales utilisée pour la création et le contrôle du matériel expérimental**

La base de données BRULEX, utilisée comme référence dans cette étude, est parfois critiquée sur trois points :

- ▷ quantité de mots : limitée à 35746 entrées, essentiellement à cause de l'absence des formes conjuguées des verbes. La position du point d'unicité théorique (PU de chaque mot) qui est fournie par cet outil (calculée sur ce corpus de 35746 mots) serait donc une approximation du PU théorique réel. Il en serait de même pour l'évaluation des fréquences d'occurrence syllabiques calculées à partir de BRULEX.
- ▷ représentativité de l'usage réel et contemporain du Français : les informations de fréquence sont issues de l'analyse (Centre de Recherche pour un Trésor de la Langue Française (Imbs, 1971)) de corpus littéraires (romans, essais, recueils de poèmes, œuvres dramatiques) publiés entre 1919 et 1964. Il ne s'agit donc pas d'une fréquence réelle d'usage quotidien et contemporain par l'adulte moyen.
- ▷ limite du nombre de facteurs renseignés

Depuis 2001, une nouvelle base de données lexicales pour le Français contemporain, LEXIQUE (New et al., 2001), est disponible aux chercheurs. Résolvant les critiques précédentes, LEXIQUE apporte un progrès certain à ce type d'outils. Cependant, la base de données BRULEX a été utilisée pour les raisons suivantes :

- ▷ à l'époque de sa publication, LEXIQUE contenait beaucoup trop d'erreurs pour apparaître comme un outil crédible et professionnel, et inciter au changement de base de référence. Par contre, BRULEX, utilisée dans de nombreuses études psycholinguistiques sur le Français, avait acquis le statut d'un standard

- ▷ le travail avait été commencé avec BRULEX. Il était difficile de justifier d'un changement de référence en cours de recherche :
  - ∩ qui aurait induit une hétérogénéité dans l'étude
  - ∩ alors que BRULEX est en adéquation suffisante avec les besoins du travail (et puisque son contenu représente un vocabulaire 3 fois supérieur au vocabulaire de l'adulte moyen, selon Content et col. (1990)).

Les remarques suivantes permettent d'appuyer cette position :

- Position du PU : les positions du PU données par BRULEX et LEXIQUE s'écartent probablement de la position réelle pour un mot et un individu donnés, du fait qu'il s'agit d'une position théorique qui dépend du lexique analysé, alors que la position réelle du PU d'un mot dépend du lexique connu par un individu (impossible à déterminer). La présence de formes conjuguées dans un corpus décale les PU vers la fin des mots, par rapport au PU calculé sur un corpus qui ne renferme pas de formes conjuguées. Il reste cependant à quantifier cet écart et à savoir si psychologiquement, les verbes conjugués sont traités comme leurs lemmes (cette question fait toujours l'objet de recherches en morphologie), avant de se prononcer sur les bienfaits expérimentaux de cette précision nouvelle. Ce point est donc considéré comme non crucial pour justifier un changement de référence.
- Fréquence d'occurrence et la représentativité de l'usage contemporain des mots : le travail développé dans l'Annexe 1 page 423 a montré que l'influence de la fréquence d'usage d'un mot n'est pas un prédicteur fiable du temps nécessaire à la reconnaissance du mot. L'effet de la fréquence disparaît lorsque l'âge d'acquisition estimé et la familiarité conceptuelle subjective d'un mot sont intégrés au modèle d'analyse, conjointement à la fréquence d'occurrence (corrélations et interactions). La fréquence aurait une influence prédictive parce qu'elle explique un peu de l'âge d'acquisition et un peu de la familiarité conceptuelle. L'âge d'acquisition et la familiarité conceptuelle ont été utilisés dans les analyses statistiques en tant que covariables, conjointement à la fréquence phonologique d'occurrence calculée à partir des fréquences fournies par BRULEX. Cette solution offre un bien meilleur contrôle que l'utilisation de fréquences plus contemporaines, telles que celles fournies dans LEXIQUE
- Fréquence phonologique d'occurrence d'une syllabe : elle dépend de 2 facteurs : le nombre de mots dans lequel la syllabe se trouve et la fréquence d'usage de chacun de ces mots. Le premier facteur, à cause de sa nature structurale, pourrait avoir une implication importante sur l'activation lexicale. Il est vrai que pour ce facteur, BRULEX conduit peut-être à une évaluation approximative du lexique (inventaire des syllabes peut-être incomplet ?). Toutefois, quelle est la proportion de syllabes manquantes, en dehors des syllabes correspondant à des affixes ? Peut-elle être si importante que BRULEX ne permette pas une approximation suffisante ? Il est vrai aussi que l'estimation du nombre de mots renfermant une syllabe donnée est biaisée s'il manque des formes conjuguées dans cette comptabilisation (cf. remarques précédentes concernant le traitement de ces formes et leur représentation en mémoire). Concernant le second facteur (fréquence d'usage du mot), il est vrai que la fréquence d'occurrence des mots dans BRULEX est



certainement moins représentative de l'usage contemporain moyen du Français que celle de LEXIQUE. Cependant, il faut considérer l'emploi qui a été fait de la fréquence phonologique des syllabes : opposer des syllabes rares à des syllabes fréquentes. Cela est certainement possible avec l'approximation donnée par BRULEX. Même si des différences existent entre BRULEX et LEXIQUE (mots nouveaux ou ayant un usage nouveau, employés fréquemment aujourd'hui alors qu'ils ne l'étaient pas dans la littérature entre 1919 et 1964 ; mots tombés en désuétude alors qu'ils étaient d'un usage plus fréquent ; lexique différent), rien n'autorise à dire qu'elles sont suffisantes pour inverser les caractéristiques d'une condition portant sur un ensemble d'items. Si des changements existent dans l'utilisation des mots d'une langue, des points demeurent constants parce que les mots sont ancrés dans la réalité, au sujet de laquelle ils servent à communiquer.

### 3.4.1.2 Enregistrement et traitements numériques

Une fois les items linguistiques sélectionnés / créés en fonction des caractéristiques voulues, ils doivent prendre corps dans le monde physique pour devenir les stimuli délivrés aux sujets lors des tests expérimentaux. Cette opération nécessite un locuteur, du matériel d'enregistrement, et une procédure d'enregistrement, trois points qui sont détaillés ci-dessous.

#### 3.4.1.2.1 Le locuteur

- ▷ Une locutrice, naïve quant aux buts de l'expérience, a été choisie pour lire les items de la partie sur la prosodie (phrases construites, sémantiquement structurées en courts textes), sur la base de la clarté et de l'expressivité de sa voix. Sa consigne était de prononcer les phrases de la manière la plus naturelle et vivante possible. Sa naïveté était un point essentiel pour la création du matériel utilisé dans cette partie du travail, puisque la relation entre les facteurs topologiques / morphosyntaxiques et la prosodie, devait, si elle existait, émerger naturellement.
- ▷ La naïveté du locuteur n'était pas un point théorique essentiel pour l'enregistrement des items utilisés dans les expériences sur l'activation lexicale. Cet aspect apparaissait même insignifiant en regard des impératifs techniques pour produire des items ayant les caractéristiques acoustiques et prosodiques précisément définies, et surtout, homogènes.

Faire appel à une personne extérieure et inexpérimentée est très coûteux en temps et en énergie :

- ∞ obtenir des productions ayant les caractéristiques désirées nécessite un temps d'enseignement puis d'entraînement
- ∞ il faut parfois procéder à plusieurs sessions d'enregistrement, lorsque le matériel produit contient des défauts ou n'est pas homogène. Agencer de nouvelles sessions prend du temps, sans garantie de la disposition du locuteur ni du résultat de l'enregistrement
- ∞ d'une expérience à l'autre, le locuteur peut être différent, introduisant un paramètre aléatoire qui limite la possibilité de comparer les résultats à travers les différentes étapes de l'étude.

Pour ces raisons, le locuteur utilisé a été l'auteur du travail. Les avantages de ce choix sont considérables :

- ∩ disponibilité à volonté
- ∩ connaissance des caractéristiques acoustiques souhaitées du matériel
- ∩ connaissance de la procédure de production et d'enregistrement
- ∩ homogénéité des corpus à travers l'étude

L'inconvénient, qui peut-être discuté, est que ce locuteur ne dispose pas toujours de l'accent local. Ceci peut avoir interféré dans la reconnaissance de certains items par certains sujets. Il faut cependant préciser qu'à Genève, l'accent est peu marqué dans la population du fait de la proximité de la France, du fait qu'une communauté française importante vit et travaille sur le canton, et du fait que beaucoup d'étudiants viennent de l'étranger et en particulier de la France.

#### **3.4.1.2.2 Le matériel d'enregistrement**

Le Laboratoire Parole et Langage (Aix-en-Provence) et le Laboratoire de Psycholinguistique Expérimentale (Genève) mettent à disposition des chercheurs les outils techniques pour produire un matériel expérimental irréprochable sur le plan acoustique :

- ▷ local insonorisé (chambre sourde à Aix-en-Provence, caisson insonorisé à Genève)
- ▷ micro de qualité professionnelle (SHURE Bêta 58A)
- ▷ enregistreur numérique DAT (Sony ZA5ES)
- ▷ dispositif de transfert de l'information DAT → ordinateur, sous un format numérique (carte son (Turtle Beach) munie de connecteurs numériques (Aix-en-Provence), module son externe (Creative Sound Blaster « Extigy ») muni de connecteurs numériques et optiques (Genève))
- ▷ logiciels d'édition du son (Sound Forge® (Sonic Foundry, 2000) et Cool Edit Pro (Syntrillium Software Corporation, 2003))
- ▷ logiciels d'analyse acoustique (Praat / Windows® (Boersma & Weenink, 2000), MES Signaux / UNIX (Espesser, 1994) et ESPS / UNIX (Entropic Research Laboratory, 1998))

#### **3.4.1.2.3 La procédure d'enregistrement**

La procédure suivie est standard. Elle est détaillée ci-dessous.

##### **3.4.1.2.3.1 Organisation de la liste d'items à lire**

Le but de la structuration du matériel à lire est de favoriser une production homogène à travers les conditions expérimentales.

- ▷ Les mots sont successivement sélectionnés un par un dans chacune des conditions expérimentales, pour former une liste de lecture. Le processus est réitéré jusqu'à épuisement des items. Ceci permet de répartir les modifications imperceptibles de diction, qui existent entre le début et la fin d'un enregistrement (rythme, mélodie, intensité, timbre de voix), de façon homogène sur l'ensemble des conditions.
- ▷ Les items monosyllabiques sont présentés séparément des items bisyllabiques. Ceci permet l'adoption plus aisée par le locuteur d'un profil prosodique constant.

#### **3.4.1.2.3.2 Installation du locuteur et réglages du volume**

Le but est ici d'assurer une qualité optimale de la prise de son (volume adapté, absence de bruit dans le signal, etc.).

- ▷ Le locuteur est installé dans la pièce insonorisée.
- ▷ La position du micro et la distance entre sa bouche et le micro est vérifiée (éviter l'enregistrement de souffles). Elle doit être constante pendant toute la durée de l'enregistrement pour limiter les variations d'intensité du signal
- ▷ Une phase de test sert à déterminer les niveaux optimaux d'enregistrement, et à « échauffer » le locuteur

#### **3.4.1.2.3.3 Production, enregistrement et transfert des données**

- ▷ Lecture des items de l'étude sur la prosodie :
  - ∩ présentation sur écran du groupe de phrases relatives à un item (phrase d'introduction sémantique, phrase porteuse)
  - ∩ lecture vivante et modulée (donner l'impression que les phrases sont formulées spontanément lors d'une discussion)
- ▷ Lecture des items des études sur la reconnaissance des mots parlés :
  - ∩ présentation sur papier
  - ∩ le locuteur est attentif à produire les items avec un débit constant et une intonation identique
  - ∩ le rythme de production est lent : chaque item est lu comme s'il était seul, c'est-à-dire suivi d'une pause assez longue. Ceci est voulu afin d'éviter une prosodie de liste
  - ∩ pour disposer d'un corpus à l'acoustique homogène, la liste est lue deux fois dans la foulée. Les items utilisés dans les expériences sont sélectionnés à partir du milieu de la première lecture jusqu'au milieu de la deuxième lecture, cette zone d'enregistrement s'avérant être en général la plus homogène (le locuteur a trouvé son rythme et n'est pas encore lassé). Une troisième lecture est faite après une pause, afin de disposer de 3 enregistrements d'un même item, au cas où la production serait bruitée ou non utilisable sur la zone sélectionnée pour une raison ou une autre (souffle, intensité trop forte, bruit de mouvement, bruit de bouche, etc.)
  - ∩ lorsque le locuteur se rend compte qu'un item est défectueux, il le répète immédiatement
- ▷ Enregistrement sur DAT, échantillonnage à 44100 Hz

- ▷ Transfert des données par connexion numérique, du DAT à l'ordinateur, où le corpus est enregistré (logiciels d'édition du son) dans un fichier son au format PCM du standard Windows® (.wav ; échantillonnage à 44100 Hz, 16 bits, stéréo. Le fichier subit immédiatement un ré-échantillonnage : passage en format mono ; passage en 16000 Hz pour le corpus utilisé dans la partie prosodie pour des raisons techniques relatives à la resynthèse – tous les autres corpus sont laissés en 44100 Hz (qualité sonore meilleure à l'écoute)).

### 3.4.1.3 Création des stimuli

#### 3.4.1.3.1 Stimuli en voix naturelle

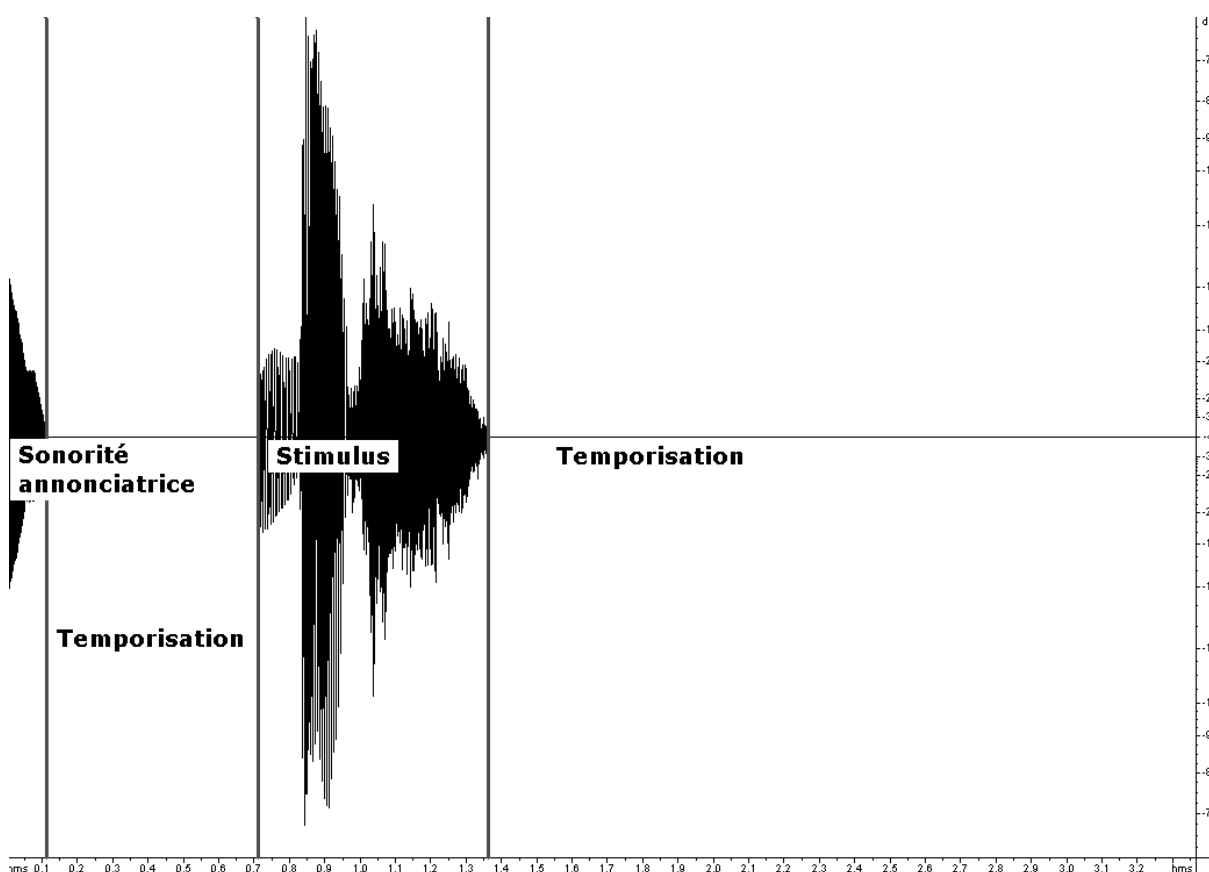
Une fois le corpus disponible sur ordinateur au format audionumérique voulu, la phase de création des stimuli commence. Elle se décompose en 4 étapes détaillées ci-dessous : nettoyage, sélection, marquage, découpage. Au terme de cette procédure, on dispose d'autant de fichiers son que le corpus renferme d'items.

- ▷ Nettoyage
  - ∩ le corpus est écouté attentivement
  - ∩ à l'aide d'un logiciel d'édition du son, il est débarrassé :
    - des portions inutiles (pauses longues, début et fin de session, toux, bruits divers)
    - des items posant un problème (parasitage, variabilité acoustique, etc.).
- ▷ Sélection
  - ∩ la version de chaque item qui sera utilisée dans l'expérience est sélectionnée dans le fichier nettoyé, de manière à ce que l'ensemble soit le plus homogène possible (intensité, mélodie, débit, articulation)
  - ∩ les sélections sont conservées tandis que le reste du fichier est effacé
- ▷ Marquage
  - ∩ chaque item contenu dans le fichier son résultant de l'étape de sélection est marqué manuellement, à l'aide du logiciel Praat.
  - ∩ cette phase de marquage consiste en deux procédures :
    - localisation précise des frontières initiale et finale de chaque item (début et fin de l'oscillation acoustique correspondant à l'item)
    - labellisation de chaque séquence sonore ainsi délimitée
- ▷ Découpage
  - ∩ le fichier son marqué et labellisé est découpé en autant de fichiers son qu'il renferme d'items. La procédure est effectuée automatiquement par un script (C. Lachaud), sous Praat.
  - ∩ ce script fonctionne de la manière suivante :
    - sélection séquentielle de chaque item dans le fichier son d'entrée, selon les informations contenues dans le fichier texte associé (position des frontières des items et label de la zone délimitée)
    - extraction de l'item sélectionné

- création de chaque stimulus par la concaténation des éléments suivants, dans cet ordre :
  - sonorité annonçant la présentation de l’item (« bip », 100 ms.)
  - 500 ms. sans signal (temporisation)
  - item extrait
  - 2 secondes sans signal (temporisation pour laisser au sujet le temps de répondre)
- écrit le bloc obtenu dans un fichier .wav, sous le nom « label » donné lors du marquage de la portion correspondant à l’item extrait
- passe à l’item suivant. Continue jusqu’à ce qu’il ne reste plus d’items à traiter.

Le résultat est un fichier son dont l’architecture est présentée dans la Figure 32.

Figure 32 : Architecture générale de la séquence acoustique utilisée comme stimulus



### 3.4.1.3.2 Stimuli particuliers

Les expériences de la partie relative à la prosodie utilisent toutes des stimuli ayant un format particulier :

- ▷ items numériques (Simulations 1 et 2)
- ▷ items désémantisés par synthèse (Expérience 1)

- ▷ items déségmentés par substitution du signal avec des trains d'impulsions (Expérience 2)
- ▷ items en parole de synthèse (Expérience 3)

Deux expériences, les Expériences 9 et 10, relatives à l'étude des superpositions lexicales, utilisent des stimuli particuliers également (Items en voix naturelle dans lesquels des cibles sonores non linguistiques ont été insérées).

Tous ces items particuliers ont été obtenus par des procédures partiellement ou totalement différentes de celles décrites précédemment. Elles sont présentées ci-dessous.

- ▷ Items numériques :

La phase de découpage est remplacée par une phase d'analyse de la  $f_0$  sur les portions de phrases marquées comme items (MES-signaux, scripts R. Espesser). En sortie, une liste de valeurs numériques décrivant chaque item est obtenue (1 valeur de  $f_0$  toutes les 10 ms., durée de chaque syllabe).

- ▷ Items désémantisés par synthèse :

Cette méthode permet de changer l'information segmentale tout en préservant les caractéristiques prosodiques (durée et  $f_0$ ) de la parole d'origine. Les syllabes marquées comme items dans les phrases du corpus sont synthétisées avec les phonèmes [m] et [a] en utilisant la mélodie et la durée d'origine (MES-signaux, scripts R. Espesser). L'opération est effectuée à partir de la base de diphtongues du système de synthèse MBROLA (Dutoit, Pagel, Pierret, Bataille, & Van der Vreken, 1996), et des données de  $f_0$  et de durée syllabique (obtenues lors de la création des items numériques) appliquées aux diphtongues grâce à l'algorithme PSOLA<sup>96</sup> ( $f_0$  extrapolée sur les portions de parole non voisées). Les items obtenus sont analysés pour vérification par comparaison de leurs caractéristiques prosodiques avec celles des items d'origine (évaluation des écarts entre courbes mélodiques : l'écart moyen n'excède pas 1.1% sur l'ensemble du corpus et est artefactuel<sup>97</sup>). La qualité de la reproduction des caractéristiques mélodiques du signal original dans le signal synthétisé avoisine donc les 100%, excepté sur les portions extrapolées. Les items (bisyllabes /mama/) sont ensuite extraits automatiquement (coupure de l'onde sonore selon la localisation des étiquettes, avec atténuation de la pente de la sinusoïde (30dB sur 10 ms., au début et à la fin de l'item), pour éviter la perception d'un « clic »), et enregistrés sous forme de fichiers son (nomenclature : [type de frontière][n° phrase d'origine][.wav] – exemple : FD5.wav).

- ▷ Items déségmentés par substitution du signal avec des trains d'impulsions :

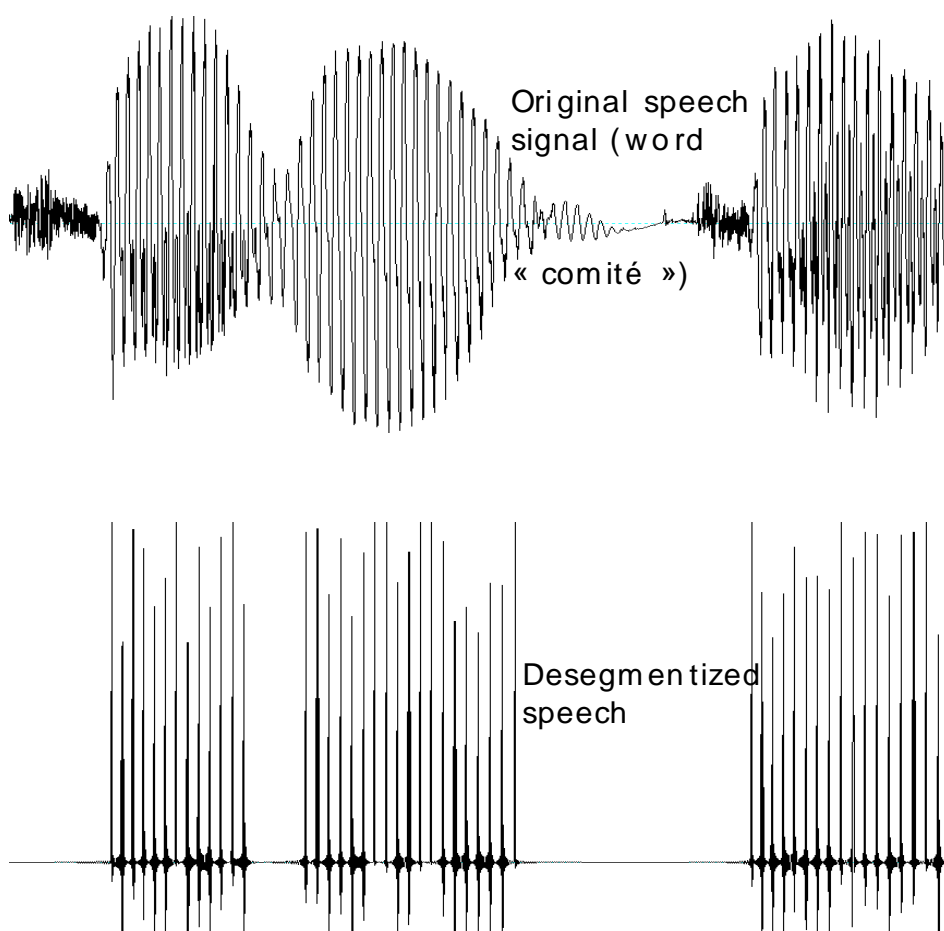
---

<sup>96</sup> PSOLA : Pitch Synchronous Overlap Add Method (algorithme permettant une modulation de la  $f_0$  dans le signal).

<sup>97</sup> Cette différence artefactuelle est liée au comportement des algorithmes de détection de la  $f_0$  aux frontières des zones non voisées (les courbes de  $f_0$  réelles présentent une transition rapide mais pas instantanée, impliquant une zone d'incertitude pour la localisation du début et de la fin du voisement. Les courbes de  $f_0$  extrapolées sont continues, d'où des écarts locaux pouvant atteindre 47%).

Cette méthode permet d'éliminer l'information segmentale tout en préservant la sensation mélodique, accentuelle et rythmique de la parole d'origine (sonne comme un bourdonnement d'insecte, modulé en mélodie et en rythme). Elle consiste à remplacer les portions voisées du signal (noyau vocalique et consonnes voisées) par des trains d'impulsions, chaque impulsion marquant le début d'une période dans l'oscillation. Après localisation manuelle et étiquetage (sonagramme) des futurs items dans les phrases (portions de 3 syllabes), ils sont extraits dans un fichier son (1 par fichier – nomenclature : [condition\_][séquence des 3 syllabes][.wav]. exemple : FD1\_tauxsuccu.wav). Un algorithme de détection du voisement localise les périodes contenues dans le signal, pour une zone de fréquence à explorer déterminée par l'utilisateur, dont il marque le début par une impulsion. Après vérification automatique de la détection (des sauts d'octave se produisent parfois sur les portions voisées périphériques), un nouveau signal est synthétisé (le signal d'origine est éliminé, seul demeure le train d'impulsions), et écrit dans un fichier .wav (Praat, scripts C. Lachaud).

Exemple :



▷ Items en parole de synthèse :

La démonstration interactive du système de synthèse vocale « Text-to-Speech » Elan Tempo™ (<http://www.elanspeech.com/demos/interactive.html>), a été utilisée pour générer les fichiers .wav du matériel expérimental (voix : « Robert » ; vitesse : lent / rapide ; hauteur : bas / très haut ; format acoustique : 16000 Hz, mono, 16 bits). Ces fichiers ont été utilisés à des fins

d'évaluation non commerciale, respectant ainsi la charte d'utilisation de ce site internet.

- ▷ Items renfermant des cibles sonores non linguistiques :

Un script rédigé pour le logiciel Praat insère la cible sonore dans les fichiers son correspondant aux stimuli déjà construits (il s'agit des stimuli des Expériences 4 et 7), à une position définie automatiquement au cours du processus d'insertion (position relative à la durée de l'item pour l'Expérience 9, calculée par le script), ou à une position définie manuellement par l'expérimentateur par une série de mesures acoustiques et d'évaluations perceptives (Expérience 10). Le procédé itératif aboutit à l'écriture d'autant de fichiers son que déterminé par le (nombre de stimuli d'origine \* nombre de positions \* nombre de types de cibles).

#### **3.4.1.4 Analyse acoustique des stimuli**

Cette étape vise à se doter de données acoustiques décrivant en moyenne les items produits, afin de disposer des possibilités de contrôle d'éventuelles hétérogénéités du matériel lors des analyses statistiques (covariables).

Il se peut en effet que malgré toutes les précautions prises, l'homogénéité acoustique voulue à travers les conditions ne soit pas au rendez-vous, et que les effets mesurés soient attribués à un traitement psychologique alors qu'ils sont de nature sensorielle. Le seul moyen de contrôle avec de la voix naturelle est d'intégrer les facteurs acoustiques dans les modèles d'analyse.

Les mesures portent sur :

- ▷ la durée (ms.) de l'item (monosyllabe) ou de ses syllabes constitutives (bisyllabes)
- ▷ la f0 moyenne par syllabe (Hz)
- ▷ l'intensité moyenne par syllabe (dB)

Elles sont réalisées automatiquement à l'aide d'un script (C. Lachaud), sous Praat, au cours de la phase de découpage lors de la genèse des stimuli.

#### **3.4.2 Tâches utilisées**

7 tâches expérimentales ont été utilisées dans l'étude :

- ▷ des tâches qualitatives sans contrainte de temps dans la partie relative à la prosodie (sauf Expérience 3) : discrimination d'items, détection et localisation de frontière
- ▷ des tâches relatives au paradigme de la chronométrie mentale dans le reste de l'étude (détection de mots et de non- / pseudo-mots (TDL go, no go), de syllabes (TDL go, no go ; word spotting + TDL go, no go), de clics, classification (TDL)).



### 3.4.2.1 Présentation des tâches qualitatives et de leur but

- ▷ Discrimination (comparaison) :

Deux items sont présentés auditivement l'un après l'autre au participant, qui doit décider s'ils sont assez similaires pour être rangés dans une même catégorie, ou s'ils doivent être rangés dans deux catégories différentes (Expérience 1).

Le but est de faire organiser le matériel expérimental à l'insu des participants, sur la base d'une évaluation perceptive des items, grâce à une tâche simple et faisable par des sujets naïfs (une tâche de catégorisation consciente serait infaisable).

- ▷ Détection de frontière :

L'auditeur entend une séquence de 3 syllabes successives, extraite d'une phrase puis déségmentée. Il doit décider si la séquence contient ou non une frontière lexicale (Expérience 2)

- ▷ Localisation de frontière :

Dans la même séquence que pour la tâche précédente, l'auditeur doit localiser la position de la frontière lexicale s'il en a détecté une (Expérience 2)

Le but des tâches de détection et de localisation de frontière est de tester la capacité consciente des auditeurs à détecter et localiser une frontière lexicale en fonction de facteurs prosodiques liés à la position des syllabes dans un mot et au type de mot dont elles sont extraites.

### 3.4.2.2 Présentation des tâches chronométriques et de leur but

- ▷ Détection de mots ou de non- / pseudo-mots (tâche de décision lexicale dite en « go, no go »)

On présente à l'auditeur, un par un, des mots et des non- / pseudo-mots. Sa tâche est de presser sur une touche dès qu'il reconnaît qu'un item est un mot, ou dès qu'il reconnaît qu'un item n'est pas un mot, selon ce qu'il lui a été spécifié de faire. Il doit produire sa réponse sous la double contrainte de rapidité et d'exactitude (Expériences 4, 5, 6, 7, 8, 12, 13, 14, 15)

Le but est de sonder les processus d'accès au lexique (mots) ou d'interférence lexicale (pseudo-mots). Pour mesurer les processus d'activation lexicale, la tâche de décision lexicale en « go, no go » est supposée être plus appropriée que la tâche de décision lexicale classique (classification). Moins sensible aux processus post-lexicaux du fait qu'elle est plus simple, et donc, effectuée plus rapidement, elle permettrait d'obtenir des données plus précises lorsqu'elle est combinée avec une méthode de filtrage adaptée (Perea, Rosa, & Consolacion, 2002).

- ▷ Détection de syllabes :

Le participant entend un item bisyllabique et doit faire une tâche de détection comparable à la précédente sur la dernière syllabe de l'item. Il lui faut donc procéder à une méta analyse de l'item (Expériences 11 et 16). Cette tâche s'apparente à du « word spotting » dans le cas de l'Expérience 16 (syllabes lexicales).

Le but est de sonder le processus de segmentation lexicale d'une séquence de syllabes, en fonction de facteurs lexicaux et acoustiques (prosodie).

## ▷ Détection de clics :

L'auditeur entend des items (mots ou pseudo-mots) qui renferment une sonorité (cible ou distractrice) pouvant apparaître à n'importe quel moment (avant le début du mot ou du pseudo-mot, à différentes positions définies dans le mot ou le pseudo-mot, après la fin du mot ou du pseudo-mot). Leur tâche est d'appuyer sur un bouton le plus vite possible et en faisant le moins d'erreurs possible dès que la sonorité cible est entendue, sans attendre la fin du mot ou du pseudo-mot, et d'inhiber leur réponse motrice si la sonorité distractrice est entendue à la place de la sonorité cible (tâche en « go, no go » - Expériences 9 et 10).

Le but est de sonder la charge de calcul dans le système à différentes étapes de l'activation lexicale, en mesurant l'interférence produite par le traitement lexical, irrépressible, avec la tâche demandée.

## ▷ Classification (tâche de décision lexicale classique) :

L'auditeur entend un par un des mots et des pseudo-mots. Il dispose de 2 touches à presser pour fournir sa réponse, l'une dédiée aux mots et l'autre aux pseudo-mots. Il procède sous la double contrainte de rapidité et d'exactitude (Expériences 3 et 7).

Le but de cette tâche est de sonder les processus d'activation lexicale à un niveau de traitement plus avancé que celui qui est mesuré avec la tâche de décision lexicale en « go, no go » (la tâche de classification étant plus complexe que la tâche de détection, elle prend plus de temps pour être réalisée). Cette tâche a été utilisée dans l'Expérience 7 (comparaison avec la TDL go, no go à des fins méthodologiques), et dans l'Expérience 3 (avant que l'on ait réfléchi à ce problème de méthode).

Remarque :

Les tâches chronométriques se prêtent habituellement à une double analyse, des temps de réponse et des erreurs. Il a cependant été décidé de ne pas procéder à l'analyse et à l'interprétation des erreurs, la justification étant la suivante :

- ▷ L'origine des erreurs peut être multiple (traitement décisionnel lié à la tâche, traitement lexical, inattention et fatigue, etc.), mais comme il est impossible de faire une distinction entre ces différents cas, on ignore quelle est précisément l'information véhiculée par cette variable.
- ▷ Si les erreurs ne véhiculent pas d'information supplémentaire par rapport aux temps de réponse, l'analyse des TR suffit. Par contre, si les erreurs apportent une information supplémentaire, elles doivent être analysées avec les temps de réponse dans un modèle à 2 variables dépendantes (erreur + TR). La mise en œuvre de cette technique est trop lourde compte tenu des enjeux et des incertitudes.

Les taux d'erreur par condition sont cependant renseignés pour chaque expérience à la section « Méthode d'analyse », au point relatif aux valeurs manquantes (pourcentage de valeurs manquantes par condition ou « taux d'erreur »).

### **3.4.3 Population testée et obtention des participants**

Les participants aux expériences sont majoritairement des étudiants de psychologie (1<sup>er</sup>, 2<sup>ème</sup> et 3<sup>ème</sup> cycle), des universités d'Aix-Marseille (France) et de Genève (Suisse). Ils sont sélectionnés selon les critères suivants :

- ▷ francophone
- ▷ monolingue
- ▷ sans troubles cognitifs connus
- ▷ sans troubles auditifs ou visuels, en fonction de la modalité sensorielle utilisée pour délivrer les stimuli (correction de la vue acceptée dans l'Expérience 8)

#### **3.4.3.1 Obtention des participants à Aix-en-Provence**

Ils sont patiemment chassés un par un dans les couloirs de l'université, et rendus coopérants grâce à de subtiles techniques d'asservissement.

#### **3.4.3.2 Obtention des participants à Genève**

Un système a été mis en place pour le recrutement et la gestion des participants aux expériences du Laboratoire de Psycholinguistique Expérimentale à la rentrée universitaire 2002-2003 (conception : Bagou O., Frauenfelder U. et Lachaud C.). Le but de ce système est de libérer les chercheurs de contraintes matérielles idiotes, auxquelles il est possible de remédier efficacement par de l'organisation. La conséquence est un gain de temps, d'énergie, et de productivité. Le recrutement s'effectue grâce à un formulaire Internet, conçu par P. Demierre, C. Lachaud et F. Serena (développement : Demierre et Serena (2002)). Une base de données des participants est alors générée. Elle contient des informations sur le profil linguistique des inscrits, ainsi que leurs coordonnées électroniques. Elle est utilisée pour contacter les étudiants de façon ciblée tout au long de l'année, et sélectionner des populations linguistiques spécifiques en fonction des besoins de la recherche (exemple : étude sur les bilingues français / italien).

### **3.4.4 Recueil des données**

#### **3.4.4.1 Données comportementales**

##### **3.4.4.1.1 Systèmes utilisés pour le recueil des données à Aix-en-Provence**

Les expériences sont pilotées par ordinateur. Elles se déroulent à l'université de Provence, dans une pièce isolée et calme, sur une station Intel® fonctionnant sous MS-DOS® (processeur cadencé à 166 MHz, 96 Mo de mémoire vive, carte son Creative

SoundBlaster, vitesse du disque dur à 5200 tours par minute). Elles sont gérées grâce à un script rédigé par le Pr. J. Pynte (présentation des stimuli, enregistrement des réponses, des temps de réponse, de l'ordre de présentation).

Les items sont présentés auditivement au participant, à l'aide d'un casque de studio. Le recueil des réponses se fait par l'intermédiaire du clavier. Cette interface est acceptable pour l'Expérience 1, dans laquelle la variable dépendante n'a aucune dimension chronométrique. Elle pose par contre un problème de précision de la mesure avec l'Expérience 4, que le Pr. J. Pynte a réduit autant que possible :

- ▷ en faisant utiliser la touche « echap » pour répondre (gérée différemment au niveau de la mémoire tampon du clavier. Permet de réduire l'incertitude de la mesure)
- ▷ par l'emploi d'une tâche ne nécessitant pas l'utilisation de plusieurs touches (décision lexicale en « go, no go »).

Cette solution diminue l'imprécision mais ne l'élimine pas : des effets psychologiques ont pu être noyés dans la variance introduite par l'outil de mesure. L'Expérience 4 a cependant été incluse dans ce rapport, car :

- ▷ il s'agit d'une expérience importante qui est à la base des développements ultérieurs de l'étude sur le rôle des mots enchâssés dans l'activation lexicale
- ▷ les effets qu'elle met en évidence sont retrouvés dans les expériences réalisées à Genève, dans des conditions techniques pour lesquelles le biais dû à la résolution du système de recueil ne se pose plus. Les résultats de l'Expérience 4 sont donc peut-être entachés d'une imprécision, mais ils ne sont pas faux ni douteux, ce qui a pu être déterminé grâce à cette réplication a posteriori.

#### **3.4.4.1.2 Systèmes utilisés pour le recueil des données à Genève**

Les expériences sont pilotées par ordinateur. Les passations se déroulent à l'université de Genève, au laboratoire, parallèlement sur deux PC Dell® Dimension™ XPS T500 (le laboratoire dispose de deux caissons d'expérimentation et offre la possibilité de faire passer les participants par 2 - individuellement et en parallèle), strictement identiques (même éléments matériels : processeur Intel® PIII cadencé à 500MHz, 256 Mo de mémoire vive, vitesse du disque dur à 7200 trs./min, carte son soundblaster (Creative SB Live !), moniteur couleur Dell® UltraScan P990 ; clonage logiciel : système d'exploitation Windows98® 2<sup>ème</sup> édition<sup>98</sup>, logiciel d'expérimentation E-prime™ v1.1 SP2 (Psychology Software Tools, 2002)). Ces stations sont exclusivement dédiées à l'expérimentation. Elles ont été testées pour connaître leur précision temporelle lors

---

<sup>98</sup> Les premières versions de E-Prime ne fonctionnaient pas sous un système Windows® NT.

d'une simulation d'expérience chronométrique en modalité auditive, où « l'oreille du sujet » est une clé vocale (boucle directe : production – récupération du signal à la clé vocale - script E-prime™ : M. Grzeskowiak, FPSE, Université de Genève). Leur précision, mesurée sur 100 essais, est de l'ordre de 3 à 4 ms. (moyennes respectives des stations : 2.96 et 3.68 ms. ; écart-types respectifs : 0.49 et 0.79. Les différences entre les deux stations sont négligeables et d'origine inconnue (différences de composants ?)).

Les expériences sont pilotées par des scripts (C. Lachaud) rédigés pour E-prime™ (enregistrement des caractéristiques des sujets (n° de code, profil linguistique, sexe, latéralité) et de la condition passée (groupe), présentation des consignes, entraînement, présentation des items (enregistrement dans le fichier réponse du nom de l'item, de l'ordre d'origine des items dans la liste expérimentale, de l'ordre de présentation des items au participant), enregistrement des réponses des participants et des temps mis pour fournir une réponse, mesurés depuis le début de présentation de chaque item).

Les stimuli sont présentés au participant à l'aide d'un casque de studio (Beyerdynamic DT100).

Les réponses sont recueillies grâce au boîtier fourni avec E-prime™, qui autorise une précision publicitaire de l'ordre de la milliseconde.

Afin de limiter les risques de biais temporels liés au fonctionnement du système d'exploitation, les items sont pré-chargés en mémoire vive et le nombre de processus actifs sur la machine est limité au minimum possible :

- ▷ stations non raccordées physiquement au réseau
- ▷ couche logicielle du réseau non installée
- ▷ pas de logiciel antivirus
- ▷ présence des seuls logiciels nécessaires à l'expérimentation

#### **3.4.4.1.3 Description d'une séance-type à Genève**

Le participant est installé dans un caisson insonorisé contenant un casque audio et un boîtier de réponse à 5 touches. Après vérification de la mise en veille de son téléphone cellulaire, il lit la présentation de l'expérience, qui lui est donnée sur le moniteur de la station expérimentale, à travers la vitre du caisson. Il utilise le boîtier de réponse qui se trouve devant lui pour faire défiler les écrans de consigne.

Après avoir pris connaissance des consignes, il récapitule à l'expérimentateur les points qui lui ont été exposés. Le but de cette procédure est de s'assurer que les aspects

essentiels du protocole soient compris. Lorsque ce n'est pas le cas, l'expérimentateur peut corriger les représentations du participant.

Un entraînement a ensuite lieu. Il porte sur une quinzaine d'items dédiés à la familiarisation du participant avec la situation et la tâche. Ils sont présentés auditivement (au casque) ou visuellement (à l'écran), selon la modalité sensorielle utilisée dans l'expérience. Le participant fournit ses réponses en appuyant sur la ou les touches actives du boîtier.

À la fin de l'entraînement, le participant peut demander des clarifications à l'expérimentateur, et éventuellement refaire l'entraînement. Sinon, l'expérience commence. Les tests durent entre 5 et 10 minutes en moyenne, et se déroulent sans pause. Pour des tests plus longs, la passation se fait par blocs d'items pour réduire la baisse des performances due à la fatigue.

À Aix-en-Provence, les séances se sont déroulées selon un protocole technique similaire.

#### **3.4.4.2 Données non comportementales**

Les analyses statistiques nécessitent d'avoir recours à un ensemble de données dont les origines sont diverses, relatives aux caractéristiques des items (physiques, psychologiques, etc.) ou des participants.

##### **3.4.4.2.1 Données relatives aux items**

- ▷ Caractéristiques physiques
  - ∩ Données fournies par les analyses acoustiques du matériel expérimental (durée,  $f_0$  et intensité moyennes)
  - ∩ Données issues d'analyses secondaires sur ces données acoustiques (indice  $dtw$ , calculé par P. Courrieu)
- ▷ Caractéristiques distributionnelles et linguistiques

Données obtenues à partir des bases de données lexicales (BRULEX essentiellement) ou de leur analyse (analyse de BRULEX par Lachaud, par Goslin)
- ▷ Caractéristiques psychologiques des items
  - ∩ Déduites des bases de données lexicales
  - ∩ Obtenues par enquête sur Internet auprès d'une population francophone (voir Annexe 1 page 423)

##### **3.4.4.2.2 Données relatives aux participants**

- ▷ Caractéristiques physiques

Questionnaire d'inscription, passation de l'expérience (latéralité manuelle et auditive, sexe, âge, santé (prise de médicaments, troubles moteurs, organes sensoriels défectueux, etc.))

▷ Caractéristiques linguistiques

Questionnaire d'inscription (origine linguistique et langues parlées, niveau de maîtrise pour chaque langue connue, etc.)

### 3.4.5 Analyse des données

Les données brutes sont préalablement filtrées, éventuellement recodées, et organisées :

- ▷ Filtrage : élimination des réponses qui s'écartent trop du centre de la distribution normale (« outliers »). Se reporter à la section 3.3.2 de cette partie pour plus de détails.
- ▷ Codage : certaines réponses sont de nature catégorielle et non numérique. Elles doivent donc subir un codage selon un format utilisable par le logiciel d'analyse (exemple : réponse même / différent). Les valeurs manquantes (non-réponses et valeurs non normales filtrées à l'étape précédente) doivent être codées également.
- ▷ Organisation : structuration des données en tables selon un format propre au logiciel d'analyse.

Les techniques d'analyse statistique suivantes ont été utilisées à travers cette étude :

- ▷ ANOVA (Statistica™, (Statsoft, 2001)) - Simulation 1, Expérience 5 (analyse acoustique des stimuli)
- ▷ analyse de régression linéaire simple et multiple, bootstrap (Gauss™, (Aptech Systems, 1998) – Scripts : Y. Bourgeois) - données de l'Expérience 1, en Annexe 11 page 482.
- ▷ analyse des effets fixes (effets principaux et interactions) par régression multiniveaux. Tests de significativité des effets et des contrastes par test du Chi<sup>2</sup> (MLwiN (Rasbash, Browne, Healy, Cameron, & Charlton, 2001)) - Toutes les expériences. Pour les Expériences 1 et 2 : utilisation de modèles logistiques uni et multi-variés. Dans la présentation des résultats :
  - ∩ Les valeurs moyennes estimées par le modèle de régression sont fournies (effet fixe du facteur).
  - ∩ Les valeurs mesurées (moyennes par condition) accompagnent les valeurs moyennes estimées
  - ∩ Les tests de significativité sont joints (effets fixes des facteurs, tests de contrastes entre conditions)
  - ∩ Le pourcentage de variance expliquée est également présenté

Le détail sur ces méthodes est disponible à la Section 3.3.4 page 123, ainsi que dans les manuels spécialisés (théorie, logiciels) et les articles donnés en bibliographie.

### **3.5 Chapitre 5 - Plan synthétique récapitulant les aspects techniques et présentant les aspects théoriques pour chaque expérience**

Les deux tableaux synoptiques suivants présentent un plan synthétique des aspects théoriques et techniques développés à travers les expériences effectuées dans ce travail de thèse.

Leur but est de récapituler ce qui vient d'être exposé tout en introduisant la partie expérimentale. Ce plan synthétique aide à se représenter l'organisation du travail et à se situer au sein de l'organisation lors de la lecture du manuscrit.

Le premier tableau concerne les aspects théoriques sous-jacents à chaque expérience (buts des expériences et enchaînement logique des idées, facteurs manipulés).

Le second tableau offre une récapitulation des aspects techniques mis en œuvre pour chaque expérience (matériel expérimental, tâche, population, recueil des données, analyses).



Tableau 17 : Aspects théoriques sous-jacents aux expériences

Nom de l'expérience	Facteurs manipulés	Conditions expérimentales	Buts	Enchaînement logique	Partie
Analyse acoustique	Catégorie syntaxique du mot porteur (Lexical L/Grammatical G); Nombre de syllabes du mot porteur (1/3); Position de la syllabe dans le mot porteur (Début D/Fin F)	4 catégories de syllabes: [G (mot 1 syl)], [L (mot 1 syl)], [F (mot 3 syl)], [D (mot 3 syl)]	Rechercher l'existence de régularités suprasegmentales au niveau de la syllabe, en lien avec leurs caractéristiques lexico-syntaxiques et topologiques	Etude de faisabilité première étape: vérification du matériel expérimental. Les caractéristiques acoustiques des syllabes (prosodie) sont bien en correspondance avec des facteurs topologiques et lexico-syntaxiques	Influence de la prosodie sur le guidage sensoriel
Simulation 1 (classification)	En plus des facteurs précédents (Analyse acoustique), l'association des types de syllabes (G/L/F*D/G/L) pour déterminer le type de frontière (bisyllabe)	9 catégories de frontières: [GD], [LD], [FD], [GG], [LG], [FG], [GL], [LL], [FL]	Préciser l'analyse précédente: montrer l'existence de formes mélodiques au niveau de la syllabe, permettant de distinguer 9 types de frontières		
Simulation 2 (discrimination)	En plus des facteurs précédents (Simulation 1), l'association des types de bisyllabes pour déterminer le type de comparaison 2 à 2: ((GD)[LD][FD][GG][LG][FG][GL][LL][FL])*((GD)[LD][FD][GG][LG][FG][GL][LL][FL])	81 types de comparaisons 2 à 2 entre ces 9 catégories de frontières	Vérifier si les catégories théoriques de frontières peuvent être retrouvées à partir de l'information prosodique, par comparaison d'items 2 à 2	Etude de faisabilité deuxième étape: vérification de l'adéquation de la tâche. Elle permet bien de faire ce qui en est attendu	
Expérience 1			Montrer que les auditeurs humains organisent le matériel, perceptivement et inconsciemment, selon une procédure identique (comparaison), à partir de la prosodie	Validation, première étape: les auditeurs humains sont capables de distinguer perceptivement les catégories théoriques de frontières à partir de la prosodie	
Expérience 2	Position de la frontière dans l'item trisyllabique (hors de la séquence de syllabes, entre les deux premières/dernières syllabes); Type de frontière (entre (F/L/G), et D <sub>2</sub> ou entre F <sub>2</sub> et (D/L/G) <sub>3</sub> )	3 conditions: [Pas de frontière]; [Une frontière entre S <sub>1</sub> et S <sub>2</sub> ]; [Une frontière entre S <sub>2</sub> et S <sub>3</sub> ]	Montrer que les auditeurs humains détectent et localisent consciemment une frontière lexicale à partir de l'information prosodique seule	Validation, deuxième étape: les auditeurs humains sont capables de détecter et de localiser consciemment les frontières lexicales à partir de la prosodie	
Expérience 3	Durée de l'item (débit Rapide [R]/Lent [L]); Hauteur de la f <sub>0</sub> (Haut [H]/Bas [B])	4 conditions: [RH], [RB], [LH], [LB]	Etudier la manière dont l'information prosodique guide le traitement linguistique (modulation de l'information acoustique)	Validation de l'idée d'un guidage sensoriel par la prosodie	
Expérience 4	Présence d'un mot enchâssé initial (E/nE); L'item peut être enchâssé à l'initiale d'un autre mot (l/nl)	4 conditions: [EI], [EnI], [nEI], [nEnI]	Tester la validité psychologique du modèle TRACE lors du traitement de superpositions lexicales (compétition)	La présence de superpositions lexicales initiales, réelles ou virtuelles, induit-elle une compétition mesurable ?	
Expérience 5	Présence d'un mot enchâssé à l'initiale de l'item (E/nE)	2 conditions: [E], [nE]	Reproduire l'effet obtenu dans l'Expérience 4 avec un matériel mieux contrôlé (items appariés phonologiquement)	Vérification que les effets de l'Expérience 4 ne soient pas dus à un biais	
Expérience 6	Lexicalité de S <sub>1</sub> (M/nM); Lexicalité de S <sub>2</sub> (M/nM)	4 conditions: [M <sub>1</sub> M <sub>2</sub> ], [M <sub>1</sub> nM <sub>2</sub> ], [nM <sub>1</sub> M <sub>2</sub> ], [nM <sub>1</sub> nM <sub>2</sub> ]	Répliquer l'effet des expériences 4 & 5 avec des mots bisyllabiques pour déterminer s'il existe une différence entre un enchâssement à l'initiale et à la fin du mot	Vérification de l'effet d'une superposition lexicale dans les mots bisyllabiques; distinction du rôle positionnel des superpositions lexicales	
Expérience 7					

Expérience 8	Présence d'un mot enchâssé à l'initiale de l'item (E/nE)	2 conditions: [E], [nE]	Vérifier si l'effet est lié à la modalité sensorielle	L'absence de compétition visible dans les Expériences 4 à 7 est inattendue et doit faire l'objet d'une enquête plus poussée (cachée par le traitement graduel de l'information acoustique?)	Influence de l'organisation du lexique mental sur la segmentation lexicale	
Expérience 9	Position de la cible à détecter (5 modalités); facteurs lexicaux de l'Expérience 4	20 conditions: (5 positions) * (4 types de mots)	Tester la compétition lexicale avec une tâche mieux adaptée (mesure d'une interférence du traitement lexical sur une tâche non-linguistique)			
Expérience 10	Position de la cible à détecter (4 modalités); facteurs lexicaux de l'Expérience 7	16 conditions: (4 positions) * (4 types de mots)				
Expérience 11	Fréquence d'occurrence de la syllabe en initiale de mot (i: Rare/Fréq.); Fréquence d'occurrence de la syllabe en position non-initiale (f: Rare/Fréq.)	4 conditions: [RiRf], [RiFf], [FiRf], [FiFf]	Mettre en évidence un biais dans la détection d'une syllabe non lexicale en fonction de sa fréquence positionnelle d'occurrence dans les mots du lexique			
Expérience 12	Existence de la syllabe non lexicale en Français (O/N); Fréquence d'occurrence des syllabes-pseudo-mot Fr. (Rare/Fréq.)	3 conditions: [Syllabe illégale], [Syllabe rare], [Syllabe fréquente]	Tester l'origine fréquentielle de ce biais pour des syllabes non lexicales (fréquence ou fréquence positionnelle ?)			Proposer et explorer l'idée d'une origine infralexicale syllabique à cet effet de facilitation. La compétition pourrait exister à un niveau sémantique plutôt que formel
Expérience 13	Fréquence d'occurrence de la syllabe lexicale en position initiale (i: Rare/Fréq.)	2 conditions: [Ri], [Fi]	Mettre en évidence une relation complexe entre les mots du lexique en fonction de la fréquence positionnelle d'occurrence de mots monosyllabiques dans d'autres mots			
Expérience 14	Fréquence d'occurrence de la syllabe non lexicale en position finale (ni: Rare/Fréq.)	2 conditions: [Rni], [Fni]				
Analyse exploratoire	Reprend les données des Expériences 13 & 14		Tester l'hypothèse d'un niveau de représentation syllabique, concurrentement à l'hypothèse d'un effet phonotactique lié à la séquence de phonèmes			
Expérience 15	Lexicalité de S <sub>2</sub> (M/nM); Prosodie (mot bisyllabique {} ou frontière {}); Lexicalité de la séquence bisyllabique	4 conditions: [{M <sub>1</sub> M <sub>2</sub> }; [{M <sub>1</sub> nM <sub>2</sub> }; [M <sub>1</sub>  M <sub>2</sub> ]; [M <sub>1</sub>  nM <sub>2</sub> ]	Tester l'influence cohésive de la prosodie et de la lexicalité des syllabes sur le percept d'un mot bisyllabique (guidage à la segmentation)	Explorer comment les 2 catégories de facteurs étudiées distinctement dans les 2 parties précédentes influencent la segmentation lexicale	Influence combinée de la prosodie et de l'organisation du lexique mental sur la segmentation lexicale	
Expérience 16		4 conditions: [mot {M <sub>1</sub> M <sub>2</sub> }], [mot M <sub>1</sub>  M <sub>2</sub> ], [pseudo-mot {M <sub>1</sub> M <sub>2</sub> }], [pseudo-mot M <sub>1</sub>  M <sub>2</sub> ]	Tester l'influence morcellante de la prosodie sur le percept d'un item porteur (guidage à la segmentation)			

Tableau 18 : Aspects techniques mis en œuvre par expérience

Nom de l'expérience	Matériel expérimental					Tâche		Population		Recueil de données		Analyse	
	Format linguistique it. expérim. (rempliss.)	Particularité	Format articuloire	Format physique	Nombre d'items expérim. (rempliss.)	Type de test	Tâche	Type de population testée	N Group.*N suj. (N suj. tot. dans analyses)	Lieu (université)	Système de recueil	Méthode d'analyse	Logiciel
Analyse acoustique	mot, pseudo-mot	-	syllabe	audio, voix naturelle	162	-	-	-	-	Aix	MES Signaux (UNIX)	Analyse multiniveaux	MLwiN
Simulation 1		items renfermant une frontière lexicale entre les 2 syllabes	seq. de 2 syllabes (it. anal. acoust.)	valeurs f0 (1/10ms.) durée	81	Simulation	Classific.	Réseaux de neurones	3*1 (3)	-	SNNS® (UNIX)	ANOVA Analyse multiniveaux	Statistica™ MLwiN
Simulation 2			combinais. de 2 bisyllabes (items de la Simulation 1)		3321				6*1 (6)			Analyse multiniveaux	MLwiN
Expérience 1	mot, portion de mot	items renfermant ou non une frontière lexicale	seq. de 3 syllabes (extraits corpus anal. acoust.)	audio, déségmenté (train d'impulsions)	81	Test perceptif	Discrimin.	Etud. Francoph. monolingues sans troubles auditifs	6*7 (42)	Aix	Scripts (MSDOS®)	Analyse multiniveaux Analyse de régression et bootstrap	MLwiN Gauss™
Expérience 2			syllabe (it. Expérience 4)	audio, déségmenté (train d'impulsions)	384				Détection+ localisation de frontière	1*50 (50)	Genève	E-prime™ (Windows®)	
Expérience 3	mot (pseudo-mot)	-	syllabe	audio, voix naturelle	60 - (180)	Test perceptif	DétECTION de mot (TDL gng)	Etud. Francoph. monolingues sans troubles auditifs	4*11 (44)	Genève	Scripts (MSDOS®)	Analyse multiniveaux	MLwiN
Expérience 4					48 - (48)				1*40 (40)	Aix			
Expérience 5					88 - (88)				2*16 (32)	-			
Expérience 6			80 - (80)	1*31 (31)									
Expérience 7	mot (pseudo-mot)		seq. de 2 syllabes	80 - (80)	DétECTION de mot (TDL gng); Classific. (TDL)	4*18 incomplet (71)							
Expérience 8			syllabe	visuel	36 - (18)	DétECTION de mot (TDL gng)	2*13 (26)						
Expérience 9	mot pseudo-mot		items renfermant cible ou sonorité distractive	syllabe (it. Expérience 4)	-	300 - (2100)	Test perceptif	DétECTION de clics	-	10*15 incomplet (140)	-		
Expérience 10		seq. de 2 syllabes (it. Expérience 7)		243 - (915)		6*15 incomplet (81)							
Expérience 11	S <sub>2</sub> pseudo-mot (S <sub>2</sub> mot)	l'item est un pseudo-mot	seq. de 2 syllabes	audio, voix naturelle	52 - (52)	Test perceptif	DétECTION de pseudo-mot sur S <sub>2</sub> (word spotting+TDL gng)	Etud. Francoph. monolingues sans troubles auditifs	1*16 (16)	Genève	-	-	
Expérience 12	non-mot, pseudo-mot (mot)	-	syllabe		18 - (54)				DétECTION de non/pseudo-mot (TDL gng)	1*11 (11)			
Expérience 13					48 - (48)				DétECTION de mot (TDL gng)	1*15 (15)			
Expérience 14	mot (pseudo-mot)				48 - (48)					1*17 (17)			
Analyse exploratoire									48 & 48	-			1*15 (15) & 1*17 (17)
Expérience 15	mot (pseudo-mot)	matériel identique dans les 2 tests	seq. de 2 syllabes	-	160 - (160)	Test perceptif	DétECTION de mot (TDL gng)	-	4*14 (56)	-	E-prime™ (Windows®)	-	
Expérience 16	S <sub>2</sub> mot (S <sub>2</sub> pseudo-mot)				DétECTION de mot sur S <sub>2</sub> (word spotting+TDL gng)				4*15 incomplet (58)				

## 4 Partie IV – Expérimentation

### Introduction de la partie expérimentale

La partie expérimentale expose une par une les 16 expériences (Expériences 1 à 16) et les 4 analyses spécifiques (Analyse acoustique, Simulations 1 et 2, Analyse exploratoire) qui ont été réalisées dans cette étude. Leur présentation suit le plan général qui a été donné dans les 2 tableaux synoptiques fournis au Chapitre 5 de la Partie III (Tableau 17 page 156 et Tableau 18 page 158).

Les expériences et les analyses spécifiques sont réparties dans trois chapitres, chacun relatif à un aspect du problème étudié (prosodie, lexique, prosodie + lexique). Un chapitre supplémentaire est ajouté pour explorer la relation entre prosodie et facteurs lexicaux, et entre prosodie et réactivité des participants (Chapitre 4). Un bref chapitre de conclusion (Chapitre 5) récapitule synthétiquement les résultats présentés dans la quatrième partie. La partie expérimentale est donc organisée comme suit :

- ▷ Chapitre 1 – Exploration du rôle des paramètres suprasegmentaux dans le guidage sensoriel de l'activation lexicale
- ▷ Chapitre 2 – Linéarisation d'éléments superposés : quelle organisation pour le système de reconnaissance des mots parlés ?
- ▷ Chapitre 3 – Exploration de l'influence combinée des aspects suprasegmentaux du signal et architecturaux du lexique mental sur le traitement lexical
- ▷ Chapitre 4 – Exploration des données issues des Chapitres 2 et 3 : relation entre caractéristiques linguistiques et prosodiques des items, et entre prosodie et temps de réponse
- ▷ Chapitre 5 – Récapitulation des résultats obtenus

Chaque expérience présentée est documentée selon l'organisation générale suivante :

- ▷ Aspects théoriques (hypothèses et contexte théorique)
- ▷ Aspects techniques :
  - ∩ Facteurs expérimentaux et conditions
  - ∩ Hypothèses opérationnelles
  - ∩ Matériel expérimental utilisé
  - ∩ Organisation de l'expérience
  - ∩ Population testée

- ∩ Tâche
- ∩ Procédure de recueil des données
- ∩ Méthode d'analyse
- ∩ Paramètres contrôlés
- ▷ Résultats obtenus
  - ∩ Rapport des analyses statistiques
  - ∩ Présentation des résultats
- ▷ Conclusions

Certains des aspects techniques ont été détaillés dans la Partie III du manuscrit. Le lecteur est donc fréquemment invité à se référer au chapitre correspondant pour trouver l'information, qui n'est pas répétée.

## **4.1 Chapitre 1 - Exploration du rôle des paramètres suprasegmentaux dans le guidage sensoriel de l'activation lexicale**

Le Chapitre 1 a pour but d'explorer l'organisation des caractéristiques suprasegmentales des mots en fonction de l'organisation théorique correspondant à la morphosyntaxe et à la topologie lexicale, et le rôle de ces caractéristiques suprasegmentales dans la segmentation du flux sonore en mots.

### **4.1.1 Présentation du contenu et de l'organisation du Chapitre 1**

L'étude commence par rechercher l'existence d'un code prosodique au niveau de la syllabe (« Étude préalable »). Ce code « prososyllabique » est supposé remplir un rôle de balisage du flux de parole, et permettre de localiser les éléments lexicaux lors du traitement (mots, frontières de mots).

L'utilité de ce code prososyllabique est ensuite testée dans deux expériences auprès d'auditeurs francophones. La première vise à montrer que les formes sonores prosodiques permettent d'organiser le matériel expérimental en catégories théoriques sur la base de la perception de participants naïfs. La deuxième vise à montrer que les auditeurs sont capables de détecter et de localiser consciemment les frontières lexicales sur la seule base de la prosodie.

Une dernière phase du travail tente de comprendre comment la prosodie peut être utilisée pour guider le système récepteur de façon ascendante. Une troisième et dernière expérience est effectuée dans ce but. Elle vise à montrer que la prosodie sert à moduler la résolution acoustique de la ligne segmentale.

Le Chapitre 1 est donc organisé de la manière suivante :

- ▷ Introduction (présentation de l'origine de ce travail)
- ▷ Étude préalable : construction du matériel, recherche de l'existence d'un code prososyllabique (Analyses acoustiques statiques, Analyses acoustiques dynamiques (Simulation 1, Simulation 2)), évaluation de la faisabilité du premier test perceptif (Simulation 2)
- ▷ Premier test perceptif (Expérience 1), discrimination (perceptibilité des formes prosodiques)

- ▷ Deuxième test perceptif (Expérience 2), détection et localisation consciente de frontières lexicales (utilité des formes prosodiques pour la localisation des frontières lexicales)
- ▷ Troisième test perceptif (Expérience 3), rôle de la durée et de la f0 sur la reconnaissance des mots (résolution acoustique de la ligne segmentale par la prosodie)
- ▷ Conclusions

### 4.1.2 Étude préalable

L'étude préalable comporte 2 volets techniques. Le premier volet concerne l'élaboration du matériel expérimental (Section 4.1.2.1 ci-dessous), le deuxième concerne la vérification du contenu du matériel expérimental (Analyses acoustiques statiques (Section 4.1.2.2.1, page 166), Analyses acoustiques dynamiques (Simulations 1 et 2, Section 4.1.2.2.2, page 175)), dans le but de s'assurer qu'il existe bien un rapport entre prosodie et facteurs théoriques, donc que le matériel est utilisable dans un test perceptif. La vérification du corpus est effectuée en deux étapes. La première étape (Analyse acoustique statique) est une analyse acoustique simple des syllabes, qui permet de s'assurer s'il existe des différences acoustiques moyennes (grossières) remarquables entre les catégories théoriques de syllabes. La deuxième étape (Analyse acoustique dynamique par des réseaux neuromimétiques) est une analyse acoustique des syllabes (Simulation 1) et de la différence entre items d'une paire comparée (Simulation 2). Elle permet d'estimer de façon plus précise que par l'analyse statique, la présence de formes sonores en relation avec les facteurs théoriques dans le matériel. Incidemment, la Simulation 2 sert aussi à estimer la faisabilité du premier test perceptif.

#### 4.1.2.1 Élaboration du matériel expérimental

##### 4.1.2.1.1 Facteurs manipulés

Trois facteurs sont manipulés dans le matériel expérimental :

- ▷ La catégorie morphosyntaxique du mot duquel provient la syllabe d'intérêt (de classe ouverte (L pour « lexical ») vs. de classe fermée (G pour « grammatical »))
- ▷ Le nombre de syllabes du mot porteur (1 vs. 3)
- ▷ La position de la syllabe dans le mot porteur (début (D) vs. fin (F))

Ces 3 facteurs permettent de déterminer 6 catégories de syllabes, mais seulement 4 sont retenues, étant donné que peu de mots grammaticaux ont 3 syllabes. Ces 4 conditions expérimentales sont récapitulées dans le Tableau 19.

Tableau 19 : Catégories de syllabes utilisées dans le matériel expérimental

N = 162 syllabes	Mot grammatical		Mot lexical	
	Syllabe de début	Syllabe de Fin	Syllabe de début	Syllabe de Fin
Mot monosyllabique	G (N = 54)		L (N = 54)	
Mot trisyllabique	-	-	D (N = 27)	F (N = 27)



Ces 4 catégories de syllabes se combinent 2 par 2 et définissent 9 catégories de frontières possibles (items), présentées dans le Tableau 20.

Tableau 20 : Catégories de frontières lexicales déterminées par la combinaison 2 à 2 des 4 types de syllabe

<b>N = 81 items</b>	<b>G<sub>(syllabe 2)</sub></b>	<b>L<sub>2</sub></b>	<b>D<sub>2</sub></b>	<b>F<sub>2</sub></b>
<b>G<sub>(syllabe 1)</sub></b>	G <sub>1</sub> G <sub>2</sub> (N = 9)	G <sub>1</sub> L <sub>2</sub> (N = 9)	G <sub>1</sub> D <sub>2</sub> (N = 9)	-
<b>L<sub>1</sub></b>	L <sub>1</sub> G <sub>2</sub> (N = 9)	L <sub>1</sub> L <sub>2</sub> (N = 9)	L <sub>1</sub> D <sub>2</sub> (N = 9)	-
<b>D<sub>1</sub></b>	-	-	-	-
<b>F<sub>1</sub></b>	F <sub>1</sub> G <sub>2</sub> (N = 9)	F <sub>1</sub> L <sub>2</sub> (N = 9)	F <sub>1</sub> D <sub>2</sub> (N = 9)	-

#### 4.1.2.1.2 **Élaboration du matériel expérimental**

L'élaboration du matériel expérimental comporte 4 étapes :

- ▷ Création des phrases porteuses
- ▷ Lecture et enregistrement des phrases
- ▷ Localisation des items dans le signal enregistré
- ▷ Extraction de l'information prosodique

##### 4.1.2.1.2.1 **Création des phrases porteuses**

Des phrases construites spécifiquement pour l'expérience sont utilisées afin de :

- ▷ mieux contrôler les sources de biais
- ▷ manipuler plus facilement les facteurs

54 petits paragraphes de 2 phrases ont été créés :

1. une phrase introductive (exemple : « Je suis très content du matériel expérimental que nous avons créé pour le prochain test. »), qui propose un contexte sémantique (amène les idées et le sens principal, prépare l'apparition de la phrase qui suit).
2. la deuxième phrase renferme l'item (phrase porteuse - exemple : « Hormis les gabarits différents des phrases lues, tout y est très homogène. »).

Les phrases porteuses se composent de deux groupes syntaxiques, l'item étant contenu dans le premier groupe. La structure du premier groupe de la phrase porteuse est la suivante : [préposition « *Hormis* »] [préposition / article 1 (exemple : « *les* »)] [nom 1 (exemple : « *gabarits* »)] [adjectif ou préposition / article 2 (exemple : « *différents* »)] [préposition / article 2 ou 3 (exemple : « *des* »)] [nom 2 ou groupe nominal (exemple : « *phrases lues* »)]. Les futurs items :

- ▷ sont localisés entre « préposition / article 1 » inclus et « préposition / article 2 ou 3 » inclus (exemple : « *les ga-* »).
- ▷ aucun item n'est proche d'une fin groupe majeur (pause plus ou moins marquée) ni proche d'une pause discursive

- ▷ tous sont insérés dans un contexte prosodique général identique (début de phrase, premier groupe majeur).

Les phrases porteuses sont présentées dans l'Annexe 2 page 457.

#### **4.1.2.1.2.2 Lecture et enregistrement des phrases**

La locutrice qui a été sélectionnée pour enregistrer les phrases possède une voix claire, une intonation vivante, et une lecture fluide et naturelle. De plus, elle est totalement naïve pour ce qui concerne les hypothèses expérimentales et le contenu théorique du corpus.

Les enregistrements se déroulent dans une chambre sourde au Laboratoire Parole et Langage (Université de Provence – Aix-Marseille I), selon la procédure décrite dans la Section 3.4.1.2, page 140.

#### **4.1.2.1.2.3 Localisation des items dans le signal enregistré**

Elle se fait par un étiquetage manuel au sonagramme (frontières initiale et finale – MES / SIGNAIX).

#### **4.1.2.1.2.4 Extraction de l'information prosodique**

L'information acoustique décrivant les caractéristiques prosodiques des items est utilisée pour :

- ▷ Les analyses acoustiques du matériel
- ▷ Les Simulations 1 et 2
- ▷ La synthèse d'items désémantisés (Expérience 1)

Pour l'obtenir, les items sont analysés (logiciel MES / Signaix et scripts de R. Espesser (1994)) pour la durée de chaque syllabe (ms.), la  $f_0^{99}$  (Hz, 1 valeur toutes les 10 ms.), l'intensité (rms converti en dB, 1 valeur toutes les 10 ms.). Cette étape permet de connaître les caractéristiques physiques du signal de parole d'origine pour en analyser les caractéristiques en lien avec les catégories théoriques (analyses acoustiques, Simulations 1 et 2).

---

<sup>99</sup> Trois méthodes sont utilisées concurremment : méthode du peigne, AMDF et auto corrélation. La courbe de  $f_0$  obtenue résulte d'un consensus entre ces 3 méthodes. La procédure est très fiable avec le signal naturel de parole.

### 4.1.2.2 Vérification du contenu du matériel

Cette section est subdivisée en 2 :

- ▷ Analyse acoustique statique des syllabes
- ▷ Analyse acoustique dynamique des syllabes et des items, par des réseaux neuromimétiques

#### 4.1.2.2.1 Analyse acoustique statique des syllabes

Le but des analyses statistiques est de vérifier que la catégorie d'une syllabe, attribuée théoriquement à partir de facteurs morphosyntaxiques et topologiques, possède une réalité physique.

Pour cela, une analyse de régression multiniveaux a été effectuée pour la durée et la f0 (items, syllabes – Corrélation intra classe = 11.35%).

##### 4.1.2.2.1.1 Rapport des analyses statistiques

Pour chacune des 2 dimensions, un modèle d'analyse est proposé. Ces 2 modèles sont présentés en Annexe 3 page 463, dans les Équation 20 et Équation 21.

Ces modèles révèlent des valeurs moyennes de durée et de f0 qui diffèrent en fonction du type de syllabe.

##### 4.1.2.2.1.2 Présentation des résultats

Le Tableau 21 présente les valeurs moyennes observées et prédites par les modèles, de durée et de f0 pour chacune des 4 catégories de syllabes.

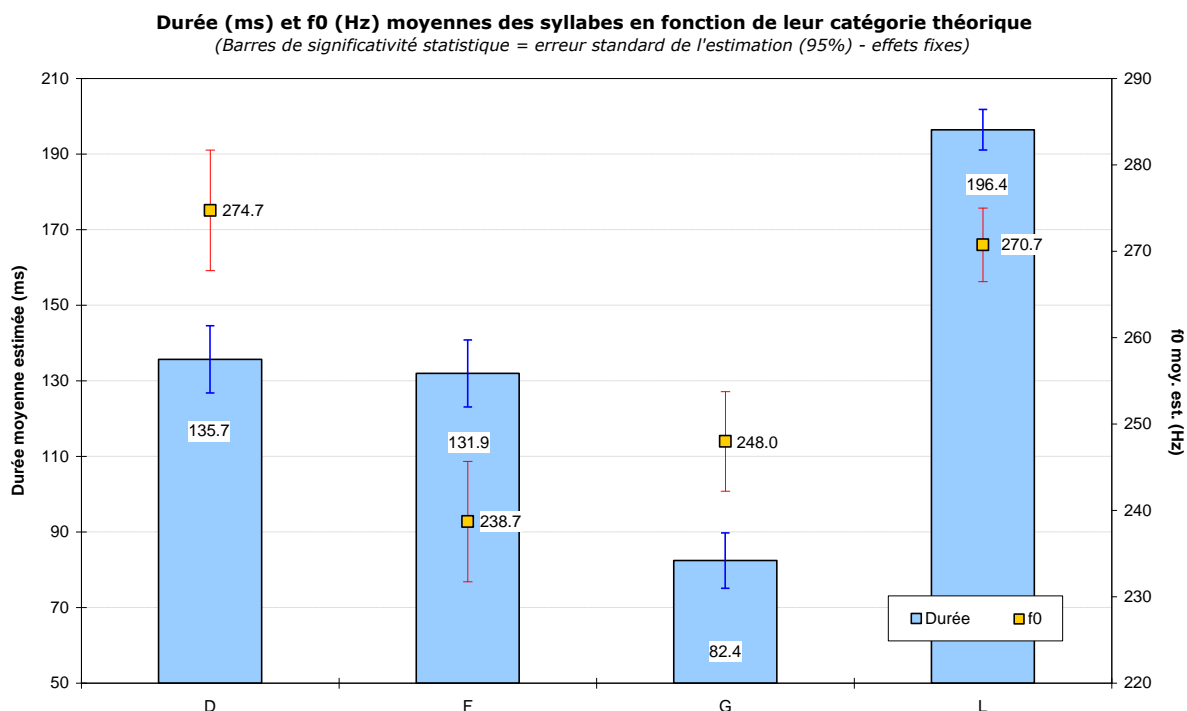
Tableau 21 : Valeurs moyennes de durée et de f0 par catégorie de syllabe (observé et prédit)

		Durée (ms)	Écart-type	f0 (Hz)	Écart-type
<b>Observé</b>	<b>D</b>	132.25	26.49	276.58	23.66
	<b>F</b>	135.22	24.24	237.94	19.47
	<b>G</b>	82.09	31.51	247.29	35.44
	<b>L</b>	196.81	55.16	270.90	34.75
		Durée (ms)	E.S. modèle	f0 (Hz)	E.S. modèle
<b>Prédit</b>	<b>D</b>	135.66	8.87	274.73	6.98
	<b>F</b>	131.95	8.87	238.70	6.98
	<b>G</b>	82.42	7.32	247.99	5.76
	<b>L</b>	196.41	5.38	270.74	4.26

Avec E.S. pour erreur standard du modèle.

Les valeurs prédites par les modèles d'analyse sont présentées dans la Figure 33.

Figure 33 : Durée et f0 moyenne de chaque catégorie de syllabe



Afin d'évaluer si les écarts entre catégories de syllabes sont significatifs, un test des contrastes a été effectué, dont les résultats sont présentés dans le Tableau 22 ci-dessous.

Tableau 22 : Tests de contrastes

contraste	Durée (ms)		f0 (Hz)	
	p<	Chi <sup>2</sup> (1, N=162)	p<	Chi <sup>2</sup> (1, N=162)
D/F	.72	0.1	<b>.0001</b>	20.1
D/G	<b>.0001</b>	34.6	<b>.0002</b>	14.1
D/L	<b>.0001</b>	45.1	.57	0.3
F/G	<b>.0001</b>	29.9	.19	1.7
F/L	<b>.0001</b>	50.7	<b>.0001</b>	20.2
G/L	<b>.0001</b>	49.2	<b>.0002</b>	15.0

Les analyses font ressortir l'existence de :

- ▷ 3 catégories de syllabes pour le facteur durée ([D & F], [L], [G]), qu'il est possible de distinguer 2 à 2.
- ▷ 2 catégories de syllabes pour le facteur f0 ([D & L], [F & G]), qu'il est possible de distinguer sur la base de la f0 moyenne.

Par combinaison des deux facteurs, durée et f0, il est donc statistiquement possible de distinguer les 4 types de syllabes utilisés dans le matériel expérimental.

#### 4.1.2.2.1.3 Conclusions relatives aux analyses acoustiques

Le matériel expérimental utilisé dans les tests perceptifs a été vérifié et validé par une analyse acoustique portant sur deux paramètres de la prosodie : la durée et la mélodie. Il a ainsi pu être montré qu'un profil prosodique particulier existe bien pour chacune des 4 catégories de syllabes expérimentales. En d'autres termes, ce matériel offre la possibilité de retrouver les 4 catégories théoriques de syllabes testées, à partir de l'information prosodique.

Les analyses effectuées sont cependant simplistes, puisqu'elles ne prennent en compte que des valeurs acoustiques moyennes par syllabe pour la  $f_0$ . Il en résulte un manque de précision dû à la perte d'information relative à la dimension dynamique du signal, qui est probablement importante dans les processus perceptifs et cognitifs. Pour cette raison, l'analyse acoustique doit être complétée par deux analyses à l'aide de réseaux neuromimétiques (Simulations 1 et 2). Leur utilisation semble justifiée par l'observation des courbes de  $f_0$  de chaque type de syllabe, présentées dans les Figure 34 à Figure 37, et par l'observation des courbes moyennes de  $f_0$  par type de frontière (Figure 39, page 171). Ces courbes laissent penser que si les profils acoustiques de frontières lexicales existent, ils sont dynamiques et dépendent de la manière dont la  $f_0$  évolue dans le temps.

Figure 34 : Courbes de  $f_0$  des syllabes D du corpus

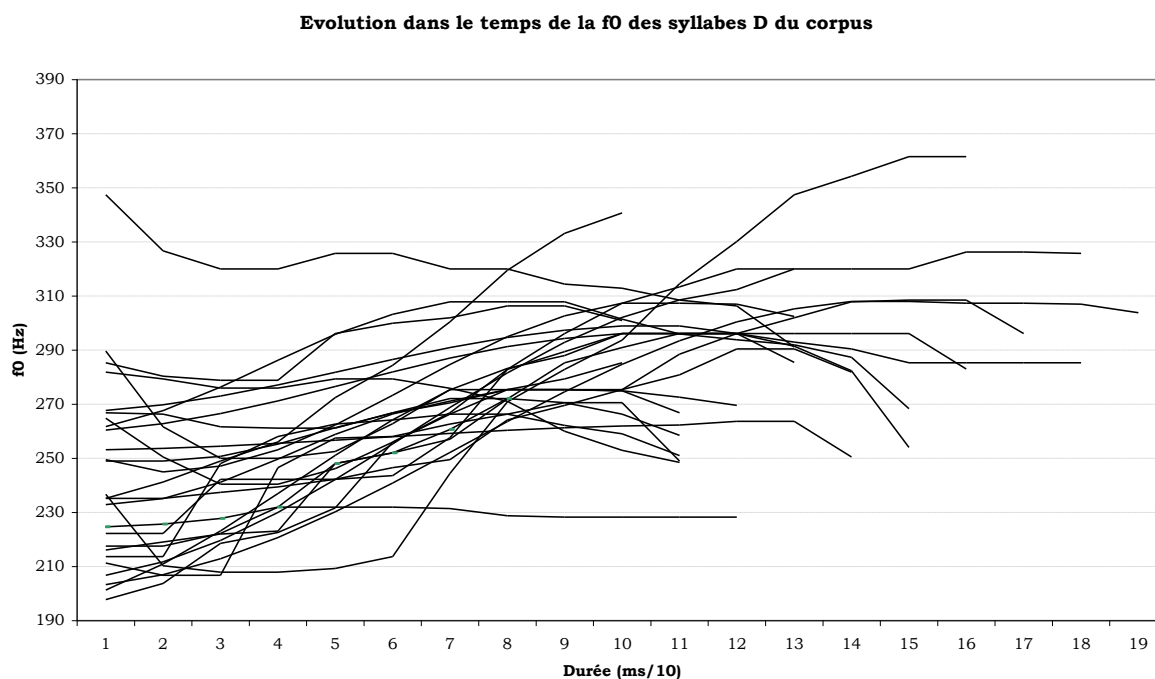


Figure 35 : Courbes de f0 des syllabes F du corpus

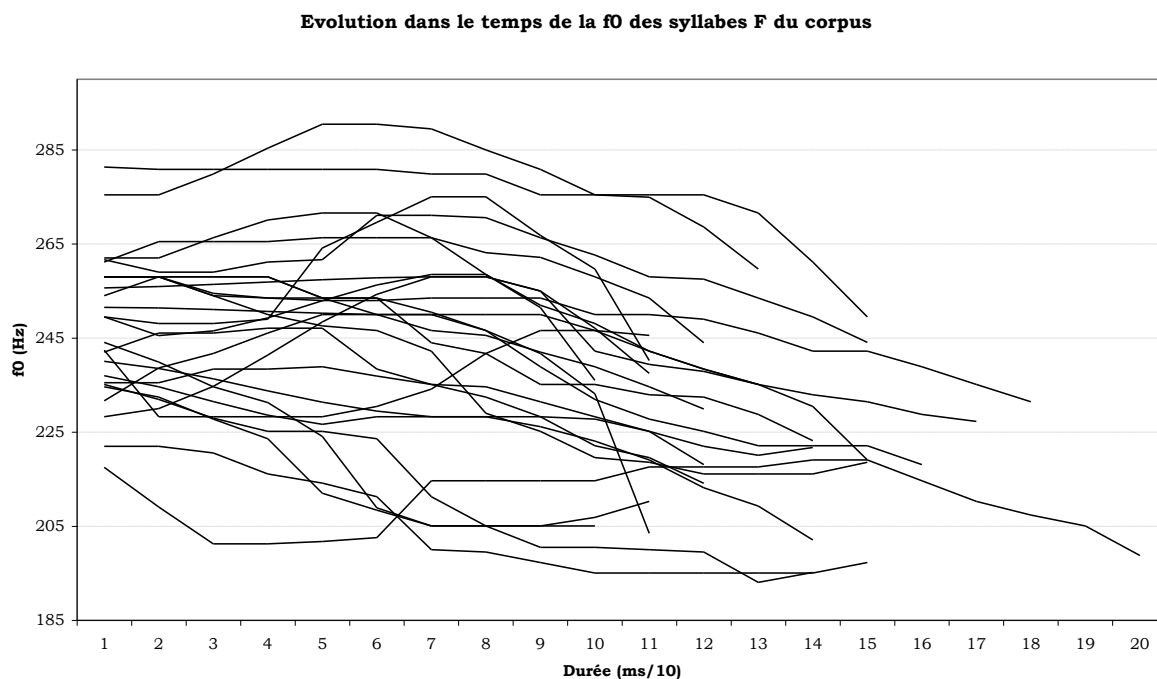


Figure 36 : Courbes de f0 des syllabes L du corpus

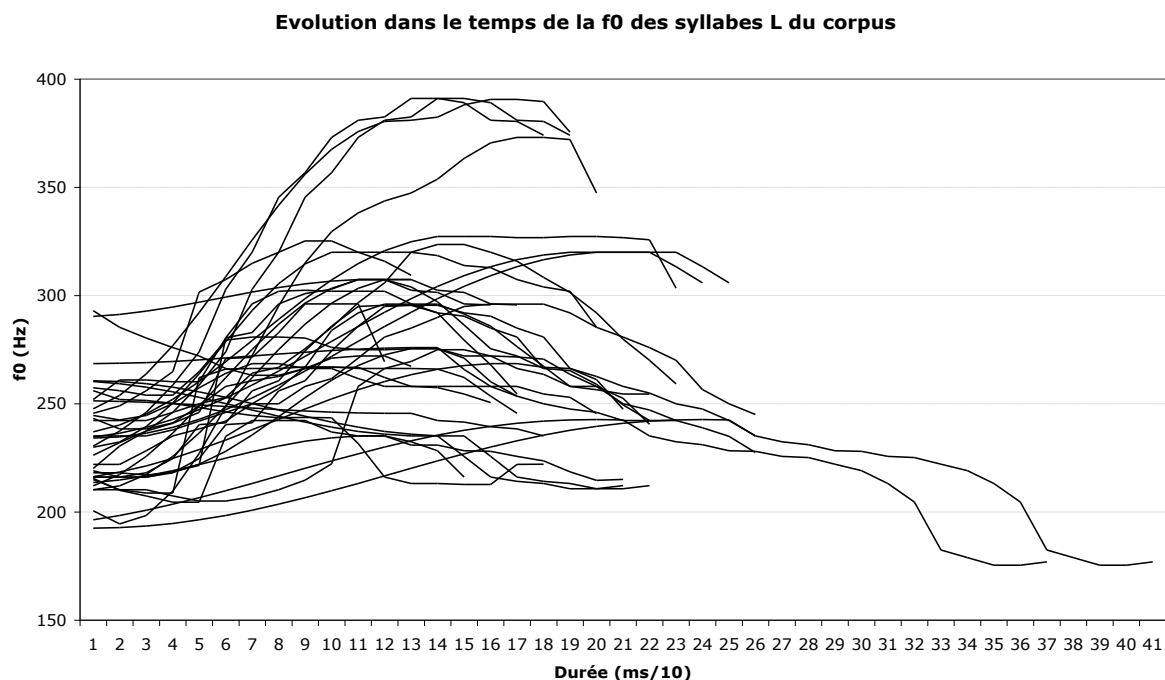
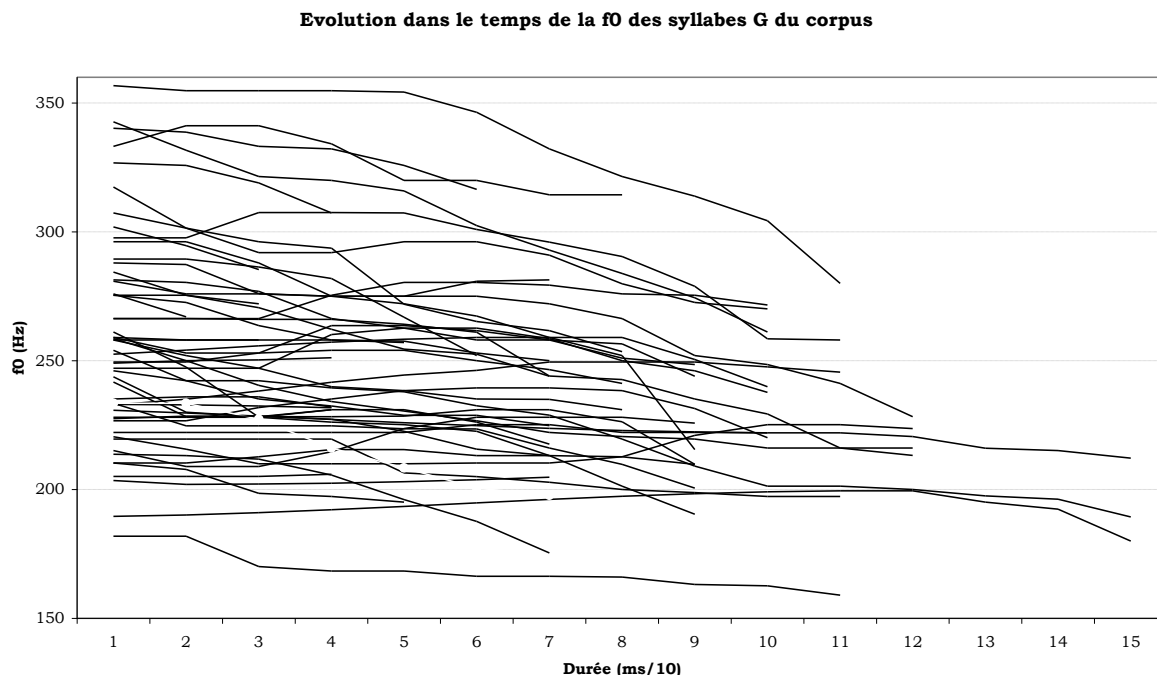
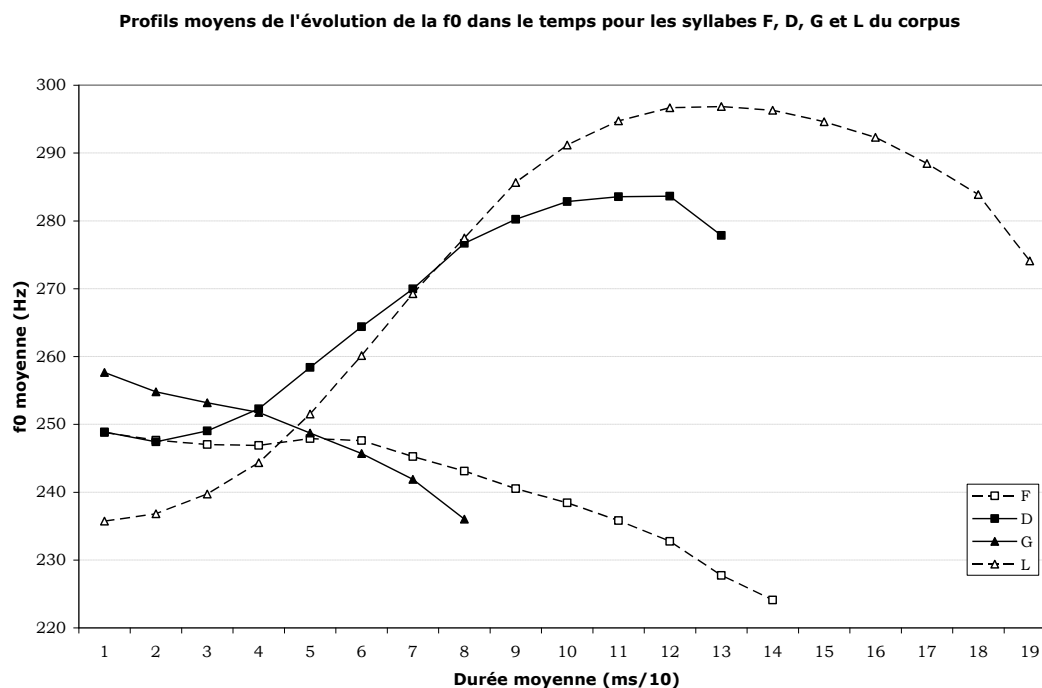


Figure 37 : Courbes de f0 des syllabes G du corpus



Un profil moyen des syllabes a pu être calculé dans les dimensions f0 et durée (courbes prototypes obtenues après normalisation sur la valeur maximum de f0 et de durée, par catégorie). Ces profils sont présentés dans la Figure 38.

Figure 38 : Profils moyens de f0 / durée pour les 4 catégories de syllabes du corpus

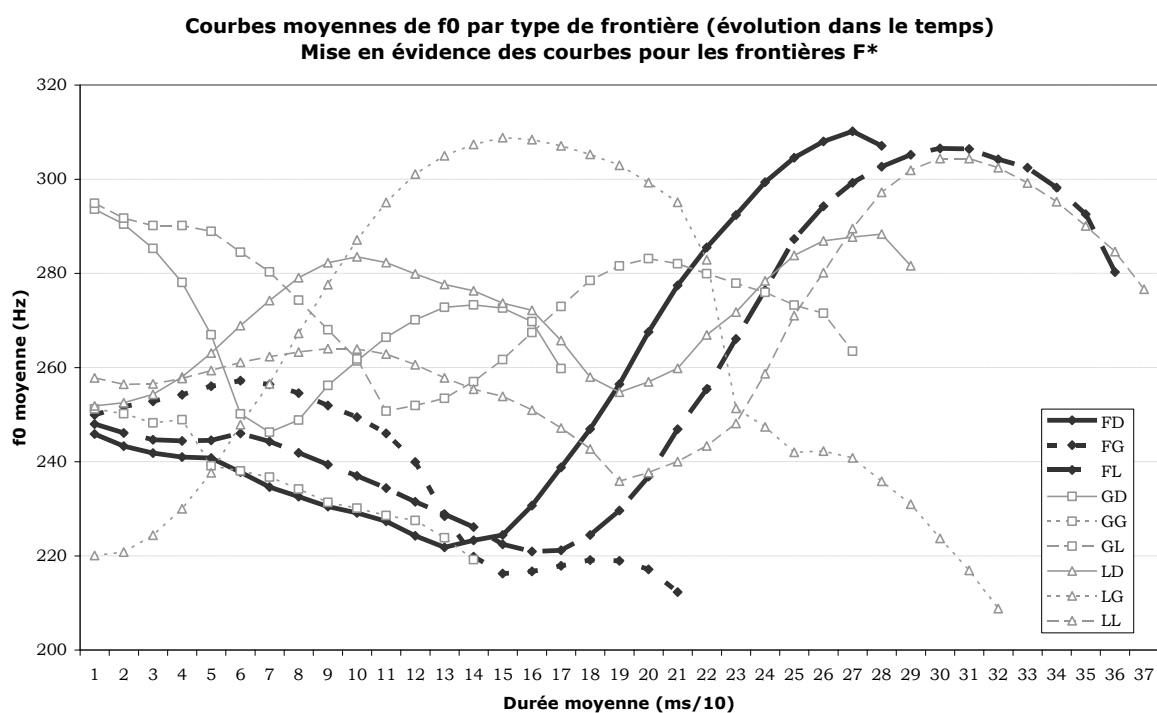


On retrouve visuellement les résultats des analyses acoustiques :

- ▷ 2 catégories de syllabes sont définies par la  $f_0$  (catégorie [D & L] :  $f_0$  montante ; catégorie [F & G] :  $f_0$  descendante)
- ▷ 3 catégories de syllabes sont définies par la durée (catégorie [L] : durée importante ; catégorie [G] : durée courte ; catégorie [F & D] : durée intermédiaire)

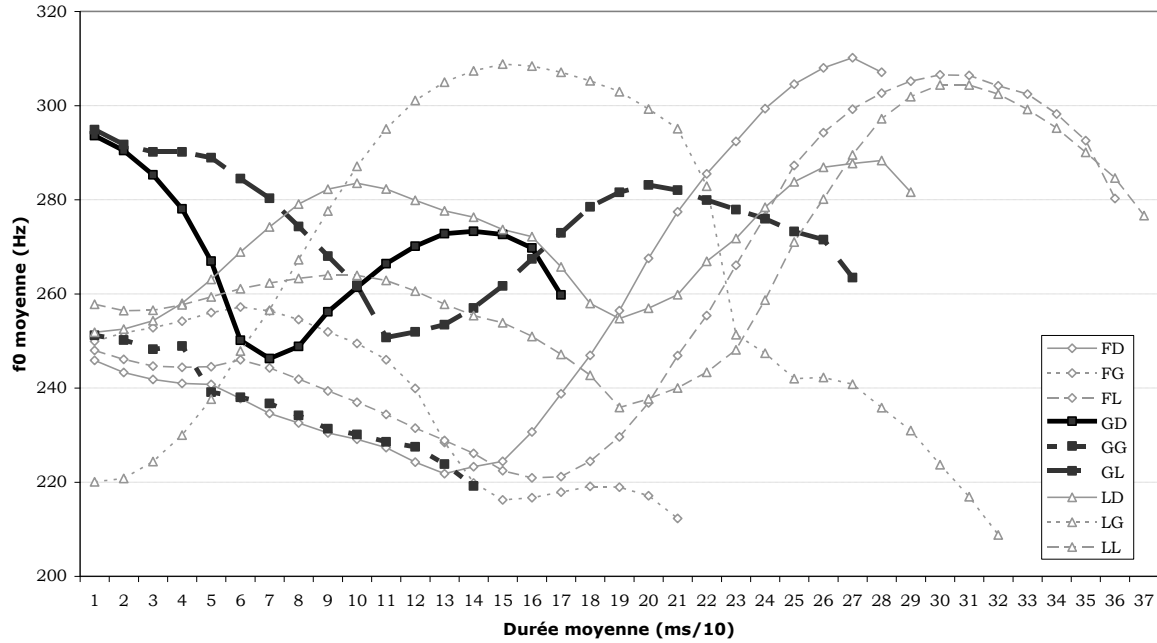
Le même procédé a été appliqué pour calculer un profil moyen de  $f_0$  / durée pour chacune des 9 frontières étudiées (Tableau 20, page 164). Ces profils sont fournis à travers la série de 6 graphiques de la Figure 39 (chaque graphique met en évidence les 3 profils dans lesquels une syllabe est impliquée à une position donnée).

Figure 39 : Profils moyens de  $f_0$  / durée pour les 9 types de frontières étudiés

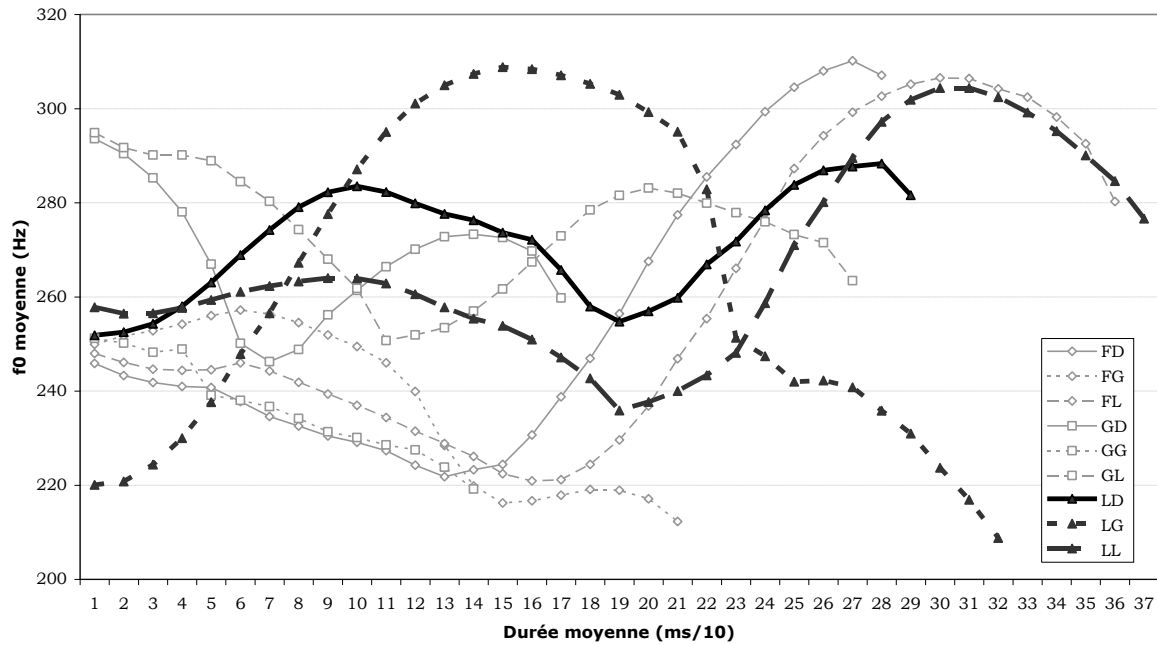




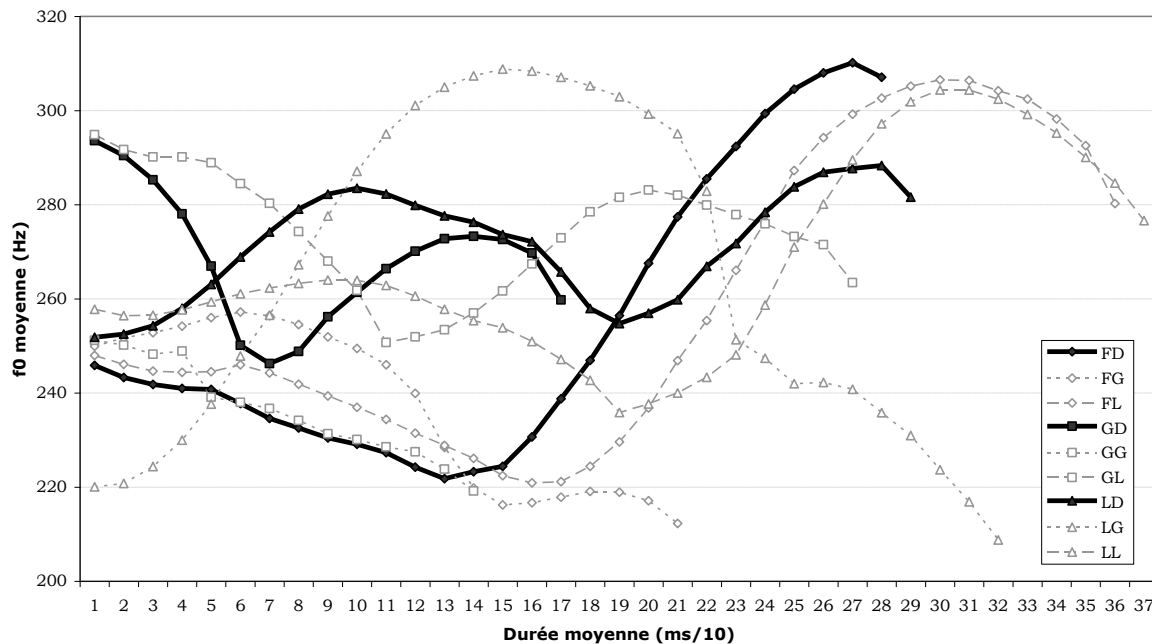
**Courbes moyennes de f0 par type de frontière (évolution dans le temps)**  
**Mise en évidence des courbes pour les frontières G\***



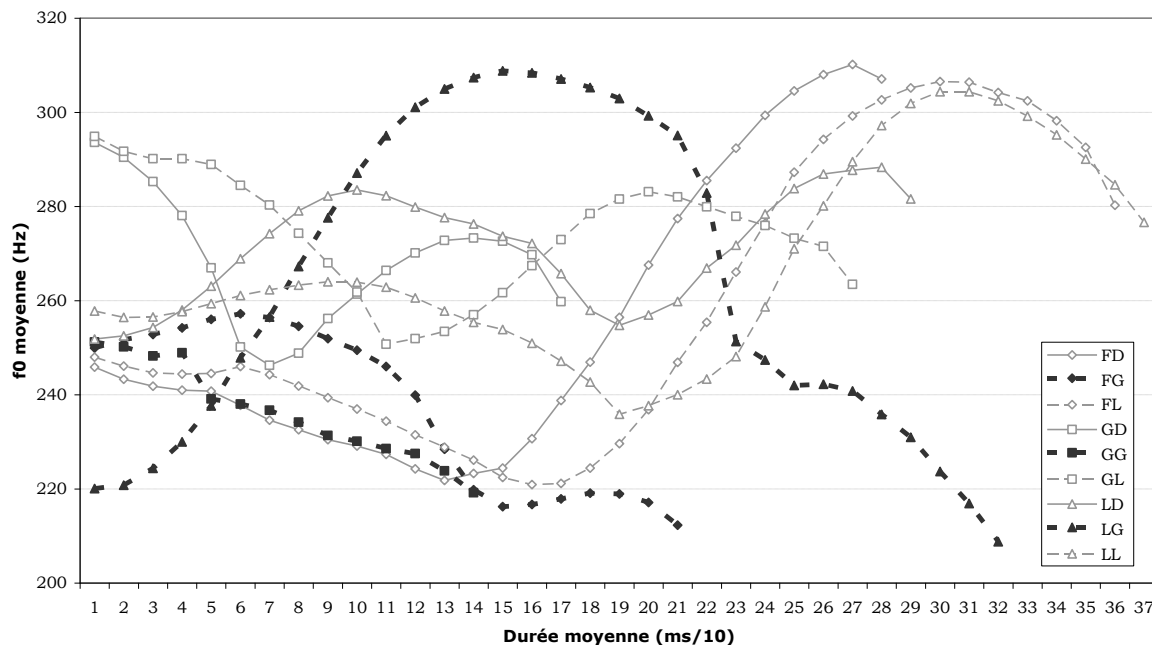
**Courbes moyennes de f0 par type de frontière (évolution dans le temps)**  
**Mise en évidence des courbes pour les frontières L\***

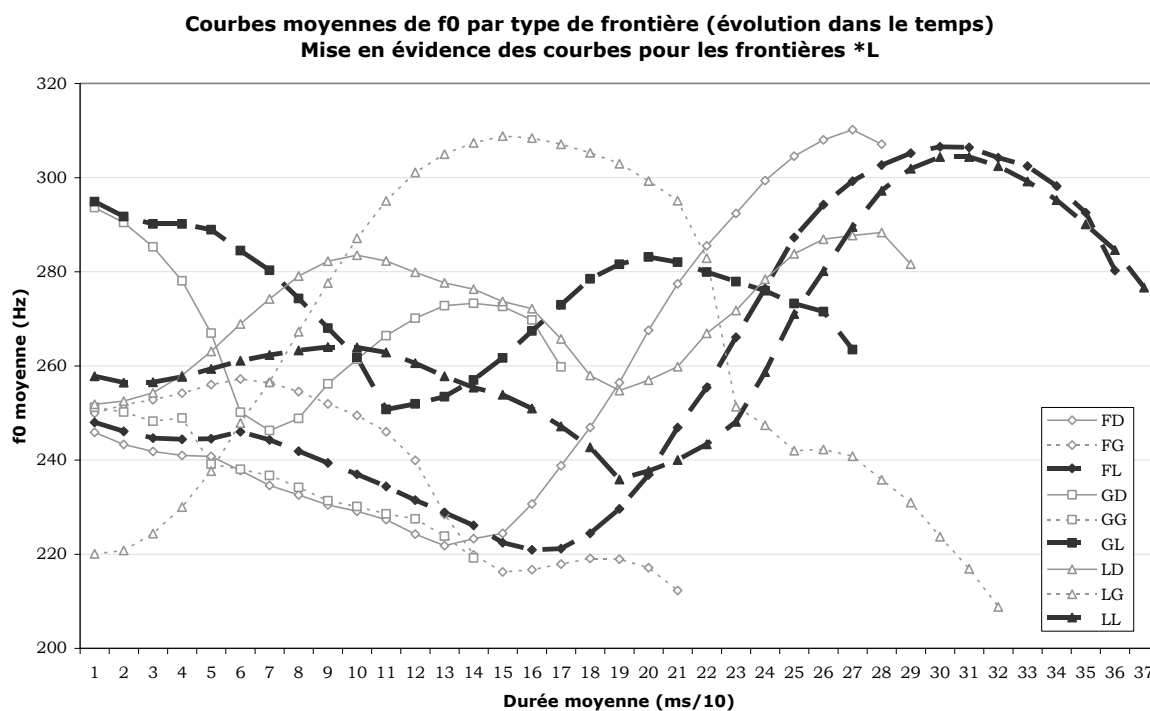


**Courbes moyennes de f0 par type de frontière (évolution dans le temps)**  
**Mise en évidence des courbes pour les frontières \*D**



**Courbes moyennes de f0 par type de frontière (évolution dans le temps)**  
**Mise en évidence des courbes pour les frontières \*G**





Ces courbes mélodiques montrent que toutes les frontières impliquant un mot monosyllabique de classe ouverte (FL, GL, LL) ou un début de mot polysyllabique de classe ouverte (FD, GD, LD) en position de deuxième syllabe dans l'item, sont marquées par un profil en V.

Inversement, toutes les frontières impliquant un mot monosyllabique de classe fermée en position de deuxième syllabe dans l'item (FG, GG, LG), ne sont marquées par aucun contraste mélodique.

Les frontières FD, FL et LL ont un profil acoustique moyen assez proche, et on peut donc s'attendre à ce que ces frontières soient confondues lors de la Simulation 2 (Tableau 34, page 192) et lors de l'Expérience 1 (Tableau 38, page 205). De même pour les frontières GD et GL, et pour les frontières FG et GG.

La frontière LG est probablement celle qui présente le profil le plus différent de toutes les autres, et c'est donc aussi celle qui devrait être la mieux distinguée, dans la Simulation 2 comme dans l'Expérience 1.

#### **4.1.2.2.2 Analyse acoustique dynamique des syllabes et des différences physiques entre items, par des réseaux neuromimétiques**

Deux analyses par réseau neuromimétique sont effectuées. La première (Simulation 1) porte sur l'analyse des profils prosodiques des syllabes du matériel expérimental. La deuxième (Simulation 2) porte sur l'analyse de la différence des profils prosodiques entre items. Les deux analyses correspondent à la résolution mathématique d'un problème de classification. Toutefois, la deuxième analyse est appelée « discrimination » afin de faire le parallèle entre la Simulation 2 et le test perceptif de l'Expérience 1. En effet, la classification effectuée par le réseau correspond alors à une répartition des items en deux catégories (même / différent), ces deux catégories étant générées chez le sujet humain par un processus de comparaison. La Simulation 2 est donc aussi une simulation du test perceptif de l'Expérience 1, ses résultats servant à estimer la faisabilité du test projeté.

##### **4.1.2.2.2.1 Qu'est-ce qu'un réseau de neurones artificiels ?**

Un réseau de neurones artificiels est un calculateur composé d'unités simples massivement connectées entre elles, chacune effectuant un même calcul simple, défini par 3 fonctions (entrée, intégration, sortie (Haykin, 1999)). Il s'agit d'un dispositif artificiel se voulant une analogie du système nerveux.

Tout problème, quelle que soit sa complexité, peut être résolu par un réseau neuromimétique, dans la mesure où une régularité existe entre l'état initial (les données, le stimulus) et l'état final (la réponse), que cette régularité soit linéaire ou non. Cette résolution correspond à l'ajustement mathématique d'un réseau (apprentissage) de manière à ce que la réponse attendue soit fournie à partir des données d'entrée. "A neural network (...) has a natural propensity for storing experiential knowledge and making it available for use" (Haykin, 1999).

Ici, les réseaux sont employés pour résoudre un problème de reconnaissance de forme sur des items inconnus, après extraction des régularités à partir d'exemples traités pendant une phase d'apprentissage supervisé<sup>100</sup>.

---

<sup>100</sup> L'apprentissage supervisé consiste à fournir au réseau 2 points d'ancrage. Le premier correspond aux données du problème à traiter, le deuxième à la solution attendue relative à ces données. L'apprentissage supervisé modifie la force des liens entre unités de manière à réduire l'erreur du réseau (écart entre solution qu'il fournit en sortie et solution attendue).

**4.1.2.2.2 Logiciel de simulation utilisé**

SNNS© v4.1 pour UNIX (Stuttgart Neural Network Simulator (Zell et al., 1989)).

**4.1.2.2.3 Caractéristiques des items utilisés dans les deux simulations**

Les items numériques sont obtenus à partir des mesures acoustiques du matériel original (une valeur de durée par syllabe, 1 valeur d'intensité et de f0 par 10 ms de durée du stimulus). Les valeurs de durée et de f0 initialement mesurées sont normalisées pour chaque dimension par rapport à la valeur maximale rencontrée à travers l'ensemble des items (listings de valeurs comprises entre 0 et 1). La structure des items numériques est légèrement différente pour chaque simulation. Le détail est donné ci-dessous.

## ▷ Structure des items numériques dans la Simulation 1

[61 valeurs {f0}, 1 {séparateur 0}, 61 {dB}, 1 {séparateur 0}, 1 {durée de la syllabe 1}, 1 {durée de la syllabe 2}, 1 {f0 moyenne de la syllabe 1}, 1 {f0 moyenne de la syllabe 2}, (8 {code de la catégorie de la syllabe, pour la sortie désirée})]<sup>101</sup>

Chacune des 4 catégories de syllabe est codée sur 4 unités. Ce codage est défini dans le Tableau 23 ci-dessous. Un item de 2 syllabes est donc codé sur 8 unités.

Tableau 23 : Code utilisé pour définir la catégorie de la syllabe

Type de syllable	Codage
<b>D</b>	1 0 0 0
<b>F</b>	0 1 0 0
<b>L</b>	0 0 1 0
<b>G</b>	0 0 0 1

## ▷ Structure des items numériques dans la Simulation 2

[61 valeurs {f0}; 1 {séparateur 0}; 1 {durée Syllabe<sub>1</sub>}; 1 {durée Syllabe<sub>2</sub>}; 1 {f0 moyenne Syl<sub>1</sub>}; 1 {f0 moyenne Syl<sub>2</sub>}]<sub>item1</sub>; [61 {f0}; 1 {séparateur 0}; 1 {durée Syllabe<sub>1</sub>}; 1 {durée Syllabe<sub>2</sub>}; 1 {f0 moySyl<sub>1</sub>}; 1 {f0 moySyl<sub>2</sub>}]<sub>item2</sub>; 1 {sortie désirée}

Codage de sortie : 0 = les 2 items comparés sont de catégories différentes ; 1 = les 2 items comparés sont de catégorie identique

Remarque :

L'information d'intensité est utilisée dans la Simulation 1 alors qu'elle ne l'est pas dans la Simulation 2, parce que la première correspond à une analyse dynamique des caractéristiques acoustiques du matériel expérimental, tandis que la deuxième simule le

<sup>101</sup> La structure de l'item varie légèrement en fonction de l'utilisation qui en est faite (apprentissage ou généralisation) :

Items d'apprentissage : listing de valeurs de f0, d'intensité, durée + réponse attendue

Items de généralisation : listing de valeurs de f0, d'intensité, durée, sans la réponse attendue

test perceptif de l'Expérience 1, dans lequel l'intensité des items dépend essentiellement des 2 phonèmes utilisés dans la synthèse (elle n'est donc pas en rapport avec les facteurs expérimentaux).

#### **4.1.2.2.2.4 Simulation 1 (classification des syllabes)**

##### **4.1.2.2.2.4.1 Hypothèse**

Les items seront classés dans leur catégorie théorique (morphosyntaxe \* topologie) par les réseaux avec un taux supérieur au hasard grâce à l'information prosodique (durée, f0, intensité).

##### **4.1.2.2.2.4.2 Caractéristiques et paramètres du réseau employé pour la Simulation 1 (classification)**

- ▷ Couche d'entrée : 128 unités
- ▷ Couche de sortie : 8 unités
- ▷ Unité de traitement : 1
- ▷ Connexions : 1160
- ▷ Apprentissage supervisé. Structuration du réseau par algorithme des corrélations en cascade

Paramètres : erreur maximale unités de sortie = 0.2 (arrêt apprentissage si erreur inférieure à cette valeur), quickprop (fonction de maximisation de la covariance ou minimisation erreur globale du réseau), changement minimal de covariance des unités candidates = 0.04 (si cette valeur pas atteinte, apprentissage arrêté et l'unité candidate possédant le maximum de covariance est incluse dans le réseau), nombre d'étapes séparant deux évaluations de la covariance = 25 (significatif si supérieur à la valeur du changement minimal de covariance des unités candidates), nombre maximal d'étapes de calcul de la covariance = 200 (après ce nombre d'étapes, l'unité candidate ayant une covariance maximale est intégrée au réseau), nombre maximal d'unités candidates entraînés à chaque cycle = 8, fonction d'activation des unités candidates = Act\_logsym, changement minimal de covariance de sortie des unités candidates = 0.01, nombre d'étapes séparant deux évaluations de la covariance en sortie = 50, nombre maximal d'étapes de calcul de la covariance en sortie = 200. Ordre aléatoire d'utilisation des items.

- ▷ Fonction d'apprentissage : Quickpropagation  
(CC -  $\eta_1 = 0.0001$ , paramètre de minimisation de l'erreur du réseau,  $\mu_1 = 2.0$ , paramètre de croissance maximale,  $\nu = 0.0001$ , terme de minimisation des poids,  $\eta_2 = 0.0007$ , paramètre de maximisation de la covariance,  $\mu_2 = 2.0$ , id  $\mu_1$ ).
- ▷ Fonction d'activation par propagation topologique dans le réseau (CC\_order)
- ▷ Fonction d'initialisation par mélange aléatoire des poids (CC\_weights)

##### **4.1.2.2.2.4.3 Organisation de la simulation**

Il est nécessaire de répartir les items en deux lots :

- ▷ Un lot d'apprentissage (pour la structuration du réseau) : 2 / 3 des items, soit  $N = 54$  (quantité choisie pour des raisons relatives à l'organisation décrite ci-dessous)
- ▷ Un lot de généralisation (pour tester les performances du réseau au cours des différentes étapes de sa structuration, sur des items inconnus) : le tiers restant ( $N = 27$ ).

Cette organisation soulève deux problèmes :

- ▷ Ne disposant que de 81 items, la quantité utilisable à chaque étape est encore réduite par cette répartition en 2 lots (apprentissage et test). Il risque donc de ne pas y avoir suffisamment d'exemples pour que le réseau présente de bonnes performances.
- ▷ Comment choisir quels items assigner à l'un des deux lots ? Un choix arbitraire peut entraîner des performances hétérogènes du réseau, liées aux capacités inégales des données à représenter toutes les dimensions du problème (manque de fiabilité du réseau à cause des exemples d'apprentissage, ou performances non représentatives liées aux items du lot de généralisation).

Une solution à ces deux problèmes a été proposée en testant trois réseaux, chacun sur une répartition aléatoire des 81 items en 9 lots, comportant chacun 1 exemple de chaque catégorie de frontière (détail de la répartition présenté dans le Tableau 24).

Tableau 24 : Répartition des items pour la structuration et le test des réseaux

Réseau	Apprentissage	Test
$R_1$	1 à 6	7 à 9
$R_2$	4 à 9	1 à 3
$R_3$	1 à 3 + 7 à 9	4 à 6

Les chiffres correspondent au numéro des items, qui est arbitrairement assigné par ordre de sélection (le premier item tiré aléatoirement dans la catégorie  $x$  porte le numéro 1).

#### **4.1.2.2.2.4 Procédure d'apprentissage et de recueil des données**

Chaque réseau est entraîné sur cinq cycles d'apprentissage (les performances ne sont pas significativement améliorées avec un nombre supérieur de cycles d'apprentissage). À la fin de chaque cycle d'apprentissage, le réseau est testé sur le lot d'exemples de généralisation (test qui ne modifie pas le réseau). Les résultats de l'évolution de la performance des réseaux sont présentés dans le Tableau 25.

Tableau 25 : Performance des 3 réseaux (erreur globale du réseau)

		Cycle 0	cy 1	cy 2	cy 3	cy 4	cy 5
Apprentissage	Réseau 1	108	1.005	0.307	0.146	0.084	0.06
	Réseau 2	108	0.968	0.241	0.091	0.061	0.044
	Réseau 3	108	4.071	0.267	0.115	0.088	0.077
Moyenne		108	2.01	0.27	0.12	0.078	0.06
Généralisation	Réseau 1	54	33.14	35.21	35.53	35.4	35.67

	<b>Réseau 2</b>	54	30.64	31.55	32.32	32.54	32.63
	<b>Réseau 3</b>	54	24.77	24.7	24.46	23.99	23.48
	<b>Moyenne</b>	54	29.52	30.49	30.77	30.64	30.59

La meilleure performance de généralisation est atteinte entre les cycles 1 et 2 (erreur globale du réseau).

Les 3 réseaux présentent une erreur globale différente, liée à la répartition des items en lots d'apprentissage / test (représentation inégale des dimensions du problème).

L'analyse statistique porte sur les valeurs moyennes fournies par les 3 réseaux pour chacune des 4 unités de sortie de chaque syllabe. Ces valeurs sont celles recueillies lors des cycles 1 et 2 (test sur les items de généralisation). Le but de l'utilisation de valeurs moyennes est de réduire le bruit lié à l'échantillonnage des items dans les lots.

#### **4.1.2.2.4.5 Rapport des analyses statistiques et présentation des résultats**

Deux types d'analyses statistiques ont été effectués sur les données :

- ▷ ANOVA, pour vérifier rapidement si le réseau parvient à retrouver les catégories théoriques des syllabes à partir de l'information prosodique (exploration des effets).
- ▷ Analyses de régression multiniveaux (syllabe, item), pour estimer le degré de caractérisation des catégories de syllabes (homogénéité d'une classe, distance entre catégories) par une analyse plus fine des performances du réseau. Cette analyse teste :
  - ∩ l'effet du type de syllabe (D, F, G, L) en fonction de sa position dans l'item ( $S_1$  ou  $S_2$ )
  - ∩ l'effet du type de syllabe en fonction du type de frontière



**4.1.2.2.4.5.1 Rapport des analyses de variance**

Avec : I : item ; R : activation correcte vs. incorrecte ; C : catégorie de la syllabe S<sub>1</sub> ou S<sub>2</sub>

- ▷ Analyse de la différence entre taux d'activation correcte et incorrecte

$$I_{81} * R_2$$

Écart significatif ( $F_{(1,80)} = 175.6, p < .0001$ ) – Taux d'activation correcte (65%) ; taux d'activation incorrecte (11%).

- ▷ Analyse de la différence entre taux d'activation correcte et incorrecte, pour la syllabe 1

$$I_{27} < C_3 > * R_2$$

Effet significatif ( $F_{(1,78)} = 141.14 ; p < .0001$ ) – Taux d'activation correcte (74.2%) ; taux d'activation incorrecte (10.4%)

- ▷ Analyse de la différence entre taux d'activation correcte et incorrecte, pour la syllabe 2

$$I_{27} < C_3 > * R_2$$

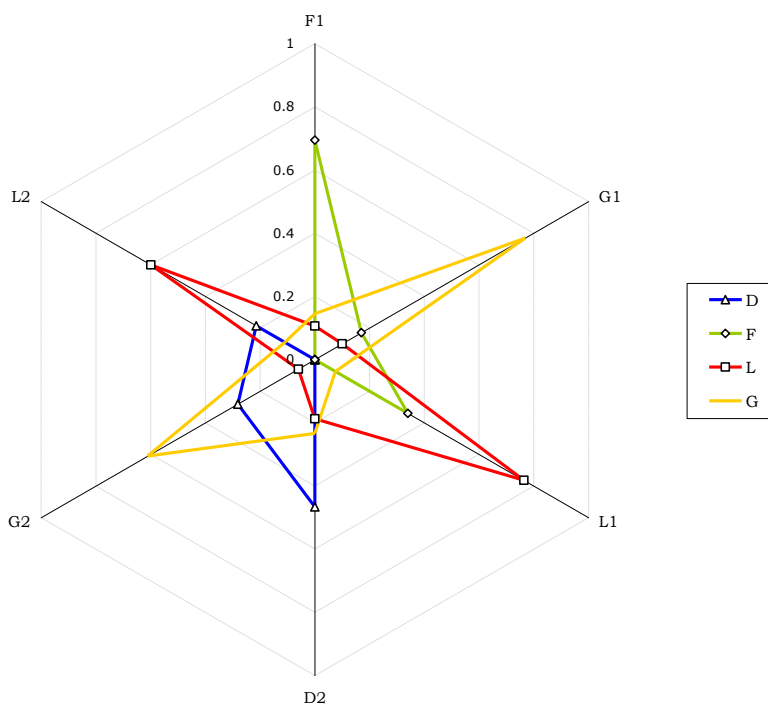
Effet significatif ( $F_{(1,78)} = 52.6 ; p < .0001$ ) – Taux d'activation correcte (55.8%) ; taux d'activation incorrecte (12.1%)

Le réseau présente un comportement classificateur avéré, qui lui permet, à partir de l'information prosodique, de retrouver les catégories théoriques dans une proportion significativement différente du hasard. L'hypothèse à la base de la Simulation 1 est donc validée par ces premiers résultats. Le graphique présenté dans la Figure 40 montre les taux de classification correcte pour les différentes catégories de syllabes (avec distinction des positions).

Figure 40 : Performances moyennes du classificateur

**Répartition des taux moyens d'activation des unités de sortie en fonction du type de syllabe et de sa position dans l'item**

	D	F	L	G
F1	0	0.69	0.11	0.15
G1	0	0.17	0.1	0.77
L1	0	0.34	0.76	0.07
D2	0.47	0	0.19	0.23
G2	0.28	0	0.06	0.61
L2	0.21	0	0.6	0.11



## 4.1.2.2.4.5.2 Rapport des analyses multiniveaux et présentation des résultats

Trois analyses sont effectuées pour explorer séparément :

- ▷ L'activation de sortie correspondant à une réponse correcte
- ▷ L'activation de sortie correspondant à une réponse incorrecte<sup>102</sup>
- ▷ La comparaison spécifique des taux d'activation de sortie correspondant aux confusions entre catégories attendues et catégories effectivement reconnues

Ces 3 analyses sont effectuées à l'aide de 2 modèles. Dans chacun des 2 modèles, les trois types suivants d'information sont analysés :

- ▷ L'activation des unités de sortie en rapport avec le type de syllabe et leur position dans l'item (dans le modèle :  $F_1, L_1, G_1, D_2, L_2, G_2$ )
- ▷ L'activation des unités de sortie en rapport avec le type de frontière (dans le modèle : FD, FG, FL, GD, GG, GL, LD, LG, LL)
- ▷ L'activation des unités de sortie en rapport avec l'interaction entre les deux facteurs précédents (exemple :  $FD.F_1$ )

Cependant, tous ces points ne seront pas détaillés dans le développement, l'information étant disponible dans les équations des 2 modèles.

1. Analyse de l'activation des unités de sortie du réseau, pour les réponses correctes et pour les réponses incorrectes

Les équations du modèle étudié (Équation 22 et Équation 23) peuvent être consultées en Annexe 4, page 464 et 465.

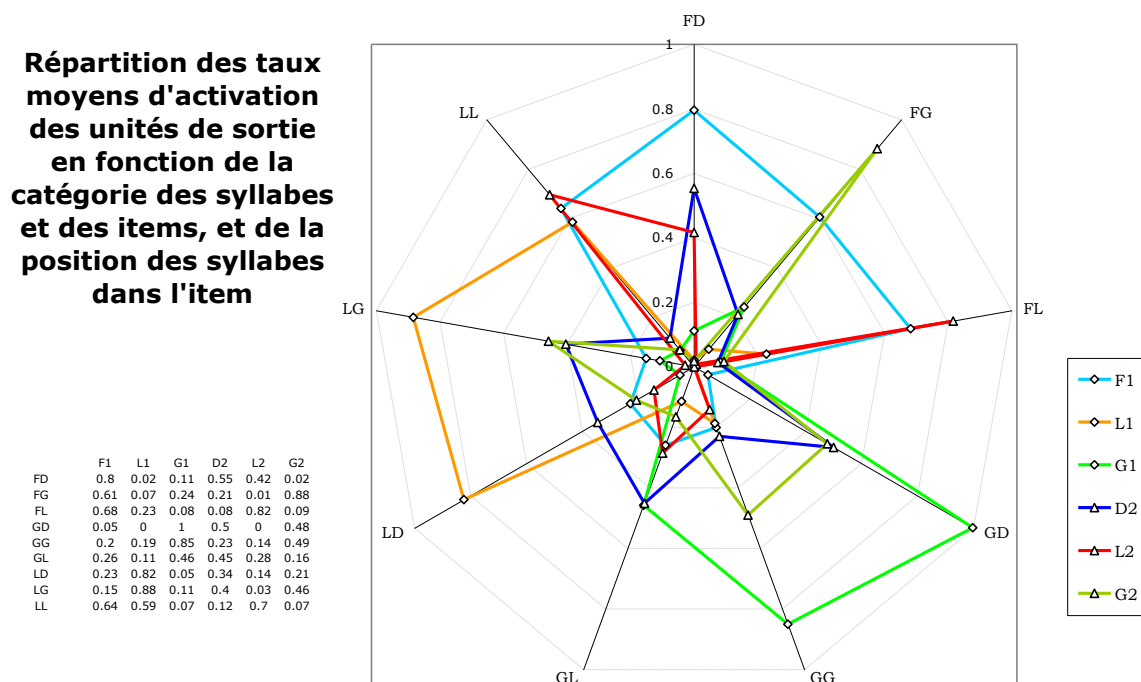
Tableau 26 : Taux moyens d'activation des unités de sortie (réponses de classification)

Frontière	Catégories correctes						Moy item	Catégories incorrectes						Moy item
	F <sub>1</sub>	G <sub>1</sub>	L <sub>1</sub>	D <sub>2</sub>	G <sub>2</sub>	L <sub>2</sub>		F <sub>1</sub>	G <sub>1</sub>	L <sub>1</sub>	D <sub>2</sub>	G <sub>2</sub>	L <sub>2</sub>	
<b>FD</b>	<b>0.8</b>	-	-	<b>0.55</b>	-	-	0.675	-	0.11	0.02	-	0.02	0.42	0.143
<b>FG</b>	<b>0.61</b>	-	-	-	<b>0.88</b>	-	0.745	-	0.24	0.07	0.21	-	0.01	0.133
<b>FL</b>	<b>0.68</b>	-	-	-	-	<b>0.82</b>	0.75	-	0.08	0.23	0.08	0.09	-	0.12
<b>GD</b>	-	<b>1</b>	-	<b>0.5</b>	-	-	0.75	0.05	-	0	-	0.48	0	0.133
<b>GG</b>	-	<b>0.85</b>	-	-	<b>0.49</b>	-	0.67	0.2	-	0.19	0.23	-	0.14	0.19
<b>GL</b>	-	<b>0.46</b>	-	-	-	<b>0.28</b>	0.37	0.26	-	0.11	<b>0.45</b>	0.17	-	0.248
<b>LD</b>	-	-	<b>0.82</b>	<b>0.35</b>	-	-	0.585	0.23	0.05	-	-	0.21	0.14	0.158
<b>LG</b>	-	-	<b>0.88</b>	-	<b>0.46</b>	-	0.67	0.15	0.11	-	0.4	-	0.03	0.173
<b>LL</b>	-	-	<b>0.59</b>	-	-	<b>0.7</b>	0.645	<b>0.64</b>	0.07	-	0.12	0.07	-	0.225

Rouge gras : catégorie attribuée par le réseau ; surligné : erreur ; autre : activité des autres unités de sortie

<sup>102</sup> Ces 2 analyses sont regroupées dans une même section afin de mettre en regard leurs résultats.

Figure 41 : Taux moyens d'activation des unités de sortie en fonction du type de frontière



Le réseau classe correctement les syllabes pour chaque type de frontière, sauf pour deux catégories de syllabes ( $L_1$  et  $L_2$ ), attribuées à deux autres :

- ▷  $L_1$  de LL est confondu avec  $F_1$ , LL tendant à être classé FL.
- ▷  $L_2$  de GL est confondu avec  $D_2$ , GL tendant à être classé GD.

Ce phénomène peut être dû :

- ▷ à la quantité moindre d'exemples  $F_1$  et  $D_2$ , qui cause une défaillance de performance pour ces catégories
- ▷ à une variabilité plus élevée de L
- ▷ à la combinaison de ces 2 facteurs, conduisant le réseau à assimiler  $L_1$  à  $F_1$  et  $L_2$  à  $D_2$ .

Tableau 27 : Test des contrastes entre taux de classification des syllabes d'une même catégorie à une même position, pour les réponses correctes uniquement

	p<	Chi <sup>2</sup> (1, N=648)		p<	Chi <sup>2</sup> (1, N=648)
<b>F<sub>1</sub> (FD)/F<sub>1</sub> (FG)</b>	.26	1.3	<b>D<sub>2</sub> (FD)/D<sub>2</sub> (GD)</b>	.87	0.0
<b>F<sub>1</sub> (FD)/F<sub>1</sub> (FL)</b>	.60	0.3	<b>D<sub>2</sub> (FD)/D<sub>2</sub> (LD)</b>	.20	1.6
<b>F<sub>1</sub> (FG)/F<sub>1</sub> (FL)</b>	.55	0.4	<b>D<sub>2</sub> (GD)/D<sub>2</sub> (LD)</b>	.27	1.2
<b>G<sub>1</sub> (GD)/G<sub>1</sub> (GG)</b>	.26	1.3	<b>G<sub>2</sub> (FG)/G<sub>2</sub> (GG)</b>	<b>.01</b>	7.1
<b>G<sub>1</sub> (GD)/G<sub>1</sub> (GL)</b>	<b>.0001</b>	15.4	<b>G<sub>2</sub> (FG)/G<sub>2</sub> (LG)</b>	<b>.01</b>	7.8
<b>G<sub>1</sub> (GG)/G<sub>1</sub> (GL)</b>	<b>.01</b>	7.8	<b>G<sub>2</sub> (GG)/G<sub>2</sub> (LG)</b>	.90	0.0
<b>L<sub>1</sub> (LD)/L<sub>1</sub> (LG)</b>	.69	0.2	<b>L<sub>2</sub> (FL)/L<sub>2</sub> (GL)</b>	<b>.0002</b>	15.1
<b>L<sub>1</sub> (LD)/L<sub>1</sub> (LL)</b>	.09	3.0	<b>L<sub>2</sub> (FL)/L<sub>2</sub> (LL)</b>	.29	1.1
<b>L<sub>1</sub> (LG)/L<sub>1</sub> (LL)</b>	<b>.03</b>	4.5	<b>L<sub>2</sub> (GL)/L<sub>2</sub> (LL)</b>	<b>.005</b>	7.9

Selon les résultats d'analyse présentés dans le Tableau 27 ci-dessus, il existe des variations de détection d'une catégorie donnée de syllabe à une position donnée dans l'item en fonction du type de frontière. Trois catégories de frontières (FG, GL et LL) sont responsables de variations significatives sur la performance de détection des catégories de syllabes.

## 2. Analyse des confusions du réseau

L'équation du modèle (Équation 24) peut être consultée en Annexe 4, page 466.

Tableau 28 : Récapitulatif des confusions observées

	<b>F<sub>1</sub></b>	<b>G<sub>1</sub></b>	<b>L<sub>1</sub></b>	<b>D<sub>2</sub></b>	<b>G<sub>2</sub></b>	<b>L<sub>2</sub></b>
<b>FD</b>	-	-	-	<b>0.55</b>	-	<i>0.42</i>
<b>FG</b>	-	-	-	-	-	-
<b>FL</b>	-	-	-	-	-	-
<b>GD</b>	-	-	-	<b>0.5</b>	<i>0.48</i>	-
<b>GG</b>	-	-	-	-	-	-
<b>GL</b>	-	-	-	<i>0.45</i>	-	<b>0.28</b>
<b>LD</b>	-	-	-	<b>0.35</b>	<i>0.21</i>	-
<b>LG</b>	-	-	-	<i>0.4</i>	<b>0.46</b>	-
<b>LL</b>	<i>0.64</i>	-	<b>0.59</b>	-	-	-

Gras : taux d'activation pour la réponse correcte ; italique : taux d'activation pour une autre réponse.

Tableau 29 : Test de significativité des contrastes entre conditions

	<b>p&lt;</b>	<b>Chi<sup>2</sup><sub>(1, N=648)</sub></b>
<b>D<sub>2</sub> (FD)/L<sub>2</sub> (FD)</b>	.17	1.9
<b>D<sub>2</sub> (GD)/G<sub>2</sub> (GD)</b>	.35	0.9
<b>L<sub>2</sub> (GL)/D<sub>2</sub> (GL)</b>	.13	2.3
<b>D<sub>2</sub> (LD)/G<sub>2</sub> (LD)</b>	.13	2.3
<b>G<sub>2</sub> (LG)/D<sub>2</sub> (LG)</b>	.60	0.3
<b>L<sub>1</sub> (LL)/F<sub>1</sub> (LL)</b>	.89	0.0

Le réseau confond certaines catégories de syllabes, puisque les 6 contrastes testés sont non significatifs. Les catégories suivantes sont surtout confondues :

- ▷ D<sub>2</sub> avec G<sub>2</sub> (50% des cas de confusion – écart moyen d'activation : 0.07)
- ▷ D<sub>2</sub> avec L<sub>2</sub> (33% des cas – 0.15)

La syllabe D serait donc plus facile à confondre avec d'autres syllabes, tandis que les catégories de syllabes seraient plus difficiles à détecter en deuxième position (une seule confusion apparaît sur S<sub>1</sub>, contre 5 sur S<sub>2</sub>).

Toutefois, sur la base des données à disposition, l'erreur de classification ne se produit que dans 2 cas (FL au lieu de LL, GD au lieu de GL).

Le phénomène de confusion, impliquant F et D, pourrait être lié au nombre 2 fois moindre d'exemples de ces 2 catégories dans le matériel expérimental. Cependant, sur la

base de la Figure 40, page 180, la comparaison des scores d'activation de  $F_1$  et  $D_2$  révèle une moins bonne performance pour les syllabes  $D_2$  que pour les syllabes  $F_1$ , à nombre équivalent d'exemples.

Ce phénomène de confusion doit donc combiner 2 événements adverses : une quantité moindre d'exemples et la structure des données (position variable du début de la 2<sup>ème</sup> syllabe sur la couche d'entrée du réseau).

On ne peut donc conclure à une moins bonne caractérisation des syllabes D.

### Synthèse des données analysées

Pour interpréter les réponses du réseau (niveaux d'activation entre 0 et 1) en termes de catégorie selon une procédure permettant la comparaison avec les 2 recherches présentées en introduction (Tableau 6, page 58), une règle de codage doit être définie. Les réponses sont codées 0/1, avec 1 attribué à la valeur la plus élevée des 4 unités de sortie pour chaque syllabe, 0 pour les 3 autres.

Tableau 30 : Présentation des taux de classification, fiabilité et certitude par syllabe et par item, obtenus dans la Simulation 1

		<b>Taux de classification</b> [Nombre d'exemples de la catégorie correctement classés] / [Nombre total d'exemples présentés de cette catégorie]	<b>Taux de fiabilité</b> [N exemples de la catégorie correctement classés] / [(N exemples de la catégorie correctement classés) + (N éléments classés par erreur dans catégorie considérée)]	<b>Certitude</b> [Taux de classification] * [Taux de fiabilité] / 100
<b>Syllabe 1</b>	<b>F<sub>1</sub></b>	77.78	77.78	60.49
	<b>G<sub>1</sub></b>	74.07	74.07	54.87
	<b>L<sub>1</sub></b>	74.07	74.07	54.87
	<b>Moyenne (%)</b>	<b>75.31</b>	<b>75.31</b>	<b>56.74</b>
<b>Syllabe 2</b>	<b>D<sub>2</sub></b>	55.56	55.56	30.86
	<b>G<sub>2</sub></b>	66.67	66.67	44.44
	<b>L<sub>2</sub></b>	66.67	66.67	44.44
	<b>Moyenne (%)</b>	<b>62.97</b>	<b>62.96</b>	<b>39.92</b>
<b>Item</b>	<b>FD</b>	66.67	63.64	42.43
	<b>FG</b>	66.67	75.00	50.00
	<b>FL</b>	55.56	38.46	21.37
	<b>GD</b>	55.56	55.56	30.87
	<b>GG</b>	44.44	36.36	16.16
	<b>GL</b>	22.22	40.00	8.89
	<b>LD</b>	44.44	44.44	19.75
	<b>LG</b>	44.44	40.00	17.78
	<b>LL</b>	22.22	66.67	14.81
	<b>Moyenne (%)</b>	<b>46.91</b>	<b>51.13</b>	<b>24.67</b>

On relève que :

- ▷ Les taux de classification sont meilleurs et plus fiables pour les syllabes en première position dans les items, que pour les syllabes en deuxième position.
- ▷ Les performances sont meilleures avec les catégories FG, FD et GD (mieux caractérisées)
- ▷ Les performances sont les pires pour les catégories GG, LL et GL (moins caractérisées)
- ▷ Les performances du réseau ne sont globalement pas si bonnes (38 items classés correctement sur 81 (46.91%)), avec une certitude de 25% environ. Toutefois, la source d'information supplémentaire que fournit la prosodie n'est pas négligeable, et il se peut qu'elle suffise à lever les ambiguïtés de segmentation lexicale. Deux remarques doivent être ajoutées :
  - ∩ les items ne proviennent pas de régions ambiguës, d'où peut-être, ces scores de classification bas. On peut en effet s'attendre à ce que la prosodie ait des caractéristiques démarcatives plus accentuées en cas d'ambiguïté. Les performances seraient alors meilleures.
  - ∩ La règle utilisée pour attribuer une catégorie à partir des scores d'activation n'est probablement pas celle qui est utilisée pour exploiter l'information prosodique dans le système naturel (plus vraisemblablement intégrée comme une pondération supplémentaire dans les réseaux de traitement).

#### **4.1.2.2.2.4.6 Conclusions relatives à la Simulation 1**

- ▷ Le réseau présente un comportement classificateur avéré à partir de l'information prosodique (taux moyen d'activation correcte : 0.65 ; taux moyen d'activation incorrecte : 0.11). L'hypothèse de l'étude est donc validée par les résultats.
- ▷ Puisque le réseau retrouve les catégories théoriques des syllabes à partir de l'information prosodique dans une proportion différente du hasard, on peut conclure que les profils prosodiques du matériel expérimental sont assez caractérisés (éloignés les uns des autres, homogènes au sein d'une catégorie).
- ▷ Les performances du réseau pour la classification d'un type de syllabe sont modulées en fonction du type d'item (exemple : unité G activée à 100% pour GD, à 46% pour GL).
- ▷ Elles pourraient certainement être améliorées avec l'augmentation du nombre d'exemples (faible : 6 par catégorie pour l'apprentissage), et l'utilisation d'un nombre homogène de cas pour toutes les catégories lors de l'apprentissage (voir ci-dessous).
- ▷ La dégradation des performances parfois observée dans les analyses (excitabilité différente de  $F_1$  et de  $D_2$  par rapport aux autres syllabes (se reporter aux résultats complémentaires d'analyse présentés dans l'Annexe 5 page 467), l'asymétrie des performances entre  $S_1$  et  $S_2$  (plus de confusions impliquant  $S_2$  et  $D_2$ ), trouve une explication dans 2 biais :
  - ∩ Le nombre d'exemples d'apprentissage, 2 fois moindre pour les syllabes de type F (seulement en  $S_1$ ) et D (seulement en  $S_2$ ), par rapport aux deux autres catégories de syllabes (en  $S_1$  et en  $S_2$ ).
  - ∩ La structure des données, avec une position variable du début de la seconde syllabe sur les unités de la couche d'entrée, contre une position fixe pour le début des  $S_1$ , qui complique le problème de classification.

Les 2 études du CLIPS - IMAG & IPC présentées en introduction (Section 2.2.3 page 56) et la présente étude, réalisée sur un autre matériel expérimental, montrent qu'environ la moitié des syllabes peut être catégorisée correctement à partir de leur prosodie, avec certitude pour environ 1 syllabe sur 4. Les résultats sont récapitulés dans le Tableau 31.

Tableau 31 : Performance de la classification des syllabes par leur prosodie, selon différentes méthodes

		Détection	Fiabilité	Certitude
<b>CLIPS - IMAG &amp; IPC</b>	<b>Algorithmes</b>	32.7	58.3	20.8
	<b>Expert humain</b>	43.7	42.9	25.2
<b>Simulation 1</b>		46.9	51.1	24.7

#### **4.1.2.2.2.5 Simulation 2 (discrimination d'items bisyllabiques)**

La Simulation 2 a pour buts de :

- ▷ définir mathématiquement le profil de réponse auquel on peut s'attendre avec une tâche de comparaison à partir d'un traitement de type « sensoriel »
- ▷ déterminer de quels facteurs dépend la performance dans la tâche de comparaison

Les réseaux utilisés ici sont parfois dénommés « Compateur » dans le texte, mais il s'agit réellement de classificateurs. La situation est structurée de manière à reproduire artificiellement un comportement de discrimination d'items, tel que celui mis en œuvre par les auditeurs humains dans le test perceptif de comparaison (Expérience 1). Toutefois, mathématiquement, il s'agit toujours d'un problème de classification (attribuer une paire d'items à la catégorie « même » ou à la catégorie « différent »).

##### **4.1.2.2.2.5.1 Hypothèse**

Les items de même catégorie théorique (morphosyntaxe \* topologie) sont considérés comme similaires avec un taux supérieur aux items de catégories différentes à partir de l'information prosodique (durée, f0).

##### **4.1.2.2.2.5.2 Caractéristiques et paramètres du réseau employé pour la Simulation 2**

- ▷ Couche d'entrée : 132 unités
- ▷ Couche de sortie : 1 unité
- ▷ Unités de traitement : 10
- ▷ Connexions : 1507
- ▷ Apprentissage supervisé. Mêmes algorithmes et paramètres que pour la Simulation 1 (voir Section 4.1.2.2.2.4.2, page 177).

**4.1.2.2.2.5.3 Organisation de la simulation**

Les réseaux sont structurés et testés avec les combinaisons 2 à 2 des 81 items du corpus. Cette combinatoire s'effectue à 2 niveaux :

- ▷ Niveau des items (matrice générale)
- ▷ Niveau des catégories d'items (sous matrices)

Au niveau des items, toutes les combinaisons ne sont pas considérées, puisque l'ordre de présentation n'est pas jugé pertinent ( $item_1 / item_2 = item_2 / item_1$ ). En effet, dans le test perceptif, les auditeurs peuvent réécouter à volonté une combinaison (l' $item_1$  se retrouve alors après l' $item_2$ ). Ceci élimine 3240 doublons ; on utilise donc  $81 * 81 - 3240 = 3321$  combinaisons d'items.

Au niveau des catégories d'items, il existe 45 combinaisons possibles ( $9 * 9 - 36$  doublons), dont 36 correspondent à des comparaisons inter catégories (exemple : FD / FG) et 9 à des comparaisons intra catégorie (exemple : FD / FD).

Pour balayer l'ensemble des combinaisons d'items de manière équilibrée, 6 réseaux sont nécessaires, tel qu'expliqué ci-dessous. Pour  $n = 36$  comparaisons d'items<sup>103</sup>,  $N = n / 4$  est le nombre de combinaisons à répartir dans 4 listes (4 est l'un des diviseurs optimaux de 36 – Tableau 32).

**Tableau 32 : Répartition des combinaisons d'items dans chaque liste**

<b>L<sub>1</sub></b>	<b>L<sub>2</sub></b>	<b>L<sub>3</sub></b>	<b>L<sub>4</sub></b>
1_2	1_3	1_4	1_5
1_6	1_7	1_8	1_9
2_3	2_4	2_5	2_6
2_7	2_8	2_9	3_4
3_5	3_6	3_7	3_8
3_9	4_5	4_6	4_7
4_8	4_9	5_6	5_7
5_8	5_9	6_7	6_8
6_9	7_8	7_9	8_9

Avec [ $n^{\circ}$  item 1\_  $n^{\circ}$  item 2], l'attribution initiale d'un numéro à chaque item étant aléatoire (ordre de tirage de l'item).

Ces 4 groupes, employés en combinatoire (2 groupes pour l'apprentissage, 2 groupes complémentaires pour la généralisation), conduisent à 6 possibilités, chacune explorée avec 1 réseau selon le plan développé dans le tableau ci-dessous.

<sup>103</sup>  $9 * 9 - 36$  (doublons par l'ordre de présentation) – 9 (comparaison d'un item avec lui-même) = 36. Ce chiffre (36) est lié à une limitation imposée par l'existence de sous matrices triangulaires dans la matrice générale (comparaisons intra catégorie).



Tableau 33 : Répartition des listes de comparaisons d'items par réseau

Réseau	Apprentissage	Test
R <sub>1</sub>	L <sub>1</sub> L <sub>2</sub>	L <sub>3</sub> L <sub>4</sub>
R <sub>2</sub>	L <sub>3</sub> L <sub>4</sub>	L <sub>1</sub> L <sub>2</sub>
R <sub>3</sub>	L <sub>1</sub> L <sub>4</sub>	L <sub>2</sub> L <sub>3</sub>
R <sub>4</sub>	L <sub>2</sub> L <sub>3</sub>	L <sub>1</sub> L <sub>4</sub>
R <sub>5</sub>	L <sub>1</sub> L <sub>3</sub>	L <sub>2</sub> L <sub>4</sub>
R <sub>6</sub>	L <sub>2</sub> L <sub>4</sub>	L <sub>1</sub> L <sub>3</sub>

Avec R pour réseau, L pour liste de combinaisons d'items.

On obtient donc :

- ▷ 6 ensembles de 1863 combinaisons pour l'apprentissage (1620 comparaisons inter catégories + 81 comparaisons d'un item avec lui-même + 162 comparaisons intra catégorie).
- ▷ 6 ensembles de 1458 combinaisons pour la généralisation (1296 comparaisons inter catégories complémentaires des 1620 utilisées dans le fichier d'apprentissage + 162 comparaisons intra catégorie complémentaires).

#### **4.1.2.2.2.5.4 Procédure d'apprentissage et de recueil des données**

Après localisation de l'optimum de généralisation de chaque réseau et de son point de convergence dans le cycle d'apprentissage, chaque comparateur est élaboré sur 10 cycles d'apprentissage. Théoriquement, l'optimum de performance est atteint lorsque l'erreur globale du réseau sur les données de validation ou de généralisation cesse de diminuer ou se remet à croître (selon le lot de données utilisé ici, entre 3 et 5 cycles d'entraînement). Pratiquement, les performances apparaissent meilleures entre 6 et 10 cycles (observation détaillée des réponses de généralisation), donc, avec un léger surentraînement (la convergence du réseau a lieu vers 16 cycles (carré moyen des erreurs désiré = 0.0001)).

À partir du 7<sup>ème</sup> cycle d'apprentissage et après chaque cycle d'apprentissage (1 entraînement sur la totalité des données), le réseau est testé sur les exemples de généralisation (le test ne modifie pas le réseau).

Les données recueillies sont donc :

- ▷ 4 réponses (cycles 7 à 10) par réseau pour une même combinaison d'items (un comportement non linéaire, rare et ponctuel, a pu être remarqué d'un cycle à l'autre pour certaines combinaisons d'items. Ce bruit est amoindri grâce à plusieurs relevés par combinaison)
- ▷ 3 réponses de 3 réseaux différents pour une même combinaison d'item (bruit lié à répartition des items par réseau amoindri grâce à 3 relevés par combinaison d'items fournis par 3 réseaux différents)

Les données qui sont testées dans les modèles d'analyse correspondent donc à une moyenne par combinaison d'item pour les  $4 * 3 = 12$  recueils.

#### **4.1.2.2.2.5 Rapport des analyses statistiques et présentation des résultats**

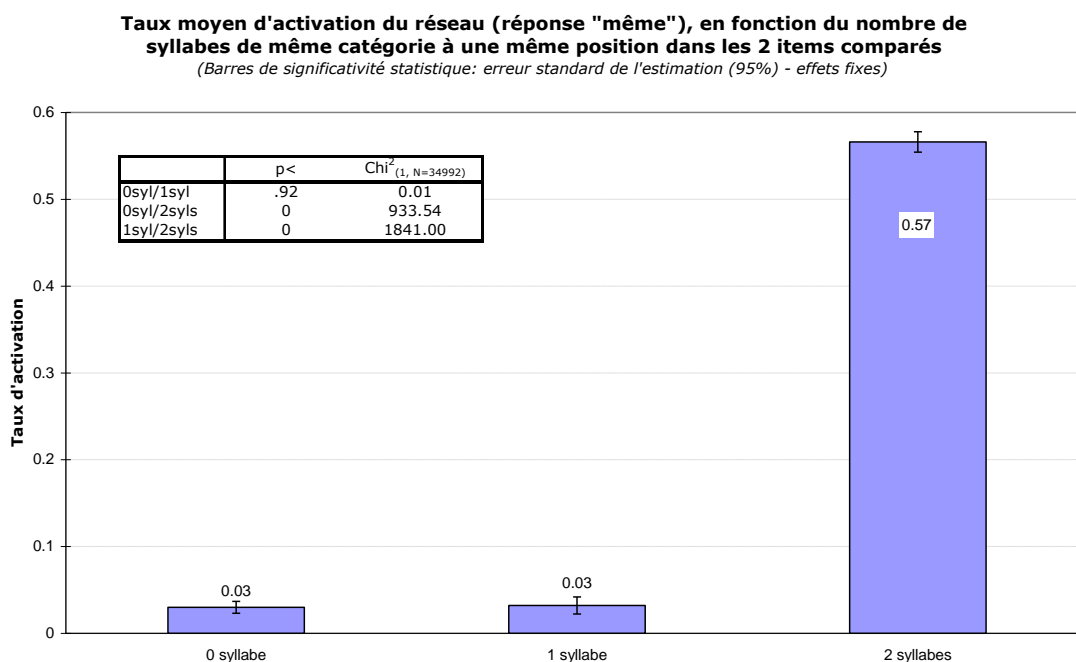
Le but des analyses statistiques présentées dans cette section est de déterminer les facteurs dont dépend la performance discriminatoire des réseaux. Une analyse de régression multiniveaux (mesures, combinaisons) sur les données fournies par les comparateurs est effectuée à l'aide de 3 modèles. L'effet de 3 types d'information sur la performance discriminatoire des réseaux, est analysé dans les 3 modèles (1 information par modèle) :

1. 1<sup>er</sup> modèle : quantité de syllabes similaires à une même position dans les deux items comparés
2. 2<sup>ème</sup> modèle : type de syllabe en commun à une même position dans les deux items comparés
3. 3<sup>ème</sup> modèle : type d'item 1 et 2 (type de frontière) comparés

1. Effet de la quantité de syllabes similaires (similitudes à une même position) dans les deux items comparés sur la performance de discrimination des items par les réseaux

L'équation du modèle (Équation 25) peut être consultée en Annexe 6, page 470. La Figure 42 présente le degré d'activation moyen des réseaux sur l'unité de sortie pour chaque cas, et la significativité des écarts obtenus lors des tests de contraste.

**Figure 42 : Taux moyen d'activation des unités de sortie estimés par le modèle en fonction du nombre de similitudes catégorielles et positionnelles**



Lorsque la similitude entre les 2 items comparés est maximale (même catégorie de syllabe à la même position dans les deux items, pour les 2 syllabes constitutives de chaque item), les taux d'activation sont significativement plus élevés que lorsqu'une ou 2 des syllabes diffèrent entre les 2 items comparés. Dans ces 2 derniers cas, aucune différence significative d'activation n'est obtenue en sortie.

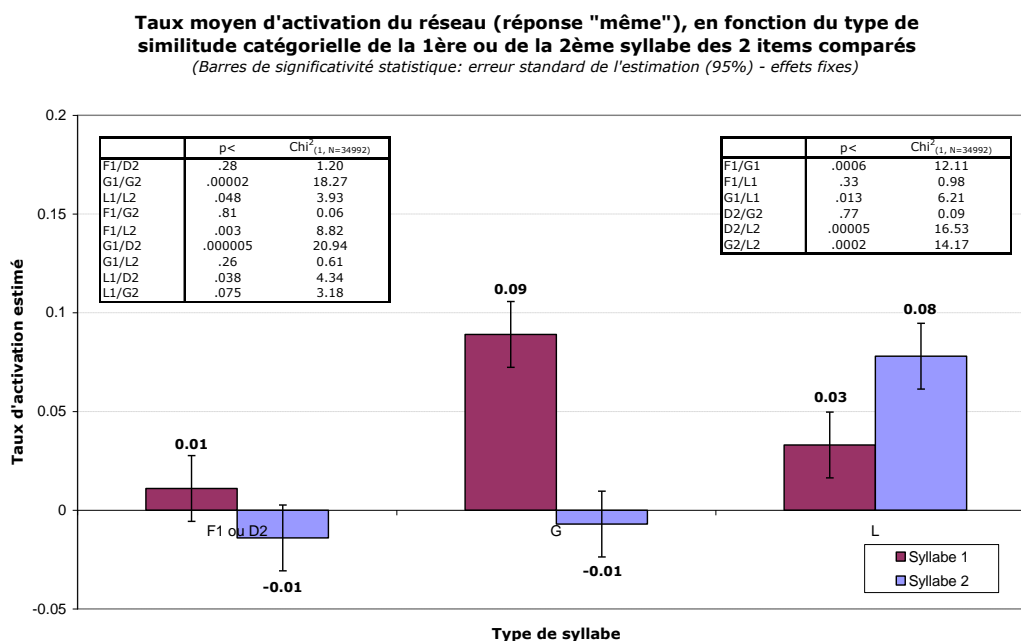
Le réseau a donc bien résolu le problème de comparaison au niveau de l'item, et non au niveau de ses composants syllabiques. Le comparateur réussit à effectuer une classification correcte des catégories théoriques à partir de l'information prosodique selon un simulacre de la tâche de comparaison. Ce résultat est à considérer comme une validation mathématique de la faisabilité de la tâche dans un test perceptif, avec ce matériel expérimental.

2. 2<sup>ème</sup> modèle : Effet du type de syllabe en commun à une même position dans les deux items comparés, sur la performance de discrimination des items par les réseaux

L'équation du modèle (Équation 26) peut être consultée en Annexe 6, page 470.

- ▷ Une seule syllabe en commun : La Figure 43 présente la part d'activation de sortie estimée due à la similitude catégorielle des syllabes à une même position dans les 2 items comparés, et les seuils de significativité des écarts obtenus lors des tests de contraste.

Figure 43 : Taux d'activation de sortie estimé par le modèle, lié à l'effet de la similitude catégorielle d'une syllabe à la même position dans les 2 items comparés

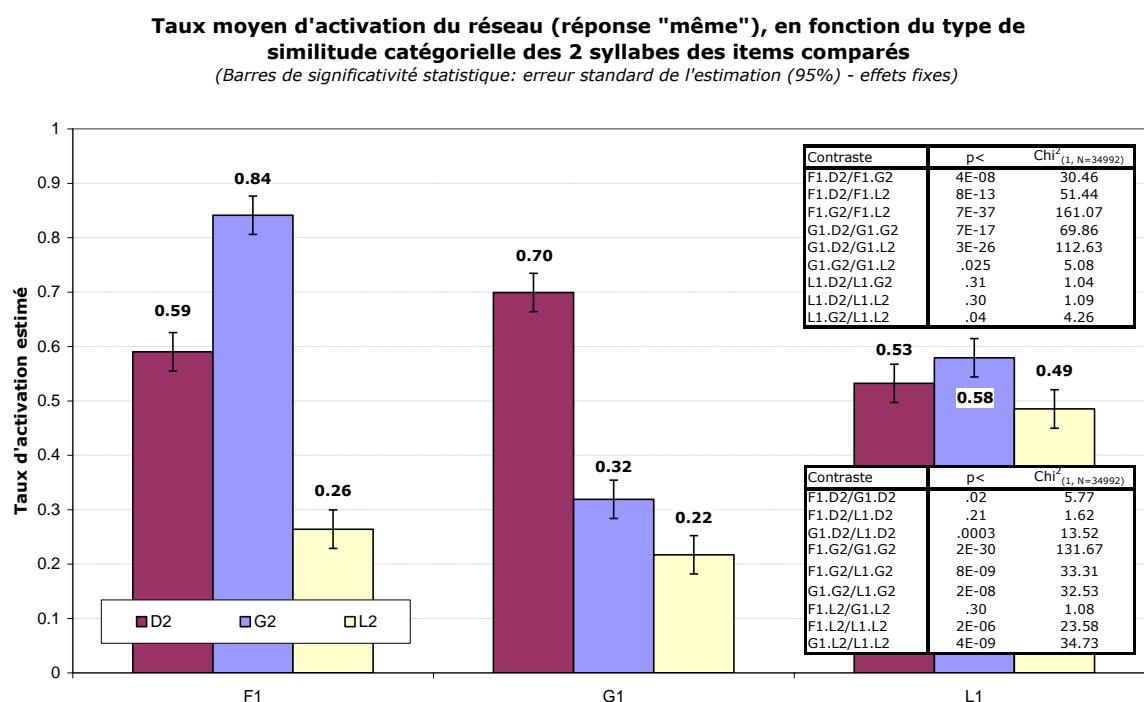


Les taux d'activation de sortie sont en moyenne augmentés lorsque la première syllabe des items comparés est de catégorie G, et lorsque la deuxième syllabe des items comparés est de catégorie L. Toutefois, ces écarts sont quantitativement très faibles.

▷ Deux syllabes en commun (interaction  $\text{syllabe}_1 * \text{syllabe}_2$ ) :

La Figure 44 présente simultanément la part d'activation de sortie estimée due à la similitude des 2 syllabes entre items selon leur catégorie, et les seuils de significativité des écarts obtenus lors des tests de contraste.

Figure 44 : Taux d'activation de sortie estimé par le modèle, lié à l'effet de la similitude catégorielle d'une syllabe à la même position dans les 2 items comparés



Le taux d'activation de sortie varie en fonction de la similitude catégorielle des 2 syllabes des items comparés. Certaines catégories d'items sont en conséquence bien catégorisées lors de la comparaison (F<sub>1</sub>.G<sub>2</sub> et G<sub>1</sub>.D<sub>2</sub>), d'autres sont correctement catégorisées (F<sub>1</sub>.D<sub>2</sub>, L<sub>1</sub>.D<sub>2</sub>, L<sub>1</sub>.G<sub>2</sub>, L<sub>1</sub>.L<sub>2</sub>), et d'autres sont mal catégorisées (F<sub>1</sub>.L<sub>2</sub>, G<sub>1</sub>.G<sub>2</sub> et G<sub>1</sub>.L<sub>2</sub>). Ces résultats indiquent que les syllabes possèdent donc une caractérisation acoustique plus ou moins ciblée en fonction de leurs catégories morphosyntaxiques et topologiques.

### 3. 3<sup>ème</sup> modèle : Effet du type de frontières 1 et 2 sur la performance de discrimination des items par les réseaux

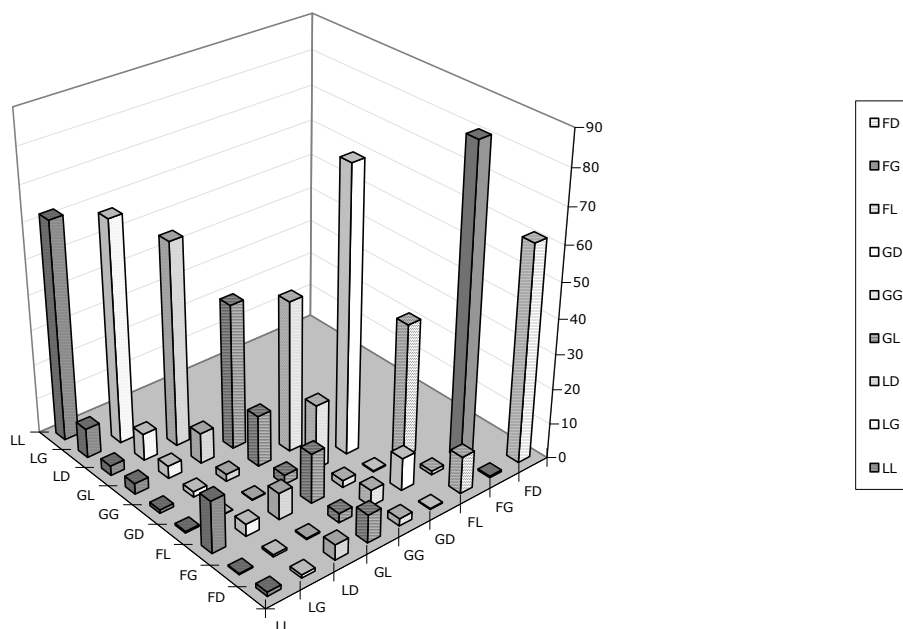
L'équation du modèle (Équation 27) peut être consultée en Annexe 6, page 471.

Tableau 34 : Récapitulation des taux moyens d'activation de l'unité de sortie du réseau estimés par le modèle, par comparaison de catégorie de frontière

	FD	FG	FL	GD	GG	GL	LD	LG	LL
FD	<b>61.7</b>								
FG	0.5	<b>87.5</b>							
FL	10.6	1.1	<b>38.2</b>						
GD	0.3	9.3	0.3	<b>80.5</b>					
GG	2.1	4.7	2	18.5	<b>43.1</b>				
GL	8	2.7	14.3	2.7	14.6	<b>41.4</b>			
LD	4.4	0.4	7.6	0.2	2.1	8.7	<b>58.1</b>		
LG	0.9	0.6	3.5	0.1	1.6	3.6	7.7	<b>63.5</b>	
LL	1.2	0.3	14.9	0.5	1	3.1	2.8	8.4	<b>62.5</b>

Figure 45 : Représentation graphique des données du Tableau 34

Taux moyens d'activation (%) par comparaison de catégorie de frontière (simulation 2)



Les taux d'activation sont toujours supérieurs dans le cas de comparaisons intra catégorie, que dans le cas de comparaisons inter catégories, confirmant la validité de l'hypothèse de l'étude. Cela signifie que les 9 frontières sont assez bien caractérisées et relativement homogènes, pour être à la fois regroupées par catégorie, et distinguées les unes des autres.

Le Tableau 35 présente les résultats de l'analyse des contrastes entre taux d'activation pour les comparaisons intra catégorie.

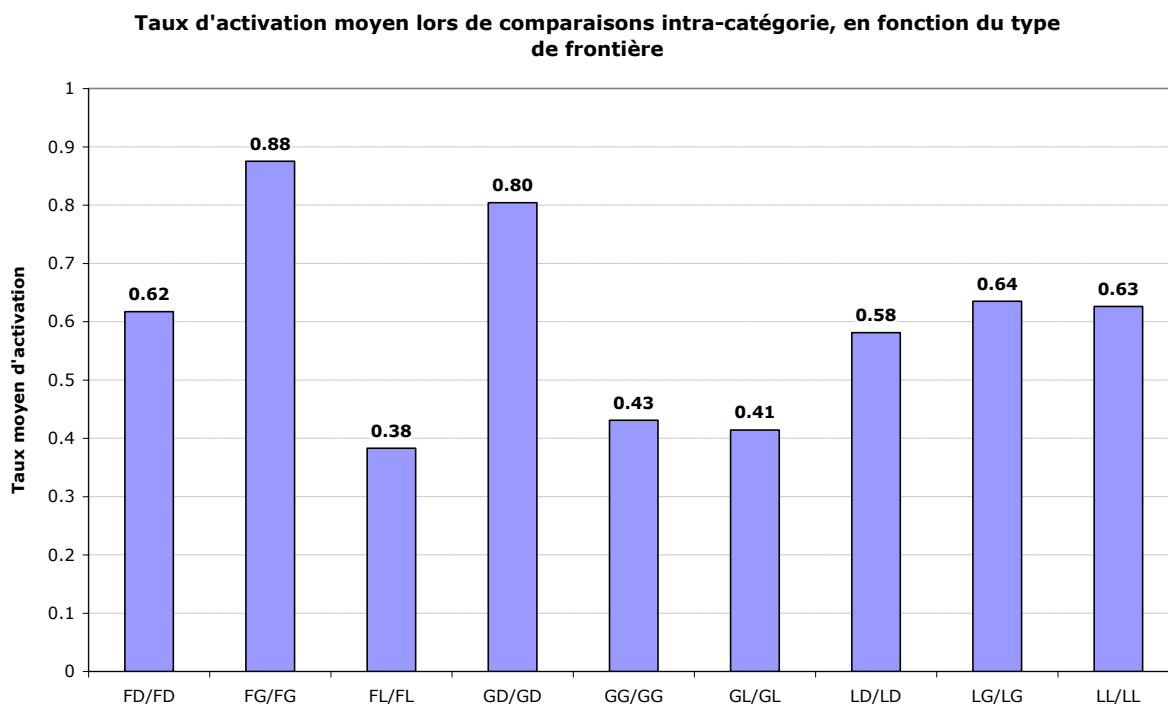
Tableau 35 : Test de significativité des contrastes entre conditions pour les comparaisons intra catégorie

Contraste	p<	Chi <sup>2</sup> (1, N=34992)	Contraste	p<	Chi <sup>2</sup> (1, N=34992)
FDFD/FGFG	<b>.0001</b>	44.3	FLFL/LDLD	<b>.0001</b>	26.4

Contraste	p<	Chi <sup>2</sup> <sub>(1, N=34992)</sub>	Contraste	p<	Chi <sup>2</sup> <sub>(1, N=34992)</sub>
FDFD/FLFL	.0001	36.8	FLFL/LGLG	.0001	42.5
FDFD/GDGD	.0001	23.5	FLFL/LLLL	.0001	39.4
FDFD/GGGG	.0001	23	GDGD/GGGG	.0001	93
FDFD/GLGL	.0001	27.7	GDGD/GLGL	.0001	102
FDFD/LDLD	.48	0.85	GDGD/LDLD	.0001	33.3
FDFD/LGLG	.65	0.21	GDGD/LGLG	.0001	19.3
FDFD/LLLL	.83	0.05	GDGD/LLLL	.0001	21.4
FGFG/FLFL	.0001	162	GGGG/GLGL	.64	0.21
FGFG/GDGD	.07	3.29	GGGG/LDLD	.0002	15
FGFG/GGGG	.0001	131	GGGG/LGLG	.0001	27.6
FGFG/GLGL	.0001	142	GGGG/LLLL	.0001	25.1
FGFG/LDLD	.0001	57.5	GLGL/LDLD	.0001	18.8
FGFG/LGLG	.0001	38.5	GLGL/LGLG	.0001	32.6
FGFG/LLLL	.0001	41.5	GLGL/LLLL	.0001	30
FLFL/GDGD	.0001	119	LDLD/LGLG	.17	1.89
FLFL/GGGG	.21	1.6	LDLD/LLLL	.26	1.29
FLFL/GLGL	.42	0.65	LGLG/LLLL	.81	0.06

Les taux d'activation pour les comparaisons intra catégorie varient d'une frontière à l'autre. Cette variation implique que certaines frontières sont mieux caractérisées que d'autres et/ou que les items de certaines frontières sont moins variables. Le profil de réponses qui peut être attendu pour le test perceptif (au cas où les participants répondent sur la base d'un traitement sensoriel de l'information) est renseigné dans la Figure 46.

Figure 46 : Taux d'activation en sortie de réseau, estimé par le modèle pour les comparaisons intra catégories, en fonction du type de frontière considéré



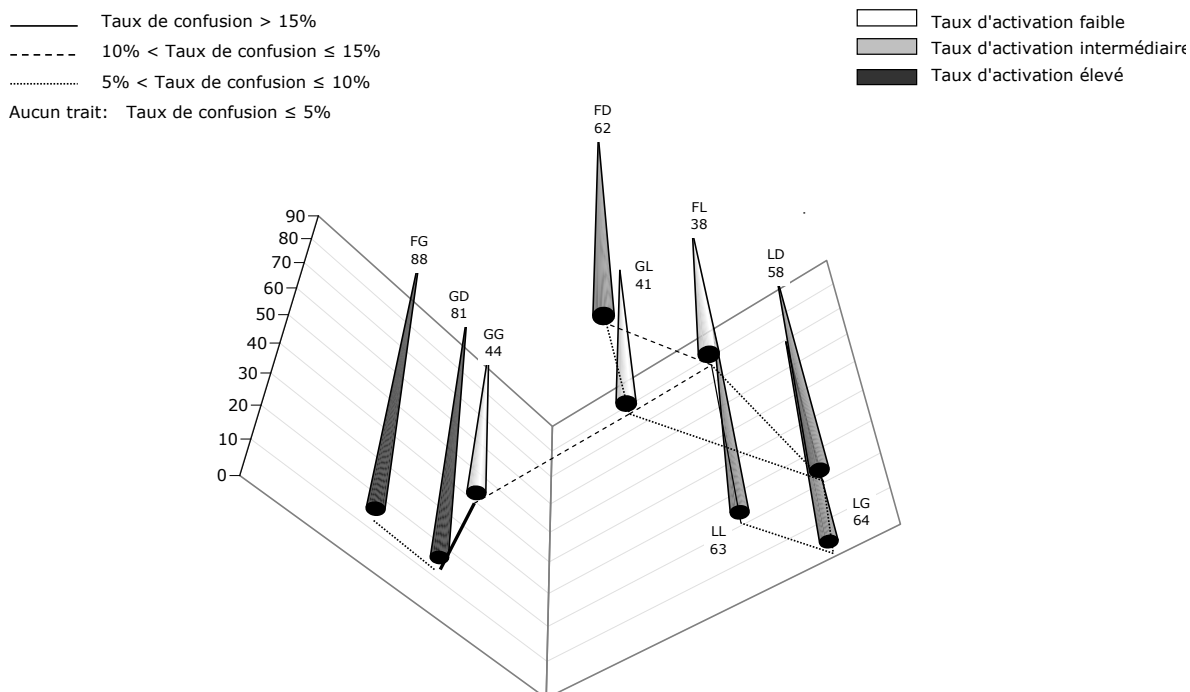
Les performances sont meilleures pour les comparaisons intra catégorie FG (87.5) et GD (80.5%).

FL (38.2%), GG (43.1%) et GL (41.4%) sont les frontières qui entraînent les moins bonnes performances.

Concernant les comparaisons inter catégories (analyses et graphiques non présentés), les performances varient également en fonction des frontières comparées (de 0.1% pour GD / LG à 18.5% pour GG / GD). Cette variation implique que certaines frontières sont apparentées (exemple : [GD / FG / GG]), ou encore, que certaines frontières sont mal caractérisées et présentent une variabilité importante transgressant leur champ catégoriel (FL et GL), alors que d'autres catégories sont fortement distinctes les unes des autres (plus fortement caractérisées, meilleure homogénéité interne).

La Figure 47 est une synthèse représentant une organisation spatiale de la caractérisation prosodique des catégories de frontières à partir de la performance du réseau pour regrouper les items d'une même catégorie (cônes), et des rapports de confusion entre catégories (traits reliant la base des cônes).

**Figure 47 : Organisation spatiale de la caractérisation prosodique des 9 catégories de frontière**



Sur la base de l'information combinée des taux de regroupement correct et des taux de confusion des catégories de frontière, il est possible de faire émerger 3 familles de frontières :

- ▷ Les frontières mal caractérisées (classe hétérogène et voisinage important)
  - ∩ FL et GL : taux de regroupement assez bas (respectivement 38.2 et 41.4%), confusion élevée entre elles (14.3%) et avec les autres catégories (jusqu'à 14.9%)
  - ∩ GG : taux de regroupement faible (43.1%), confusion élevée avec GD (18.5%) et GL (14.6%)
- ▷ Les frontières bien caractérisées (classe homogène et isolée)
  - FG et GD : taux de regroupement les plus élevés (respectivement 87.5 et 80.5%), taux de confusion bas. Proximité relative entre elles (taux de confusion de 9.3%).
- ▷ Les autres frontières, à score moyen : FD, LD, LG, LL

#### **4.1.2.2.5.6 Conclusions relatives à la Simulation 2**

- ▷ Ces résultats indiquent que la tâche de comparaison d'items est théoriquement faisable par des auditeurs humains.
- ▷ La tâche de comparaison permet de meilleures performances qu'une tâche de classification (respectivement : [max. = 88%, min. = 38%] et [max. = 67%, min. = 22%]). Son avantage avec une quantité faible d'exemples provient du fait qu'elle démultiplie les items par combinatoire.
- ▷ Les items des 9 frontières étudiées ont un profil prosodique caractéristique de la classe théorique à laquelle ils appartiennent (les 9 frontières sont distinctes les unes des autres), mais toutes les frontières ne sont pas aussi clairement caractérisées (proximité des moyennes, dispersion plus ou moins étendue autour de la moyenne).

#### **4.1.2.2.3 Conclusions relatives à la vérification du matériel expérimental**

- ▷ Les catégories définies théoriquement par la morphosyntaxe et la topologie lexicale ont été encodées par la locutrice naïve dans les productions enregistrées. Il existe donc bien un rapport relativement transparent entre niveau théorique, niveau cognitif, et niveau acoustique.
- ▷ Le matériel expérimental renferme bien une information prososyllabique de type code. Ce matériel apparaît donc comme valide et adéquat pour être utilisé dans un test perceptif.



### 4.1.3 Expérience 1 : Test perceptif de discrimination – Comparaison d’items désémantisés par synthèse

La phase de vérification du matériel expérimental a montré l’existence d’une corrélation entre les caractéristiques théoriques des syllabes, leurs caractéristiques acoustiques, et la possibilité de discriminer entre elles les 9 catégories de frontières étudiées. L’Expérience 1, réalisée en collaboration avec les Pr. G. Caelen-Haumont et J. Pynte, va donc s’attacher à montrer que les profils prosodiques de frontières mis en évidence lors de la vérification du matériel expérimental sont distingués les uns des autres par les auditeurs humains, et peuvent théoriquement être utilisés à des fins de guidage de la segmentation du flux de parole en mots.

Le test perceptif utilisé consiste à faire organiser par les auditeurs, à leur insu, l’ensemble du matériel expérimental en catégories perçues à partir d’un comportement local de jugement de similitudes pour chaque paire présentée. Si la relation entre caractéristiques acoustiques et caractéristiques théoriques (morphosyntaxe \* topologie) est suffisamment transparente, les groupements d’items obtenus après organisation perceptive devraient correspondre aux groupements théoriques de frontières.

#### 4.1.3.1 Aspects théoriques

Les écarts au hasard dans la répartition des réponses peuvent être justifiés par des facteurs physiques (degré de différence acoustique des items comparés) comme par des facteurs psychologiques (sensibilité à des dimensions utiles dans le traitement de la parole).

Le problème est donc plus compliqué à analyser qu’une simple comparaison des performances avec le hasard théorique : chaque réponse de l’auditeur peut correspondre à l’une des 8 situations décrites dans le Tableau 36 ci-dessous.

Tableau 36 : Logique de la situation expérimentale

	Jugement de similitude (OUI)		Jugement de différence (NON)	
	Items physiquement ressemblants	Items physiquement différents	Items physiquement ressemblants	Items physiquement différents
<b>Comparaison intra catégorie</b>	VV (1)	VF (2)	FF (3)	FV (4)
<b>Comparaison inter catégories</b>	FV (5)	FF (6)	VF (7)	VV (8)

Avec : V pour vrai, F pour faux. La première lettre fait référence à une réponse vraie ou fausse selon la catégorie théorique des items, la seconde selon leurs caractéristiques physiques. Ce qui lie une catégorie

théorique aux réponses des participants est la validité psychologique de cette catégorie théorique, sans laquelle elle ne peut être encodée dans le message acoustique.

Ci-dessous, la situation est examinée en détail.

#### Dans le cas des comparaisons intra catégorie

- ▷ Une réponse Oui (cas 1 et 2) est théoriquement correcte, mais elle est fautive dans le cas 2 par rapport à réalité physique.
- ▷ Une réponse Non (cas 3 et 4) est théoriquement incorrecte, mais elle est juste dans le cas 4 par rapport aux caractéristiques acoustiques des items comparés.

#### Dans le cas des comparaisons inter catégories

- ▷ Une réponse Oui (cas 5 et 6) est théoriquement incorrecte, mais elle est exacte dans le cas 5 par rapport aux caractéristiques acoustiques des items.
- ▷ Une réponse Non (cas 7 et 8) est théoriquement correcte, mais elle est fautive pour le cas 7 par rapport à réalité physique.

La réponse de l'auditeur peut correspondre à l'une des 4 catégories de réponse suivantes :

1. réponse justifiée d'un point de vue théorique et physique (1 et 8).
2. réponse justifiée d'un point de vue théorique seulement (2 et 7).
3. réponse justifiée d'un point de vue physique seulement (4 et 5).
4. réponse non justifiée (3 et 6).

Les réponses de type 1, 2, 7 et 8 impliquent une validation perceptive des catégories théoriques, alors que les réponses de type 3, 4, 5 et 6 impliquent une validation perceptive des catégories acoustiques. Il y a donc, en théorie, deux manières d'interpréter les données recueillies :

- ▷ comme étant en rapport avec la réalité physique (corrélation entre réponses et aspects physiques).
- ▷ comme étant en rapport avec le niveau représentationnel (mode de réponse selon un codage discret (présence / absence), éventuellement nuancé ( $\pm$  fortement présent)).

Les deux hypothèses théoriques suivantes devraient donc être formulées :

- ▷ hypothèse psychologique : le comportement des auditeurs s'explique par les caractéristiques théoriques assignées aux items, donc par des traitements descendants impliquant des modules de reconnaissance de forme dédiés à la détection de motifs spécifiques, et liés à la pratique de la langue (représentations mentales).
- ▷ hypothèse physique : le comportement des auditeurs s'explique par les caractéristiques physiques des items, donc par des traitements sensoriels ascendants.

Toutefois, comme ces 2 hypothèses sont difficiles à départager dans la pratique, une hypothèse plus simple a été adoptée. Elle consiste à tester la performance des auditeurs en fonction des caractéristiques théoriques des items sans se préoccuper de la cause qui permet un lien entre performance et caractéristiques théoriques des items (traitement ascendant, traitement descendant).

### **4.1.3.2 Aspects techniques de l'expérience**

#### **4.1.3.2.1 Facteurs expérimentaux et conditions**

Les mêmes facteurs expérimentaux que dans la Simulation 2 sont utilisés, et les conditions expérimentales sont identiques (se reporter à la Section 4.1.2.1.1, page 163 et au Tableau 20, page 164).

#### **4.1.3.2.2 Hypothèse opérationnelle**

Les items de même catégorie théorique sont classés comme similaires avec un taux supérieur aux comparaisons inter catégories, à partir de l'information prosodique disponible (durée,  $f_0$ ).

#### **4.1.3.2.3 Matériel expérimental utilisé**

Stimuli de synthèse (voir le détail de la procédure de désémantisation par synthèse dans la Section 3.4.1.3.2 commençant page 144).

Il s'agit de combinaisons de bisyllabes /mama/ dont la prosodie varie, ces bisyllabes provenant des phrases porteuses décrites à la Section 4.1.2.1.2.1, page 164. Le test comporte 1620 de ces combinaisons (voir les détails de l'organisation dans la Section suivante). Tous les items sont expérimentaux, l'expérience ne comporte pas de distracteurs.

#### **4.1.3.2.4 Organisation de l'expérience**

La comparaison d'un item avec lui-même est jugée non pertinente à explorer ( $N = 81$ ), car la réponse sera forcément « identique », de même que les comparaisons qui diffèrent sur l'ordre de présentation des 2 items ( $item_1 / item_2$  et  $item_2 / item_1$ , considérées comme doublons). Intégrer ces combinaisons dans l'expérience reviendrait à alourdir inutilement le test, qui comporte déjà une quantité très élevée d'items pour les participants. Le nombre de comparaisons considérées est donc initialement de  $((81 * 81) - 81) / 2 = 3240$ .

À ce stade, deux problèmes d'organisation apparaissent :

- ▷ La quantité élevée de comparaisons rend le test excessivement coûteux (nombre de participants, temps).
- ▷ L'existence de 2 types de sous matrices (carrées (comparaisons inter catégories :  $N = 36$ ) et triangulaires (comparaisons intra catégorie :  $N = 9$ )), avec un nombre différent de combinaisons d'items dans chaque type de sous matrice ( $N = 81$  pour sous matrices carrées (1 item apparaît 9 fois),  $N = 36$  pour sous matrices triangulaires (1 item apparaît 8 fois)), peut poser problème pour les analyses statistiques (le nombre de combinaisons d'items devrait être identique dans comparaisons intra catégorie et inter catégories, de même que le nombre de fois où un item apparaît dans une combinaison).

Cette difficulté a été résolue grâce à un artifice inspiré de la technique des carrés latins. L'organisation adoptée permet de :

- ▷ répartir les 36 combinaisons des sous matrices triangulaires de façon homogène sur les 72 possibilités des sous matrices carrées (81 possibilités sans les 9 de la diagonale, éliminés par souci de symétrie avec les sous matrices triangulaires).
- ▷ tester 4 fois chaque item ( $36 / 9$ ), dans 4 combinaisons différentes.
- ▷ tester tous les items.

Cependant, cette organisation conduit à éliminer du test une partie des combinaisons. Ainsi, toutes les combinaisons d'items des comparaisons inter catégories ne sont pas explorées, contrairement aux comparaisons intra catégories, le but étant que la quantité de combinaisons mesurées soit identique dans les 2 types de sous matrices. Le test comporte donc [36 combinaisons d'items] \* [45 combinaisons de catégories d'items] = 1620 comparaisons.

Un participant ne peut être testé sur 1620 cas : il est nécessaire de répartir les items dans des listes, de façon homogène pour l'ensemble des 45 comparaisons de catégories d'items. 6 listes sont nécessaires, la répartition des items dans ces listes est récapitulée dans le Tableau 37.

Tableau 37 : Répartition des combinaisons d'items dans les 6 listes expérimentales

	1	2	3	4	5	6	7	8	9
1			3/1		5/1		7/1		9/1
2	1/2			4/2		6/2		8/2	
3		2/3			5/3		7/3		9/3
4	1/4		3/4			6/4		8/4	
5		2/5		4/5			7/5		9/5
6	1/6		3/6		5/6			8/6	
7		2/7		4/7		6/7			9/7
8	1/8		3/8		5/8		7/8		
9		2/9		4/9		6/9		8/9	

Chaque liste est distinguée par la combinaison couleur \* casse des caractères.

Les chiffres font référence à la numérotation des items indépendamment de leur catégorie.

Chaque participant est donc testé avec  $1620 / 6 = 270$  comparaisons, et ne fournit qu'une réponse par comparaison.

#### 4.1.3.2.5 Population testée

Les données de 42 participants (7 par liste) ont été utilisées pour les analyses. Il s'agit d'étudiants francophones monolingues de l'université d'Aix-Marseille, sans troubles auditifs, cognitifs et moteurs, testés en janvier 2000.

#### 4.1.3.2.6 Tâche

Tâche de discrimination, modalité auditive.

L'auditeur entend séquentiellement le premier item (/mama/), puis le second (/mama/) de la paire à comparer. Il doit estimer si les deux formes sonores sont assez voisines pour être regroupées dans une même catégorie ou non, sachant par avance que ces deux formes ne sont jamais identiques (variations prosodiques). La tâche de comparaison perceptive trouve ses fondements dans l'évaluation de contrastes de durée et/ou de fréquence fondamentale entre les items comparés.

#### 4.1.3.2.7 Réponses obtenues

Réponses binaires (Oui/Non), et non plus linéaires (de 0 à 1) comme dans la Simulation 2.

#### 4.1.3.2.8 Procédure de recueil des données

Expérience réalisée à l'université d'Aix-Marseille I. Détails techniques sur le recueil dans la Partie III, Chapitre 4<sup>104</sup>.

$6 * 270 * 7 = 11340$  mesures sont effectuées.

Cela correspond à 7 mesures par comparaison de bisyllabes, et à 252 mesures par comparaison de frontières.

---

<sup>104</sup> Complément d'information : un essai correspond aux 3 étapes suivantes :

- Présentation séquentielle au casque des 2 items à comparer (/mama/<sub>1</sub> - /mama/<sub>2</sub>), automatiquement répété une fois.

- Présentation à l'écran d'une invite de réponse (OUI/NON). Le participant se prononce alors à l'aide de 2 touches du clavier sur son impression de proximité des 2 items (OUI = suffisamment semblables pour être regroupés dans une même catégorie ; NON = insuffisamment semblables pour appartenir à une même catégorie). Il peut réécouter à volonté la séquence s'il éprouve une difficulté à se prononcer.

- La réponse est enregistrée et l'item suivant est automatiquement présenté.

La phase test de l'expérience se déroule en 6 blocs de 45 comparaisons, avec pause entre chaque bloc.

Le participant pilote lui-même le déroulement de l'expérience, et n'est pas sous contrainte temporelle pour fournir sa réponse.

La totalité du recueil dure entre 45 minutes et 1 heure selon la vitesse des participants.

#### 4.1.3.2.9 Méthode d'analyse

L'analyse des données a été effectuée à l'aide de 4 modèles multiniveaux de régression logistique (réponses binaires). Le but des analyses effectuées est :

- ▷ de déterminer quels facteurs influencent la performance des auditeurs dans la tâche de comparaison :
  - ∩ effet de la quantité de syllabes de même catégorie à la même position dans les 2 items comparés (1<sup>er</sup> modèle, Équation 28, page 472)
  - ∩ effet du type de syllabe en commun à une même position dans les 2 items (2<sup>ème</sup> modèle, Équation 29, page 473)
- ▷ d'estimer le profil de réponse des participants obtenu sur le matériel expérimental avec une tâche de comparaison d'items, par type de frontières comparées (3<sup>ème</sup> modèle, Équation 30, page 473) :
  - ∩ effet du type d'item pour les comparaisons intra catégorie
  - ∩ effet du type d'item pour les comparaisons inter catégories

Les modèles utilisés sont à 2 niveaux (comparaison, sujet), malgré une corrélation intra classe logiquement faible (6.02% - les comparaisons ne sont pas représentées mentalement).

La distribution présente 7 valeurs manquantes (0.06% de non-réponses - enregistrement d'un code correspondant à une autre touche que les deux désignées pour répondre).

#### 4.1.3.2.10 Paramètres contrôlés<sup>105</sup>

Se reporter à la Section 4.1.2.1.2.1, page 164, pour la création des phrases porteuses, et à la Section 3.4.1.3.2 page 144, pour la création des items désémanés.

Les stimuli sont présentés aléatoirement.

---

<sup>105</sup> Le facteur liste n'est pas directement modélisé, mais il est intégré au modèle implicitement, par la prise en compte simultanée des sujets et des items.

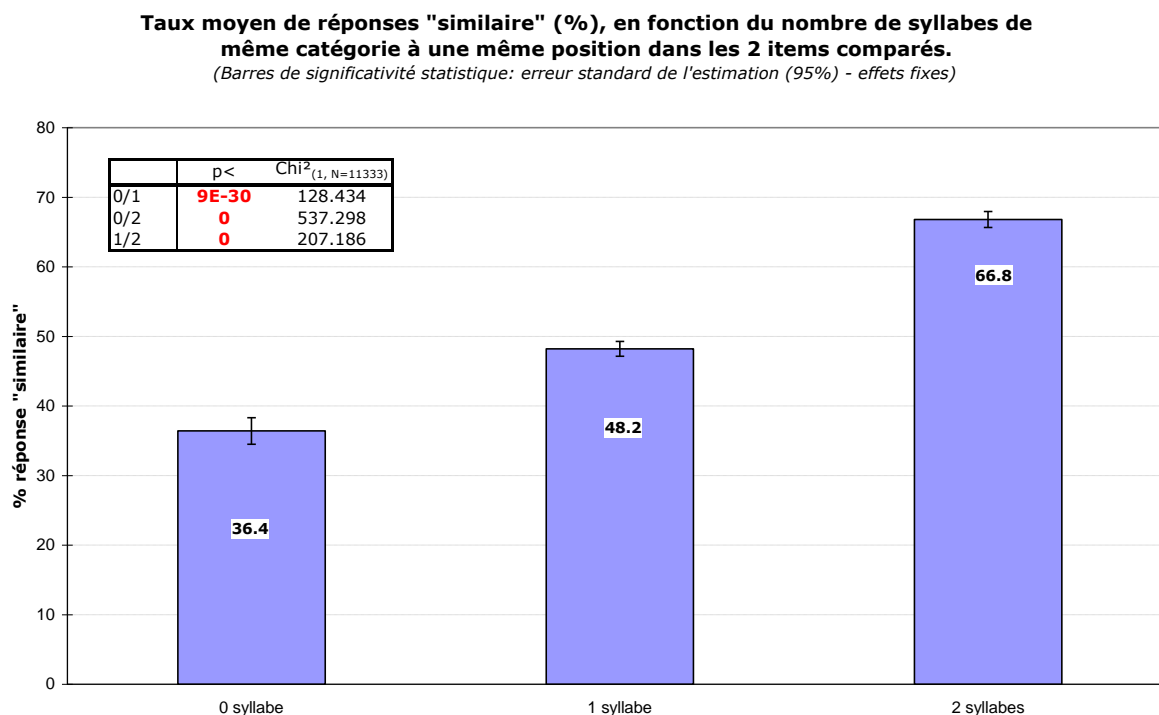
### 4.1.3.3 Résultats obtenus

#### 4.1.3.3.1 Effet de la quantité de syllabes similaires à une même position dans les deux items comparés, sur la performance de discrimination des items par les auditeurs (taux de réponses « similaire »)

L'équation du modèle étudié (Équation 28) peut être consultée en Annexe 7, page 472.

La Figure 48 présente les taux moyens de réponse « similaire » fournis par les participants, en fonction de la quantité de syllabes de même catégorie théorique à une même position dans les 2 items comparés. Elle présente également les résultats de l'analyse des contrastes entre conditions (table insérée).

Figure 48 : Nombre moyen de réponses « similaire », en fonction du nombre de syllabes de même catégorie entre les deux items comparés



Plus il existe de similitudes catégorielles entre les syllabes des deux items, plus les taux de réponse d'identité augmentent (aucun type de syllabe en commun : 36.4% ; 1 type de syllabe en commun : 48.2% ; 2 types de syllabes en commun : 66.8%). Cette règle paraît être linéaire, mais l'écart entre les conditions 0 et 1 syllabes de même catégorie (11.8%) est inférieur à l'écart entre les conditions 1 et 2 syllabes de même catégorie (18.6%).

La discrimination porterait à la fois sur une estimation de la distance globale des 2 objets comparés, et sur une estimation de la distance entre leurs composants (analyse syllabique de la prosodie), contrairement au traitement effectué par un comparateur neuromimétique.

Les auditeurs semblent donc avoir été capables de réaliser une classification du matériel à l'aide de la tâche de comparaison.

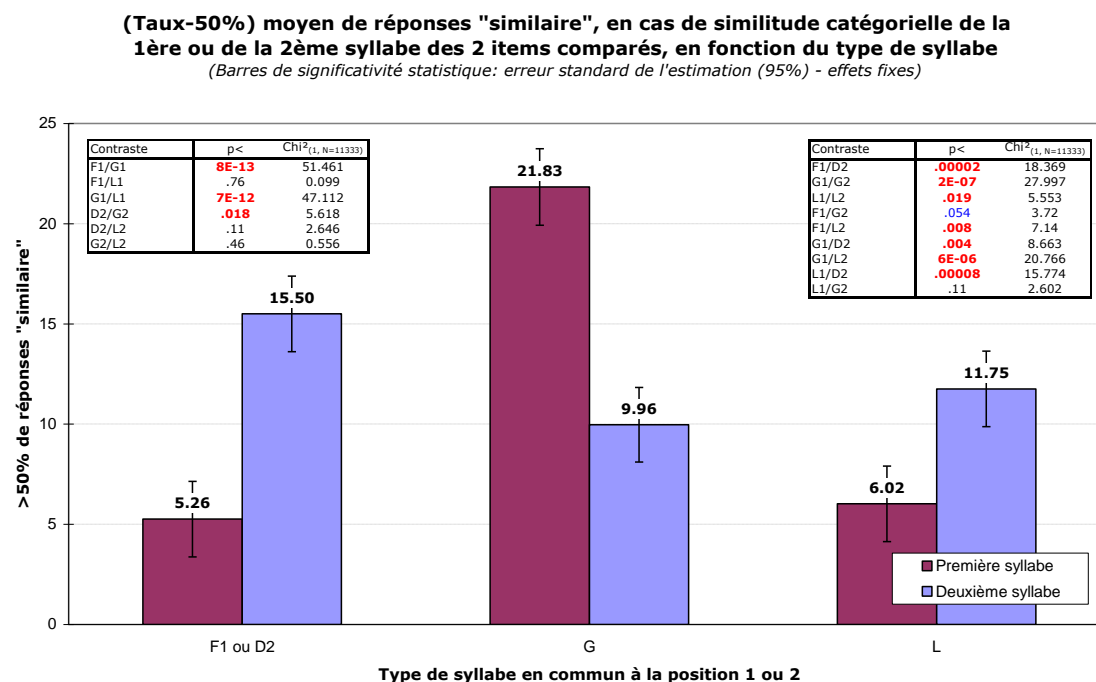
### 4.1.3.3.2 Effet du type de similitude à une position donnée dans les deux items comparés, sur la performance de discrimination des items par les auditeurs (taux de réponses « similaire »)

L'équation du modèle étudié (Équation 29) peut être consultée en Annexe 7, page 473.

#### Effets avec une seule syllabe en commun :

La Figure 49 présente les taux moyens de réponse « similaire » au-delà du hasard, fournis par les participants en fonction du type de syllabe similaire à une position donnée. Les résultats des analyses des contrastes  $\gamma$  sont également présentés (tables insérées).

Figure 49 : Taux de réponses « similaire » en fonction d'une similitude catégorielle de la première ou la deuxième syllabe





La quantité moyenne de réponses « similaire » varie significativement en fonction du type de syllabe en commun, à la position 1 comme à la position 2.

Les taux de réponse « similaire » sont les plus élevés sous l'influence de G<sub>1</sub> (72%), puis de D<sub>2</sub> (66%).

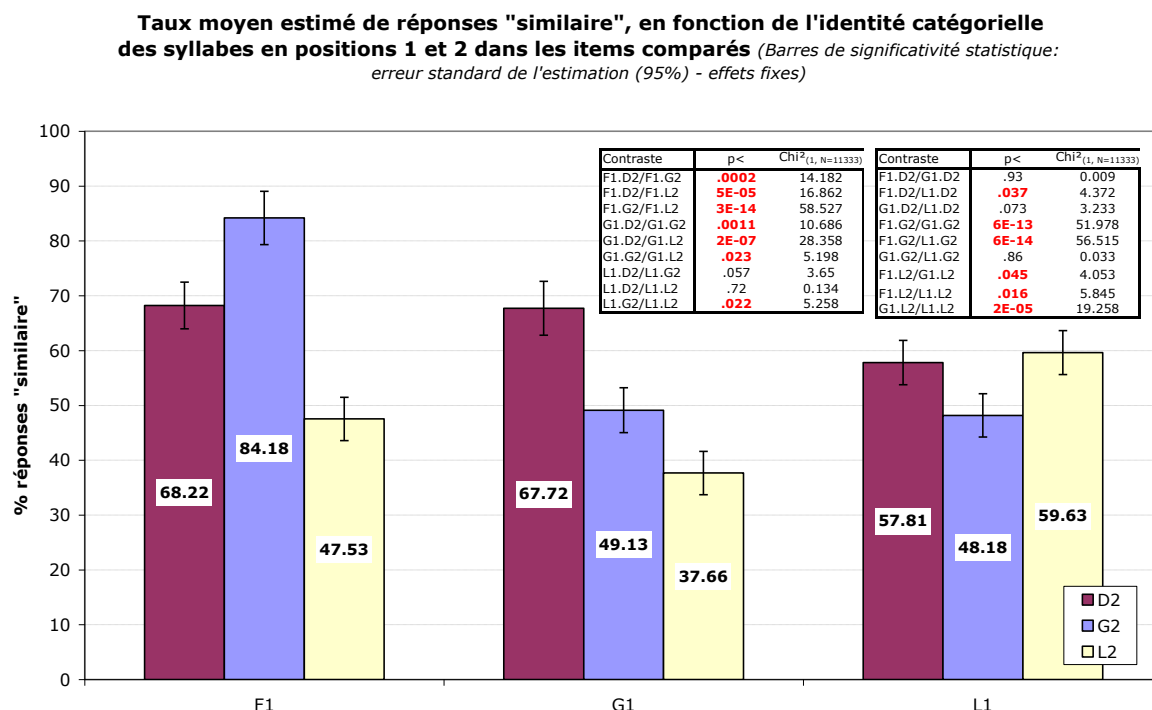
Ils sont les moins augmentés lorsque les items ont F<sub>1</sub> ou L<sub>1</sub> en commun (respectivement 55 et 56%). Lorsque les items ont en commun G<sub>2</sub> ou L<sub>2</sub>, des taux intermédiaires d'augmentation du pourcentage de réponses « similaire » sont obtenus (respectivement 60% et 62%).

Le système perceptif serait donc plus réactif avec les syllabes D et G, et la position de la syllabe dans la séquence serait importante (pour G, première syllabe et non deuxième).

Effets avec deux syllabes en commun (interaction syllabe<sub>1</sub> \* syllabe<sub>2</sub>) :

La Figure 50 présente la proportion de réponses d'identité (centrée sur le hasard) lorsque les deux items comparés ont leurs 2 syllabes de même catégorie à la même position. Les seuils de significativité des contrastes sont également présentés sous forme de tableaux insérés.

Figure 50 : Taux de réponses « similaire » en fonction de l'identité catégorielle des syllabes en première et deuxième positions dans les 2 items comparés



La proportion de réponses d'identité varie en fonction de la similitude catégorielle des 2 syllabes des items comparés. Comme lors de la Simulation 2, certaines catégories d'items sont bien catégorisées lors de la comparaison (F<sub>1</sub>.G<sub>2</sub>, F<sub>1</sub>.D<sub>2</sub> et G<sub>1</sub>.D<sub>2</sub>), d'autres sont correctement catégorisées (L<sub>1</sub>.D<sub>2</sub> et L<sub>1</sub>.L<sub>2</sub>), et d'autres sont mal catégorisées (F<sub>1</sub>.L<sub>2</sub>, G<sub>1</sub>.G<sub>2</sub>, G<sub>1</sub>.L<sub>2</sub> et L<sub>1</sub>.G<sub>2</sub>). Un profil de réponse voisin de celui obtenu lors de la Simulation 2 peut être observé (voir Figure 44 page 191).

#### 4.1.3.3.3 Effet du type de frontière sur la performance de discrimination des items par les auditeurs (taux de réponses « similaire »)

L'équation du modèle étudié (Équation 30) peut être consultée en Annexe 7, page 473.

Tableau 38 : Récapitulation des moyennes par comparaison de catégorie d'item (matrice de comparaison - taux moyens de réponses « similaire »)

	FD	FG	FL	GD	GG	GL	LD	LG	LL
FD	<b>74.2</b>								
FG	44.8	<b>84.9</b>							
FL	54	25.4	<b>50.8</b>						
GD	48.4	67.1	24.7	<b>85.3</b>					
GG	29.8	69.4	20.6	73	<b>67.8</b>				
GL	52.4	49.2	41.3	63	42.1	<b>58.7</b>			
LD	65.9	36.5	46.8	41.9	35	51.2	<b>65.5</b>		
LG	30.6	40.1	19.8	28.6	29	44.4	43.9	<b>50.4</b>	
LL	52	26.2	58.3	30.8	22.2	31.8	51	31.8	<b>63.5</b>

Les taux de regroupement lors des comparaisons intra catégorie sont supérieurs dans 5 cas sur 9 aux taux de confusion. L'hypothèse de l'étude est donc partiellement validée, ces résultats ne la remettant pas en question.

La validation statistique des écarts de performance entre frontières (comparaisons intra catégorie uniquement) peut être trouvée dans le Tableau 39.

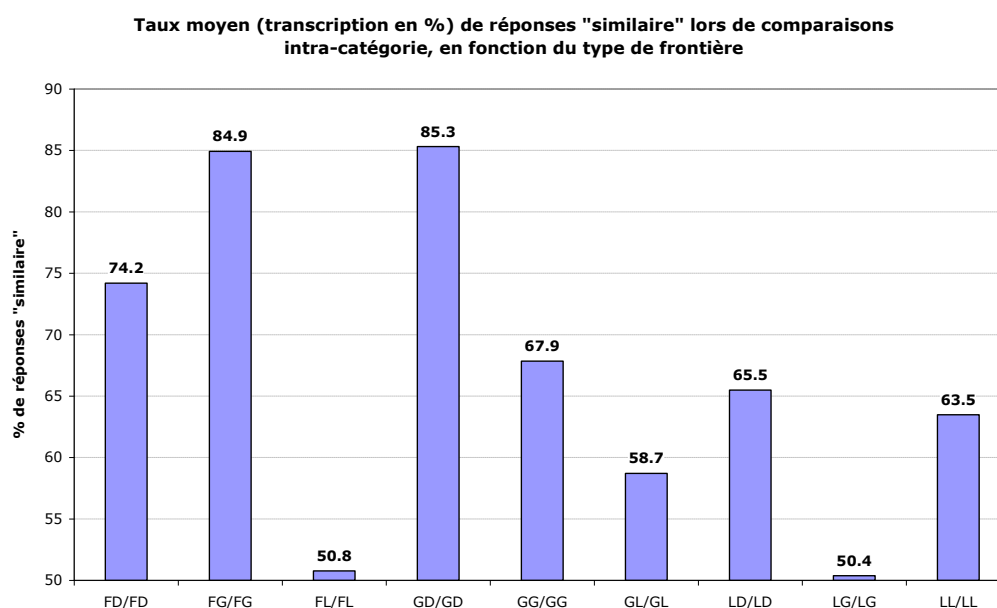
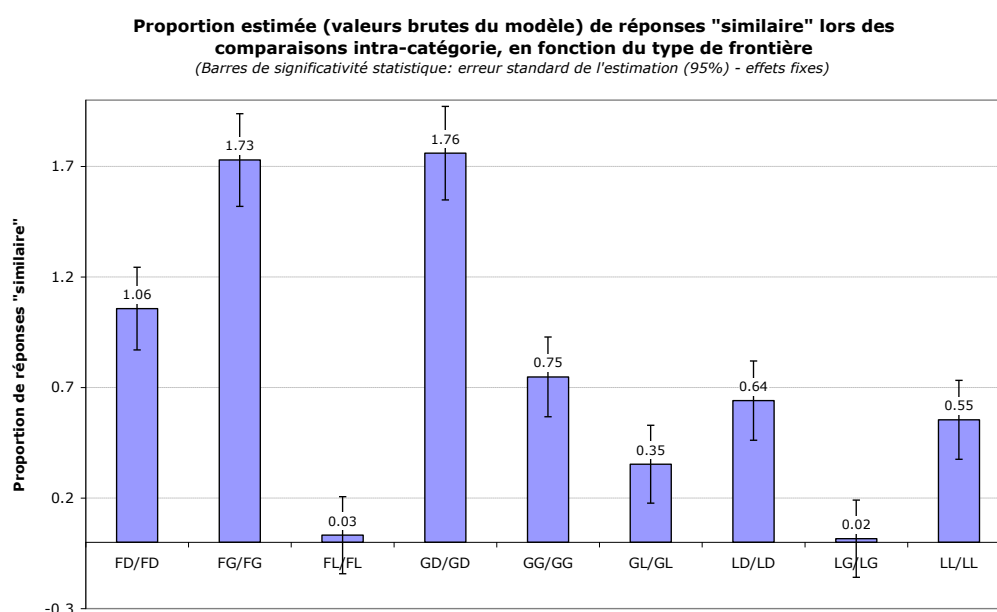
Tableau 39 : Test de significativité des contrastes entre conditions (contrastés entre comparaisons intra catégorie)

	p<	Chi <sup>2</sup> <sub>(1, N=11333)</sub>		p<	Chi <sup>2</sup> <sub>(1, N=11333)</sub>
FDFD/FGFG	<b>.003</b>	8.8	FLFL/LDLD	<b>.001</b>	11.1
FDFD/FLFL	<b>.0001</b>	28.7	FLFL/LGLG	.93	0.0
FDFD/GDGD	<b>.002</b>	9.5	FLFL/LLLL	<b>.004</b>	8.2
FDFD/GGGG	.12	2.5	GDGD/GGGG	<b>.0001</b>	20.7
FDFD/GLGL	<b>.0003</b>	13.4	GDGD/GLGL	<b>.0001</b>	41.4
FDFD/LDLD	<b>.033</b>	4.5	GDGD/LDLD	<b>.0001</b>	25.6
FDFD/LGLG	<b>.0001</b>	29.6	GDGD/LGLG	<b>.0001</b>	64.3
FDFD/LLLL	<b>.01</b>	6.7	GDGD/LLLL	<b>.0001</b>	30.0
FGFG/FLFL	<b>.0001</b>	61.7	GGGG/GLGL	<b>.034</b>	4.5
FGFG/GDGD	.90	0.0	GGGG/LDLD	.57	0.3
FGFG/GGGG	<b>.0001</b>	19.7	GGGG/LGLG	<b>.0001</b>	15.7

FGFG/GLGL	.0001	40.1	GGGG/LLLL	.30	1.1
FGFG/LDLD	.0001	24.5	GLGL/LDLD	.12	2.4
FGFG/LGLG	.0001	62.8	GLGL/LGLG	.061	3.5
FGFG/LLLL	.0001	28.8	GLGL/LLLL	.27	1.2
FLFL/GDGD	.0001	63.1	LDLD/LGLG	.0007	11.7
FLFL/GGGG	.0002	15.0	LDLD/LLLL	.64	0.2
FLFL/GLGL	.075	3.2	LGLG/LLLL	.003	8.8

La Figure 51 illustre cette variation de performances en présentant les taux moyens de réponse « similaire » par comparaison intra catégorie.

Figure 51 : Taux de réponse « similaire » en fonction du type de frontière, lors de comparaisons intra catégorie



Les meilleures performances s'observent pour les frontières FG (84.9% de réponses « similaire ») et GD (85.3%). Ces deux catégories de frontières sont perçues comme étant les mieux caractérisées et/ou les plus homogènes.

Les moins bonnes performances s'observent pour les frontières FL (50.8%) et LG (50.4%). Ces frontières sont perçues comme étant les moins caractérisées et/ou les plus hétérogènes.

Concernant les comparaisons inter catégories, certaines frontières sont perçues comme voisines (GD et GG, 73%), alors que d'autres sont clairement identifiées comme distinctes (FD et GG, 29.8%).

Certaines catégories de frontières présentent des taux de confusion de leurs items avec ceux d'une autre classe supérieurs au taux de regroupement, conduisant à une confusion des catégories (se produit 4 fois sur 9 – voir le Tableau 40).

Tableau 40 : Récapitulation des moyennes par comparaison de catégorie d'item (matrice de comparaison - taux moyens de réponses « similaire »)

	FD	FG	FL	GD	GG	GL	LD	LG	LL
FD	<b>74.2</b>								
FG	44.8	<b>84.9</b>							
FL	54	25.4	50.8						
GD	48.4	67.1	24.7	<b>85.3</b>					
GG	29.8	69.4	20.6	<b>73</b>	67.8				
GL	52.4	49.2	41.3	<b>63</b>	42.1	58.7			
LD	<b>65.9</b>	36.5	46.8	41.9	35	51.2	65.5		
LG	30.6	40.1	19.8	28.6	29	44.4	43.9	<b>50.4</b>	
LL	52	26.2	<b>58.3</b>	30.8	22.2	31.8	51	31.8	<b>63.5</b>

Gras : réponse statistique de similitude perçue.

Grisé : erreur / réponse voisine

Ce sont :

- ▷ FL (FLFL (50.8%), alors que FLLL (58%) et FLFD (54%))
- ▷ GG (GGGG (67.8%), alors que GGGD (73%) et FG GG (69.4%))
- ▷ GL (GLGL (58.7%), alors que GLGD (63%))
- ▷ LD (LDLD (65.5%), alors que LDFD (65.9%))

Dans tous ces cas, la confusion porte sur des catégories de frontières voisines, ne différant que par une syllabe. Ce type d'erreur ne remet donc pas en question les résultats précédents. Il traduit probablement un traitement intégratif diffus de l'information pour réaliser la tâche de comparaison.

#### 4.1.3.4 Conclusions de l'Expérience 1

L'hypothèse de l'Expérience 1 est validée par les données. Les items de même catégorie théorique sont considérés comme similaires avec un taux supérieur aux comparaisons inter catégories, sur la base de la  $f_0$  et de la durée. Les différences acoustiques des syllabes et des frontières lexicales sont donc exploitées pour organiser le matériel en catégories perçues, correspondant dans une proportion supérieure au hasard aux catégories théoriques.

Certaines catégories de frontières conduisent à de meilleures performances que d'autres (taux de regroupement élevé des items de la catégorie, taux de confusion faible de la catégorie avec d'autres).

On retiendra la réactivité plus importante du système pour les syllabes  $G_1$  et  $D_2$ , révélée dans la Figure 49, page 203, qui ne trouve pas de justification acoustique satisfaisante (Simulation 1, Simulation 2 : effet significatif mais d'amplitude relative).

Cette différence peut s'expliquer par une propension du système naturel à assimiler les frontières GG et GL à la frontière GD, d'une part, et d'autre part, à identifier correctement cette frontière (meilleures performances - Tableau 40, page 207). Il est à noter que les catégories GG et GL ne tendent pas à être confondues entre elles (ce point laisse donc penser qu'il s'agit d'une assimilation à une super-catégorie, et non d'une confusion entre catégories).

Le phénomène ne se produit pas avec les frontières  $F^*$  et  $L^*$ , et les profils prosodiques moyens des frontières  $G^*$  sont distincts (Figure 39, page 171).

Un biais cognitif, lié au traitement des messages linguistiques, aurait donc pu interférer dans les réponses des participants.

Le traitement des frontières GD semble particulièrement crucial, puisque le locuteur crée une modulation fortement marquée sur une durée courte, ce qui nécessite probablement un effort plus important, et un ciblage plus précis d'une forme « prototypique » (GD présente un des profils les plus contrastés des 9 frontières étudiées - Figure 39, page 171).

Il semble que le système, lorsqu'il perçoit un monosyllabe grammatical, enclenche une association forte selon une règle du type « dès que G : D ». Une telle règle viserait à activer la recherche lexicale sur les mots de classe ouverte plurisyllabique, et donc, le

système ne favoriserait pas l'accès lexical des mots monosyllabiques, mais celui des mots polysyllabiques (l'accès lexical pour les mots monosyllabiques se ferait plus facilement, simplement, naturellement, et ne nécessiterait pas de stratégie de facilitation ?).

Les mots du lexique renferment une majorité de mots polysyllabiques avec un ou plusieurs mots enchâssés à l'initiale. Une telle configuration complique le processus, car les mots enchâssés initiaux sont les premiers à être activés. Une procédure visant à les défavoriser pourrait donc être utile pour éviter au système de s'engager dans une direction erronée, qui requiert ensuite des procédures correctives.

Cette première expérience confirme donc que les auditeurs humains sont capables de traiter l'information prosodique encodée au niveau de la syllabe. Théoriquement, cela implique que les auditeurs peuvent l'utiliser lors de la perception, à la manière d'un code de guidage pour la segmentation du flux de parole en mots.

#### **4.1.4 Expérience 2 : Test perceptif de détection et de localisation consciente de frontières lexicales**

L'Expérience 1 ayant montré que l'information prosodique encodée dans le signal est suffisamment marquée pour être exploitée par le système sensoriel, que certaines frontières sont éventuellement responsables de biais perceptifs et/ou cognitifs, une seconde expérience est réalisée pour savoir si :

- ▷ les motifs prosodiques sont à même de permettre à un auditeur de détecter et de localiser consciemment la présence d'une frontière entre deux mots.
- ▷ des différences de performance se produisent dans la détection et la localisation des frontières en fonction des catégories syllabiques impliquées.

Cette partie de l'étude explore plus directement l'utilité de tels motifs dans la segmentation du flux de parole en mots, en cherchant à déterminer s'ils sont perçus comme des frontières lexicales ou non.

##### **4.1.4.1 Aspects théoriques**

Étant donné que les frontières entre mots sont prosodiquement caractérisées (résultats de l'étude préalable), est-ce qu'un auditeur francophone entend ces motifs prosodiques comme des frontières lexicales ? On peut a priori supposer que oui au vu des résultats précédents.

Étant donné que les frontières peuvent être distinguées les unes des autres, à la perception, par leur prosodie, en fonction de facteurs morphosyntaxiques et topologiques (Simulation 2, Expérience 1), est-ce que les performances d'un auditeur francophone pour détecter / localiser les frontières varient d'une catégorie à l'autre ? On peut également supposer que oui, d'une part parce que les frontières ne sont pas marquées de la même manière sur le plan acoustique, et d'autre part parce que la réactivité des auditeurs peut varier en fonction des traitements cognitifs habituellement associés aux différentes catégories de frontière lors de la perception d'un message parlé (exemple dans l'Expérience 1, pour la catégorie GD).

Ces 2 hypothèses théoriques n'ont pas été précisées plus, étant donné la complexité de la situation étudiée et le caractère exploratoire de ce type d'expérience.

#### 4.1.4.2 Aspects techniques

##### 4.1.4.2.1 Facteurs expérimentaux et conditions

L'Expérience 2 utilise des séquences de 3 syllabes afin de permettre la réalisation de la tâche de détection / localisation. En plus des 3 facteurs utilisés dans les expériences précédentes (se reporter à la Section 4.1.2.1.1, page 163), 2 facteurs supplémentaires sont employés :

- ▷ existence d'une frontière lexicale dans l'item (2 modalités : oui / non)
- ▷ position de cette frontière dans l'item, lorsqu'elle existe (2 modalités : entre les 2 premières syllabes de l'item vs. entre les 2 dernières)

Le second facteur étant conditionnel du premier, il a été choisi d'assimiler le premier facteur au deuxième afin de créer un nouveau facteur « position », à 3 modalités :

- ▷ Position 0 ou  $P_0$  = contrôle (aucune frontière dans l'item, qui correspond à un mot trisyllabique).
- ▷  $P_1$  = existence d'une frontière lexicale entre la 1<sup>ère</sup> ( $F_1$ ,  $G_1$  ou  $L_1$ ) et la 2<sup>ème</sup> ( $D_2$ ) syllabe de l'item
- ▷  $P_2$  = frontière lexicale entre la 2<sup>ème</sup> ( $F_2$ ) et la 3<sup>ème</sup> ( $D_3$ ,  $G_3$  ou  $L_3$ ) syllabe de l'item.

L'Expérience 2 manipule donc les 4 facteurs suivants :

- ▷ la catégorie morphosyntaxique du mot dont provient chaque syllabe de la séquence stimulus (2 modalités : lexical vs. grammatical)
- ▷ le nombre de syllabes du mot porteur (2 modalités : 1 vs. 3)
- ▷ la position des syllabes dans les mots porteurs trisyllabiques (2 modalités : début vs. fin)
- ▷ la position de la frontière lexicale dans l'item (3 modalités :  $P_0$ ,  $P_1$ ,  $P_2$ ).

L'expérience ne renferme que 7 conditions, récapitulées dans le Tableau 41, car les 4 facteurs ne se laissent pas croiser selon un plan complet.

Tableau 41 : Structure de l'item selon la condition expérimentale

$P_0$	$P_1$		$P_2$	
Mot (contrôle)	$F_1-$	$-D_2-M_3$	$M_1-F_2-$	$-D_3$
	$L_1-$			$-L_3$
	$G_1-$			$-G_3$

Avec [FLGDM] le type de syllabe (M syllabe de milieu de mot trisyllabique), [123] la position de la syllabe dans l'item.

Exemples d'extraits de phrase ayant servi à créer les items (Annexe 2 page 457) :

- ▷  $F_1D_2M_3$  : « -gercompé- » (boulangier compétent – Items G / D et F / D, phrase 9)



- ▷  $L_1D_2M_3$  : « *champlabou-* » (champ labouré – Items L / D et G / L, phrase 1)
- ▷  $G_1D_2M_3$  : « *lechoco-* » (le chocolat – Items G / D et F / D, phrase 2)
- ▷  $M_1F_2D_3$  : « *-baretcha-* » (cabaret chaleureux – Items G / D et F / D, phrase 5)
- ▷  $M_1F_2L_3$  : « *-credisaint* » (mercredi saint – Items F / L et L / G, phrase 1)
- ▷  $M_1F_2G_3$  : « *-cumentdu* » (document du – Items F / G, phrase 8)

#### 4.1.4.2.2 Hypothèses opérationnelles

L'Expérience 2 comporte 2 hypothèses opérationnelles :

- ▷ Les frontières seront détectées / localisées à un taux supérieur au hasard grâce à la seule information prosodique présente dans le signal acoustique (durée,  $f_0$ ).
- ▷ Les performances de détection / localisation varieront en fonction du type de frontière considérée.

#### 4.1.4.2.3 Matériel expérimental utilisé

Stimuli en parole de synthèse (items trisyllabiques désegmentés par remplacement du signal avec un train d'impulsions – voir les détails techniques dans la Section 3.4.1.3.2 page 144).

Ces stimuli proviennent des mêmes phrases et des mêmes enregistrements que ceux qui ont été utilisés dans les études précédentes. Les portions de parole sont cependant un peu plus larges (3 syllabes au lieu de 2 syllabes), mais sont extraites des mêmes régions des phrases porteuses (ces items trisyllabiques renferment les items bisyllabiques précédents. Toutes les conditions précédentes de frontière ne sont cependant pas reprises et étudiées).

L'expérience utilise 81 items expérimentaux (pas d'items distracteurs), répartis dans les 7 conditions de l'expérience (6 catégories de frontières + 1 contrôle (absence de frontière)) comme suit :

- ▷  $P_0$  :  $N = 27$
- ▷  $P_1$  :  $N = 27$ , 9 items pour chacune des 3 catégories de frontière ( $F_1D_2-$ ,  $L_1D_2-$ ,  $G_1D_2-$ )
- ▷  $P_2$  :  $N = 27$ , 9 items pour chacune des 3 catégories de frontière ( $-F_2D_3$ ,  $-F_2L_3$ ,  $-F_2G_3$ )

#### 4.1.4.2.4 Organisation de l'expérience

Un participant est testé avec tous les items ( $N = 81$ ), et fournit une réponse par item.

#### 4.1.4.2.5 Population testée

Les données de 50 francophones monolingues, sans troubles auditifs, cognitifs et moteurs, étudiant à l'université de Genève, ont été utilisées pour les analyses. Les tests se sont déroulés en mai 2002.

#### 4.1.4.2.6 Tâche

D'un point de vue logique, la tâche se subdivise en 2 procédures, détection et localisation, l'exécution du deuxième aspect étant conditionnelle de la première :

- ▷ 1<sup>ère</sup> procédure : détection d'une frontière.
  - ∩ Si l'auditeur ne détecte pas de frontière lexicale dans l'item, il fournit une réponse négative et passe à l'item suivant (réponse P<sub>0</sub>).
  - ∩ Si l'auditeur détecte une frontière lexicale dans l'item, il effectue la 2<sup>ème</sup> procédure de la tâche.
- ▷ 2<sup>ème</sup> procédure : localisation de la frontière dans l'item
  - ∩ Si l'auditeur localise la frontière lexicale entre les deux premières syllabes de l'item, il répond P<sub>1</sub>.
  - ∩ Si l'auditeur localise la frontière lexicale entre les deux dernières syllabes de l'item, il répond P<sub>2</sub>.

Toutefois, d'un point de vue pratique, ces 2 aspects de la tâche sont indissociables, détection impliquant localisation et inversement. Les participants ne fournissent donc pas deux réponses mais une seule, relative à la fois à la détection et à la localisation d'une frontière.

Les items sont présentés au casque, en modalité auditive.

L'auditeur peut réécouter un item à volonté avant de se prononcer. Il réalise une décision sur la base de la seule information prosodique (rythmique et mélodique).

#### 4.1.4.2.7 Procédure de recueil des données

Expérience réalisée à l'université de Genève. Détails dans la Partie III, Chapitre 4.

50 \* 81 = 4050 mesures effectuées (aucune valeur manquante). La phase test de l'expérience dure 15 à 20 minutes environ, et se déroule sans pause.

Les réponses recueillies sont de nature catégorielle :

- ▷ 0 : pas de frontière, l'item est un mot trisyllabique
- ▷ 1 : une frontière entre les deux premières syllabes de l'item
- ▷ 2 : une frontière entre les deux dernières syllabes de l'item

#### 4.1.4.2.8 Méthode d'analyse

Les données étant catégorielles, des modèles logistiques uni et multi-variés sont utilisés lors d'analyses de régression multiniveaux (item, sujet).

Deux séries d'analyses sont effectuées :

- ▷ Analyse générale de la répartition de chaque catégorie de réponse (0, 1, 2 – réponse obtenue) par rapport à la situation présentée (0, 1, 2 - réponse attendue – modèles logistiques uni variés). Cette première série d'analyses vise à estimer la capacité des participants à détecter une frontière lexicale et à la localiser dans l'item.
- ▷ Analyse détaillée de la répartition des réponses en fonction du type de frontière testé (modèles logistiques multi-variés). Cette seconde série d'analyses vise à connaître la performance des participants en fonction des facteurs lexicaux et topologiques manipulés.

#### 4.1.4.2.9 Paramètres contrôlés

Il s'agit des facteurs contrôlés lors de la création du matériel expérimental. Les détails de la procédure sont fournis à la Section 4.1.2.1.2.1, page 164 (création des phrases porteuses) et à la Section 3.4.1.3.2 page 144 (création des items déségmentés), en plus des facteurs classiquement contrôlés dans une expérience (nombre d'items équivalents par modalité de facteur, présentation aléatoire des items lors de la passation, etc.).

#### 4.1.4.3 Résultats obtenus

##### 4.1.4.3.1 Rapport des analyses statistiques

###### 4.1.4.3.1.1 Analyse générale de la répartition de chaque catégorie de réponse obtenue par rapport à la situation présentée (réponse attendue)

Cette analyse nécessite l'utilisation de 3 modèles logistiques (1 par type de réponse). Les équations de ces modèles sont présentées sous le label Équation 31, à l'Annexe 8 page 474 (ainsi que le détail du mode de calcul des taux de réponse à partir des informations fournies par les modèles).

Figure 52 : Répartition des réponses obtenues en fonction des réponses attendues

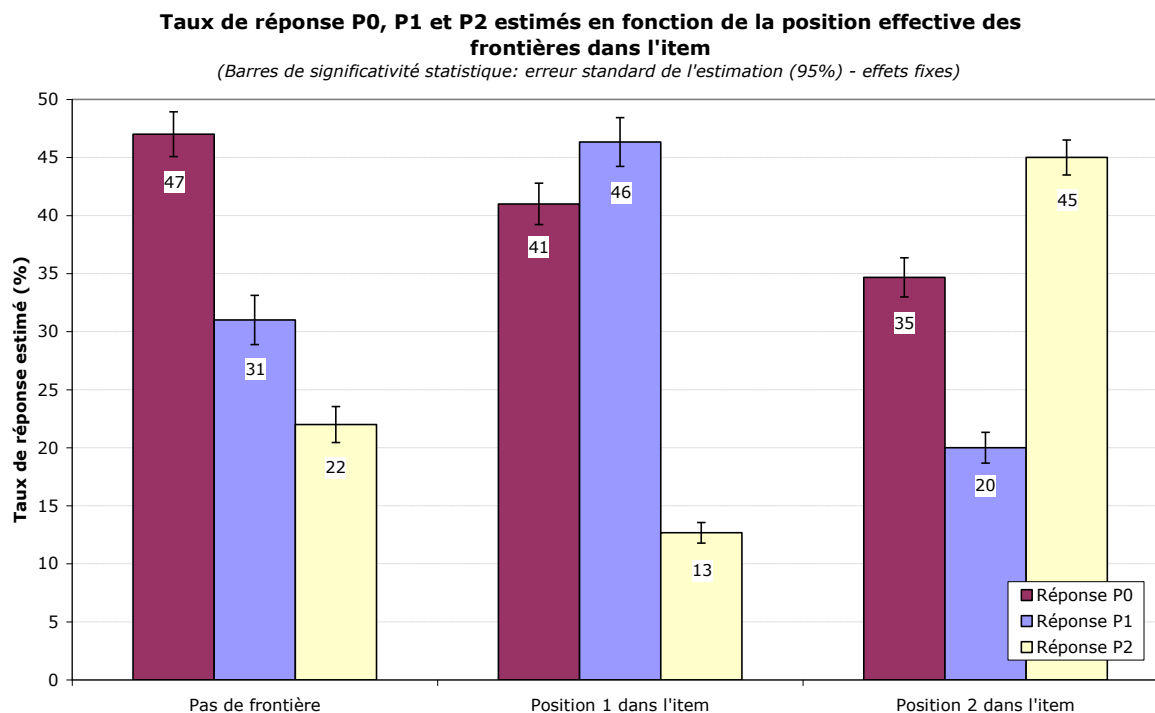


Tableau 42 : Test de significativité des contrastes entre conditions

	Position	Réponse P <sub>0</sub>	Réponse P <sub>1</sub>	Réponse P <sub>2</sub>
<b>Chi<sup>2</sup><sub>(1, N=4050)</sub></b>	<b>P<sub>0</sub>/P<sub>1</sub></b>	10.1	64.3	37.7
	<b>P<sub>0</sub>/P<sub>2</sub></b>	41.1	42.6	159.5
	<b>P<sub>1</sub>/P<sub>2</sub></b>	11.1	312.1	309
<b>p&lt;</b>	<b>P<sub>0</sub>/P<sub>1</sub></b>	<b>.0015</b>	<b>.0001</b>	<b>.0001</b>
	<b>P<sub>0</sub>/P<sub>2</sub></b>	<b>.0001</b>	<b>.0001</b>	<b>.0001</b>
	<b>P<sub>1</sub>/P<sub>2</sub></b>	<b>.001</b>	<b>.0001</b>	<b>.0001</b>

Tous les écarts sont significatifs.

Les auditeurs détectent l'absence de frontière (47%) à un seuil significativement supérieur aux confusions (P<sub>0</sub> au lieu de P<sub>1</sub> : 41% ; P<sub>0</sub> au lieu de P<sub>2</sub> : 35%). De même, ils localisent correctement les frontières en P<sub>1</sub> (46%) et en P<sub>2</sub> (45%), et ce avec de meilleurs taux qu'ils ne leur attribuent une mauvaise réponse (P<sub>1</sub> au lieu de P<sub>0</sub> : 31% ; P<sub>1</sub> au lieu de P<sub>2</sub> : 35% ; P<sub>2</sub> au lieu de P<sub>0</sub> : 22% ; P<sub>2</sub> au lieu de P<sub>1</sub> : 13%).

Dans 2 cas cependant, il semble y avoir eu une difficulté pour faire la distinction entre frontière / pas de frontière, lorsqu'une frontière est localisée entre les deux :

- ▷ premières syllabes de l'item (P<sub>0</sub> = 41% / P<sub>1</sub> = 46%)
- ▷ dernières syllabes de l'item (P<sub>0</sub> = 35% / P<sub>2</sub> = 45%)

L'analyse détaillée explique la cause de cette difficulté (section suivante).

Par contre, les participants n'ont aucun problème pour localiser la frontière une fois détectée (frontière entre les deux premières syllabes de l'item :  $P_1 = 46\%$  /  $P_2 = 13\%$  ; frontière entre les deux dernières syllabes de l'item :  $P_1 = 20\%$  /  $P_2 = 45\%$ ).

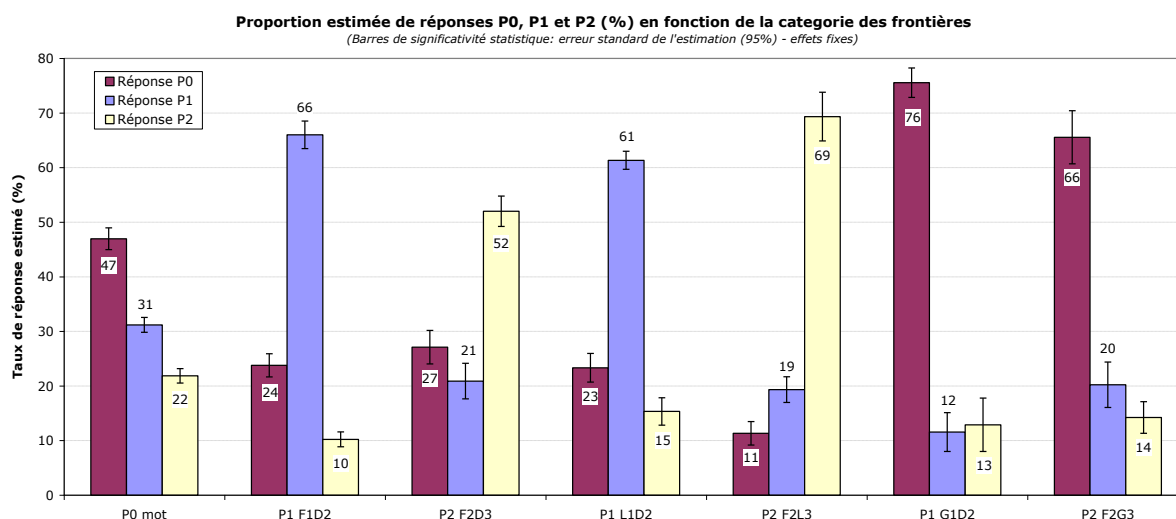
Cette première analyse confirme donc la capacité des auditeurs à détecter et localiser consciemment les frontières entre mots grâce à la seule information prosodique (durée,  $f_0$ ). Il s'agit donc d'une indication que les frontières entre mots sont prosodiquement marquées.

#### 4.1.4.3.1.2 Analyse détaillée de la répartition des réponses en fonction du type de frontière testé

Cette analyse nécessite l'utilisation de 3 modèles logistiques multi-variés, chacun contrastant 2 types de réponse avec le troisième. Ces 3 modèles sont présentés sous le label « Équation 33 » à l'Annexe 8 page 476.

Le Tableau 43 présente les taux estimés de réponse obtenus par type de réponse attendue pour chaque catégorie de stimulus.

Tableau 43 : Répartition des réponses  $P_0$ ,  $P_1$  et  $P_2$  par catégorie d'item



Étant donné qu'un problème de complexité se pose pour contraster les éléments de la figure précédente, le codage présenté dans le Tableau 44 ci-après est utilisé pour nommer les contrastes.

Pour une plus grande clarté, les résultats des contrastes entre types d'items par réponse (contrastés en ligne dans le Tableau 44) et entre réponses par type d'items

(contrastes en colonne dans le Tableau 44) sont présentés dans deux tableaux (respectivement, Tableau 45 et Tableau 46).

Tableau 44 : Codes adoptés pour la nomenclature des tests de contraste

Codes	0 Mot	1 F <sub>1</sub> D <sub>2</sub>	1 L <sub>1</sub> D <sub>2</sub>	1 G <sub>1</sub> D <sub>2</sub>	2 F <sub>2</sub> D <sub>3</sub>	2 F <sub>2</sub> L <sub>3</sub>	2 F <sub>2</sub> G <sub>3</sub>
Réponse P <sub>0</sub>	a	b	c	d	e	f	g
Réponse P <sub>1</sub>	h	i	j	k	l	m	n
Réponse P <sub>2</sub>	o	p	q	r	s	t	u

Tableau 45 : Test de significativité des contrastes pour les comparaisons entre type d'item par réponse

Réponse P <sub>0</sub>			Réponse P <sub>1</sub>			Réponse P <sub>2</sub>		
Code contraste	p<	Chi <sup>2</sup> (1, N=4050)	Code contraste	p<	Chi <sup>2</sup> (1, N=4050)	Code contraste	p<	Chi <sup>2</sup> (1, N=4050)
a/b	.0001	71.4	h/i	.0001	158.5	o/p	.0001	28.3
a/c	.0001	74.1	h/j	.0001	122.4	o/q	.0031	8.8
a/d	.0001	104.1	h/k	.0001	61.4	o/r	.0001	16.8
a/e	.0001	52.9	h/l	.0001	17.3	o/s	.0001	138.5
a/f	.0001	149.3	h/m	.0001	22.9	o/t	.0001	295.4
a/g	.0001	45.7	h/n	.0001	19.6	o/u	.0006	12.1
b/c	0.88	0.0	i/j	0.15	2.1	p/q	.025	5.2
b/d	.0001	216.5	i/k	.0001	230.1	p/r	.22	1.6
b/e	0.26	1.3	i/l	.0001	170.4	p/s	.0001	153.2
b/f	.0001	23.2	i/m	.0001	181.2	p/t	.0001	257.6
b/g	.0001	148.0	i/n	.0001	175.0	p/u	.069	3.3
c/d	.0001	219.7	j/k	.0001	200.3	q/r	.30	1.1
c/e	0.20	1.7	j/l	.0001	140.9	q/s	.0001	123.0
c/f	.0001	21.8	j/m	.0001	151.2	q/t	.0001	231.2
c/g	.0001	150.9	j/n	.0001	79.8	q/u	.64	0.2
d/e	.0001	192.6	k/l	.0002	14.1	r/s	.0001	138.1
d/f	.0001	297.2	k/m	.0014	10.2	r/t	.0001	245.8
d/g	.0011	10.8	k/n	.0005	12.4	r/u	.56	0.3
e/f	.0001	34.2	l/m	.57	0.3	s/t	.0001	28.0
e/g	.0001	126.4	l/n	.81	0.1	s/u	.0001	129.9
f/g	.0001	228.2	m/n	.74	0.1	t/u	.0001	238.1

Tableau 46 : Test de significativité des contrastes entre conditions pour les comparaisons entre réponses par type d'item

Contraste réponses P <sub>0</sub> /P <sub>1</sub>			Contraste réponses P <sub>0</sub> /P <sub>2</sub>			Contraste réponses P <sub>1</sub> /P <sub>2</sub>		
Code contraste	p<	Chi <sup>2</sup> (1, N=4050)	Code contraste	p<	Chi <sup>2</sup> (1, N=4050)	Code contraste	p<	Chi <sup>2</sup> (1, N=4050)
a/h	.0001	25.5	a/o	.0001	61.3	h/o	.0001	18.4
b/i	.0001	78.8	b/p	.0001	15.9	i/p	.0001	164.2
c/j	.0001	115.3	c/q	.027	5.0	j/q	.0001	121.0
d/k	.0001	166.8	d/r	.0001	160.8	k/r	.62	0.3
e/l	.12	2.5	e/s	.0001	28.1	l/s	.0001	59.8
f/m	.015	6.1	f/t	.0001	142.3	m/t	.0001	137.1
g/n	.0001	91.1	g/u	.0001	115.1	n/u	.051	3.8

#### 4.1.4.3.2 Présentation des résultats

En règle générale, les frontières sont correctement détectées / localisées dans les items grâce à la prosodie, ce qui permet d'affirmer que la première hypothèse de l'étude est validée.

Les performances de détection / localisation varient en fonction des frontières considérées, appuyant la deuxième hypothèse. On remarque en particulier que :

- ▷ les frontières entre mots de classe ouverte sont correctement détectées et localisées (contrastes b/i et i/p, e/s et l/s, c/j et j/q, f/t et m/t).
- ▷ les frontières entre mots de classe ouverte et mots de classe fermée sont manquées la plupart du temps : les auditeurs répondent  $P_0$  à un taux très élevé (76 et 66%), confondant de telles frontières avec un mot trisyllabique (absence de frontière – contrastes d/k et d/r, g/n et g/u). L'attribution d'une réponse  $P_0$  est dans ce cas largement plus élevée que pour une véritable situation  $P_0$  (47% - contrastes a/d et a/g), indiquant qu'un mot trisyllabique de classe ouverte paraît moins cohérent à la perception qu'une séquence de 3 syllabes comportant un monosyllabe de classe fermée, suivi ou précédé de 2 syllabes d'un mot trisyllabique de classe ouverte.

Les items trisyllabiques sont distingués des situations où une frontière existe entre deux mots lexicaux (47% de  $P_0$  contre 31%  $P_1$  et 22%  $P_2$  – contrastes a/h et a/o).

La difficulté à distinguer entre présence et absence d'une frontière mise en évidence dans l'analyse générale (Section 4.1.4.3.1.1, page 214 : taux élevé de réponses  $P_0$  pour une situation  $P_1$  ou  $P_2$ ) trouve sa justification avec la perception des frontières  $G_1D_2$  et  $F_2G_3$  comme un mot trisyllabique.

#### 4.1.4.4 Conclusions de l'Expérience 2

Les 2 hypothèses de l'Expérience 2 sont globalement validées par les résultats obtenus :

- ▷ Détection des frontières lexicales avec un taux supérieur au hasard, grâce à la seule information de  $f_0$  et de durée.
- ▷ Variation des performances de détection / localisation en fonction du type de frontière considéré.

Les analyses statistiques effectuées apportent les détails suivants :

- ▷ L'information prosodique marque les frontières entre mots de classe ouverte. Ce résultat est toujours vrai dans l'expérience, quelle que soit la catégorie de syllabe lexicale impliquée (D, F, L), et quelle que soit la position de la frontière dans l'item ( $F_1D_2$  et  $F_2D_3$ , ou  $L_1D_2$  et  $F_2L_3$ ). Il semble donc que l'information prosodique soit structurée de manière à dissocier les items supportant la trame sémantique du message. Un tel balisage peut s'avérer utile pour éviter une segmentation incorrecte dans les cas d'ambiguïté locale (exemple : mots enchâssés à l'initiale de mots plus longs, possibilités de confusion entre statut lexical et grammatical d'un mot monosyllabique).
- ▷ L'information prosodique ne marque pas les frontières entre mots lexicaux plurisyllabiques et monosyllabes de classe fermée, mais au contraire, paraît renforcer la perception de la séquence comme une unité (séquences GDM et MFG, taux de réponse  $P_0$  plus important (respectivement 76 et 66%) que pour la séquence DMF (47%  $P_0$ )). Ce phénomène ne s'explique pas en termes de similarité acoustique entre syllabes G et [F ou D] (voir analyses acoustiques, Figure 33, page 167, Figure 38, page 170). Il signifie que les frontières  $G_1D_2$

et  $F_2G_3$  (qui étaient les plus fortement caractérisées et distinguées dans les expériences de comparaison (Simulation 2, Expérience 1)), sont ici assimilées à un même groupe (absence de frontière). Or, les courbes moyennes de  $f_0$  (Figure 39, page 171) montrent que les frontières GD et FG ne correspondent pas à une même forme sonore. De plus, il faut se souvenir que dans l'Expérience 2, les frontières GL et GG, bien qu'ayant un profil mélodique différent de celui des frontières GD, avaient tendance à être assimilées aux frontières GD. On a donc, sur deux expériences, confirmation d'un phénomène particulier touchant la perception des frontières GD, et pouvant être interprété comme un biais cognitif lié aux traitements de haut niveau lors de la perception d'un message parlé (le fort marquage de frontière, qui est perçu (d'après l'Expérience 1), pourrait permettre au système de connaître l'existence d'une frontière et d'identifier les éléments composant cette frontière, afin de les rattacher comme une seule unité fonctionnelle grâce à l'intervention d'une représentation / d'un module particulier de traitement).



### 4.1.5 Expérience 3 : Effets du débit d'élocution et de la hauteur de la voix sur la perception

La partie précédente du travail (analyses acoustiques, Simulations 1 et 2, Expériences 1 et 2) a montré que la dimension suprasegmentale du signal de parole n'était pas amorphe, mais structurée en fonction de paramètres morphosyntaxiques et topologiques. Elle a aussi montré que cette structuration pouvait être utile pour guider le traitement cognitif. On doit donc s'interroger sur les mécanismes psychophysiques qui permettent ce guidage. C'est le cœur de l'Expérience 3.

#### 4.1.5.1 Aspects théoriques

Soit le guidage est de nature sensorielle (guidage par les données), soit il est de nature sensorielle et cognitive (guidage par les données et par des représentations (réactivité du système à des formes prosodiques particulières)).

Au vu des résultats de la Simulation 2 et des Expériences 1 et 2, il est difficile d'affirmer que le guidage n'est que sensoriel, puisque les auditeurs humains présentent un biais de leurs réponses pour la catégorie GD (Expérience 1), et savent reconnaître la « sonorité » des frontières lexicales avec un taux qui diffère fortement du hasard (Expérience 2). On est cependant certain qu'il est au moins sensoriel. La question posée dans l'Expérience 3 est donc relative aux mécanismes psychophysiques par lesquels ce guidage sensoriel pourrait avoir lieu. Au moins deux mécanismes peuvent être envisagés :

- ▷ un fenêtrage temporel des éléments linguistiques infra-lexicaux (phonèmes, syllabes), qui permet de moduler leur résolution dans l'espace perceptif grâce au paramètre durée de la prosodie. Un élément linguistique infra-lexical impliqué dans une information d'importance pour le traitement serait décrit par une quantité plus grande d'information acoustique afin que son identification soit assurée et facilitée. Le but serait de garantir l'optimisation des traitements cognitifs du message reçu.
- ▷ une modulation de la force de la stimulation qui agit directement sur le niveau vigile du système récepteur (modulation de l'attention - terme considéré dans son acception psychophysologique). Le but serait de modifier la capacité du système à traiter de l'information en fonction des impératifs cognitifs, grâce aux paramètres fréquence fondamentale et intensité de la prosodie. Face à un élément linguistique infra-lexical impliqué dans une information d'importance pour le traitement, le système récepteur serait plus fortement stimulé, ce qui lui permettrait de mobiliser plus de ressources attentionnelles en vue d'un traitement plus difficile ou plus délicat.

Le sens de ces deux mécanismes est donc d'assurer une optimisation du traitement du message, en désaturant le système grâce à la durée (donner plus

d'information acoustique pour décrire une même quantité d'éléments abstraits) et en augmentant sa disponibilité à traiter de l'information grâce à la fréquence fondamentale et l'intensité. Dans l'Expérience 3, seulement deux des trois paramètres prosodiques sont manipulés, la durée (débit d'élocution) et la fréquence fondamentale (hauteur de voix), techniquement plus faciles à maîtriser que l'intensité.

#### 4.1.5.2 Aspects techniques

##### 4.1.5.2.1 Facteurs expérimentaux et conditions

Deux facteurs prosodiques à 2 modalités chacun :

- ▷ Débit d'élocution (rapide (R) vs. lent (L))
- ▷ Hauteur de la voix (haute (H) vs. basse (B))

Ces deux facteurs sont croisés entre eux selon un plan complet. On a donc 4 conditions prosodiques (R.H, R.B, L.H, L.B).

Afin de faire faire une tâche de classification aux participants (décision lexicale classique), une partie des items (mots et pseudo-mots) de l'Expérience 4 (présentée au Chapitre 2 de la Partie IV (se reporter à la Section 4.2.2.2.1, page 236 pour plus de détails sur les items)) est utilisée.

Dans ce matériel linguistique, cinq facteurs lexicaux sont manipulés :

- ▷ Lexicalité de l'item (mot (M) vs. pseudo-mot (nM))

Pour les mots : 2 facteurs à 2 modalités chacun

- ∩ Présence d'un mot enchâssé à l'initiale du mot stimulus (oui (E) vs. non (nE))
- ∩ Enchâssabilité du mot stimulus à l'initiale d'un mot plus long (oui (I) vs. non (nI))

Ces deux facteurs lexicaux sont croisés entre eux selon un plan complet. On a donc 4 types de mots (E.I, E.nI, nE.I, nE.nI – des exemples de mots se trouvent à la Section 4.2.2.2.1, page 236).

Pour les pseudo-mots : 2 facteurs à 2 modalités chacun

- ∩ Appartenance de la syllabe au Français (oui (voire le point suivant ci-dessous) vs. non (« syllabe » illégale en Français selon BRULEX (III) – exemple : /mti/))
- ∩ Pour les syllabes du Français seulement, leur position d'occurrence dans les mots du lexique (syllabe de début de mot (D – exemple : /kli/ existe à l'initiale du mot « climat ») vs. syllabe de position non initiale (nD – exemple : /3εk/ existe en position non initiale dans le mot « éjection »))

Ces deux facteurs ne sont pas croisés entre eux, mais partiellement emboîtés. Ils permettent de distinguer 3 types de pseudo-mots (D, nD, II).

La situation expérimentale comporte donc 4 conditions prosodiques et 7 conditions lexicales, croisées entre elles selon un plan complet, donné dans le Tableau 47, soient un total de 28 conditions expérimentales.

Tableau 47 : Plan de l'expérience

	Mot				Pseudo-mot		
	E.I	E.nI	nE.I	nE.nI	D	nD	III
<b>R.H</b>	R.H.E.I	R.H.E.nI	R.H.nE.I	R.H.nE.nI	R.H.D	R.H.nD	R.H.III
<b>R.B</b>	R.B.E.I	R.B.E.nI	R.B.nE.I	R.B.nE.nI	R.B.D	R.B.nD	R.B.III
<b>L.H</b>	L.H.E.I	L.H.E.nI	L.H.nE.I	L.H.nE.nI	L.H.D	L.H.nD	L.H.III
<b>L.B</b>	L.B.E.I	L.B.E.nI	L.B.nE.I	L.B.nE.nI	L.B.D	L.B.nD	L.B.III

Toutefois, les hypothèses étant relatives à la prosodie et non à des facteurs lexicaux ou à l'interaction des facteurs prosodiques et lexicaux, le plan d'analyse ne prend en compte que les 4 conditions prosodiques, présentées dans le Tableau 48.

Tableau 48 : Plan d'analyse

	Hauteur haute (H)	Hauteur basse (B)
<b>Débit rapide (R)</b>	R.H	R.B
<b>Débit lent (L)</b>	L.H	L.B

#### 4.1.5.2.2 Hypothèses opérationnelles

- ▷ Les auditeurs catégoriseront plus rapidement les syllabes produites avec un débit lent que celles produites avec un débit rapide (il est plus facile de reconnaître un item s'il est décrit par une grande quantité d'information).
- ▷ Les auditeurs catégoriseront plus rapidement les syllabes produites avec une voix haute que celle produites avec une voix basse (il est plus facile de reconnaître un item si le système récepteur est dans un état vigile plus élevé).

On s'attend donc à ce que les auditeurs présentent les temps de réponse les plus rapides lorsque les syllabes sont produites avec une voix haute et un débit lent.

#### 4.1.5.2.3 Matériel expérimental utilisé

Stimuli monosyllabiques en voix de synthèse (se reporter à la Section 3.4.1.3.2 page 144, pour les détails techniques relatifs à la synthèse des items, et à l'Annexe 9 page 479 pour la liste des items).

- ▷ Mots expérimentaux : 48 mots monosyllabiques (12 items par type de mot)
- ▷ Pseudo-mots expérimentaux : 48 pseudo-mots (16 items par type de pseudo-mot)

L'expérience ne comporte pas d'items distracteurs, étant donné la tâche demandée et le propos de l'étude.

Chaque item est produit avec les 4 prosodies différentes (R.H, R.B, L.H, L.B). L'Expérience 3 comporte donc un total de 384 items.

#### **4.1.5.2.4 Organisation de l'expérience**

Chaque participant est testé une fois avec tous les mots et tous les pseudo-mots. Chaque participant est testé avec les 4 prosodies, de manière à ce que la quantité d'exemples prosodiques soit équilibrée dans les différentes conditions lexicales. L'expérience nécessite donc 4 groupes des participants / listes d'items (répartition selon une organisation en carré latin).

#### **4.1.5.2.5 Population testée**

Les données de 44 francophones monolingues, sans troubles auditifs, cognitifs et moteurs, étudiant à l'université de Genève, ont été utilisées pour les analyses. Le recueil des données a été effectué en novembre 2002.

#### **4.1.5.2.6 Tâche**

Tâche de catégorisation (décision lexicale classique). Modalité auditive.

#### **4.1.5.2.7 Procédure de recueil des données**

Expérience réalisée à l'université de Genève. Détails dans la Partie III, Chapitre 4.

44 \* 96 = 4224 mesures effectuées.

La phase test de l'expérience ne dure pas plus de 10 minutes.

#### **4.1.5.2.8 Méthode d'analyse**

- ▷ Détail des procédures pour le tri des données et les analyses dans le Chapitre 3, Partie III.
- ▷ Analyses de régression multiniveaux. Utilisation d'un modèle à deux niveaux (sujets, items. Corrélation intra classe : 13.99%).
- ▷ 990 valeurs manquantes (23.44% de non-réponses et de valeurs filtrées (détail par condition : R.B 22.3% ; R.H 22.3% ; L.B 26.4% ; L.H 22.8%)).

#### **4.1.5.2.9 Paramètres contrôlés**

À la construction du matériel et à la structuration de l'expérience

- ▷ Détails relatifs à la maîtrise des aspects linguistiques : Expérience 4, Section 4.2.2.2.9, page 237.

- ▷ Afin de standardiser mathématiquement la prosodie pour chaque condition, les stimuli auditifs sont générés automatiquement par synthèse vocale grâce à un système de synthèse par règles. Détails à la Section 3.4.1.3.2 page 144.

#### Lors des analyses statistiques

- ▷ La lexicalité des items et les interactions que ce facteur fait avec les 2 facteurs prosodiques sont contrôlés.
- ▷ La position acoustique (en ms.) du point de reconnaissance (pour les mots) ou du point de déviation<sup>106</sup> (pour les pseudo-mots) est contrôlée.
- ▷ Listes : se reporter à la note de bas de page n°105, page 201.

### 4.1.5.2.10 Résultats obtenus

#### 4.1.5.2.11 Rapport des analyses statistiques

L'équation du modèle (Équation 34) et les tests de significativité relatifs aux covariables peuvent être consultés à l'Annexe 10 page 481. Le Tableau 49 présente l'effet des facteurs étudiés.

Tableau 49 : Tests de significativité des facteurs étudiés

	Coefficient	Chi <sup>2</sup> <sub>(1, N=3234)</sub>	p<	% variance expliquée
<b>R</b>	25.433	5.674	<b>.018</b>	0.115
<b>H</b>	13.799	2.246	.14	0.039
<b>R.H</b>	-32.574	6.036	<b>.015</b>	0.134

L'analyse des effets fixes de la prosodie fait ressortir un effet principal du débit d'élocution. Plus un item est prononcé rapidement, plus il est difficile à traiter (écart moyen estimé entre les 2 modalités de 25 ms.). L'hypothèse relative à la vitesse d'élocution est donc validée par ces résultats.

La hauteur de la voix est sans effet significatif. On note toutefois que la probabilité approche le seuil tendanciel, le facteur agissant apparemment dans le sens d'une augmentation des temps de traitement lorsque la fréquence fondamentale est haute. L'hypothèse relative à la fréquence fondamentale n'est donc pas validée par ces résultats.

Une interaction existe entre les deux facteurs. Elle est responsable d'une diminution significative des temps de réponse, de 33 ms. L'action combinée des deux facteurs étudiés est donc plus complexe qu'un simple cumul des effets de chacun d'eux, tel que supposé. Les hypothèses proposées seraient donc trop simples.

<sup>106</sup> Phonème à partir duquel une séquence de phonèmes ne peut plus être un mot.

Tableau 50 : Tests de significativité des contrastes

Contraste		Chi <sup>2</sup> (1, N=3234)	p<
R.H	R.B	3.867	<b>.049</b>
	L.H	0.455	.50
	L.B	0.382	.54
R.B	L.H	1.229	.27
	L.B	5.674	<b>.018</b>
L.H	L.B	2.246	.14

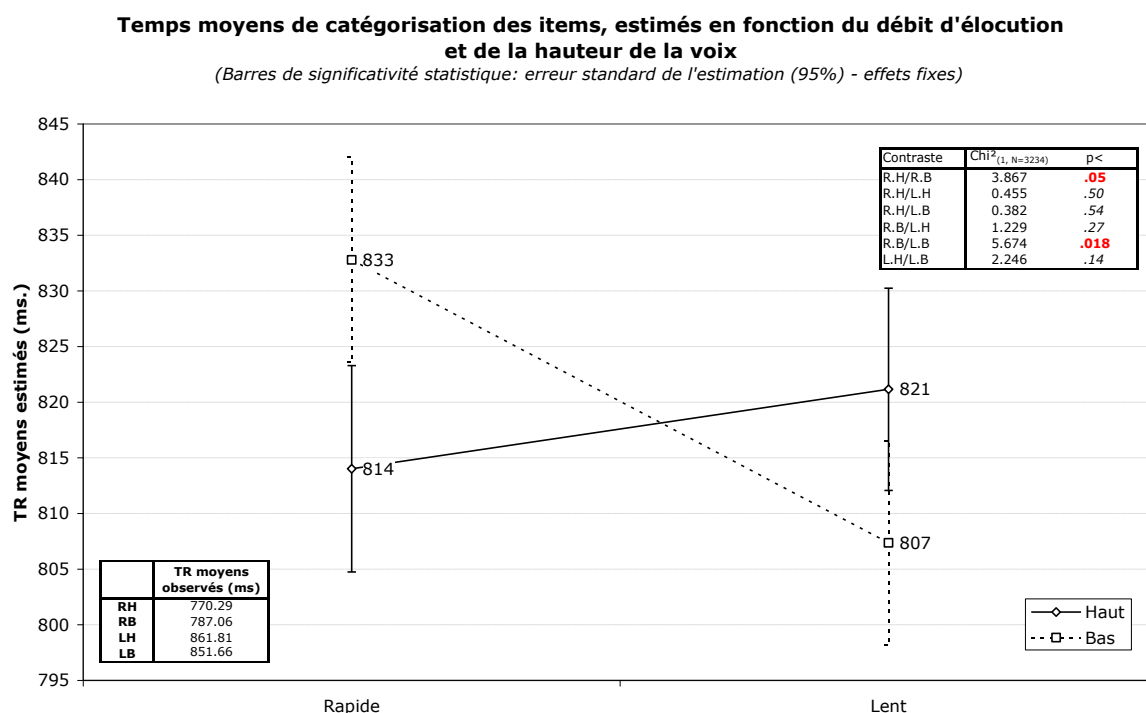
L'analyse des contrastes révèle un écart dans les temps de réponse, lié :

- ▷ au débit lorsque la voix est basse (R.B / L.B - pas d'effet avec une voix haute (R.H / L.H))
- ▷ à la hauteur de la voix lorsque le débit d'élocution est rapide (R.H / R.B - pas d'effet dans le cas d'un débit lent (L.H / L.B)).

Les autres contrastes sont non significatifs.

### 4.1.5.3 Présentation des résultats

Figure 53 : Temps de réponse moyens estimés en fonction du débit d'élocution et de la hauteur de la voix



Le motif de résultats obtenu est difficile à interpréter dans le cadre des hypothèses proposées pour l'Expérience 3. En effet, on comprend mal pourquoi l'effet du débit s'inverse en fonction de la hauteur de la voix. L'explication suivante est proposée :

- ▷ Étant donné que les stimuli sont en voix de synthèse, leur audibilité pourrait ne pas être identique dans les 4 conditions prosodiques<sup>107</sup>.
- ▷ Le système disposerait en mémoire de représentations mentales de formes sonores « prototypiques » (infra-lexicales et/ou lexicales), qui seraient activées proportionnellement à leur degré de correspondance avec la forme sonore entendue.

Une forme sonore se définit aussi par le nombre de vibrations qui la composent, et pas seulement par la durée, la fréquence, l'intensité, ou le spectre correspondant à un segment particulier. Ainsi, un item qui est produit :

- ▷ avec un débit lent se compose d'un nombre de vibrations périodiques et aperiodiques plus élevé qu'un item produit avec un débit rapide.
- ▷ avec une voix haute se compose d'un nombre de périodes plus élevé qu'un item produit avec une voix basse (essentiellement pour les régions voisées du signal).

On a donc, commun à la durée et à la  $f_0$ , un facteur sous-jacent qui est le nombre de périodes ou d'oscillations qui composent un son<sup>108</sup>.

Ainsi, l'audibilité d'un item et le degré d'adéquation entre forme sonore et prototype mental pourraient toutes deux dépendre du nombre de vibrations qui composent le stimulus (résolution acoustique). Trop peu d'oscillations rendent l'item difficilement reconnaissable, tandis que trop de vibrations créent un éloignement entre le stimulus et la représentation mentale ou prototype qui lui est associé en mémoire à long terme. En observant la structure acoustique des stimuli en fonction des deux facteurs expérimentaux, on s'aperçoit qu'un item qui est produit avec :

- ▷ un débit lent et une fréquence fondamentale élevée, possède la plus grande quantité d'information sonore.
- ▷ un débit rapide et une fréquence fondamentale élevée, ou un débit lent et une fréquence fondamentale basse, présente une quantité d'information sonore intermédiaire.
- ▷ un débit rapide et une fréquence fondamentale basse, présente une quantité d'information sonore minimale.

Si l'audibilité était le seul paramètre en cause, le motif de résultats suivant devrait être observé dans les temps de traitement : [L.H] < [R.H et L.B] < [R.B]. Or, L.H n'est

---

<sup>107</sup> L'audibilité générale de ces stimuli en voix de synthèse est déjà très inférieure à celle de stimuli en voix naturelle, malgré la haute qualité du système de synthèse vocale utilisé. Une preuve est apportée dans ce sens par la quantité élevée de valeurs manquantes (près du quart des mesures).

<sup>108</sup> Ce facteur pourrait être en relation avec l'augmentation de durée subjective d'une voyelle lorsqu'une modulation mélodique se produit, comparativement à une voyelle qui présente une  $f_0$  statique (Lehiste, 1976).

pas la condition qui présente les temps de traitement les plus brefs. On observe en effet plutôt le motif de résultats suivant : [L.B] < [R.H] < [L.H] < [R.B].

Il peut s'expliquer par le phénomène de résolution acoustique décrit plus haut (L.B et R.H auraient une résolution optimale, la durée et la hauteur se combinant pour fournir un nombre d'oscillations voisin du prototype mental de la forme sonore ; L.H aurait une résolution trop élevée (parfaitement audible (lent), mais trop éloigné du prototype (trop grande quantité d'oscillations)) ; R.B aurait une résolution trop basse (mal audible car trop rapide par rapport au nombre d'oscillations qui décrivent l'item)).

#### 4.1.5.4 Conclusions de l'Expérience 3

Des 2 hypothèses de l'Expérience 3, l'hypothèse relative au débit d'élocution est validée par les résultats, l'hypothèse relative à la hauteur de la voix ne l'est pas.

Les résultats obtenus dans l'Expérience 3 appuient l'idée d'un rôle prosodique dans la résolution acoustique des éléments linguistiques infra-lexicaux qui constituent le message parlé<sup>109</sup>. La durée, ou la durée et la f0, moduleraient la résolution acoustique des segments.

Une résolution optimale serait la seule à même de favoriser le traitement, cette optimalité pouvant correspondre à une quantité d'information voisine de celle définie dans la représentation mentale formelle de l'unité considérée. Afin de faciliter le traitement du message, le locuteur devrait donc produire une onde ciblant précisément les prototypes infra-lexicaux. Le traitement perceptif pourrait aussi utiliser des stratégies basées sur l'alternance entre zones faciles / difficiles à traiter. Rendre certaines régions du signal plus difficilement perceptibles pourrait permettre de différer un traitement, afin d'éviter au système de s'engager dans une fausse direction puis de devoir corriger ses hypothèses avec l'arrivée d'une information désambiguïsante plus tardive.

Aucun phénomène en rapport avec une modulation du niveau vigile dans le système réceptif, tel qu'initialement supposé pour la fréquence fondamentale, n'a pu être mis en évidence. Cet échec peut tenir aux caractéristiques acoustiques des stimuli (synthèse vocale), ou au fait que la fréquence fondamentale n'est pas le facteur impliqué dans ces modulations du niveau vigile. L'intensité, seule ou combinée à la fréquence fondamentale, pourrait être un candidat plus adapté pour remplir une telle fonction.

---

<sup>109</sup> Ceux-ci sont probablement de gabarit phonémique. L'effet est indépendant de la lexicalité des stimuli alors que les syllabes pseudo-mot renferment aussi des « syllabes » illégales, ce qui devrait induire une distinction mot/pseudo-mot si le gabarit était syllabique.



#### 4.1.6 Conclusions relatives aux études présentées dans le Chapitre 1, Partie IV

Les analyses acoustiques statiques et dynamiques ont révélé que la dimension suprasegmentale du signal acoustique est fortement organisée en fonction de la morphosyntaxe et de la topologie lexicale. Un code prososyllabique semble exister, dont le principe fondamental (contrastes syllabiques) est universellement applicable à toute langue.

L'Expérience 1 a montré que ce code prososyllabique est perceptible.

L'Expérience 2 a montré qu'il est utilisable consciemment pour détecter et localiser une frontière lexicale avec une probabilité largement supérieure au hasard.

L'Expérience 3 a révélé l'une des pistes possibles pour expliquer comment ce code prososyllabique pourrait agir en bas niveau sur le guidage des traitements lexicaux : en modulant la résolution acoustique de la ligne segmentale.

Ces résultats appuient donc l'hypothèse d'un guidage ascendant par la prosodie du découpage du flux de parole en mots. Ils permettent par ailleurs de tirer les conclusions suivantes :

##### Concernant les caractéristiques du signal de parole

- ▷ La caractérisation prosodique des syllabes forme un code utile pour baliser le flux sonore et localiser les frontières des mots. Les frontières de mots ne sont pas contenues dans le mot lui-même (à son début ou à sa fin), mais dans la transition dynamique d'un mot à l'autre, sous forme de contrastes localisés sur au moins 2 syllabes. Pour cette raison, le code prososyllabique est qualifié de relatif. Les frontières lexicales ne seraient donc pas marquées prosodiquement de manière simple, d'où la difficulté pour mesurer leur présence acoustique dans la prosodie avec des procédés linéaires.
- ▷ L'Expérience 3 a montré que le guidage ascendant des traitements pouvait se faire via une modulation de la résolution acoustique des segments, celle-ci étant à même de modifier la vitesse de l'activation lexicale. Dans les zones du signal importantes pour le traitement, la prosodie optimiserait la résolution acoustique pour faciliter les traitements perceptifs et les traitements abstraits du message. La résolution acoustique pourrait aussi être impliquée dans des stratégies de traitement plus complexes, comme de retarder un processus pour éviter l'engagement du système dans une direction erronée (exemple : segmentation lexicale suite au traitement d'un mot enchâssé initial).
- ▷ Si la prosodie orchestre les traitements perceptifs par guidage sensoriel, c'est que le système émetteur encode à dessein des informations nécessaires à leur décodage. Les analyses acoustiques et la Simulation 1 ont montré que cet encodage était relatif à l'information morphosyntaxique et topologique, et qu'il traduisait probablement les mécanismes mentaux de manipulation du langage lors de la production. Par conséquent, la prosodie permettrait une mesure des

plus directes et des plus naturelles des processus mentaux impliqués dans la production du langage, d'une part, et d'autre part, l'existence de catégories prosodiques de syllabes en fonction du type de mot porteur et en fonction de la position de la syllabe dans un mot correspondrait à une validation psychologique de ces facteurs.

#### Concernant le traitement du signal de parole

- ▷ Les indices prosodiques seraient traités de manière multidimensionnelle et dynamique lors de la perception, puisque les analyses de régression (Annexe 11 page 482) échouent à montrer une relation simple entre le comportement et la réalité physique, alors qu'une simulation de la tâche par réseaux neuromimétiques (Simulation 2) produit un profil artificiel de performances comparable à celui des auditeurs humains.
- ▷ Si l'étude semble appuyer l'idée d'un guidage sensoriel du traitement perceptif par la prosodie, elle met aussi en évidence quelques phénomènes pouvant traduire l'influence de représentations mentales en lien avec une implication cognitive de la prosodie, ayant interféré dans la perception des participants (assimilation des frontières GG et GL à la frontière GD (Expérience 1) en dépit d'une divergence entre profils mélodiques (Figure 39, page 171), détection de la frontière GD comme un mot (Expérience 2) en dépit de son fort marquage acoustique (Simulation 2) et d'un motif mélodique de frontière (motif en V - Figure 39, page 171)).

#### Conséquences pour la segmentation du flux de parole en mots

- ▷ L'information prososyllabique est théoriquement exploitable dans la segmentation naturelle du signal de parole en mots, parce qu'elle induit des contrastes de place en place entre syllabes, certains de ces contrastes apparaissant comme naturellement associés à des frontières lexicales (Expérience 2). L'information prosodique est donc aussi exploitable pour le traitement automatique de la parole, et cette étude a montré une manière possible de procéder efficacement dans ce but (comparaisons de proche en proche pour évaluer les zones de rupture de cohésion sur la séquence).
- ▷ Aucune des 3 frontières impliquant un mot de classe fermée en deuxième position dans l'item n'est marquée (Figure 39, page 171). Au contraire, toutes les frontières entre deux mots de classe ouverte, ainsi que toutes celles entre un mot de classe fermée en première position dans l'item, et un mot de classe ouverte en deuxième position dans l'item, sont spécifiées par un motif mélodique en V (Figure 39, page 171). Ceci pourrait signifier que, le traitement étant orienté de « gauche à droite » (temporalité), les mots de classe fermée n'ont pas besoin d'être spécifiés par la prosodie vu qu'ils le sont par la trame syntaxique du message. Par contre, les mots de classe ouverte le sont de manière assez systématique pour lever toute ambiguïté (nature lexicale ou grammaticale du mot, mot libre ou enchâssé). Une telle distinction pourrait s'avérer représenter une aide très importante à la segmentation lexicale.
- ▷ Deux frontières (GD et FG) induisent les performances les plus élevées dans les tests de discrimination (Simulation 2, Expérience 1). Toutes deux impliquent la conjonction de la trame syntaxique (G) avec la trame sémantique (mot polysyllabique de classe ouverte - D ou F), mais dans le cas de mots lexicaux monosyllabiques (GL et FL), une telle conjonction ne conduit pas à de telles performances. Il faut donc conclure que dans le matériel expérimental utilisé, une caractérisation plus ou moins forte (ciblage plus

précis d'une forme distincte des autres) existe en rapport avec un traitement lexical ou cognitif plus difficile ou plus complexe.

- ▷ Les frontières impliquant un mot monosyllabique de classe ouverte entraînent généralement des performances moins bonnes que les frontières impliquant un début de mot plurisyllabique de classe ouverte (comparaison des performances entre FD et FL, GD et GL, LD et LL – Expérience 1). Les frontières avec un début de mot plurisyllabique de classe ouverte auraient une forme acoustique moins variable et plus ciblée que les frontières avec un monosyllabe de classe ouverte (peut-être parce qu'une unité plurisyllabique possède une cohérence acoustique plus importante). Toutefois, l'aspect acoustique n'explique pas l'assimilation constatée dans les réponses des participants (Expérience 1), des frontières GL avec les frontières GD, et dans une moindre mesure, des frontières FL avec les frontières FD. Ce phénomène traduit peut-être l'existence d'un biais de traitement pour les initiales de mot de classe ouverte plurisyllabique, qui pourrait être en lien avec une stratégie visant à pallier à l'ambiguïté organisationnelle du lexique mental (nombre élevé de mots plurisyllabiques commençant par un mot enchâssé monosyllabique). Cela permettrait au système de disposer d'une procédure pour distinguer les mots libres monosyllabiques de classe ouverte (à segmenter), des mots monosyllabiques enchâssés à l'initiale de mots plurisyllabiques de classe ouverte (à ne pas segmenter). Les résultats de ce travail indiquent qu'il ne serait pas nécessaire de marquer les mots monosyllabiques (reconnaissance facile de mots simples), mais de favoriser l'accès lexical des mots porteurs (reconnaissance plus difficile de mots complexes).

## **4.2 Chapitre 2 - Linéarisation d'éléments superposés : quelle organisation pour le système de reconnaissance des mots parlés ?**

Le premier chapitre de la partie expérimentale était consacré à la recherche d'un code prososyllabique en rapport avec les caractéristiques morphosyntaxiques et topologiques des syllabes, et supposé servir à guider de manière ascendante le traitement de segmentation du flux de parole en mots.

Le deuxième chapitre de la partie expérimentale est consacré à l'étude des conséquences de l'organisation superposante du lexique sur la reconnaissance des mots parlés, et donc, incidemment, sur la segmentation du flux de parole en mots selon un mode descendant.

### **4.2.1 Introduction**

La question centrale abordée dans ce deuxième chapitre est : est-il vraiment plus difficile de reconnaître un mot impliqué dans une superposition lexicale qu'un mot qui ne l'est pas, comme le proposent la plupart des théories psycholinguistiques ? Cette question doit être posée, car l'idée d'une difficulté accrue du traitement face à une configuration universelle (se reporter à la Section 2.3 page 61) apparaît a priori paradoxale. En effet, si le traitement des superpositions lexicales représentait un véritable problème, le langage serait structuré différemment, ou il ne se serait peut-être pas développé jusqu'à un stade d'élaboration aussi poussé.

Le corps de ce deuxième chapitre se compose de 11 expériences et d'une Analyse exploratoire, qui reflètent implicitement 3 étapes dans l'évolution de la recherche :

- ▷ découverte d'un phénomène singulier et confirmation de son existence en modalité auditive (Expériences 4 à 7).
- ▷ vérification de son existence dans une autre modalité sensorielle (Expérience 8) et avec une autre tâche expérimentale (Expériences 9 et 10).
- ▷ recherche d'une origine infra-lexicale possible à ce phénomène : exploration de l'hypothèse syllabique (Expériences 11 à 14), de l'hypothèse phonémique (Analyse exploratoire), et comparaison des deux hypothèses (Analyse exploratoire).

Le chapitre se termine par une récapitulation des effets obtenus dans ces expériences et d'une brève conclusion.

Deux types de superpositions lexicales existent :

- ▷ Les superpositions lexicales réelles font référence à la présence d'un mot enchâssé dans le mot stimulus (Exemple : /la/ dans /plas/).
- ▷ Les superpositions lexicales virtuelles font référence à la possibilité qu'un mot stimulus soit lui-même enchâssé dans un autre mot qui n'est pas produit, mais qui est représenté en mémoire à long terme (Exemple : /plas/ dans /kõtāplasjō/).

Les superpositions lexicales peuvent de plus être distinguées topologiquement en fonction de leur position dans un mot :

- ▷ Elles peuvent être initiales (mot enchâssé aligné avec le mot stimulus - Exemple : /pla/ dans /plas/)
- ▷ Elles peuvent être non initiales (mot enchâssé non aligné avec le mot stimulus - Exemple : /las/ dans /plas/).

Par conséquent, il est possible de définir 4 catégories de superpositions de mots :

- ▷ Les superpositions lexicales réelles initiales
- ▷ Les superpositions lexicales réelles non initiales
- ▷ Les superpositions lexicales virtuelles initiales
- ▷ Les superpositions lexicales virtuelles non initiales

Afin de poser un cadre théorique à cette partie de l'étude, le modèle TRACE de la reconnaissance des mots parlés a été choisi comme référence. Ce choix est justifié par 3 arguments :

- ▷ TRACE est un modèle classique de la psycholinguistique. Il a été validé par de nombreux travaux expérimentaux au cours de ses 25 années d'existence.
- ▷ TRACE fonctionne sur la base d'un ensemble de mécanismes simples, relativement aisés à comprendre et à se représenter, à partir desquels il est possible de faire des prédictions très précises sur la réaction du système face à une configuration lexicale donnée.
- ▷ TRACE propose un mécanisme original de compétition lexicale basé sur l'inhibition latérale, implémenté pour résoudre les situations ambiguës telles que les superpositions lexicales.

TRACE prédit qu'en cas de superposition lexicale :

- ▷ initiale (réelle et virtuelle), un conflit a lieu entre représentations lexicales, qui se traduit par une augmentation des temps de reconnaissance du mot stimulus.
- ▷ non initiale (réelle et virtuelle), un conflit a également lieu entre représentations lexicales, mais son impact sur les temps de reconnaissance du mot stimulus est moindre qu'avec des superpositions lexicales initiales, et peut être non mesurable en fonction des configurations.

Pour reprendre l'organisation de ce deuxième chapitre en fonction du classement des superpositions lexicales qui vient d'être introduit (nature \* position), on a :

- ▷ L'étude de l'impact des superpositions lexicales initiales réelles sur la reconnaissance des mots dans les Expériences 4, 5, 6, 7, 8, 9, et 10
- ▷ L'étude de l'impact des superpositions lexicales initiales virtuelles sur la reconnaissance des mots dans les Expériences 4, 9, et 13
- ▷ L'étude de l'impact des superpositions lexicales non initiales réelles sur la reconnaissance des mots dans les Expériences 6, 7, et 10
- ▷ L'étude de l'impact des superpositions lexicales non initiales virtuelles sur la reconnaissance des mots dans l'Expérience 14

### 4.2.2 Expérience 4 : Effet d'une superposition lexicale initiale réelle et/ou virtuelle sur la reconnaissance auditive d'un mot monosyllabique en modalité auditive

L'Expérience 4 explore l'impact de superpositions lexicales initiales réelles et virtuelles sur la vitesse de reconnaissance de mots parlés monosyllabiques à l'aide d'une tâche de décision lexicale en « go, no-go ». TRACE prédit alors qu'une compétition aura lieu entre mots superposés, et que les temps de reconnaissance du mot stimulus seront allongés sous l'effet de cette compétition.

Logiquement, on peut cependant supposer que des effets différents apparaissent si les superpositions lexicales sont réelles ou virtuelles, puisque les premières impliquent un traitement au niveau infra-lexical et au niveau lexical, alors que les deuxièmes n'impliquent que des traitements au niveau lexical.

L'Expérience 4 teste des mots d'un gabarit monosyllabique afin de travailler avec des stimuli :

- ▷ de structure phonologique simple
- ▷ à durée minimale (concentration des processus de traitement sur une quantité réduite de temps, impliquant une probabilité plus importante de mesurer des effets. Avec un matériel plurisyllabique, les effets peuvent disparaître par dilution dans le temps ou parce qu'ils sont masqués par d'autres processus survenant plus tardivement)
- ▷ autorisant une distinction techniquement plus aisée entre superpositions lexicales réelles et virtuelles

L'Expérience 4 se concentre sur l'effet des superpositions réelles et virtuelles en initiale de mot uniquement. Ce choix pour une première expérience explorant ce problème est justifié par l'importance théorique accordée au traitement du début des mots (direction gauche / droite du flux d'information) et aux prédictions du modèle TRACE (Section ci-dessous).

#### 4.2.2.1 Aspects théoriques

Le modèle TRACE prédit que :

- ▷ En cas de superposition lexicale initiale réelle (mot enchâssé à l'initiale d'un mot porteur (exemple : « *Biais* » dans « *Bière* »)) :
  - ∩ la représentation mentale du mot stimulus est activée (*Bière*)
  - ∩ la représentation mentale du mot enchâssé (*Biais*) est transitoirement activée, en même temps que la représentation mentale du mot porteur
  - ∩ les deux représentations entrent transitoirement en conflit (compétition lexicale), conformément à l'algorithme de TRACE, jusqu'au moment où

l'information acoustique vient lever l'ambiguïté (confirmation de l'hypothèse (*Bière*), infirmation de l'autre (*Biais*). Une durée supplémentaire est encore nécessaire pour l'évacuation complète du conflit, durée proportionnelle à l'importance de celui-ci, et qui est responsable de l'allongement des temps de reconnaissance du mot stimulus.

Ce phénomène de compétition lexicale a pour conséquence un allongement du temps de reconnaissance du mot porteur, par rapport à un mot ne renfermant pas de mot enchâssé initial (exemple : « *Drap* »), pour lequel cette phase de compétition n'a pas lieu.

- ▷ En cas de superposition lexicale initiale virtuelle (mot stimulus enchâssable à l'initiale d'un autre mot du lexique (exemple : « *clé* » dans « *clématite* »)) :

- ∟ la représentation mentale du mot stimulus est activée (*clé*)
- ∟ la représentation mentale du mot à l'initiale duquel le mot stimulus peut exister est transitoirement activée (*clématite*), mais avec un niveau d'excitation moindre que le mot « *clé* » (l'activation ascendante n'est que partielle)
- ∟ les deux représentations entrent transitoirement en conflit, conformément à l'algorithme de TRACE, jusqu'au moment où l'information acoustique vient confirmer l'une des deux hypothèses et infirmer l'autre

Les mêmes conséquences que pour une superposition réelle devraient être observées sur les temps de reconnaissance des mots pour une superposition virtuelle (allongement pour un mot stimulus enchâssable à l'initiale d'un autre mot, par rapport à un mot non enchâssable à l'initiale d'un autre mot (exemple : « *Foin* »)) : TRACE ne prédit pas de distinction entre superpositions réelles et virtuelles. Cependant, la quantité de mots pouvant commencer par le stimulus dans le cas de superpositions virtuelles a plus de chances d'être supérieure au nombre de mots enchâssés à l'initiale d'un mot porteur. L'effet d'une superposition virtuelle devrait donc être plus visible que celui d'une superposition réelle (compétition plus importante, effet de « gang »).

- ▷ L'effet d'une superposition réelle est théoriquement cumulable à celui d'une superposition virtuelle (un même mot (exemple : /plas/) peut être impliqué à la fois dans une superposition réelle (contient /pla/) et dans une superposition virtuelle (peut être le début de /plastik/). Au niveau lexical, on aura une activation simultanée :

- ∟ de la représentation mentale du mot stimulus (*place*)
- ∟ de la représentation mentale du mot enchâssé (*plat*)
- ∟ de la représentation mentale du mot à l'initiale duquel le mot stimulus peut exister (*plastique*)

Les trois représentations entreront transitoirement en conflit, conformément à l'algorithme de TRACE, jusqu'au moment où l'information acoustique permet de confirmer l'une des hypothèses et infirme les autres. Il découlera de ce conflit lexical des temps de reconnaissance du mot porteur plus longs que dans les cas de superpositions lexicales simples, lorsque le conflit est moins important au sein du système.



## 4.2.2.2 Aspects techniques

### 4.2.2.2.1 Facteurs expérimentaux et conditions

Deux facteurs lexicaux relatifs aux mots expérimentaux, à 2 modalités chacun, croisés selon le plan complet présenté dans le tableau ci-dessous (4 conditions).

Tableau 51 : Facteurs expérimentaux

		Présence d'un mot français enchâssé à l'initiale du mot stimulus	
		Oui (E)	Non (nE)
Enchâssabilité du mot stimulus à l'initiale d'un autre mot français	Oui (I)	E.I (« Place », « Four », « Bac »)	nE.I (« Nièce », « Bière », « Rôle »)
	Non (nI)	E.nI (« Pont », « Clé », « Drap »)	nE.nI (« Dieu », « Foin », « Jouet »)

#### 4.2.2.2.2 Hypothèses opérationnelles

- ▷ Hypothèse E : Effet d'une superposition lexicale initiale réelle (présence d'un enchâssement initial)

Selon le modèle TRACE, les temps de latence pour reconnaître un mot renfermant un mot enchâssé initial seront significativement plus longs que ceux pour reconnaître un mot sans mot enchâssé initial.

- ▷ Hypothèse I : Effet d'une superposition lexicale initiale virtuelle (enchâssabilité du stimulus à l'initiale d'un mot)

Selon le modèle TRACE, les temps de latence pour reconnaître un mot qui peut être enchâssé à l'initiale d'un autre mot seront plus longs que ceux pour reconnaître un mot qui n'est pas enchâssable à l'initiale d'un autre mot du lexique.

- ▷ Hypothèse E + I : Cumul des effets d'une superposition lexicale initiale réelle avec ceux d'une superposition lexicale initiale virtuelle

Selon le modèle TRACE, les temps de latence pour reconnaître un mot renfermant un enchâssement initial et pouvant être enchâssé à l'initiale d'un autre mot du lexique seront plus longs encore que ceux obtenus avec un mot renfermant un enchâssement initial seul, ou que ceux obtenus pour un mot enchâssable à l'initiale d'autres mots et sans enchâssement initial.

#### 4.2.2.2.3 Matériel expérimental utilisé

Stimuli en voix naturelle.

- ▷ Items expérimentaux :  
60 mots monosyllabiques (15 par condition). Liste présentée en Annexe 9 page 479.
- ▷ Items distracteurs :  
∞ 60 pseudo-mots monosyllabiques (20 peuvent commencer un mot plurisyllabique ([gʁõ] dans « gronder »), 20 ne commencent jamais un mot

plurisyllabique selon BRULEX ([<sub>3</sub>ék] dans « éjecter »), et 20 sont des « syllabes » illégales en Français selon BRULEX ([bjy], [fky], [dɛf]).

- ∞ 60 mots bisyllabiques (30 avec enchâssement initial, 30 sans)
- ∞ 60 pseudo-mots bisyllabiques (30 avec enchâssement initial, 30 sans)

L'expérience comporte donc 240 stimuli.

#### **4.2.2.2.4 Organisation de l'expérience**

Chaque participant est testé avec les 240 items. Il traite chaque stimulus 1 fois.

#### **4.2.2.2.5 Population testée**

Les données de 40 individus francophones monolingues, sans troubles auditifs, cognitifs et moteurs, étudiant à l'université de Provence (Aix-Marseille I), ont été utilisées pour les analyses. Ces données ont été recueillies en mars 2000.

#### **4.2.2.2.6 Tâche**

Tâche de détection (décision lexicale en « go, no go » - détection des mots stimuli). Modalité auditive.

#### **4.2.2.2.7 Procédure de recueil des données**

Expérience réalisée à l'université de Provence (Aix-Marseille I). Détails dans la Partie III, Chapitre 4.

40 \* 60 = 2400 mesures utilisées. La phase test de l'expérience dure 10 minutes environ, et se déroule sans pause.

#### **4.2.2.2.8 Méthode d'analyse**

- ▷ Détail des procédures pour le tri des données et les analyses dans le Chapitre 3, Partie III.
- ▷ Analyse de régression multiniveaux, modèle à deux niveaux (sujets, items. Corrélation intra classe : 17.58%).
- ▷ 300 valeurs manquantes (12.5% de non-réponses et de valeurs filtrées (Détail par condition : E.I 10.3% ; E.nI 13.2% ; nE.I 15.2% ; nE.nI 11.3%)).

#### **4.2.2.2.9 Paramètres contrôlés**

À la construction du matériel et à la structuration de l'expérience

- ▷ Nombre moyen de phonèmes équilibré entre les conditions (Tableau 52 ci-dessous).

Tableau 52 : Nombre moyen de phonèmes par mot pour chaque condition

	<b>I</b>	<b>nI</b>
<b>E</b>	3.13	3.33
<b>nE</b>	2.67	3.2

- ▷ Fréquence formelle phonologique moyenne inter conditions (Tableau 53 ci-dessous).

Tableau 53 : Fréquence formelle phonologique moyenne par mot pour chaque condition

	<b>I</b>	<b>nI</b>
<b>E</b>	5681,6	6554,1
<b>nE</b>	6038,9	8058,7

- ▷ Équilibre des conditions pour le nombre d'items
- ▷ Population (origine linguistique et monolinguisme des participants, facultés perceptives, cognitives et motrices intactes)
- ▷ Relations sémantiques entre mots expérimentaux, entre mots expérimentaux et items distracteurs, entre mots enchâssés et mots expérimentaux (autant qu'il est possible d'en détecter)
- ▷ Présentation aléatoire des items lors de la passation

#### Lors des analyses statistiques

- ▷ Origine de la mesure des temps de réponse
- ▷ Position acoustique du point de reconnaissance du mot stimulus (ms.)
- ▷ Position acoustique du début du point d'unicité du mot stimulus (ms.)
- ▷ Durée du mot stimulus (ms.)
- ▷ Nombre de phonèmes composant le mot stimulus
- ▷ Position phonétique du point d'unicité phonologique du mot stimulus (en nombre de phonèmes)
- ▷ Fréquence formelle phonologique du mot stimulus (logarithme en base 10 du cumul des fréquences lexicales pour la forme phonologique considérée)
- ▷ Âge d'acquisition estimé du mot (moyenne par forme phonologique)
- ▷ Familiarité conceptuelle du mot (moyenne par forme phonologique)

#### Remarque :

Les deux facteurs prosodiques  $f_0$  et intensité n'ont pas été contrôlés dans le modèle, bien que l'analyse acoustique (effectuée a posteriori, par souci d'homogénéiser la procédure d'analyse sur l'ensemble des expériences) ait révélé une relation statistique forte entre ces facteurs prosodiques et les V.I. (détail des analyses dans l'Annexe 12 page 494).

La décision de ne pas contrôler exceptionnellement ces 2 facteurs est justifiée par les arguments suivants :

- ▷ L'expérience est exploratoire

- ▷ Elle est entachée d'erreurs à cause de la procédure de recueil des temps de réponse (se reporter à la Section 3.4.4.1.1 page 150).
- ▷ Le phénomène observé ici est confirmé dans d'autres expériences mieux contrôlées, y compris pour la prosodie.

### 4.2.2.3 Résultats obtenus

#### 4.2.2.3.1 Rapport des analyses statistiques

L'équation du modèle (Équation 37) et les tests de significativité relatifs aux covariables peuvent être consultés à l'Annexe 13 page 497.

Tableau 54 : Tests de significativité pour les facteurs expérimentaux

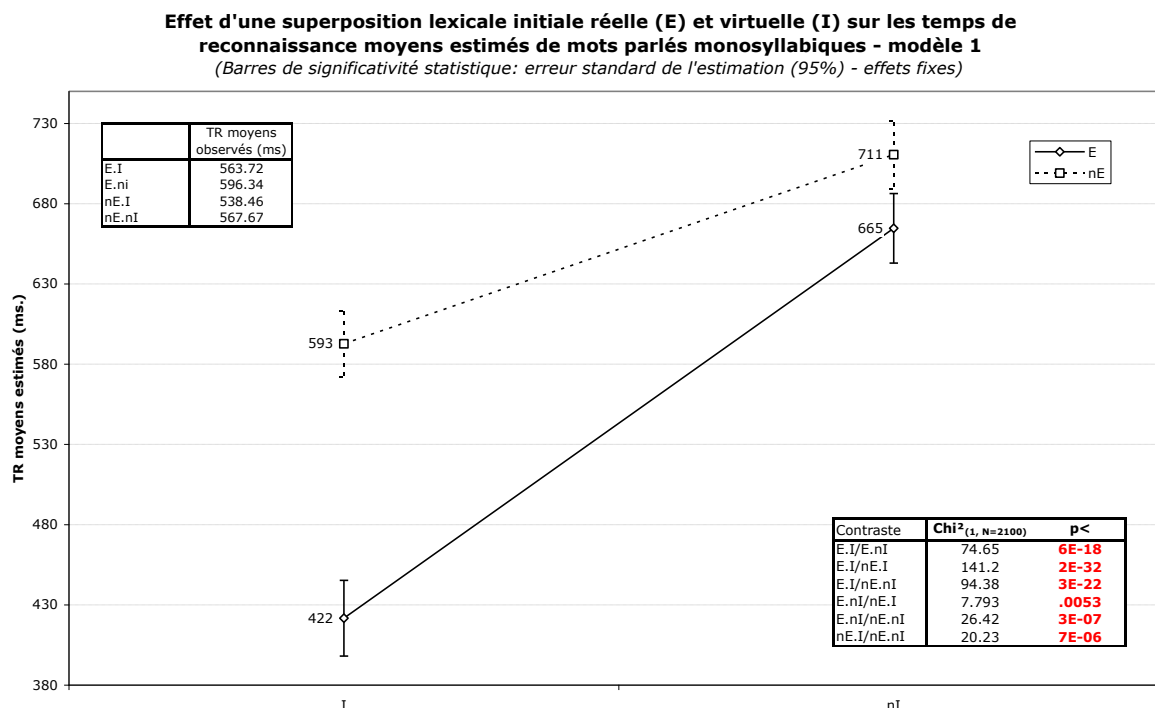
Facteur / Interaction	Coefficient	Chi <sup>2</sup> <sub>(1, N=2100)</sub>	p<	% variance expliquée
<b>I</b>	-117.88	20.231	<b>7<sup>E-06</sup></b>	0.756
<b>E</b>	-45.893	26.421	<b>3<sup>E-07</sup></b>	0.935
<b>I.E</b>	-125.1	72.668	<b>2<sup>E-17</sup></b>	2.27

Tableau 55 : Tests de significativité des contrastes entre conditions

Contraste		Chi <sup>2</sup> <sub>(1, N=2100)</sub>	p<
<b>E.I</b>	<b>E.nI</b>	74.65	<b>6<sup>E-18</sup></b>
	<b>nE.I</b>	141.2	<b>2<sup>E-32</sup></b>
	<b>nE.nI</b>	94.38	<b>3<sup>E-22</sup></b>
<b>E.nI</b>	<b>nE.I</b>	7.793	<b>.006</b>
	<b>nE.nI</b>	26.42	<b>3<sup>E-07</sup></b>
<b>n.EI</b>	<b>nE.nI</b>	20.23	<b>7<sup>E-06</sup></b>

### 4.2.2.3.2 Présentation des résultats

Figure 54 : Temps moyens de reconnaissance des mots en fonction de la présence d'une superposition réelle ou virtuelle dans la situation



- ▷ Hypothèse E (présence d'un enchâssement initial), Hypothèse I (enchâssabilité du stimulus à l'initiale d'un mot) :

Les deux facteurs expérimentaux (présence d'un mot enchâssé à l'initiale du mot stimulus et enchâssabilité du mot stimulus à l'initiale d'un autre mot) s'avèrent être responsables d'une diminution des temps de reconnaissance des items (valeurs estimées par le modèle : respectivement, de 118 ms. ( $p < .0001$ ) et de 46 ms ( $p < .0001$ ) – Tableau 54, page 239). L'hypothèse E et l'hypothèse I sont donc invalidées par ces données.

Ce résultat est inattendu. Il n'est prédit par aucun des modèles psycholinguistiques actuels.

- ▷ Hypothèse E + I (cumul de l'effet des deux hypothèses) :

L'effet des deux facteurs semble de plus cumulatif : l'interaction entre les deux facteurs est responsable d'une diminution supplémentaire des temps de reconnaissance (valeur estimée de 125 ms. ;  $p < .0001$ ). L'hypothèse E + I est donc elle aussi invalidée par les données.

L'analyse des contrastes, pour la comparaison des différences moyennes de temps de reconnaissance des mots entre conditions révèle que tous les écarts sont significatifs (Tableau 55, page 239).

#### **4.2.2.4 Conclusions de l'Expérience 4**

Les 3 hypothèses de l'Expérience 4 sont invalidées par les données.

L'influence des superpositions lexicales initiales réelles et virtuelles sur la reconnaissance d'un mot parlé donne lieu à des résultats opposés à ceux prédits par l'un des principaux modèles de la psycholinguistique (TRACE). Une superposition lexicale initiale réelle ou virtuelle produit une aisance supérieure pour reconnaître le mot stimulus, et non une difficulté. L'Expérience 4 révèle donc une contradiction entre faits et théorie.

### 4.2.3 Expérience 5 : Effet d'une superposition lexicale initiale réelle sur la reconnaissance d'un mot monosyllabique en modalité auditive – contrôle phonétique

Le matériel utilisé dans l'Expérience 4 est très hétérogène phonologiquement et physiquement, puisque les mots des 4 conditions ne pouvaient pas être appariés dans ces deux dimensions. Malgré toutes les variables contrôlées pour corriger d'éventuels biais liés à cette hétérogénéité (position du point d'unicité, nombre de phonèmes composant le mot, durée des items, etc.), les effets paradoxaux qui ont été obtenus peuvent avoir été induits par une trop grande différence acoustique entre les mots. L'Expérience 5 a pour but de contrôler la validité des résultats de l'Expérience 4 à l'aide d'un matériel phonétiquement apparié, donc acoustiquement voisin et comparable, entre les 2 conditions. Afin de faciliter ce contrôle, la situation a été simplifiée et ne comporte plus qu'un facteur expérimental (superposition lexicale initiale réelle). Ce choix est justifié par :

- ▷ l'absence de différence dans le sens de l'effet pour les deux types de superpositions dans l'Expérience 4.
- ▷ la possibilité de contrôler la situation de manière concrète avec une superposition réelle (il est possible de savoir quels mots sont présents dans le stimulus), mais pas avec une superposition virtuelle (il est difficile de savoir quels mots sont évoqués chez un participant donné).

#### 4.2.3.1 Aspects théoriques

Se reporter à la Section 4.2.2.1, page 234 pour un développement de l'hypothèse théorique relative à l'effet d'une superposition lexicale initiale réelle sur les temps de reconnaissance d'un mot parlé monosyllabique.

#### 4.2.3.2 Aspects techniques

##### 4.2.3.2.1 Facteurs expérimentaux et conditions

1 facteur expérimental (enchâssement initial) à 2 modalités (présence / absence), donnant lieu à deux conditions expérimentales :

- ▷ mots avec un enchâssement initial (exemple : le mot « *taupe* » renferme le mot « *taux* »)
- ▷ mots sans enchâssement initial appariés aux mots avec enchâssement initial (exemple : le mot « *type* » ne renferme pas de mot enchâssé à son initiale)

#### 4.2.3.2.2 Hypothèse opérationnelle

Selon le modèle TRACE, les mots monosyllabiques qui renferment un mot enchâssé initial seront reconnus plus lentement que les mots monosyllabiques qui ne renferment pas de mot enchâssé initial.

#### 4.2.3.2.3 Matériel expérimental utilisé

Stimuli en voix naturelle. La liste des items est fournie en Annexe 14 page 499.

▷ Items expérimentaux :

48 mots monosyllabiques appariés deux à deux entre les 2 conditions expérimentales, de telle manière que la différence phonémique entre les mots d'une paire ne porte que sur le noyau vocalique.

∩ 24 mots avec enchâssement initial (exemple : « *trousse* » contient « *trou* »).

∩ 24 mots sans enchâssement initial (exemple : « *transe* » commence par /tʁã/, qui n'est pas un mot).

▷ Items distracteurs :

48 « syllabes » illégales en Français selon BRULEX (la séquence de phonèmes peut toutefois exister dans cette langue), appariées aux mots expérimentaux pour le nombre de phonèmes et la structure syllabique, et variant avec le facteur expérimental manipulé (présence d'un enchâssement initial ou non (N = 24 - N = 24)). Exemples : /bʁɛj/ et /flef/ (4 phonèmes, structure CCVC, appariés aux mots expérimentaux « *trousse* » et « *transe* », et renfermant un mot enchâssé pour le premier (« *brin* »)).

#### 4.2.3.2.4 Organisation de l'expérience

Les participants sont répartis en deux groupes. Chaque groupe est testé avec la moitié des items (N = 24 : 50% des items d'une condition (N = 12) et 50% des items de l'autre condition (N = 12)). Le but est de répartir les items des deux conditions de manière à éviter les phénomènes d'amorçage formel (exemple : les mots « *type* » et « *taupe* » partagent 66% de leurs phonèmes du fait de leur appariement. Un même participant ne doit donc pas être testé sur ces deux items).

Chaque participant est testé avec 48 items (mot + distracteur), et traite chaque item de la liste 1 fois.

#### 4.2.3.2.5 Population testée

Les données de 32 participants (16 par groupe) ont été utilisées pour les analyses. Il s'agit d'étudiants francophones monolingues de l'université de Genève, sans troubles auditifs, cognitifs et moteurs. Les tests se sont déroulés en janvier 2003.



#### 4.2.3.2.6 Tâche

Tâche de détection (décision lexicale en « go, no go » - détection des mots stimuli).  
Modalité auditive.

#### 4.2.3.2.7 Procédure de recueil des données

Expérience réalisée à l'université de Genève. Détails dans la Partie III, Chapitre 4.

16 \* 2 \* 24 = 768 mesures utilisées.

La phase test de l'expérience ne dure pas plus de 5 minutes.

#### 4.2.3.2.8 Méthode d'analyse

- ▷ Détail des procédures pour le tri des données et les analyses dans le Chapitre 3, Partie III.
- ▷ Analyse de régression multiniveaux. Modèle adopté à deux niveaux (sujets, items. Corrélation intra classe : 10.15%).
- ▷ 60 valeurs manquantes (7.8% de non-réponses et de valeurs filtrées (détail par condition : nE 9.1% ; E 6.5%)).

#### 4.2.3.2.9 Paramètres contrôlés

##### À la construction du matériel et à la structuration de l'expérience

- ▷ Appariement phonologique des mots des deux conditions (nombre de phonèmes, structure phonétique)
- ▷ Appariement phonologique des pseudo-mots aux mots (nombre de phonèmes et structure phonétique)
- ▷ Position post-finale du point d'unicité du mot stimulus
- ▷ Fréquence formelle phonologique moyenne par condition
- ▷ Vérification / limitation des relations sémantiques entre mots (autant que possible)
- ▷ Mots enchâssés et mots porteurs sans lien sémantique ou morphologique (autant que possible)
- ▷ Répartition des items en listes pour éviter les effets d'amorçage formel
- ▷ Présentation aléatoire des items lors de la passation

##### Lors des analyses statistiques

- ▷ Ordre de présentation des items
- ▷ Latéralité manuelle des participants
- ▷ Sexe

- ▷ Durée des items (ms.)<sup>110</sup>
- ▷ Nombre de phonèmes composant le mot stimulus
- ▷ Position acoustique du point de reconnaissance (en ms.) du mot stimulus
- ▷ Position phonétique du point d'unicité (en nombre de phonèmes) du mot stimulus
- ▷ Fréquence formelle phonologique (logarithme en base 10 du cumul des fréquences lexicales pour la forme phonologique considérée)
- ▷ Âge d'acquisition estimé du mot stimulus (moyenne par forme phonologique)
- ▷ Familiarité conceptuelle du mot stimulus (moyenne par forme phonologique)
- ▷ Liste : se reporter à la remarque en note de bas de page n°105, page 201.

Remarque :

Des analyses acoustiques succinctes (ANOVA) ont été effectuées pour vérifier l'homogénéité du matériel entre les deux modalités du facteur expérimental. Les résultats sont présentés dans les Tableau 56 et Tableau 57 ci-dessous.

Tableau 56 : Valeurs moyennes des paramètres acoustiques par condition

	<b>f0 moy</b>	<b>dB moy syll</b>	<b>dB zones voisées</b>	<b>Durée</b>
<b>E</b>	125.44	66.51	73.74	0.67
<b>nE</b>	123.22	64.78	73.61	0.69

Tableau 57 : Résultats des ANOVA pour la différence entre les 2 conditions du facteur étudié

	<b>dl Effet</b>	<b>CM Effet</b>	<b>dl Erreur</b>	<b>CM Erreur</b>	<b>F</b>	<b>p&lt;</b>
<b>f0</b>	1	58.852	46	78.046	0.754	.39
<b>dB syll</b>	1	35.707	46	66.577	0.536	.47
<b>dB zones voisées</b>	1	0.218	46	5.183	0.042	.84
<b>Durée</b>	1	0.005	46	0.003	1.659	.21

Avec :

*f0*, la fréquence fondamentale moyenne de la syllabe (en Hz)

*dB syll*, l'intensité moyenne de la syllabe (en dB)

*dB zones voisées*, l'intensité moyenne mesurée uniquement sur les portions voisées de la syllabe (en dB)

*Durée*, la durée de la syllabe (en sec.)

Ces analyses ne révèlent aucune différence acoustique moyenne significative entre les deux conditions étudiées. Aussi, les facteurs f0 et intensité n'ont pas été intégrés au modèle d'analyse.

<sup>110</sup> La durée est intégrée au modèle d'analyse (bien que les ANOVA aient montré qu'en moyenne ce facteur était homogène entre les deux conditions), car il est nécessaire, avec d'autres facteurs (position acoustique du P.R. et du début du P.U.), pour résoudre le problème de l'origine d'une mesure chronométrique avec un stimulus temporel (voir Chapitre 2, Partie III).

### 4.2.3.3 Résultats obtenus

#### 4.2.3.3.1 Rapport des analyses statistiques

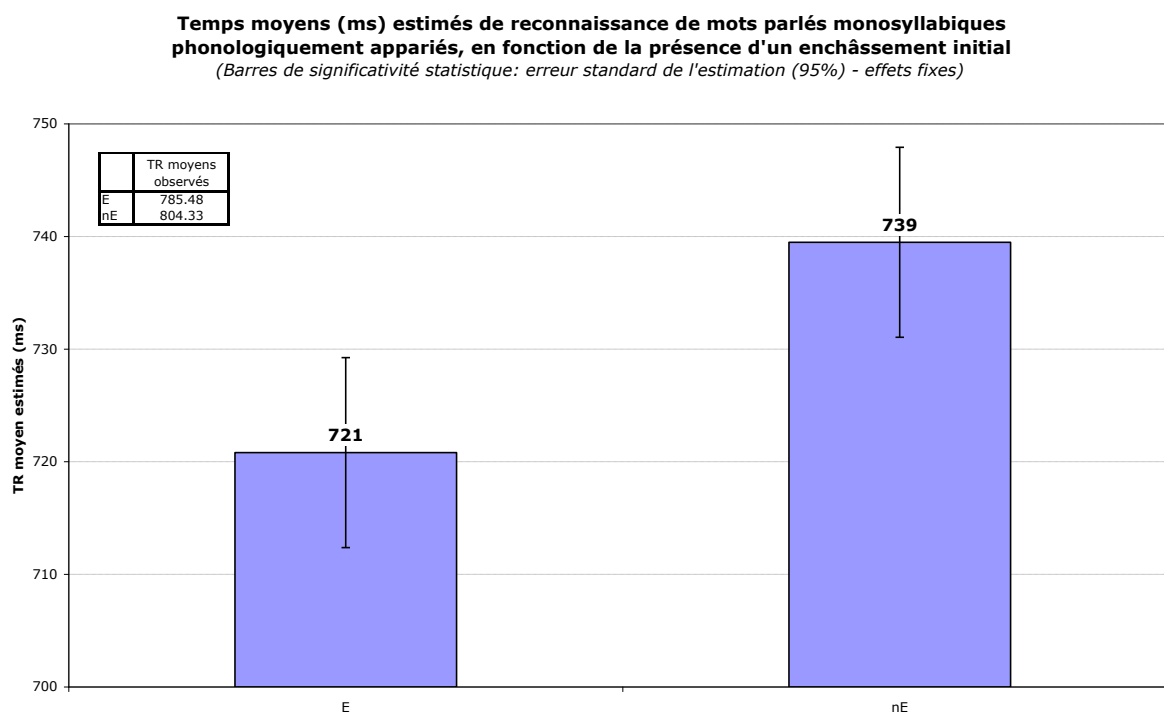
L'équation du modèle (Équation 38) et les tests de significativité relatifs aux covariables peuvent être consultés en Annexe 15 page 500.

Tableau 58 : Test de significativité du facteur expérimental

	Coefficient	Chi <sup>2</sup> <sub>(1, N=708)</sub>	p<	% variance expliquée
<b>E</b>	-18.674	4.709	<b>.031</b>	0.607

#### 4.2.3.3.2 Présentation des résultats

Figure 55 : Temps moyens de reconnaissance des mots, estimés par le modèle, en fonction de la présence d'un enchâssement initial



Les participants reconnaissent plus rapidement un mot monosyllabique renfermant un mot enchâssé initial qu'un mot monosyllabique phonologiquement apparié ne renfermant pas d'enchâssement initial ( $p < .035$ ). L'écart moyen entre les deux conditions, estimé par le modèle, est de 19 ms. environ.

L'hypothèse de l'expérience est donc invalidée.

Malgré un contrôle accru de la situation, il n'est pas possible de mettre en évidence l'effet de compétition lexicale prédit par TRACE en présence d'un mot enchâssé initial.

#### **4.2.3.4 Conclusions de l'Expérience 5**

L'hypothèse de l'Expérience 5 est invalidée par les données. Cette expérience reproduit les résultats observés dans l'Expérience 4 : les temps de reconnaissance sont significativement plus courts dans le cas d'un mot parlé impliqué dans une superposition lexicale initiale réelle que dans le cas de mots parlés phonétiquement appariés, qui ne renferment pas d'enchâssement initial. Cet effet incongru est reproduit malgré des conditions de contrôle plus poussées – il semble être robuste.

#### 4.2.4 Expérience 6 : Effet d'une superposition lexicale réelle, initiale et/ou finale, sur la reconnaissance d'un mot bisyllabique, en modalité auditive

Les deux expériences précédentes (Expériences 4 et 5) ont étudié l'impact de superpositions lexicales initiales (respectivement réelle + virtuelle, réelle seulement) sur la reconnaissance de mots parlés avec des items de gabarit monosyllabique. Ce gabarit présente deux inconvénients :

- ▷ il confond le niveau syllabique (phonologique) et le niveau lexical (lexical)
- ▷ il ne permet pas de manipuler avec une latitude suffisante la position de la superposition dans le mot afin d'en évaluer l'effet sur le processus d'activation lexicale

Pour résoudre ces inconvénients, il est nécessaire d'utiliser des mots d'un gabarit polysyllabique. Cependant, avec une méthode chronométrique de mesure comportementale, lorsque la taille des mots est conséquente (trois syllabes ou plus), l'effet d'un traitement qui survient au début du mot peut être dilué dans le temps ou masqué par d'autres traitements qui surviennent entre le début et la fin du mot. La conséquence est que les traitements initiaux peuvent ne plus être mesurables au moment où a lieu le sondage.

Pour limiter ce problème (qui donc est peut-être toujours présent) tout en solutionnant celui lié à l'usage d'un matériel monosyllabique, des mots bisyllabiques ont été utilisés. Ce gabarit est suffisant pour :

- ▷ tester la robustesse des effets d'une superposition lexicale réelle
- ▷ décorrélérer le niveau phonologique et le niveau lexical pour tenter de mieux cerner l'origine de l'effet d'une superposition lexicale réelle
- ▷ tester l'effet de la position d'un enchaînement lexical dans un mot

L'Expérience 6 explore l'impact sur la reconnaissance de mots parlés bisyllabiques de la position de superpositions lexicales réelles.

##### 4.2.4.1 Aspects théoriques

La présence d'une superposition initiale réelle est responsable, selon le modèle TRACE, d'un phénomène de compétition au niveau lexical, qui retarde la reconnaissance du mot stimulus (mot porteur). L'effet est supposé varier en fonction de la position de l'enchaînement dans le mot stimulus (au début ou à la fin), variation qui s'explique essentiellement par la dynamique de fonctionnement du modèle (voir par exemple Frauenfelder et Peeters (1990)) :

- ▷ Lorsque la superposition lexicale réelle est initiale, la reconnaissance du mot porteur est fortement perturbée (la représentation du mot stimulus est en conflit avec la représentation du mot porteur dès le début du traitement. Si les facteurs statistiques modulant l'excitabilité d'une représentation ne sont pas considérés (fréquence d'usage, voisinage, etc.), aucun des deux candidats n'est initialement privilégié, donc aucun des deux n'inhibe plus fortement l'autre. Ceci est vrai jusqu'au moment où de l'information acoustique vient départager les deux possibilités. Le décalage dans le temps du moment où le mot porteur est reconnu, est lié à la perturbation, causée initialement par le conflit entre deux possibilités, sur l'évolution de l'activation lexicale du mot porteur).
- ▷ Lorsque la superposition lexicale réelle est finale, la reconnaissance du mot porteur est peu ou pas perturbée car celui-ci est déjà en bonne voie d'être reconnu (la représentation lexicale du mot concurrent ne dispose pas d'un niveau d'activation suffisant pour influencer de façon décisive celui de la représentation lexicale du mot stimulus, tandis que l'inverse est vrai).

TRACE propose donc trois hypothèses servant de cadre théorique à l'Expérience 6 :

- ▷ Hypothèse  $M_1$  : Effet d'une superposition lexicale initiale réelle ( $S_1 = \text{mot}$ )  
un mot bisyllabique renfermant un mot enchâssé initial est plus difficile à reconnaître qu'un mot bisyllabique ne renfermant pas de mot enchâssé initial.
- ▷ Hypothèse  $M_2$  : Effet d'une superposition lexicale finale réelle ( $S_2 = \text{mot}$ )  
un mot bisyllabique renfermant un mot enchâssé en position finale n'est pas plus difficile à reconnaître qu'un mot bisyllabique ne renfermant pas de mot enchâssé en position finale.
- ▷ Hypothèse  $M_1 / M_2$  : Différence d'effet entre une superposition lexicale initiale et une superposition lexicale finale  
un mot bisyllabique renfermant un mot enchâssé initial est plus difficile à reconnaître qu'un mot bisyllabique renfermant un mot enchâssé en position finale.

#### 4.2.4.2 Aspects techniques

##### 4.2.4.2.1 Facteurs expérimentaux et conditions

Deux facteurs lexicaux à 2 modalités chacun ; croisement complet (4 conditions).  
Facteurs et conditions sont présentés dans le Tableau 59 ci-dessous.

Tableau 59 : Facteurs et conditions

		Lexicalité de la 1 <sup>ère</sup> syllabe	
		Oui ( $M_1$ )	Non ( $nM_1$ )
Lexicalité de la 2 <sup>ème</sup> syllabe	Oui ( $M_2$ )	$M_1.M_2$ « <i>Timbale</i> » (/tɛ̃/-/bal/)	$nM_1.M_2$ « <i>Climat</i> » ( $S_1$ -/ma/)
	Non ( $nM_2$ )	$M_1.nM_2$ « <i>Résine</i> » (/ʁe/- $S_2$ )	$nM_1.nM_2$ « <i>Colonne</i> » ( $S_1$ - $S_2$ )

#### 4.2.4.2.2 Hypothèses opérationnelles

- ▷ Hypothèse  $M_1$  : Effet d'une superposition lexicale réelle au début d'un mot  
Les temps de reconnaissance d'un mot bisyllabique dont la première syllabe est un mot sont plus longs que ceux d'un mot bisyllabique dont la première syllabe n'est pas un mot.
- ▷ Hypothèse  $M_2$  : Effet d'une superposition lexicale réelle à la fin d'un mot  
Les temps de reconnaissance d'un mot bisyllabique dont la deuxième syllabe est un mot sont statistiquement identiques à ceux d'un mot bisyllabique dont la deuxième syllabe n'est pas un mot.
- ▷ Hypothèse  $M_1 / M_2$  : Différence d'effet entre une superposition lexicale réelle au début et à la fin d'un mot  
Les temps de reconnaissance d'un mot bisyllabique dont la première syllabe est un mot sont plus longs que ceux d'un mot bisyllabique dont la deuxième syllabe est un mot.

#### 4.2.4.2.3 Matériel expérimental utilisé

Stimuli en voix naturelle. La liste des items est donnée en Annexe 16 page 501.

- ▷ Items expérimentaux :  
88 mots bisyllabiques (22 mots par condition – détail des conditions dans le Tableau 59, page 249)
- ▷ Items distracteurs :  
88 pseudo-mots bisyllabiques (4 catégories déterminées par la manipulation des deux facteurs expérimentaux selon le même schéma que pour les mots expérimentaux donné dans le Tableau 59, page 249 (exemples : /klãʃɔk/ - « *clan* »-« *choc* », /bakmid/ - « *bac* »- $S_2$ , /ʒasbɔt/ -  $S_1$ -« *botte* », /dʁakʃad/ -  $S_1$ - $S_2$ ) ; 22 items de chaque type)

#### 4.2.4.2.4 Organisation de l'expérience

Les 176 stimuli (dont  $N = 88$  stimuli expérimentaux) sont présentés une seule fois à chaque participant.

#### 4.2.4.2.5 Population testée

Les données de 31 étudiants de l'université de Genève, francophones, monolingues, sans troubles auditifs, cognitifs et moteurs, ont été utilisées pour les analyses. Leur recueil a eu lieu en mai 2002.

#### 4.2.4.2.6 Tâche

Tâche de détection (décision lexicale en « go, no go » - détection des mots stimuli). Modalité auditive.

#### 4.2.4.2.7 Procédure de recueil des données

Expérience réalisée à l'université de Genève, en collaboration avec une étudiante effectuant une recherche pour l'obtention de sa Licence en Psychologie (G. Niederhauser - Direction : C. Lachaud, U. H. Frauenfelder). Détails de la procédure dans la Partie III, Chapitre 4.

88 \* 31 = 2728 mesures utilisées.

La phase test de l'expérience se déroule sans pause, et dure moins de quinze de minutes.

#### 4.2.4.2.8 Méthode d'analyse

- ▷ Détail des procédures pour le tri des données et les analyses dans la Partie III, Chapitre 3.
- ▷ Analyse multiniveaux, modèle à deux niveaux (sujets, items. Corrélations intra classe : 12.92%).
- ▷ 155 valeurs manquantes (6.45% de non-réponses et de valeurs filtrées (détail par condition :  $M_1.M_2$  5.7% ;  $M_1.nM_2$  2.7% ;  $nM_1.M_2$  9.3% ;  $nM_1.nM_2$  8.2%)).

#### 4.2.4.2.9 Paramètres contrôlés

##### À la construction du matériel et à la structuration de l'expérience

- ▷ Mots et pseudo-mots respectent la structure présentée dans le Tableau 59, page 249, les deux facteurs expérimentaux étant également manipulés dans les pseudo-mots pour avoir un matériel homogène pour les V.I. (autrement, une modalité d'un facteur peut être favorisée par rapport aux autres, et produire un biais - stratégique ou d'une autre nature).
- ▷ Homogénéité structurale des syllabes (toutes commencent par une consonne)
- ▷ Degré de complexité moyen des 2 syllabes assez voisin et homogène
- ▷ La position du point d'unicité phonologique de la deuxième syllabe (donc, lorsqu'elle est lexicale), est identique à celle du point d'unicité phonologique du mot porteur (selon BRULEX)
- ▷ Dans le cas des items  $*.nM_2$ , la deuxième syllabe du mot stimulus ne commence jamais un mot (selon BRULEX), se trouve rarement au milieu d'un mot, et est statistiquement une syllabe de fin de mot dans la langue française
- ▷ Élimination des mots préfixés, composés, argots, étrangers, ayant des homophones, ressemblant à des adjectifs, ressemblant à des prénoms
- ▷ Élimination des syllabes doublon ( $S_1 = S_2$  au sein d'un mot ;  $S_1$  ou  $S_2$  d'un mot =  $S_1$  ou  $S_2$  d'un autre mot)
- ▷ Vérification de l'aisance à syllaber la séquence  $S_1S_2$  en  $S_1$  et  $S_2$
- ▷ Homogénéité moyenne des conditions expérimentales pour la fréquence formelle



- ▷ Homogénéité moyenne des conditions expérimentales pour le nombre de phonèmes
- ▷ Homogénéité moyenne des conditions expérimentales pour la position du point d'unicité (en nombre de phonèmes)
- ▷ Vérification des relations sémantiques entre mots stimuli au maximum possible
- ▷ Vérification des relations sémantiques entre mots enchâssés au maximum possible
- ▷ Vérification des relations sémantiques entre mots enchâssés et mots stimuli au maximum possible
- ▷ Élimination des mots ayant un voisinage se limitant à un rapport morphologique (« *graine* », « *grainetier* », etc.)
- ▷ Présentation aléatoire des items lors de la passation

#### Lors des analyses statistiques

Les analyses acoustiques révèlent que les V.I. de l'expérience sont significativement en lien avec les paramètres prosodiques des syllabes constituant les mots (durée, f<sub>0</sub>, intensité). Le détail de ces analyses acoustiques est fourni dans l'Annexe 17 page 502.

Ces facteurs ont donc été intégrés au modèle d'analyse afin d'être contrôlés.

#### Liste des facteurs contrôlés lors des analyses statistiques :

- ▷ Ordre de présentation des items
- ▷ Sexe des participants
- ▷ Latéralité manuelle des participants
- ▷ Fréquence fondamentale moyenne de chaque syllabe (Hz)
- ▷ Intensité moyenne de chaque syllabe (dB)
- ▷ Durée de chaque syllabe (ms.)
- ▷ Fréquence formelle (phonologique) de chaque syllabe (logarithme en base 10)
- ▷ Fréquence formelle (phonologique) du mot stimulus (id.)
- ▷ Âge d'acquisition estimé du mot stimulus (moyenne par forme phonologique)
- ▷ Familiarité conceptuelle du mot stimulus (moyenne par forme phonologique)
- ▷ Nombre de phonèmes composant le mot stimulus
- ▷ Position acoustique du début du point d'unicité du mot stimulus (ms.)
- ▷ Position acoustique du point de reconnaissance du mot stimulus (ms.)

### **4.2.4.3 Résultats obtenus**

#### **4.2.4.3.1 Rapport des analyses statistiques**

L'équation du modèle (Équation 36) et les tests de significativité relatifs aux covariables peuvent être consultés à l'Annexe 18 page 506.

Le modèle étudié est assez complexe. Il doit prendre en considération beaucoup de facteurs, ainsi que les interactions que font ces facteurs entre eux.

Les facteurs acoustiques étant en lien avec les V.I. doivent être précisés pour chaque syllabe, ainsi que les interactions qu'ils font avec les V.I. Les interactions de moindre importance ne sont cependant pas modélisées (exemple : interactions impliquant la durée d'une syllabe avec la  $f_0$  ou l'intensité, interactions impliquant un rapport entre syllabes 1 et 2 pour ces facteurs).

Âge d'acquisition, familiarité conceptuelle et fréquence d'usage du mot forment un groupe de facteurs corrélés et influents. Leurs interactions sont donc également modélisées.

Enfin, la durée des syllabes, la position du point de reconnaissance et du début du point d'unicité phonologique, ainsi que le nombre de phonèmes, sont en relation et constituent un autre groupe de facteurs dont les interactions ont été modélisées.

Tableau 60 : Tests de significativité pour les facteurs expérimentaux

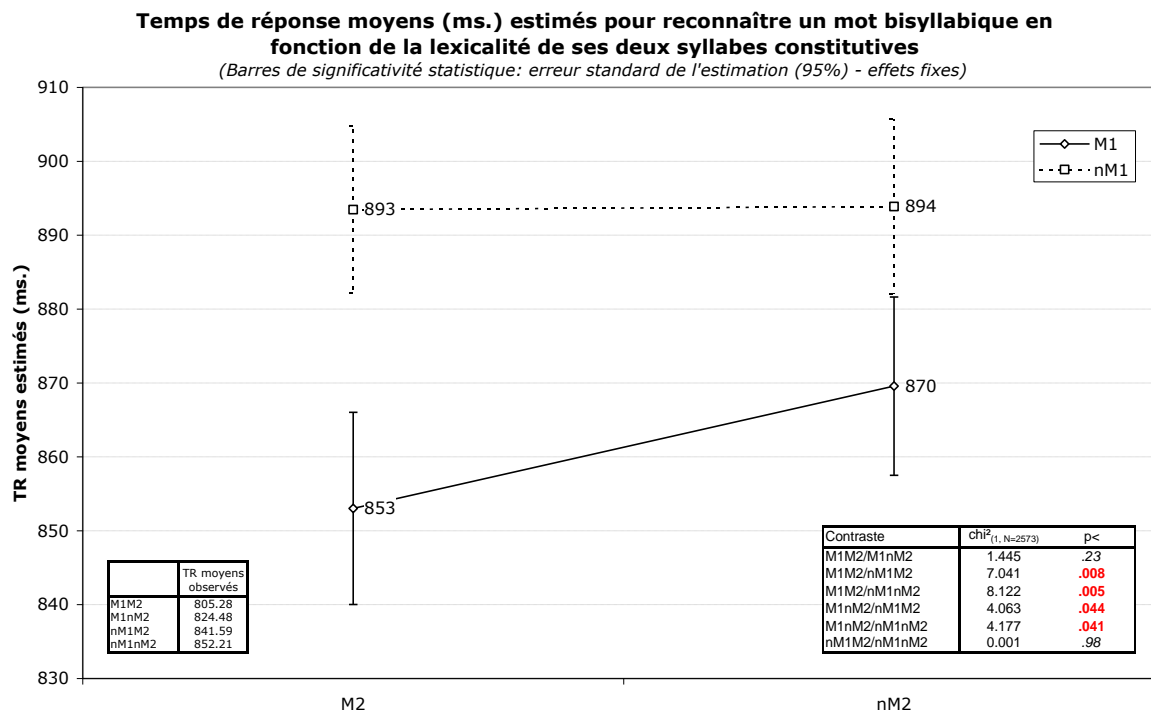
	Coefficient	Chi <sup>2</sup> <sub>(1, N=2573)</sub>	p<	% variance expliquée
<b>M<sub>1</sub></b>	-24.31	4.177	<b>.041</b>	0.130
<b>M<sub>2</sub></b>	-0.415	0.001	.98	0.0005
<b>M<sub>1</sub>.M<sub>2</sub></b>	-16.137	0.769	.39	0.027

Tableau 61 : Tests de significativité des contrastes entre conditions

Contraste		Chi <sup>2</sup> <sub>(1, N=2573)</sub>	p<
<b>M<sub>1</sub>M<sub>2</sub></b>	<b>M<sub>1</sub>nM<sub>2</sub></b>	1.445	.23
	<b>nM<sub>1</sub>M<sub>2</sub></b>	7.041	<b>.008</b>
	<b>nM<sub>1</sub>nM<sub>2</sub></b>	8.122	<b>.005</b>
<b>M<sub>1</sub>nM<sub>2</sub></b>	<b>nM<sub>1</sub>M<sub>2</sub></b>	4.063	<b>.044</b>
	<b>nM<sub>1</sub>nM<sub>2</sub></b>	4.177	<b>.041</b>
<b>nM<sub>1</sub>M<sub>2</sub></b>	<b>nM<sub>1</sub>nM<sub>2</sub></b>	0.001	.98

### 4.2.4.3.2 Présentation des résultats

Figure 56 : Temps de réponse moyens estimés en fonction de la lexicalité des syllabes du mot stimulus



#### Hypothèse M<sub>1</sub> : Effet d'une superposition lexicale réelle au début d'un mot sur le traitement du mot stimulus

Les mots stimuli dont la première syllabe est lexicale sont reconnus en moyenne 24 ms. plus rapidement que les mots dont la première syllabe n'est pas lexicale. L'hypothèse M<sub>1</sub> est donc invalidée.

#### Hypothèse M<sub>2</sub> : Effet d'une superposition lexicale réelle à la fin d'un mot sur le traitement du mot stimulus

Les mots stimuli dont la deuxième syllabe est lexicale sont en moyenne reconnus aussi rapidement que les mots dont la deuxième syllabe n'est pas lexicale. L'hypothèse M<sub>2</sub> est donc validée.

#### Hypothèse M<sub>1</sub> / M<sub>2</sub> : Différence d'effet entre une superposition lexicale réelle initiale et une superposition lexicale réelle finale, sur le traitement d'un mot

Les écarts entre les conditions M<sub>1</sub>.nM<sub>2</sub> et nM<sub>1</sub>.M<sub>2</sub> sont significatifs : les mots dont la première syllabe est un mot sont reconnus plus rapidement que les mots dont la deuxième syllabe est un mot. L'hypothèse M<sub>1</sub> / M<sub>2</sub> est donc invalidée.

#### **4.2.4.4 Conclusions de l'Expérience 6**

Des trois hypothèses de l'Expérience 6, formulées à partir de la logique de fonctionnement du modèle TRACE, seule l'hypothèse concernant l'effet d'une superposition lexicale finale réelle est vérifiée (pas de différence). Les effets obtenus pour une superposition lexicale réelle initiale sont exactement inverses à ceux prédits par TRACE (facilitation du traitement liée à la présence d'un mot enchâssé à l'initiale de l'item au lieu d'une gêne).

Ce résultat réplique l'effet trouvé dans les Expériences 4 et 5, et le généralise à un matériel plurisyllabique.

### 4.2.5 Expérience 7 : Réplication et contrôle

L'absence d'effet lié à la lexicalité de la deuxième syllabe dans l'Expérience 6 est intrigante, bien qu'un tel phénomène soit justifié d'après TRACE. Si les prédictions du modèle sont fausses pour l'effet d'un enchâssement initial, elles le sont peut-être aussi pour l'effet d'un enchâssement non initial. L'examen plus approfondi de la situation de l'Expérience 6 a révélé que l'apparition d'effets sur la fin du mot était défavorisée pour deux raisons :

- ▷ La tâche utilisée (TDL en « go, no go »), du fait de sa simplicité par rapport à la TDL classique (un seul type de réponse à fournir contre deux, un seul type de traitement (détection) contre 2 (détection + catégorisation)), peut occulter des phénomènes tardifs, qui apparaîtraient si un participant prenait plus de temps pour fournir sa réponse (le temps nécessaire pour identifier et choisir la catégorie du stimulus parmi deux possibilités devrait logiquement être plus long que le temps nécessaire pour identifier une cible de catégorie définie).
- ▷ Les pseudo-mots utilisés dans l'Expérience 6 sont construits à partir de mots bisyllabiques qui deviennent des pseudo-mots dès le premier phonème de la deuxième syllabe. Grâce à cette régularité dans le matériel, il est statistiquement possible de distinguer les mots des pseudo-mots avant que la deuxième syllabe n'ait été entièrement présentée, et donc d'effectuer la tâche même si le mot stimulus n'est pas totalement reconnu.

Il est donc apparu important de refaire une expérience comparable à l'Expérience 6 en manipulant le contexte expérimental de manière à retarder les réponses des participants, et ainsi se placer dans des conditions favorables à l'apparition d'effets tardifs.

#### 4.2.5.1 Aspects théoriques

Les trois hypothèses théoriques de l'Expérience 6 sont reprises.

- ▷ Hypothèse  $M_1$  : Effet d'une superposition lexicale initiale réelle  
un mot bisyllabique renfermant un mot enchâssé initial est plus difficile à reconnaître qu'un mot bisyllabique ne renfermant pas de mot enchâssé initial.
- ▷ Hypothèse  $M_2$  : Effet d'une superposition lexicale non initiale réelle  
un mot bisyllabique renfermant un mot enchâssé en position non initiale n'est pas plus difficile à reconnaître qu'un mot bisyllabique ne renfermant pas de mot enchâssé en position non initiale (reprise de l'idée proposée par le modèle de référence, TRACE).
- ▷ Hypothèse  $M_1 / M_2$  : Différence d'effet entre superpositions lexicales réelles initiale et non initiale  
un mot bisyllabique renfermant un mot enchâssé initial est plus difficile à reconnaître qu'un mot bisyllabique renfermant un mot enchâssé en position non initiale.

Pour manipuler le contexte expérimental, deux hypothèses théoriques supplémentaires sont introduites. Elles sont relatives aux 2 facteurs (tâche, type de pseudo-mot distracteur) introduits pour retarder la réponse des participants, afin de permettre aux phénomènes tardifs du traitement lexical (comme un conflit lié à une superposition lexicale réelle finale) de devenir mesurables :

▷ Hypothèse TDL /  $M_2$  :

la tâche de décision lexicale classique est plus à même de révéler des phénomènes lexicaux tardifs que la tâche de décision lexicale en « go, no go » du fait qu'un participant met plus de temps pour l'accomplir.

▷ Hypothèse Distracteur /  $M_2$  :

l'emploi d'un matériel distracteur favorisant l'écoute de la totalité du mot est plus à même de rendre manifeste des phénomènes lexicaux tardifs qu'un matériel distracteur qui permet d'effectuer la tâche sur la base de stratégies de réponse liées à des régularités statistiques dans le matériel (comme répondre sans avoir à écouter les items jusqu'à leur fin).

Cependant ces 2 hypothèses sont regroupées en une seule, hypothèse d'intérêt.

▷ Hypothèse Distracteur /  $M_2$  / TDL :

L'emploi simultané d'une tâche et d'un matériel distracteur qui ralentissent les participants dans la délivrance de leur réponse, permet de faire apparaître des effets lexicaux tardifs. Par conséquent, une influence de la lexicalité de la deuxième syllabe doit devenir observable (allongement des temps de reconnaissance du mot stimulus).

## 4.2.5.2 Aspects techniques

### 4.2.5.2.1 Facteurs expérimentaux et conditions

#### Deux facteurs lexicaux relatifs aux mots expérimentaux

À 2 modalités chacun, croisés selon le plan complet présenté dans le Tableau 62 ci-dessous (4 conditions).

Tableau 62 : Facteurs lexicaux et types de mots

		Lexicalité de la 1 <sup>ère</sup> syllabe	
		Oui ( $M_1$ )	Non ( $nM_1$ )
Lexicalité de la 2 <sup>ème</sup> syllabe	Oui ( $M_2$ )	$M_1.M_2$ « <i>Timbale</i> » (/tɛ̃/-/bal/)	$nM_1.M_2$ « <i>Climat</i> » ( $S_1$ -/ma/)
	Non ( $nM_2$ )	$M_1.nM_2$ « <i>Résine</i> » (/ʁe/- $S_2$ )	$nM_1.nM_2$ « <i>Colonne</i> » ( $S_1$ - $S_2$ )

#### Deux facteurs contextuels relatifs à la tâche et aux distracteurs

À 2 modalités chacun, croisés selon le plan complet présenté dans le Tableau 63 ci-dessous (4 conditions correspondant à 4 groupes de participants).

Tableau 63 : Facteurs et conditions contextuelles

		Moment où le distracteur passe d'un statut lexical au statut de pseudo-mot	
		1 <sup>er</sup> phonème de la 2 <sup>ème</sup> syllabe (DerS) /klãʃɔk/, /bakmid/, /ʒasbɔt/, /dʁakʃad/	Dernier phonème de la 2 <sup>ème</sup> syllabe (DerP) /plãʃu/, /fʁãʃiv/, /bibɛa/, /blazu/
Tâche	TDL en « go, no go » (gng)	gng.DerS	gng.DerP
	TDL classique (TDL)	TDL.DerS	TDL.DerP

#### 4.2.5.2.2 Hypothèses opérationnelles

- ▷ Hypothèse M<sub>1</sub> : Effet d'une superposition lexicale réelle au début d'un mot sur le traitement perceptif d'un mot parlé bisyllabique

Les temps de reconnaissance d'un mot bisyllabique dont la première syllabe est un mot sont plus longs que ceux obtenus pour la reconnaissance d'un mot bisyllabique dont la première syllabe n'est pas un mot.

- ▷ Hypothèse M<sub>2</sub> : Effet d'une superposition lexicale réelle à la fin d'un mot sur le traitement perceptif d'un mot parlé bisyllabique

Les temps de reconnaissance d'un mot bisyllabique dont la deuxième syllabe est un mot sont identiques à ceux obtenus pour la reconnaissance d'un mot bisyllabique dont la deuxième syllabe n'est pas un mot.

- ▷ Hypothèse M<sub>1</sub> / M<sub>2</sub> : Différence d'effet entre une superposition lexicale réelle au début et à la fin d'un mot sur le traitement perceptif d'un mot parlé bisyllabique

Les temps de reconnaissance d'un mot bisyllabique dont la première syllabe est un mot sont plus longs que ceux obtenus pour la reconnaissance d'un mot dont la deuxième syllabe est un mot.

- ▷ Hypothèse Distracteur / TDL / M<sub>2</sub> : augmentation de la taille de l'effet de la superposition lexicale finale par retardement de la réponse des participants

Des écarts entre les temps de reconnaissance d'un mot bisyllabique dont la deuxième syllabe est un mot et les temps de reconnaissance d'un mot bisyllabique dont la deuxième syllabe n'est pas un mot apparaissent avec l'utilisation combinée d'une tâche de décision lexicale classique et d'un matériel distracteur qui incite les participants à écouter les mots jusqu'à leur dernier phonème (les temps de reconnaissance d'un mot renfermant une superposition lexicale finale deviennent statistiquement plus longs que ceux d'un mot ne renfermant pas de superposition lexicale finale).

#### 4.2.5.2.3 Matériel expérimental utilisé

Stimuli en voix naturelle. La liste des items est donnée en Annexe 19 page 508.

- ▷ Items expérimentaux :

80 mots bisyllabiques (20 par condition – détail donné dans le Tableau 62, page 257). Il s'agit majoritairement des mêmes mots que ceux utilisés dans l'Expérience 6 (11 mots qui avaient posé problème à quelques participants ont été retirés et 3 ajoutés). Par contre, il s'agit de nouveaux enregistrements

pour garantir l'homogénéité acoustique entre items expérimentaux et entre items expérimentaux et items distracteurs.

▷ Items distracteurs :

Deux lots (DerS et DerP) de 80 pseudo-mots bisyllabiques chacun (4 catégories déterminées par la manipulation des deux facteurs expérimentaux (20 items de chaque type par lot) selon le même schéma que pour les mots expérimentaux (détail dans le Tableau 63, page 258)). Le lot DerS provient des pseudo-mots utilisés dans l'Expérience 6. Le lot DerP a été conçu pour l'Expérience 7. Les pseudo-mots « DerP » sont initialement des mots bisyllabiques dont le dernier phonème a été modifié pour aboutir à l'une des 4 structures d'items (définie par la lexicalité de chaque syllabe).

#### 4.2.5.2.4 Organisation de l'expérience

Les participants sont répartis en 4 groupes, déterminés par les 4 conditions contextuelles détaillées dans le Tableau 63, page 258. Chaque participant est testé avec 160 items, et traite chaque stimulus 1 fois.

#### 4.2.5.2.5 Population testée

Les données de 71 francophones monolingues ( $G_1$  : N = 17 ;  $G_2$  : N = 17 ;  $G_3$  : N = 18 ;  $G_4$  : N = 19), sans troubles auditifs, cognitifs et moteurs, étudiant à l'université de Genève, ont été utilisées pour les analyses. Les tests se sont déroulés en janvier 2003.

#### 4.2.5.2.6 Tâches

- ▷ Tâche de catégorisation (TDL classique). Modalité auditive.
- ▷ Tâche de détection (TDL en « go, no go » - détection des mots stimuli). Modalité auditive.

#### 4.2.5.2.7 Procédure de recueil des données

Expérience réalisée à l'université de Genève, en collaboration avec deux étudiantes effectuant une recherche pour l'obtention de leur Licence en Psychologie (L. Bischoff, D. Sunnier - Direction : C. Lachaud, U. H. Frauenfelder).

Détails techniques sur le recueil dans la Partie III, Chapitre 4.

71 \* 80 = 5680 mesures utilisées.

La phase test de l'expérience se déroule sans pause et dure moins de 15 minutes.

#### 4.2.5.2.8 Méthode d'analyse

- ▷ Détail des procédures pour le tri des données et les analyses dans la Partie III, Chapitre 3.



- ▷ Analyse de régression multiniveaux, modèle à 2 niveaux (sujets, items. Corrélation intra classe : 21.67%).
- ▷ 304 valeurs manquantes (5.43% de non-réponses et de valeurs filtrées (détail par condition : gng.DerS –  $M_1.M_2$  4.4% ;  $M_1.nM_2$  4.4% ;  $nM_1.M_2$  9.2% ;  $nM_1.nM_2$  10.3% / gng.DerP –  $M_1.M_2$  4.7% ;  $M_1.nM_2$  3.7% ;  $nM_1.M_2$  7.9% ;  $nM_1.nM_2$  10.5% / TDL.DerS –  $M_1.M_2$  3.8% ;  $M_1.nM_2$  2.8% ;  $nM_1.M_2$  5.9% ;  $nM_1.nM_2$  3.1% / TDL.DerP –  $M_1.M_2$  5% ;  $M_1.nM_2$  1.5% ;  $nM_1.M_2$  4.1% ;  $nM_1.nM_2$  4.1%)).

#### 4.2.5.2.9 Paramètres contrôlés

Les mêmes facteurs que dans l'Expérience 6 ont été contrôlés (à la construction du matériel et lors des analyses statistiques – voir Section 4.2.4.2.9, page 251).

Une vérification a été faite pour savoir s'il était nécessaire de contrôler les facteurs prosodiques lors des analyses statistiques (détail des analyses acoustiques présentés en Annexe 20 page 510) : ces facteurs ont été intégrés au modèle d'analyse vu qu'ils sont en corrélation avec les variables expérimentales.

En plus de ces facteurs, l'homogénéité acoustique du matériel a été contrôlée à la construction et à la structuration de l'expérience, en procédant à un nouvel enregistrement de tous les stimuli utilisés (éviter aux participants de pouvoir différencier les pseudo-mots (lot « DerP ») des mots et faire la tâche sur une base stratégique).

Le facteur liste n'a pas été directement contrôlé (se reporter à la note de bas de page n°105, page 201).

#### 4.2.5.3 Résultats obtenus

##### 4.2.5.3.1 Rapport des analyses statistiques

L'équation du modèle (Équation 44) et les tests de significativité relatifs aux covariables peuvent être consultés à l'Annexe 21 page 514.

Tableau 64 : Tests de significativité pour les 2 facteurs lexicaux

	Coefficient	Chi <sup>2</sup> <sub>(1, N=5296)</sub>	p<	% variance expliquée
<b>M<sub>1</sub></b>	-62.032	26.601	<b>3<sup>E-07</sup></b>	0.434
<b>M<sub>2</sub></b>	16.342	2.454	.12	0.029
<b>M<sub>1</sub>.M<sub>2</sub></b>	-8.254	0.313	.58	0.002

Tableau 65 : Tests de significativité pour les facteurs contextuels et les interactions entre facteurs (lexicaux et contextuels)

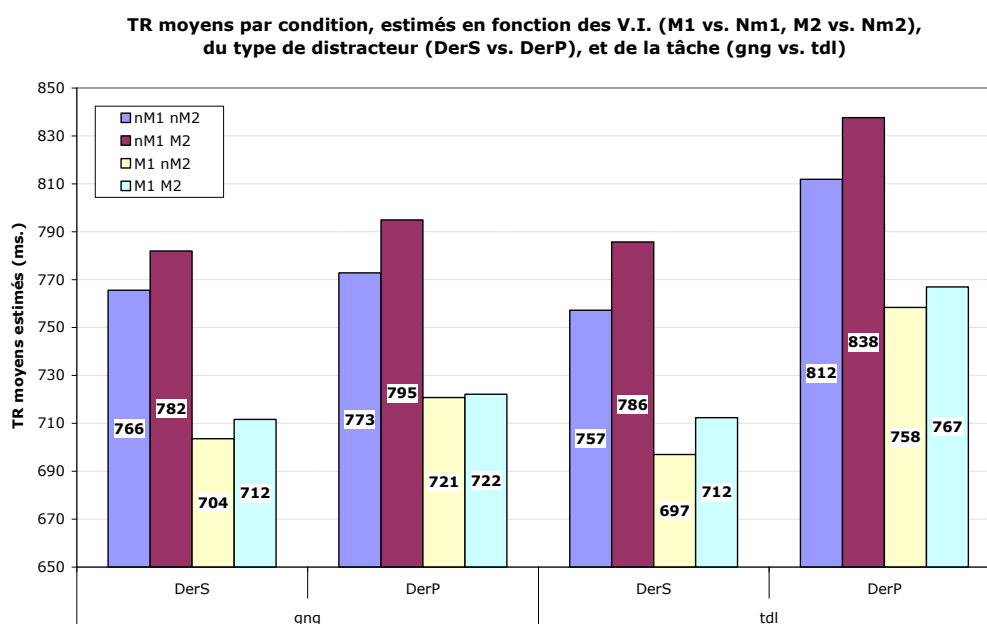
	Coefficient	Chi <sup>2</sup> (1, N=5296)	p<
TDL	-8.401	0.176	.68
DerP	7.21	0.137	.72
M <sub>1</sub> .TDL	1.829	0.026	.88
M <sub>1</sub> .DerP	10.015	0.846	.36
M <sub>2</sub> .TDL	12.147	1.141	.29
M <sub>2</sub> .DerP	5.806	0.278	.60
TDL.DerP	47.46	2.865	.091
M <sub>1</sub> .M <sub>2</sub> .TDL	-4.886	0.094	.76
M <sub>1</sub> .M <sub>2</sub> .DerP	-12.529	0.665	.42
M <sub>1</sub> .TDL.DerP	-3.278	0.044	.84
M <sub>2</sub> .TDL.DerP	-8.507	0.289	.60
M <sub>1</sub> .M <sub>2</sub> .TDL.DerP	8.4	0.143	.71

Tableau 66 : Tests de significativité des contrastes entre conditions pour les facteurs lexicaux

Contraste		Chi <sup>2</sup> (1, N=5296)	p<
M <sub>1</sub> M <sub>2</sub>	M <sub>1</sub> nM <sub>2</sub>	0.645	.43
	nM <sub>1</sub> M <sub>2</sub>	46.318	2 <sup>E-11</sup>
	nM <sub>1</sub> nM <sub>2</sub>	28.66	9 <sup>E-08</sup>
M <sub>1</sub> nM <sub>2</sub>	nM <sub>1</sub> M <sub>2</sub>	43.182	5 <sup>E-11</sup>
	nM <sub>1</sub> nM <sub>2</sub>	26.601	3 <sup>E-07</sup>
nM <sub>1</sub> M <sub>2</sub>	nM <sub>1</sub> nM <sub>2</sub>	2.454	.12

### 4.2.5.3.2 Présentation des résultats

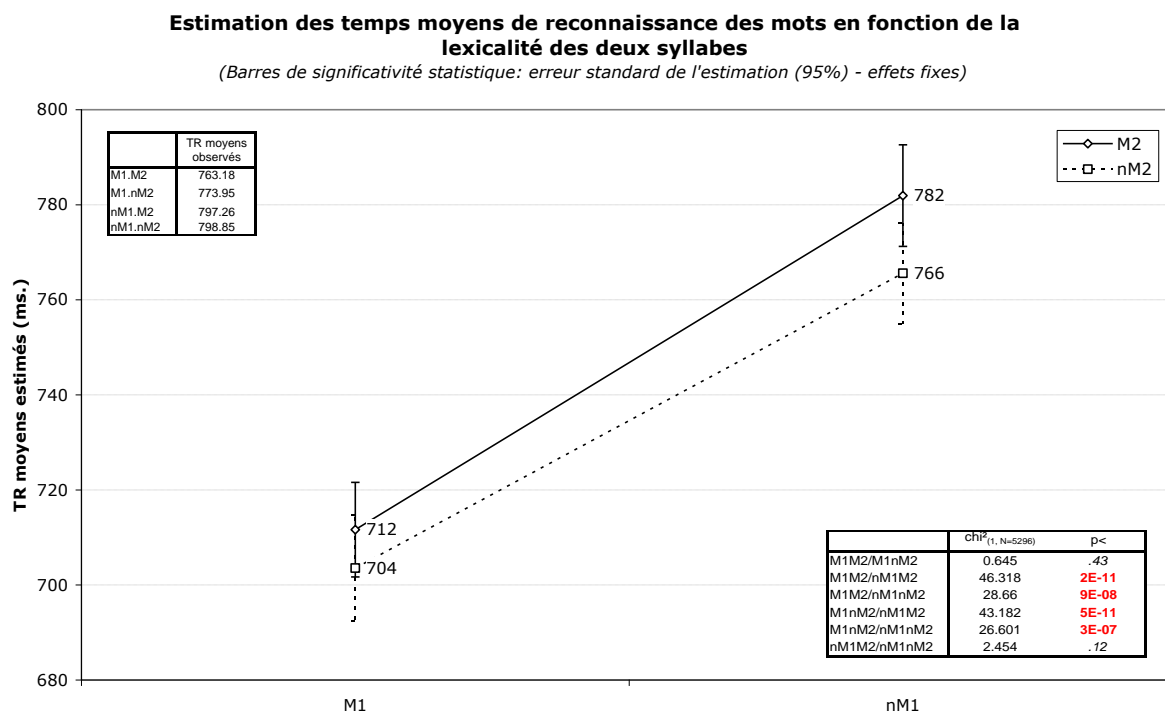
Figure 57 : Vue d'ensemble des temps de réponse moyens estimés par condition



Ce graphique est présenté afin de donner au lecteur une vision d'ensemble des moyennes estimées par condition.

Les données sont par la suite décomposées et organisées en fonction des questions posées dans les hypothèses.

Figure 58 : Temps de réponse moyens estimés en fonction de la lexicalité des syllabes du mot stimulus



Hypothèse M<sub>1</sub> : Effet d'une superposition lexicale initiale réelle sur le traitement d'un mot parlé bisyllabique

Les mots dont la première syllabe est lexicale sont reconnus en moyenne 62 ms. plus rapidement que les mots dont la première syllabe n'est pas lexicale (se reporter au Tableau 64, page 260). La première hypothèse est donc invalidée par les données recueillies.

Hypothèse M<sub>2</sub> : Effet d'une superposition lexicale finale sur le traitement d'un mot parlé bisyllabique

Les mots dont la deuxième syllabe est lexicale sont reconnus aussi rapidement que les mots dont la deuxième syllabe n'est pas lexicale (se reporter au Tableau 64, page 260). La deuxième hypothèse est donc validée.

Hypothèse  $M_1 / M_2$  : Différence d'effet entre une superposition lexicale réelle initiale et une superposition lexicale finale sur le traitement d'un mot parlé bisyllabique

Les écarts entre les conditions  $M_1.nM_2$  et  $nM_1.M_2$  sont significatifs (se reporter au Tableau 66, page 261). Les mots dont la première syllabe est lexicale sont traités plus rapidement que les mots dont la deuxième syllabe est lexicale. L'hypothèse est donc invalidée par les données recueillies.

Effet des facteurs contextuels : Hypothèse Distracteur / TDL /  $M_2$ 

Chaque facteur contextuel considéré séparément n'a pas une influence significative sur les temps moyens estimés de traitement des mots expérimentaux (TDL et DerP, Tableau 65, page 261). Par contre, lorsque les deux facteurs contextuels sont présents de façon combinée dans la situation expérimentale (TDL.DerP), les temps moyens estimés de traitement des mots expérimentaux tendent à être plus longs (en moyenne de 47 ms ;  $p < .1$ ). Cela signifie que lorsque les pseudo-mots sont des mots qui changent de statut lexical par leur dernier phonème ET que la tâche demandée aux auditeurs porte sur un choix catégoriel (condition contextuelle la plus adverse pour le traitement des stimuli), les temps de traitement des mots tendent à être plus longs que dans les 3 autres conditions contextuelles. Bien que non significatif, cet effet est intéressant car son amplitude peut être suffisante pour permettre l'émergence d'un effet de superposition lexicale réelle non initiale.

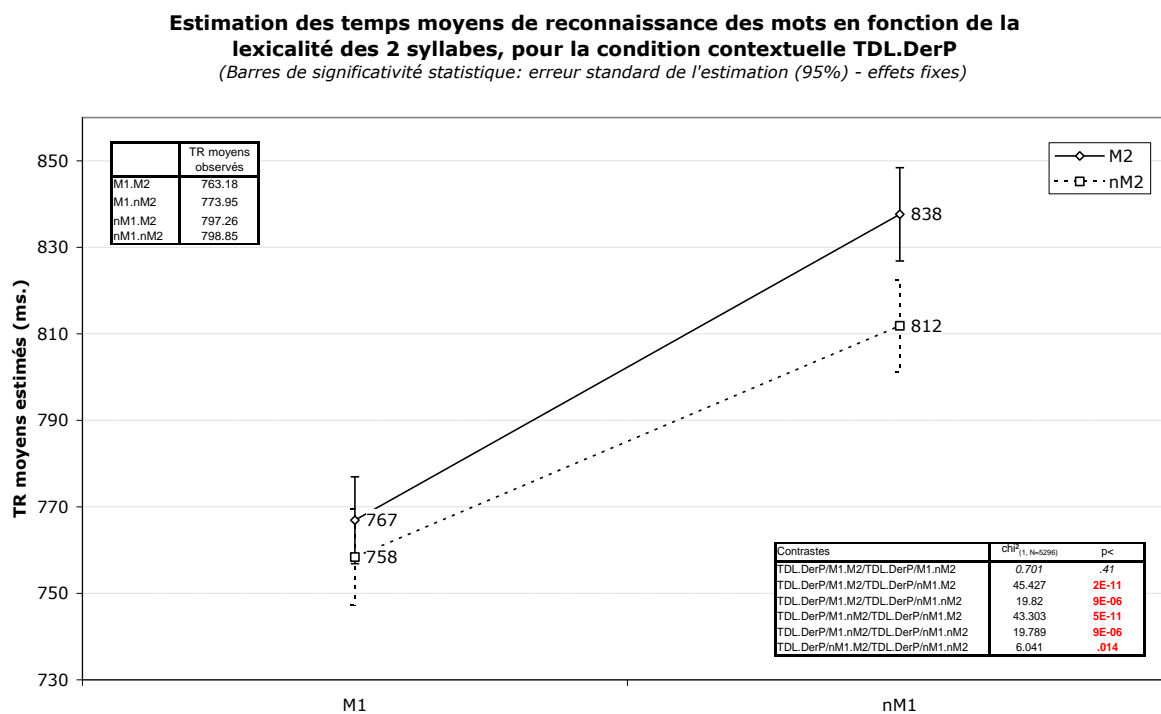
Tableau 67 : Tests de significativité des contrastes entre conditions lexicales dans la condition contextuelle TDL.DerP

Contraste		Chi <sup>2</sup> <sub>(1, N=5296)</sub>	p<
TDL.DerP / $M_1.M_2$	TDL.DerP / $M_1.nM_2$	0.701	.41
	TDL.DerP / $nM_1.M_2$	45.427	<b>2<sup>E-11</sup></b>
	TDL.DerP / $nM_1.nM_2$	19.82	<b>9<sup>E-06</sup></b>
TDL.DerP / $M_1.nM_2$	TDL.DerP / $nM_1.M_2$	43.303	<b>5<sup>E-11</sup></b>
	TDL.DerP / $nM_1.nM_2$	19.789	<b>9<sup>E-06</sup></b>
TDL.DerP / $nM_1.M_2$	TDL.DerP / $nM_1.nM_2$	6.041	<b>.014</b>

L'analyse des contrastes entre types de mots stimuli pour cette condition contextuelle spécifique révèle une différence liée à la lexicalité de la syllabe  $S_2$  lorsque le mot ne renferme pas de superposition lexicale réelle initiale. Les temps de reconnaissance des mots renfermant uniquement un enchâssement non initial sont en moyenne allongés de 26 ms par rapport à ceux de mots ne renfermant aucune superposition lexicale.

La Figure 59 ci-dessous présente les temps moyens estimés de reconnaissance des mots stimuli dans la condition contextuelle TDL.Derp, en fonction des 2 facteurs lexicaux M<sub>1</sub> et M<sub>2</sub>.

Figure 59 : Temps moyens estimés de reconnaissance des mots dans la condition contextuelle TDL.Derp



Cet effet de lexicalité de la deuxième syllabe du mot stimulus n'apparaît pas si la première syllabe du mot stimulus est lexicale. Il pourrait donc être masqué par d'autres processus lexicaux liés à la présence d'un mot enchâssé initial.

#### 4.2.5.4 Conclusions de l'Expérience 7

Des 4 hypothèses de l'Expérience 7, seule la 2<sup>ème</sup> hypothèse est franchement validée par les résultats. L'effet facilitant d'une superposition lexicale réelle initiale sur la reconnaissance d'un mot parlé a été répliqué (déjà observé dans les Expériences 4, 5 et 6). Cet effet semble robuste, et va à l'encontre des prédictions faites par le modèle TRACE.

Dans cette expérience, l'effet d'une superposition lexicale réelle varie en fonction de sa localisation dans le mot porteur :

- ▷ Une superposition lexicale initiale réelle facilite la reconnaissance du mot porteur
- ▷ Une superposition lexicale finale produit une difficulté à reconnaître le mot porteur.

Ce dernier effet est plus difficile à observer (il nécessite des conditions particulières dans la situation expérimentale ; il disparaît en cas de traitement lexical plus complexe (lexicalité de la première syllabe)). Il s'agirait donc d'un phénomène n'influençant pas la reconnaissance du mot stimulus dans des proportions importantes, et qui pourrait être relatif à des processus de segmentation de la séquence sonore en mots, plutôt qu'à des processus relatifs à l'accès lexical.

### 4.2.6 Expérience 8 : Effet d'une superposition lexicale initiale réelle sur la reconnaissance d'un mot monosyllabique, en modalité visuelle

Le matériel utilisé jusqu'ici pour tester l'effet d'une superposition lexicale réelle sur la reconnaissance du mot stimulus est de nature acoustique. L'information à traiter parvient donc au système de façon graduelle, et au fur et à mesure que le traitement s'effectue. Ceci soulève deux problèmes :

- ▷ le problème de l'origine de la mesure, qui a déjà été abordé dans la Partie III, Chapitre 2 (où les solutions techniques développées sont exposées)
- ▷ le possible masquage du conflit lexical prédit par TRACE, à cause de l'étalement du traitement dans le temps (lié à cette arrivée progressive de l'information dans le système), qui ne serait pas mesurable avec la méthode chronométrique utilisée (il resterait encore à expliquer pourquoi, suite à un conflit lexical, une facilitation est obtenue).

Il est donc nécessaire de vérifier si l'hypothèse de la compétition lexicale prédite par TRACE dans le cas d'une superposition lexicale réelle est vraie, grâce à une situation expérimentale où elle ne risque pas d'être masquée. Pour cela, le matériel linguistique est présenté en modalité visuelle. La totalité de l'information physique décrivant le mot est alors délivrée au système au même instant, et si un conflit a lieu entre la représentation mentale du mot stimulus et celle du mot enchâssé, il ne peut plus être masqué par le caractère temporel du stimulus.

#### 4.2.6.1 Aspects théoriques

Selon les prédictions du modèle TRACE, un mot monosyllabique renfermant à son initiale une superposition lexicale réelle doit être reconnu moins vite qu'un mot monosyllabique ne renfermant pas de superposition lexicale réelle initiale, à cause du phénomène de compétition lexicale. Cette prédiction est vraie en modalité auditive comme en modalité visuelle.

#### 4.2.6.2 Aspects techniques

##### 4.2.6.2.1 Facteurs expérimentaux et conditions

1 facteur expérimental (enchâssement initial) à 2 modalités (présence / absence), donnant lieu à deux conditions expérimentales :

- ▷ mots avec un enchâssement initial (exemple : le mot « *traite* » renferme le mot « *trait* »)

- ▷ mots sans enchâssement initial phonologiquement appariés aux mots de la condition précédente (exemple : le mot « *trente* » ne renferme pas de mot enchâssé à son initiale)

#### 4.2.6.2.2 Hypothèse opérationnelle

Les mots monosyllabiques qui renferment un mot enchâssé initial seront reconnus plus lentement que les mots monosyllabiques qui ne renferment pas de mot enchâssé initial.

#### 4.2.6.2.3 Matériel expérimental utilisé

Stimuli visuels. La liste des items est donnée en Annexe 22 page 516.

- ▷ Items expérimentaux :  
36 mots monosyllabiques (18 de chaque condition : « *traite* » ; « *trente* »). Les mots sont appariés phonologiquement, comme dans l'Expérience 5. La modification pour passer d'une condition à l'autre porte sur 1 ou 2 caractères (1 graphème), de manière à obtenir un changement du noyau vocalique.
- ▷ Items distracteurs :  
18 pseudo-mots monosyllabiques. 9 renferment un mot enchâssé initial (« *brinre* »), 9 non (« *cheupe* »).

Le faible nombre d'items s'explique par la difficulté à trouver des paires de mots, contrôlées à la fois graphiquement et phonologiquement.

#### 4.2.6.2.4 Organisation de l'expérience

Les participants sont répartis en 2 groupes (un participant est testé sur la moitié des items), pour les mêmes raisons que celles détaillées dans l'Expérience 5 (Section 4.2.3.2.4, page 243). Chaque participant est testé avec 36 items, et traite chacun d'eux 1 fois.

#### 4.2.6.2.5 Population testée

Les données de 26 participants (13 par groupe) francophones monolingues, sans troubles visuels (sauf si vision corrigée), cognitifs et moteurs, étudiant à l'université de Genève, ont été utilisées pour les analyses. Le recueil a été effectué en janvier 2004.

#### 4.2.6.2.6 Tâche

Tâche de détection (TDL en « go, no go » - détection des mots stimuli). Modalité visuelle.



#### 4.2.6.2.7 Procédure de recueil des données

Expérience réalisée à l'université de Genève. Les stimuli sont présentés avec des caractères en police « Verdana » de taille 16, au centre d'un écran cathodique de 19 pouces. Leur apparition est précédée d'une croix de fixation localisée sur le premier caractère de l'item. Celle-ci s'affiche 1 seconde et disparaît à l'apparition du stimulus. Chaque item reste affiché jusqu'à ce que le participant fournisse sa réponse, à l'aide d'un boîtier et selon la procédure standard suivie pour les tâches de décision lexicale dans les expériences en modalité auditive.

13 \* 2 \* 18 = 468 mesures utilisées.

La phase test de l'expérience dure 3 ou 4 minutes.

#### 4.2.6.2.8 Méthode d'analyse

- ▷ Détail des procédures pour le tri des données et les analyses dans la Partie III, Chapitre 3.
- ▷ Analyse de régression multiniveaux, modèle à 2 niveaux (sujets, items. Corrélation intra classe : 14.42%).
- ▷ 60 valeurs manquantes (12.82% de non-réponses et valeurs filtrées (détail par condition : E 14.1% ; nE 11.5%)).

#### 4.2.6.2.9 Paramètres contrôlés

##### À la construction du matériel et à la structuration de l'expérience

- ▷ Mots monosyllabiques
- ▷ Mots phonologiquement et graphiquement appariés entre conditions : la variation porte sur un seul graphème (maximum 2 lettres) correspondant au noyau vocalique de la syllabe (situation comparable à l'Expérience 5)
- ▷ La superposition lexicale initiale réelle existe à la fois en modalité visuelle et en modalité auditive
- ▷ Le mot enchâssé n'a pas de lien sémantique ou morphologique avec le mot porteur, sauf dans 2 cas sur 36 (« net » / « nette », « blet » / « blette »)
- ▷ Les pseudo-mots sont construits en fonction du facteur expérimental
- ▷ Les pseudo-mots sont appariés aux mots pour le nombre de caractères et la structure syllabique
- ▷ Proportion identique d'exemplaires de chaque condition, aussi bien pour les mots que pour les pseudo-mots (ne pas favoriser un traitement par un déséquilibre dans le nombre d'exemples).
- ▷ Présentation aléatoire des items lors de la passation

##### Lors des analyses statistiques

- ▷ Latéralité manuelle des participants

- ▷ Sexe des participants
  - ▷ Nombre de caractères composant le mot stimulus
  - ▷ Nombre de phonèmes composant le mot stimulus
  - ▷ Position du point d'unicité graphique du mot stimulus
  - ▷ Position du point d'unicité phonologique du mot stimulus
  - ▷ Fréquence formelle graphique de l'item (logarithme en base 10)
  - ▷ L'âge d'acquisition estimé et la familiarité conceptuelle n'ont pas été utilisés dans les analyses, compromis laxiste justifié par les points suivants :
    - ∩ le recueil des données par Internet pour ces deux facteurs a été effectué avant que l'Expérience 8 ne soit conçue. L'acquisition supplémentaire des données manquantes aurait nécessité une enquête « manuelle » assez coûteuse en ressources.
    - ∩ les données déjà à disposition ne renseignaient que pour 16 des 36 mots, ce qui a interdit l'intégration de l'âge d'acquisition et de la familiarité conceptuelle dans l'analyse
- Remarque : une partie de l'influence de ces deux facteurs est cependant contrôlée par la fréquence formelle (corrélations des 3 facteurs entre eux).
- ▷ Le facteur liste n'a pas été directement contrôlé (voir la note de bas de page n°105, page 201).

#### Lors d'une vérification a posteriori

Afin de s'assurer que tout effet obtenu n'est pas dû à des facteurs parasites comme le nombre de voisins orthographiques (ONTY) ou phonographique (PGNTY) par mot, des analyses de régression linéaire simple du matériel ont été effectuées a posteriori à partir des données de la base LEXOP (Peereman & Content, 1999). L'analyse ne porte que sur 78% des items de l'expérience (N = 28 sur 36), 4 items du matériel n'étant pas renseignés ou correctement renseignés par LEXOP (« *traque* », « *lu* » dans « *luge* », « *ta* » dans « *tape* », et « *spi* » dans « *spire* »), les items appariés (« *trique* », « *loge* », « *type* » et « *spore* ») ayant donc aussi dû être retirés.

L'analyse ne révèle aucune différence significative des caractéristiques du voisinage des items entre les deux conditions expérimentales (ONTY :  $F_{(1, 26)} = 0.5$ ,  $p < .49$  ; PGNTY :  $F_{(1, 26)} = 1.95$ ,  $p < .18$ ).

Une différence apparaît si les scores ONTY et PGNTY de la condition des mots avec enchâssement intègrent aussi les scores des mots enchâssés (ONTY :  $F_{(1, 26)} = 11.37$ ,  $p < .0024$  ; PGNTY :  $F_{(1, 26)} = 13.59$ ,  $p < .0011$ ). Toutefois, ce résultat est une autre expression du facteur enchâssement manipulé, et le matériel utilisé apparaît donc comme correctement équilibré pour ces aspects relatifs au voisinage lexical d'un mot.

### 4.2.6.3 Résultats obtenus

#### 4.2.6.3.1 Rapport des analyses statistiques

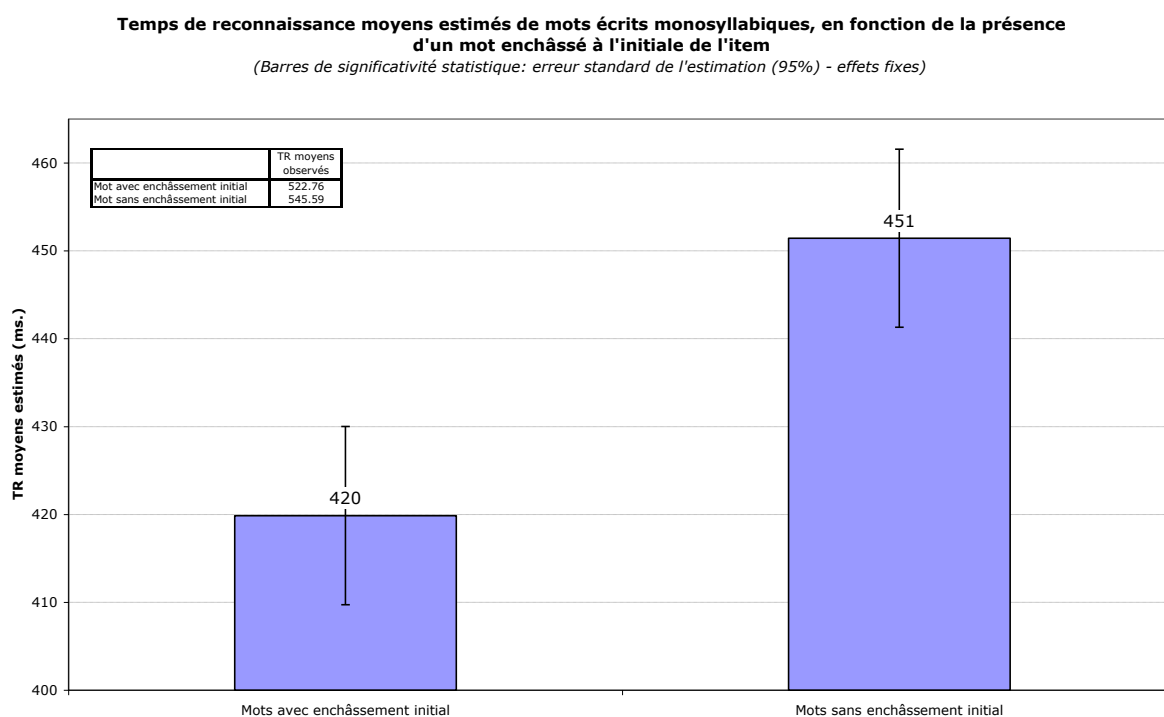
L'équation du modèle (Équation 45) et les tests de significativité relatifs aux covariables peuvent être consultés à l'Annexe 23 page 517.

Tableau 68 : Test de significativité du facteur expérimental

	Coefficient	Chi <sup>2</sup> <sub>(1, N=408)</sub>	p<	% variance expliquée
<b>E</b>	-31.572	9.318	<b>.003</b>	2.089

#### 4.2.6.3.2 Présentation des résultats

Figure 60 : Temps moyens de reconnaissance des mots, en fonction de la présence d'un mot enchâssé initial



En modalité visuelle, les participants reconnaissent 32 ms. plus rapidement un mot monosyllabique renfermant un mot enchâssé initial qu'un mot monosyllabique phonologiquement apparié et ne renfermant pas de mot enchâssé initial. Ces résultats invalident donc l'hypothèse de l'étude.

#### **4.2.6.4 Conclusions de l'Expérience 8**

L'hypothèse de l'Expérience 8 a été invalidée par les résultats obtenus. L'effet de facilitation lié à une superposition lexicale initiale réelle sur la reconnaissance d'un mot monosyllabique, est également trouvé avec un matériel linguistique présenté en modalité visuelle. Dans ce cas, le phénomène de compétition lexicale n'a pas pu être masqué par l'arrivée progressive de l'information comme c'est le cas en modalité auditive. Il faut donc en conclure que s'il existe, il n'a pas lieu à une étape précoce du traitement formel, mais à une étape plus tardive, post-lexicale ou sémantique, et que le modèle TRACE est incomplet, ou qu'il existe un biais insoupçonné dans la procédure suivie dans cette étude.

#### **4.2.7 Expérience 9 : Sondage de la charge de calcul dans le système, aux différentes étapes de la reconnaissance d'un mot monosyllabique impliqué dans des superpositions lexicales initiales réelles et virtuelles, en modalité auditive**

Bien que montrant les mêmes effets que dans les Expériences 4, 5, 6 et 7, l'Expérience 8 teste le traitement d'un matériel linguistique présenté en modalité visuelle, alors que les buts poursuivis dans cette étude concernent le traitement de mots parlés (modalité auditive). Des différences de traitement existent peut-être au niveau lexical en fonction de la modalité. Il serait donc intéressant de disposer d'une technique pour tester l'existence d'un conflit lexical à un moment donné du traitement d'un mot parlé, puisque ce conflit est prédit par un modèle validé expérimentalement, mais n'a pas pu être observé dans les mesures réalisées jusqu'ici dans ce travail.

L'Expérience 9 propose d'utiliser un procédé de sondage du traitement lexical à différentes étapes de son évolution. Ce sondage est effectué à l'aide d'une cible auditive non linguistique insérée dans le mot stimulus. L'auditeur est chargé de détecter cette cible le plus rapidement possible. Le traitement lexical étant irrépessible, il doit interférer avec le traitement nécessaire à l'effection de la tâche<sup>111</sup>. Des différences devraient alors pouvoir être mesurées en fonction de la difficulté du traitement lexical, liée aux facteurs lexicaux manipulés.

La compétition lexicale impliquant une quantité supplémentaire de calculs à effectuer, l'équation est simple :

- ▷ soit il faut mobiliser plus de ressources<sup>112</sup> pour effectuer ces calculs
- ▷ soit il faut plus de temps, à niveau constant de ressources utilisées

Étant donné que cette deuxième proposition ne semble pas en accord avec les données (qui montrent que moins de temps est nécessaire), il faut donc préférer la solution relative à la mobilisation des ressources du système pour tenter d'expliquer

---

<sup>111</sup> L'outil de sondage (la tâche) et l'objet sondé (le processus de reconnaissance du mot) correspondent à des traitements de type et de niveau différents. Le premier est effectué consciemment (nécessite une analyse attentive du stimulus), tandis que le deuxième se produit automatiquement et inconsciemment. La tâche de détection ne permet donc pas de sonder directement les phénomènes qui surviennent lors de la reconnaissance d'un mot, comme la compétition lexicale ou les processus infra-lexicaux. Par contre, elle devrait permettre de révéler les moments où le système est transitoirement « occupé ailleurs », donc moins apte à procéder à la détection de la cible (phénomène d'interférence infra-lexicale/lexicale (?) sur la détection de la cible).

<sup>112</sup> Ces ressources peuvent être biologiques (structurales, énergétiques), psychologiques (attentionnelles), ou autres. La modulation des capacités du système pour effectuer une tâche secondaire serait donc corrélée à la disponibilité des ressources dans le système perceptif.

d'éventuels effets de compétition lexicale. Si de la compétition se produit, cette technique de sondage devrait théoriquement permettre de la mesurer.

#### 4.2.7.1 Aspects théoriques

Deux hypothèses théoriques ont trait à l'évaluation de la compétition lexicale dans l'Expérience 9 :

- ▷ Hypothèse relative à la quantité de possibilités lexicales à explorer :  
la détection des cibles devrait être d'autant plus difficile que la quantité d'hypothèses lexicales est grande (superpositions lexicales réelle et virtuelle).
- ▷ Hypothèse relative au moment où plusieurs possibilités lexicales entrent en conflit :  
la détection des cibles devrait être plus difficile lorsque la cible est insérée au moment où une superposition lexicale survient.

Dans la situation expérimentale, ces deux hypothèses ne sont pas séparables, et seront réorganisées (se reporter à la présentation des hypothèses opérationnelles).

#### 4.2.7.2 Aspects techniques

##### 4.2.7.2.1 Facteurs expérimentaux et conditions

###### 4.2.7.2.1.1 Deux facteurs lexicaux

À deux modalités chacun (présence / absence d'une superposition réelle / virtuelle) - identiques à ceux de l'Expérience 4 (voir Section 4.2.2.2.1, page 236).

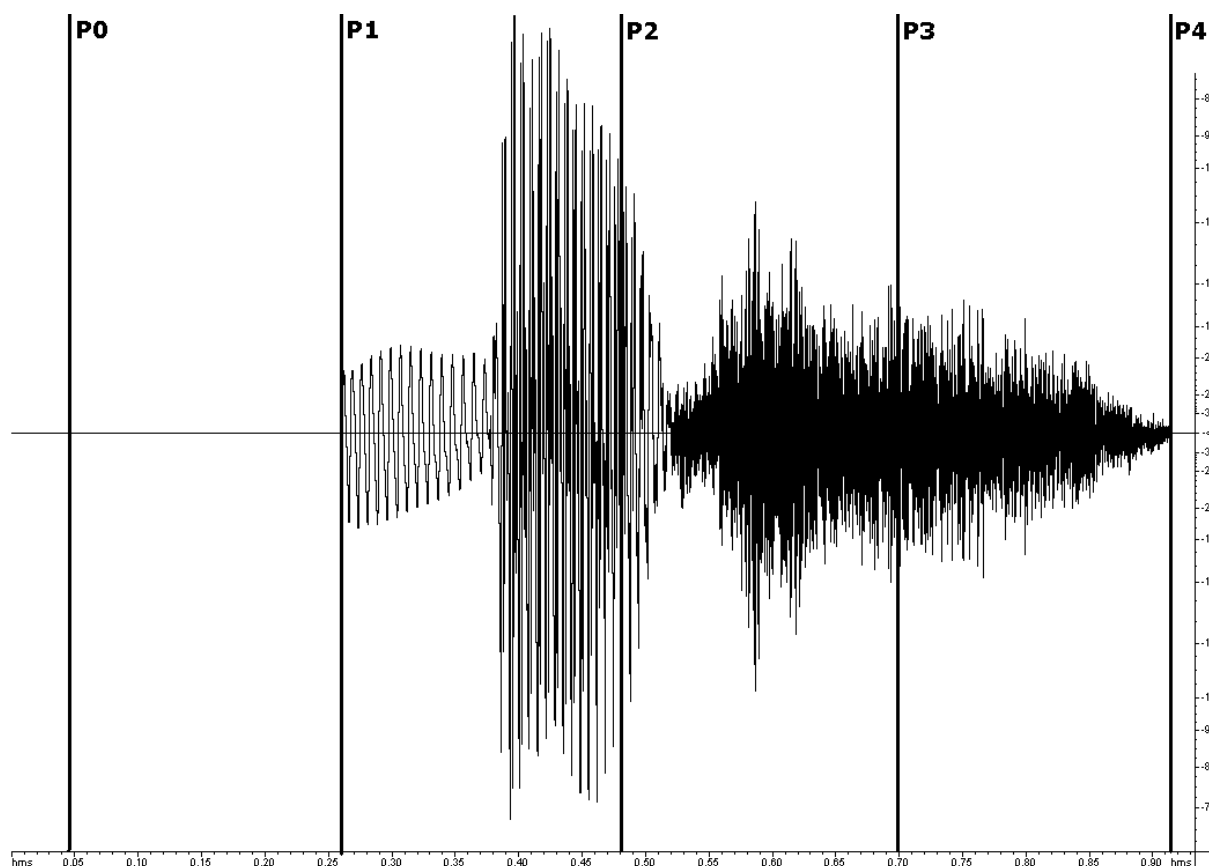
###### 4.2.7.2.1.2 Deux facteurs techniques

###### Position de la cible dans l'item

À 5 modalités (Figure 61 ci-après).

- ▷ P<sub>0</sub> : condition « Contrôle ». Cible insérée 500 ms. avant le début du mot ou du pseudo-mot (valeur fixe), et toujours à la même durée du début de l'item que le début du mot / pseudo-mot dans un item P<sub>1</sub>-P<sub>4</sub> (700 ms. – position fixe)
- ▷ P<sub>1</sub> : début du mot / pseudo-mot – Condition test « P<sub>1</sub> » (position fixe)
- ▷ P<sub>2</sub> : 33% de la durée du mot / pseudo-mot – Condition test « P<sub>2</sub> » (position relative à la durée de l'item)
- ▷ P<sub>3</sub> : 66% de la durée du mot / pseudo-mot – Condition test « P<sub>3</sub> » (position relative à la durée de l'item)
- ▷ P<sub>4</sub> : fin du mot / pseudo-mot – Condition test « P<sub>4</sub> » (position relative à la durée de l'item)

Figure 61 : Position de la cible dans l'item



#### Type de sonorité insérée dans l'item

Un facteur lié à la tâche, à 2 modalités :

- ▷ cible : « clic » (bruit blanc, durée 10 ms.)
- ▷ sonorité distractive : « bip » (son périodique perçu comme acoustiquement voisin, mais tout de même perçu comme différent, du bruit blanc - durée 10 ms.)

Les facteurs lexicaux déterminent 4 types de mots expérimentaux (E.I, E.nI, nE.I, nE.nI). Les facteurs techniques déterminent 5 situations expérimentales (P<sub>0</sub>.clic, P<sub>1</sub>.clic, P<sub>2</sub>.clic, P<sub>3</sub>.clic, P<sub>4</sub>.clic) et 5 situations distractrices (P<sub>0</sub>.bip, P<sub>1</sub>.bip, P<sub>2</sub>.bip, P<sub>3</sub>.bip, P<sub>4</sub>.bip). Au total, l'expérience comporte  $4 * 5 = 20$  conditions expérimentales (et 20 conditions distractrices).

#### **4.2.7.2.2 Hypothèses opérationnelles**

Les deux hypothèses théoriques ne sont pas séparables dans la réalité comme elles le sont logiquement, puisque ce qui est testé dans l'expérience est l'effet d'une superposition lexicale initiale sur l'évolution du traitement du stimulus. Les deux hypothèses opérationnelles suivantes sont donc proposées :

- ▷ Hypothèse E : Effet d'une superposition lexicale initiale réelle (enchâssement) :

les temps de détection des cibles seront en moyenne plus longs lorsque la cible est insérée dans le mot renfermant un mot enchâssé, au moment où survient l'enchâssement lexical (positions  $P_2$  et  $P_3$  : activation du mot enchâssé et du mot porteur), que lorsqu'elle est insérée au même moment dans un mot ne renfermant pas de mot enchâssé. En position  $P_4$ , lorsque le conflit lexical est résolu, il ne devrait pas exister de différence entre les temps de détection des cibles insérées dans les mots avec enchâssement et dans les mots sans enchâssement (ou différence résiduelle non significative). Aucune différence ne devrait être observée en  $P_0$  et  $P_1$  pour ce facteur, puisque les traitements lexicaux n'ont pas encore débuté ( $P_0$  et  $P_1$ ), ou sont trop précoces s'ils ont débuté ( $P_1$ ).

- ▷ Hypothèse I : Effet d'une superposition lexicale initiale virtuelle (enchâssabilité) :

les temps de détection des cibles seront en moyenne plus longs lorsque la cible est insérée dans un mot enchâssable à l'initiale d'un autre mot (positions  $P_2$ ,  $P_3$  et  $P_4$  : le mot peut se continuer par d'autres mots), que lorsqu'elle est insérée dans un mot non enchâssable. Aucune différence ne devrait être observée en  $P_0$  et  $P_1$ , pour les mêmes raisons que dans l'hypothèse précédente.

Pour récapituler :

- ▷ une superposition lexicale initiale réelle produit une différence en  $P_2$  et  $P_3$ , pas en  $P_4$ .
- ▷ une superposition lexicale initiale virtuelle produit une différence en  $P_2$ ,  $P_3$  et  $P_4$ . La différence entre les 2 types de superposition est donc visible en  $P_4$ .

#### 4.2.7.2.3 Matériel expérimental utilisé

Il s'agit des stimuli utilisés dans l'Expérience 4 (stimuli en voix naturelle. 60 mots expérimentaux monosyllabiques ; Distracteurs : 60 mots et 60 pseudo-mots bisyllabiques, 60 pseudo-mots monosyllabiques – liste donnée en Annexe 9 page 479). Tous les fichiers son de l'Expérience 4 sont réutilisés afin de pouvoir faire une comparaison directe entre les résultats de l'Expérience 4 et de l'Expérience 9.

Ces fichiers son ont subi un traitement numérique pour l'insertion des sonorités cible et distractrice aux positions définies (décrit dans la Section 3.4.1.3.2 page 144).

- ▷ Items expérimentaux :
  - ∩ 15 mots expérimentaux \* 4 types de mots \* 5 positions de la cible à détecter = 300
- ▷ Items distracteurs :
  - ∩ 15 mots expérimentaux \* 4 types de mots \* 5 positions de la sonorité distractrice = 300
  - ∩ 15 mots distracteurs bisyllabiques \* 4 types de mots \* 5 positions \* 2 types de sonorités = 600



- ∩ 60 pseudo-mots monosyllabiques \* 5 positions \* 2 types de sonorités = 600
- ∩ 15 pseudo-mots bisyllabiques \* 4 types de pseudo-mots \* 5 positions \* 2 types de sonorités = 600

L'expérience comporte donc un total de 2400 stimuli.

#### 4.2.7.2.4 Organisation de l'expérience

Un participant est testé avec tous les mots expérimentaux, chacun lui étant présenté une seule fois. Il est testé avec toutes les situations (P? \* Sonorité), et donc dans les 40 conditions (20 conditions expérimentales + 20 conditions distractrices). Cependant, un participant n'est pas testé avec tous les items : lorsqu'il est testé avec un mot, c'est pour l'une des 10 situations (5 situations expérimentales + 5 situations distractrices). Afin que tous les mots soient testés dans toutes les situations, il est donc nécessaire d'avoir recours à 10 groupes de participants.

Chaque participant est testé avec 240 stimuli, et fournit 30 réponses utiles (et non 60 : les 30 qui manquent sont dues à l'utilisation d'une sonorité distractrice dans la moitié des cas), 6 par position de cible. Le plan n'est donc pas complet, ni totalement équilibré :

- ▷ il n'est pas possible d'avoir le même nombre de réponses d'un participant par situation et par type de mot ( $6 / 4 = 1.5$ . Un participant peut fournir, par exemple : 2 réponses E.I, 2 réponses E.nI, 1 réponse nE.I, 1 réponse nE.nI)
- ▷ il n'est pas possible d'avoir le même nombre de réponses par type de mot pour l'ensemble de l'expérience ( $30 / 4 = 7.5$ ).

Cependant, le recueil est équilibré sur l'ensemble des listes, et la répartition inégale des items par participant est prise en considération dans les analyses.

#### 4.2.7.2.5 Population testée

Les données de 140 individus (nombre variable mais voisin par groupe : N = 14 pour les Groupes 1 à 8 ; N = 15 pour le Groupe 9 ; N = 13 pour le Groupe 10), francophones monolingues, sans troubles auditifs, cognitifs et moteurs, étudiant à l'université de Genève, ont été utilisées pour les analyses. Le recueil a été effectué en janvier et mars 2003.

#### 4.2.7.2.6 Tâche

Tâche de détection de cibles. Modalité auditive.

#### 4.2.7.2.7 Procédure de recueil des données

Expérience réalisée à l'université de Genève, en collaboration avec 3 étudiantes effectuant une recherche pour l'obtention de leur Licence en Psychologie (V. Beysard, A. Boulguy, V. Zufferey - Direction : C. Lachaud, U. H. Frauenfelder). Détails techniques dans la Partie III, Chapitre 4.

140 \* 30 = 4200 mesures utilisées.

La phase test de l'expérience se déroule sans pause, et dure une quinzaine de minutes.

#### 4.2.7.2.8 Méthode d'analyse

- ▷ Détail des procédures pour le tri des données et les analyses dans la Partie III, Chapitre 3.
- ▷ Analyse de régression multiniveaux, modèle à 2 niveaux (sujets, items. Corrélation intra classe : 35.44%).
- ▷ 557 valeurs manquantes (13.26% de non-réponses et de valeurs filtrées (détail par condition : E.I – P<sub>0</sub> 13.9% ; P<sub>1</sub> 11% ; P<sub>2</sub> 16.7% ; P<sub>3</sub> 12.9% ; P<sub>4</sub> 9% / E.nI – P<sub>0</sub> 15.3% ; P<sub>1</sub> 11% ; P<sub>2</sub> 9.5% ; P<sub>3</sub> 12.9% ; P<sub>4</sub> 11.4% / nE.I – P<sub>0</sub> 16.1% ; P<sub>1</sub> 18.6% ; P<sub>2</sub> 16.2% ; P<sub>3</sub> 10.5% ; P<sub>4</sub> 12.4% / nE.nI – P<sub>0</sub> 10.9% ; P<sub>1</sub> 21.4% ; P<sub>2</sub> 11.4% ; P<sub>3</sub> 12.4% ; P<sub>4</sub> 11.9%)).

#### 4.2.7.2.9 Paramètres contrôlés

##### À la construction du matériel et à la structuration de l'expérience

- ▷ Ensemble des facteurs contrôlés lors de l'élaboration du matériel pour l'Expérience 4 (Section 4.2.2.2.9, page 237).
- ▷ L'intégration des cibles dans les stimuli est automatisée par un script dont la procédure a été testée et validée, pour éviter les erreurs et imprécisions de localisation (cibles toujours à la même position, fixe (P<sub>0</sub>, P<sub>1</sub>) ou relative (P<sub>2</sub> à P<sub>4</sub>)).
- ▷ Les sonorités cible et distractrice sont choisies assez similaires (même durée, proximité acoustique) pour obliger les auditeurs à être très attentifs s'ils veulent discriminer les sonorités cibles des sonorités distractrices, ce afin de les forcer à traiter aussi le mot, et donc afin d'éviter les réponses issues d'un traitement stratégique ou sensoriel (environ soixante participants n'ont pas été retenus pour les analyses car ils ont été incapables d'effectuer correctement la tâche. Beaucoup de participants se sont plaints de sa difficulté).
- ▷ Les auditeurs appuient sur la touche avec leur index dominant.
- ▷ Répartition équilibrée des conditions dans l'ensemble des groupes (contre balancement).
- ▷ Présentation aléatoire des items lors de la passation.

Lors des analyses statistiques

Les effets des variables étudiées sont difficiles à mettre en évidence. Afin d'éviter un écrasement des résultats lié à une quantité trop importante de facteurs dans le modèle, les covariables qui étaient sans incidence significative sur les temps de détection des cibles n'ont pas été intégrées. La non significativité de ces facteurs a cependant été vérifiée au préalable (liste ci-dessous) :

- ▷ Latéralité manuelle des participants
- ▷ Sexe des participants
- ▷ Nombre de phonèmes composant le mot stimulus
- ▷ Position phonétique du point d'unicité phonologique du mot stimulus
- ▷ Position acoustique du début du point d'unicité phonologique du mot stimulus
- ▷ Position acoustique du point de reconnaissance du mot stimulus
- ▷ Rapport signal / bruit (indice de détectabilité de la cible par rapport au signal acoustique environnant) avec le signal avant la cible
- ▷ Rapport signal / bruit avec le signal après la cible
- ▷ Fréquence formelle phonologique pour le mot stimulus (logarithme)
- ▷ Âge d'acquisition estimé (moyenne par forme phonologique)
- ▷ Familiarité conceptuelle (moyenne par forme phonologique)

Les résultats de ces analyses préalables sont donnés dans le Tableau 69 ci-dessous.

Tableau 69 : Tests de significativité des covariables

	<b>Chi<sup>2</sup><sub>(1, N=3643)</sub></b>	<b>p&lt;</b>
<b>latéralité</b>	1.864	.18
<b>sexe</b>	0.452	.51
<b>Nombre de phonèmes</b>	0.372	.55
<b>PU</b>	2.226	.14
<b>Début du PU</b>	2.218	.14
<b>PR</b>	1.206	.28
<b>Rapport signal/bruit av</b>	3.25	.072
<b>Rapport signal/bruit ap</b>	0.375	.55
<b>Frq formelle phonol</b>	0.029	.87
<b>AoA</b>	1.247	.27
<b>Fam</b>	1.712	.20

Avec : av = avant (signal avant la sonde) ; ap = après (signal après la sonde)

Le facteur liste n'est pas intégré au modèle pour les raisons exposées dans la note de bas de page n°105, page 201. Par contre, les 3 covariables suivantes ont été contrôlées et intégrées au modèle final :

- ▷ L'ordre de présentation des items
- ▷ La position acoustique des cibles (en ms.)
- ▷ La durée des stimuli (ms.)

Les deux dernières covariables ont été intégrées au modèle malgré leur non significativité, à cause de leur importance théorique et technique :

- ▷ La position acoustique des cibles doit être renseignée dans le modèle car le facteur position qui est utilisé comme V.I. est de nature catégorielle. Il s'agit donc d'un complément d'information pour contrôler les variations de localisation temporelle des cibles à une position donnée due à leur valeur relative à la durée des items ( $P_2, P_3, P_4$ ).
- ▷ La durée des stimuli corrige un éventuel biais dans les temps de réponse, lié au fait que plus un mot est long, plus l'auditeur, s'il n'a pas encore entendu de cible ou de sonorité distractive, est prêt à répondre. Il contribue aussi à la correction des variations de position des cibles dans les items, liée à la variation de durée des items.

### 4.2.7.3 Résultats obtenus

#### 4.2.7.3.1 Rapport des analyses statistiques

L'équation du modèle (Équation 46) et les tests de significativité relatifs aux covariables peuvent être consultés à l'Annexe 24 page 519.

Tableau 70 : Tests de significativité pour les facteurs expérimentaux et leurs interactions

	Coefficient	Chi <sup>2</sup> <sub>(1, N=3643)</sub>	p<
<b>Test</b>	6.727	0.377	.54
<b>P<sub>2</sub></b>	-5.47	0.159	.70
<b>P<sub>3</sub></b>	-2.349	0.014	.91
<b>P<sub>4</sub></b>	-14.997	0.302	.59
<b>Ench</b>	8.934	0.624	.43
<b>Init</b>	-8.588	0.594	.45
<b>Test.Ench</b>	-13.129	0.719	.40
<b>Test.Init</b>	18.453	1.441	.23
<b>P<sub>2</sub>.Ench</b>	3.102	0.039	.85
<b>P<sub>2</sub>.Init</b>	1.496	0.009	.93
<b>P<sub>3</sub>.Ench</b>	1.916	0.014	.91
<b>P<sub>3</sub>.Init</b>	-19.299	1.496	.23
<b>P<sub>4</sub>.Ench</b>	-14.442	0.747	.39
<b>P<sub>4</sub>.Init</b>	-4.481	0.077	.79
<b>Ench.Init</b>	-5.054	0.102	.75
<b>Test.Ench.Init</b>	-10.833	0.251	.62
<b>P<sub>2</sub>.Ench.Init</b>	23.982	1.213	.28
<b>P<sub>3</sub>.Ench.Init</b>	18.885	0.726	.40
<b>P<sub>4</sub>.Ench.Init</b>	32.836	2.057	.16

Tableau 71 : Tests de significativité des contrastes entre conditions

		P <sub>0</sub>		P <sub>1</sub>		P <sub>2</sub>		P <sub>3</sub>		P <sub>4</sub>	
		Chi <sup>2</sup> ( <sub>1</sub> , N=3643)	p<	Chi <sup>2</sup> ( <sub>1</sub> , N=3643)	p<	Chi <sup>2</sup> ( <sub>1</sub> , N=3643)	p<	Chi <sup>2</sup> ( <sub>1</sub> , N=3643)	p<	Chi <sup>2</sup> ( <sub>1</sub> , N=3643)	p<
<b>nE.nI</b>	<b>nE.I</b>	0.594	.45	0.755	.39	1.111	.30	0.734	.40	0.228	.64
	<b>E.nI</b>	0.624	.43	0.127	.73	0.01	.92	0.043	.84	2.618	.11
	<b>E.I</b>	0.157	.70	0.698	.41	2.574	.11	0.567	.46	0.089	.77
<b>nE.I</b>	<b>E.nI</b>	1.904	.17	1.218	.27	1.143	.29	0.356	.56	3.24	<b>.072</b>
	<b>E.I</b>	0.083	.78	0.046	.83	0.335	.57	0.003	.96	0.013	.91
<b>E.nI</b>	<b>E.I</b>	1.541	.22	0.299	.59	3.106	<b>.078</b>	0.358	.55	4.255	<b>.039</b>

Ces analyses ne révèlent ni effet principal des facteurs, ni interactions (Tableau 70). Toutefois, l'analyse des contrastes entre types de mots par position révèle plusieurs tendances et écarts significatifs dans les temps de détection des cibles (Tableau 71).

### 4.2.7.3.2 Présentation des résultats

Les graphiques présentés pour cette expérience et l'expérience suivante (Expérience 10) ne comportent pas de barres de significativité statistique pour des raisons de lisibilité. L'information de significativité est disponible dans les tableaux présentant les résultats des analyses pour les facteurs et les contrastes.

Figure 62 : Présentation des temps moyens estimés de détection des cibles par condition

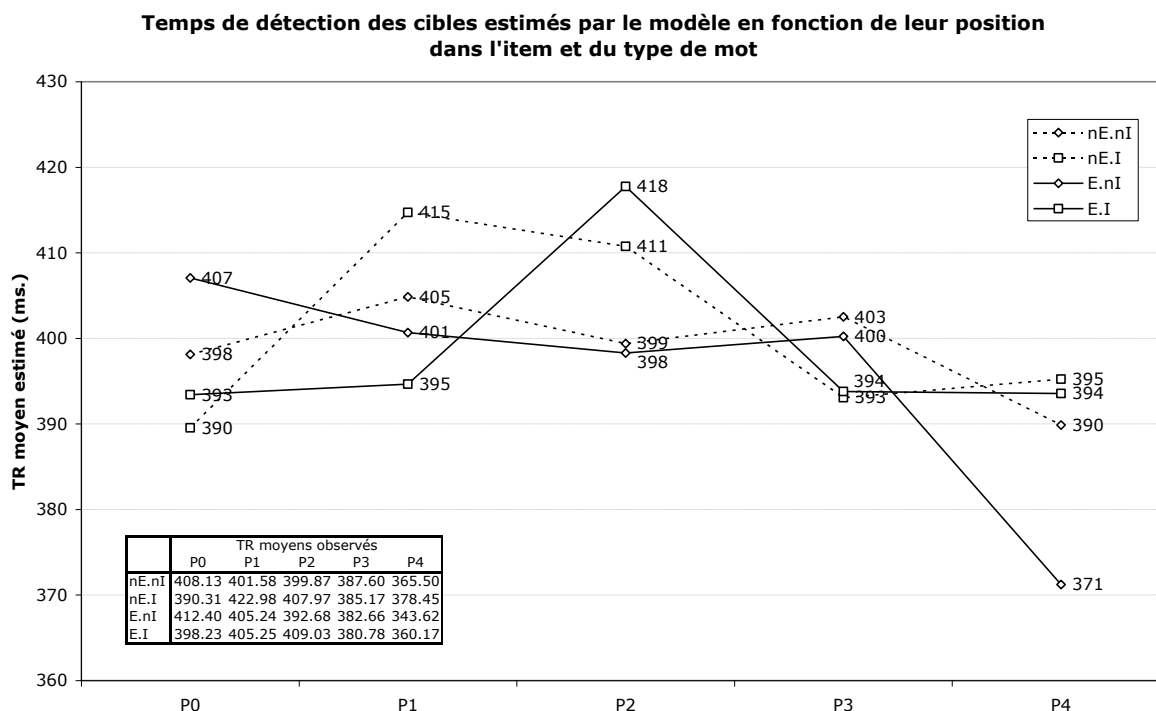
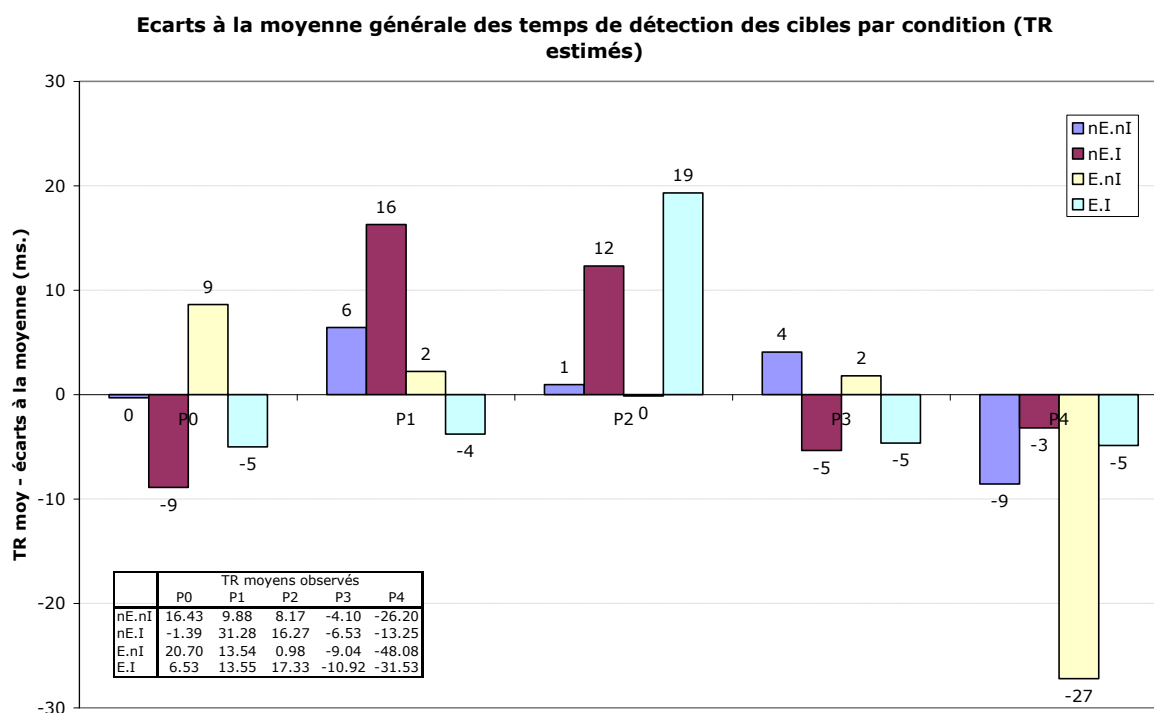


Figure 63 : Présentation des écarts à la moyenne générale, pour les temps moyens estimés de détection des cibles par condition



Ces deux graphiques montrent :

- ▷ qu'en début de mot (positions P<sub>1</sub> et P<sub>2</sub>), les cibles paraissent être plus difficiles à détecter (par rapport à la ligne de base P<sub>0</sub>) lorsqu'elles sont insérées dans un mot enchâssable (nE.I (P<sub>1</sub>), nE.I et E.I (P<sub>2</sub>)).
- ▷ que la condition E.I, pour laquelle la quantité de superpositions lexicales est la plus grande, semble aussi être celle pour laquelle les traitements diffèrent le plus (augmentation des temps de détection de la cible en P<sub>2</sub>), puisqu'une tendance à la significativité apparaît dans le test des contrastes (Tableau 71, page 280) pour cette position, entre E.I et E.nI ( $p < .08$ ).
- ▷ que les temps de détection des cibles semblent revenir à la ligne de base en P<sub>3</sub> et P<sub>4</sub>, pour toutes les conditions sauf pour E.nI, qui présente des temps de traitement beaucoup plus courts que les 3 autres conditions en P<sub>4</sub>. Le test des contrastes (Tableau 71, page 280) révèle alors un écart significatif entre E.nI et E.I ( $p < .04$ ) et une tendance entre E.nI et nE.I ( $p < .08$ ), liés au traitement de la condition E.nI.

Toutefois, il est difficile de réfléchir sur les écarts entre conditions parce que les effets des facteurs manipulés ne sont pas distingués. Les résultats présentés dans la Figure 64 et les Tableau 72 et Tableau 73 ci-après fournissent l'information relative à l'influence distincte de chaque facteur (E et I) et de leur interaction (E.I) pour chaque position de sondage.

Figure 64 : Influence de chacun des 2 facteurs expérimentaux sur les temps de détection des cibles, en fonction de la position de celles-ci dans l'item

Interférence moyenne des facteurs E, I, et de leur interaction, sur les temps de détection des cibles, pour chaque position

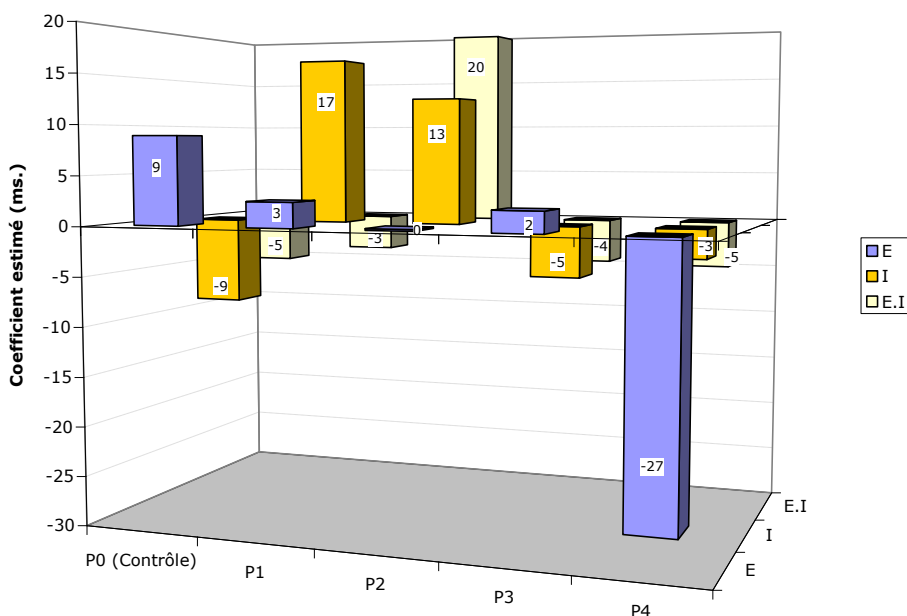


Tableau 72 : Tests de significativité des contrastes entre position contrôle et positions tests, pour chaque facteur et pour l'interaction

	E		I		E.I	
	Chi <sup>2</sup> <sub>(1, N=3643)</sub>	p<	Chi <sup>2</sup> <sub>(1, N=3643)</sub>	p<	Chi <sup>2</sup> <sub>(1, N=3643)</sub>	p<
P <sub>0</sub> /P <sub>1</sub>	0.336	.57	5.489	<b>.02</b>	0.409	.53
P <sub>0</sub> /P <sub>2</sub>	0.331	.57	2.826	<b>.093</b>	2.056	<b>.16</b>
P <sub>0</sub> /P <sub>3</sub>	0.084	.78	0.042	.84	0.126	.73
P <sub>0</sub> /P <sub>4</sub>	1.159	<b>.29</b>	0.062	.81	0.073	.79

Tableau 73 : Tests de significativité des contrastes entre facteurs et entre facteurs et leur interaction, pour chaque position

	E / I		E / E.I		I / E.I	
	Chi <sup>2</sup> <sub>(1, N=3643)</sub>	p<	Chi <sup>2</sup> <sub>(1, N=3643)</sub>	p<	Chi <sup>2</sup> <sub>(1, N=3643)</sub>	p<
P <sub>0</sub>	1.904	.17	1.541	.22	0.083	.78
P <sub>1</sub>	1.218	.27	0.299	.59	2.249	.14
P <sub>2</sub>	1.143	.29	3.106	<b>.079</b>	0.335	.57
P <sub>3</sub>	0.356	.56	0.358	.55	0.003	.96
P <sub>4</sub>	3.24	<b>.072</b>	4.255	<b>.04</b>	0.013	.91

Cette analyse permet de répondre plus aisément aux questions posées dans les hypothèses de l'expérience.

- ▷ Hypothèse E (effet d'une superposition lexicale initiale réelle sur la reconnaissance du mot stimulus, étant donné nI) :

La première hypothèse prédit des temps de détection supérieurs en positions  $P_2$  et  $P_3$  pour les mots renfermant un enchâssement initial, par rapport aux mots ne renfermant pas d'enchâssement initial, et pas de différence en position  $P_4$ . Cette hypothèse est invalidée par les données pour les positions  $P_2$  et  $P_3$ , où l'effet prédit n'apparaît pas (Tableau 72, page 282). En position  $P_4$ , un phénomène peut être remarqué (même tableau : seuil de significativité plus fort bien que non significatif, case grisée). Il est impliqué dans un phénomène tendanciel qui apparaît dans le test de significativité du contraste entre E et I pour la position  $P_4$ , et par la significativité du contraste entre E et E.I pour la position  $P_4$  (Tableau 73, page 282).

- ▷ Hypothèse I (effet d'une superposition lexicale initiale virtuelle sur la reconnaissance du mot stimulus, étant donné nE) :

La deuxième hypothèse prédit des temps de détection supérieurs en  $P_2$ ,  $P_3$  et  $P_4$  pour les mots enchâssables à l'initiale d'un mot plus long, par rapport aux mots non enchâssables en initiale. Elle est partiellement vérifiée : l'effet attendu aux positions  $P_2$  et  $P_3$  se manifeste plus précocement (positions  $P_1$  et  $P_2$ ). Par contre, le conflit lexical prévu semble être résolu assez rapidement, puisqu'on observe un retour à la ligne de base dès la position  $P_3$ .

- ▷ Autre effet non prédit par les hypothèses :

On note une augmentation des temps de détection en position  $P_2$  dans le cas de l'interaction E \* I. Bien que non significative (le seuil de probabilité, à  $p < .16$ , est tout de même d'un ordre de grandeur inférieur aux seuils pour les autres positions (Tableau 72, page 282, case grisée)), cette interaction est impliquée dans un effet tendanciel lointain lors du test de contraste entre E et E \* I pour la position  $P_2$  (Tableau 73, page 282).

Il est logique d'observer un phénomène différent selon le facteur E ou I, puisque la superposition lexicale est alors de nature différente (respectivement réelle ou virtuelle) :

- ▷ Superposition lexicale initiale réelle :

La présence d'un enchâssement initial n'induirait pas de difficulté particulière dans le traitement, au moment où l'activation simultanée des représentations du mot enchâssé et du mot porteur est supposée avoir lieu (les temps moyens estimés de détection des cibles varient peu autour de la ligne de base pour les 3 premières positions test). Contrairement à la situation de superposition lexicale initiale virtuelle, moins de représentations lexicales seraient activées. Peut-être même que l'activation de la représentation du mot enchâssé et de celle du mot porteur a lieu séquentiellement, et non simultanément, condition nécessaire pour rendre le traitement plus difficile. En  $P_4$ , un nouveau phénomène apparaît, non prévu par la théorie. Il se manifeste par une aptitude plus importante du système pour détecter les cibles. Elle pourrait s'expliquer par la libération soudaine, en fin de mot, des ressources mobilisées pour résoudre une difficulté qui finalement ne se produit pas ou est soudainement solutionnée.

- ▷ Superposition lexicale initiale virtuelle :

La possibilité qu'un mot soit le début d'un autre mot induirait une augmentation des temps de détection des cibles insérées en début de mot ( $P_1$  et  $P_2$ ). Ce phénomène pourrait être dû au fait d'une plus grande difficulté du traitement que dans le cas d'une superposition lexicale réelle, car :

- ∩ La quantité de représentations lexicales activées (de mots pouvant commencer par la séquence stimulus de phonèmes) serait plus importante



que la quantité de représentations lexicales activées dans le cas d'une superposition lexicale réelle (seulement le mot porteur et le mot enchâssé)

- ∞ L'activité serait essentiellement au niveau lexical, donc moins éloignée du seuil de conscience au-delà duquel la tâche est effectuée, et plus à même d'interférer avec la réalisation de celle-ci.

L'influence des représentations lexicales activées serait seulement mesurable avec les cibles  $P_1$  et  $P_2$ , donc probablement pour une information acoustique / lexicale plus avancée que celle dévoilée au moment où les cibles sont présentées (correspondant peut-être à l'information entendue jusqu'en  $P_3$ ). Au-delà, l'information acoustique lève l'ambiguïté lexicale avant que le mot ne se termine et désactive les candidats inadéquats (la prosodie indique qu'il s'agit d'un mot monosyllabique, et la coarticulation qu'il n'y a plus de phonèmes, donc que le mot n'est pas le début d'un autre). On assiste alors à un retour à la ligne de base (dès  $P_3$  - information entendue jusqu'en  $P_4$  ?). L'ambiguïté lexicale qui était prédite jusqu'au point d'unicité (en position  $P_4$ ) se résout donc bien avant, grâce à l'intégration d'autres sources d'information (au P.R. ?).

- ▷ Interaction entre superpositions lexicales initiales réelle et virtuelle (E.I) :

L'interaction en position  $P_2$ , au début du traitement de l'item, responsable d'une plus grande difficulté pour détecter les cibles, n'est pas présente avant, puis disparaît ensuite. Elle survient à un moment précoce du traitement, où les représentations mentales des mots impliqués dans les superpositions réelles et virtuelles s'activent probablement. Le fait qu'il y ait un pic de difficulté sous l'influence combinée des 2 facteurs, qui disparaît ensuite, appuie l'idée selon laquelle la représentation mentale du mot enchâssé et celle du mot porteur, dans le cas des superpositions lexicales initiales réelles seulement, sont activées l'une après l'autre avec un bref moment de coexistence (les 2 représentations entrent alors en conflit).

#### 4.2.7.4 Conclusions de l'Expérience 9

Des 2 hypothèses de l'Expérience 9, la 1<sup>ère</sup> est invalidée par les données, tandis que la 2<sup>ème</sup> est partiellement invalidée. Les effets sont difficiles à mettre en évidence avec la technique de détection de cibles sonores non linguistiques qui a été utilisée. Au moins quatre raisons peuvent expliquer ce fait :

- ▷ Similitude importante entre la sonorité cible et la sonorité distractive :  
certains participants se sont plaints de la difficulté de la tâche (n'arrivaient pas à distinguer les deux sonorités), responsable d'une distraction qui a pu faire disparaître les effets (une autre technique avait été considérée<sup>113</sup>. Elle a dû être écartée à cause d'autres inconvénients).
- ▷ Sensibilité de la tâche :  
le fait que les covariables habituellement contrôlées (qui ont des effets fortement significatifs dans les expériences sur la perception des mots) soient ici en majorité non significatives, semble confirmer le fait :
  - ∩ que la tâche de détection des cibles s'effectue selon un traitement non linguistique, ou
  - ∩ qu'un cloisonnement, assez fort pour limiter l'impact sur la mesure d'une interférence entre processus, existe dans le système entre ces différents types de traitement (tel que supposé en introduction de l'Expérience 9 – Section 4.2.7, page 272)

Si le système de traitement est fortement modulaire, les processus inconscients de la reconnaissance des mots et les processus conscients et inconscients pour la détection de la cible peuvent s'effectuer en parallèle sans interférer ou en interférant très peu. Cutler et collègues (1993) parviennent à une conclusion de ce type par rapport à la technique de détection de clics. Elle ne sonderait pas les traitements linguistiques, mais les traitements sensoriels de bas niveau (mais voir discussion plus loin).
- ▷ Position du sondage :
  - ∩ la localisation des cibles est choisie pour faire un échantillonnage assez fréquent (4 fois sur un gabarit monosyllabique) de l'évolution du traitement sur le continuum temporel. Il se peut que cet échantillonnage soit trop fréquent, que le modèle d'analyse renferme trop de mesures voisines, et que cela tende à écraser les effets.

---

<sup>113</sup> Détection de pauses artificielles dans le signal. Technique développée par Mattys et Herslow (2002). Cette méthode utilise des pauses au lieu de cibles acoustiques, à détecter selon une tâche identique. Les discussions avec le Dr. Mattys n'étaient pas suffisamment convaincantes quant aux avantages de la tâche de détection de pauses sur ceux de la tâche de détection de clics, par rapport au but poursuivi dans l'expérience (sonder le traitement naturel d'un mot).

Les pauses ayant une durée comprise entre 100 et 150 ms., elles interrompent l'entrée d'information auditive dans le système pendant une durée de l'ordre de grandeur d'un phonème. Par conséquent, le stimulus lexical est altéré, et le traitement ne peut plus suivre son cours naturellement (modifié ou momentanément interrompu). La technique des cibles acoustiques, au contraire, ne produit pas ces problèmes. Les cibles durent 10 ms., si bien qu'elles apparaissent comme un bruit superposé au mot. Le mot lui-même n'est pas modifié et son traitement naturel est garanti. De plus, dans le cas de mots monosyllabiques dont les plus courts durent 194 ms. (« pont »), l'insertion d'une pause de 100 ms. revient à altérer une proportion trop importante du stimulus.

- ∩ La position des cibles est relative à la durée du mot. Cela cause une variabilité de la position des cibles par rapport aux étapes psychologiques du traitement, puisque les cibles, en fonction des mots (items) et du type de mot (condition), peuvent aléatoirement tomber avant ou après qu'une étape particulière n'ait eu lieu (par exemple, en  $P_2$ , la cible peut être insérée dans un cluster, avant le noyau vocalique. Dans ce cas la représentation du mot enchâssé n'a pas encore pu être activée en mémoire, puisque l'information acoustique nécessaire n'est pas suffisante. La cible peut aussi tomber dans le noyau vocalique, auquel cas la représentation du mot enchâssé a pu être activée). La conséquence est que des effets ont pu être perdus à cause de cette variabilité, qui marque surtout les positions  $P_2$  et  $P_3$  (les conclusions liées aux positions  $P_0$ ,  $P_1$  et  $P_4$  sont plus fiables parce que ce problème n'a pas lieu).

Cependant, la technique d'exploration comportementale introduite dans cette expérience s'avère intéressante et prometteuse. Bien que des éléments indiquent que le niveau de traitement mesuré est peut-être sensoriel ou de bas niveau (absence d'effet des facteurs linguistiques dans cette étude et conclusions de Cutler et collègues (Cutler et al., 1993)), cette méthode a permis de tracer de manière subtile et précise le décours temporel du traitement de mots parlés monosyllabiques. Si les écarts observés entre types de mots ne sont pas liés au traitement linguistique de ces mots mais à leur traitement sensoriel, alors ils révèlent des différences acoustiques passées inaperçues lors de l'analyse des stimuli. Si de telles différences existent, cela implique qu'elles ont été produites lors de l'encodage par le système locuteur, en lien avec le traitement des caractéristiques manipulées des items. Ce qui signifie que si la technique de détection de clics ne mesure pas l'effet des superpositions lexicales sur le traitement des mots par le système récepteur, elle mesure indirectement l'effet des superpositions lexicales sur le traitement des mots par le système locuteur. L'interprétation des résultats est donc ambiguë, et dans certains cas, cette technique peut être inadéquate. Dans la présente étude, pourtant, elle a permis de montrer des effets lexicaux liés à la présence de superpositions lexicales.

Un deuxième intérêt de cette technique est la simplicité de sa mise en œuvre (nul besoin d'une technologie complexe comme les E.R.P.). Tout laboratoire disposant de moyens limités peut l'utiliser, sous réserve des remarques précédentes.

Les résultats qu'elle permet d'obtenir sont intéressants malgré leur faible visibilité statistique – justifiée par un ensemble de conditions adverses.

Dans l'hypothèse où ils sont liés au traitement lexical des superpositions de mots par le système récepteur, ils révèlent que la recherche des représentations à travers le lexique est utilisée tôt au cours du traitement, et qu'elle prend fin bien avant le point d'unicité, probablement aux alentours du point de reconnaissance (intégration de

diverses sources d'information acoustique – facteur I). Ils montrent aussi qu'en cas de superposition lexicale initiale réelle (facteur E), il n'y a aucun conflit, contrairement à ce qui est prévu par la théorie de la compétition lexicale.

Le phénomène de facilitation qui est mesuré dans les Expériences 4 à 8 est un phénomène tardif, qui s'explique peut-être par la libération de ressources mobilisées en prévision d'une éventuelle difficulté de traitement, qui n'a finalement pas lieu ou est rapidement résolue. Ces données appuient donc l'idée que la présence d'une superposition lexicale initiale réelle a un effet facilitateur sur le traitement. Ces observations sont difficilement interprétables par l'hypothèse d'un traitement sensoriel des différences acoustiques encodées par le système locuteur sous l'influence du traitement lexical des superpositions de mots.

#### **4.2.8 Expérience 10 : Sondage de la charge de calcul dans le système de traitement, à différentes étapes de la reconnaissance d'un mot bisyllabique renfermant une superposition lexicale réelle, initiale et/ou finale, en modalité auditive**

Les résultats de l'Expérience 9 sont à la fois peu concluants (pas d'effets flagrants) et assez encourageants (observations ouvrant des pistes de réflexion, possibilité d'améliorer la technique). Malgré la difficulté de ce type d'expérience (technique en cours de développement, incertitude théorique sur l'information mesurée, préparation difficile des items, lourdeur de la passation, difficulté des analyses statistiques, etc.), l'étude a été retenue :

- ▷ sur un matériel de gabarit bisyllabique (matériel de l'Expérience 7), pour sonder séparément le traitement de l'item (niveau lexical) et le traitement de ses composants syllabiques (niveaux infra-lexical et lexical).
- ▷ avec une technique mieux ajustée :
  - ∩ différence acoustique plus importante entre cible et sonorité distractive
  - ∩ localisation des sondes sur des sites pertinents par rapport aux étapes du traitement (au lieu d'un sondage mathématiquement déterminé et psychologiquement aléatoire, sondage psychologiquement déterminé et physiquement contrôlé lorsque possible).

##### **4.2.8.1 Aspects théoriques**

Les hypothèses théoriques proposées dans les Expériences 6 et 7 pour les facteurs lexicaux sont réutilisées, ainsi que le type d'hypothèse relatif au moment du sondage, proposé dans l'Expérience 9. Ces hypothèses conduisent à supposer qu'une cible est plus difficile à détecter lorsqu'elle est insérée au moment où se produit un enchaînement lexical, à cause du phénomène de compétition entre représentations lexicales, qui sont parallèlement activées à ce moment précis. Ceci est vrai pour un enchaînement survenant au début comme à la fin du mot porteur. Cependant, TRACE suppose que l'effet d'un enchaînement initial est plus important sur le traitement du mot porteur, et que l'effet de compétition dû à un enchaînement final est quasiment inexistant. L'interférence du traitement lexical sur la détection de la cible devrait donc être plus importante lorsque la sonde est au début du mot stimulus que lorsqu'elle est à la fin, dans le cas où des superpositions réelles ont lieu à ces positions.

## 4.2.8.2 Aspects techniques

### 4.2.8.2.1 Facteurs expérimentaux et conditions

#### 4.2.8.2.1.1 Deux facteurs lexicaux

À 2 modalités chacun ; croisement complet (4 types de mots - Tableau 74 ci-dessous).

Tableau 74 : Facteurs lexicaux et types de mots

		Lexicalité de la 1 <sup>ère</sup> syllabe	
		Oui (M <sub>1</sub> )	Non (nM <sub>1</sub> )
Lexicalité de la 2 <sup>ème</sup> syllabe	Oui (M <sub>2</sub> )	M <sub>1</sub> .M <sub>2</sub> « <i>Timbale</i> » (/tɛ̃/-/bal/)	nM <sub>1</sub> .M <sub>2</sub> « <i>Climat</i> » (S <sub>1</sub> -/ma/)
	Non (nM <sub>2</sub> )	M <sub>1</sub> .nM <sub>2</sub> « <i>Résine</i> » (/ʁe/-S <sub>2</sub> )	nM <sub>1</sub> .nM <sub>2</sub> « <i>Colonne</i> » (S <sub>1</sub> -S <sub>2</sub> )

#### 4.2.8.2.1.2 Deux facteurs techniques

##### Position de la cible dans l'item

- ▷ Positions « test » : à 3 modalités
  - ∩ P<sub>1</sub> = à la fin de la première syllabe<sup>114</sup>
  - ∩ P<sub>2</sub> = dans la deuxième syllabe et avant le point de reconnaissance du mot stimulus :
    - syllabes se terminant par le noyau vocalique : vers la fin de l'attaque de la deuxième syllabe (premier phonème pour une syllabe de structure CV, deuxième consonne du cluster pour une syllabe CCV), ou au début du noyau vocalique (dans la demi voyelle pour une syllabe CYV)
    - syllabes se terminant par une consonne : vers la fin du noyau vocalique
  - ∩ P<sub>3</sub> = au point de reconnaissance du mot stimulus
- ▷ Positions « contrôle » : à 3 modalités
  - ∩ Positions moyennes P<sub>1</sub>, P<sub>2</sub>, P<sub>3</sub> (respectivement calculées à partir de l'ensemble des positions P<sub>1</sub>, P<sub>2</sub> ou P<sub>3</sub>, des cibles insérées dans les mots expérimentaux). Ces positions moyennes sont insérées dans un item vide.
- ▷ Position « distractrice anti-stratégique » :
  - ∩ P<sub>4</sub> = 200 ms. après la fin du mot

##### Type de sonorité insérée dans l'item

Facteur lié à la tâche, à 2 modalités :

---

<sup>114</sup> Remarque : les positions sont localisées en tenant compte de la coarticulation. Si le dernier phonème de la première syllabe est coarticulé avec le premier phonème de la syllabe suivante, la position P<sub>1</sub> sera avancée de manière à ce que, jusqu'à ce point, l'auditeur n'ait pas d'information consciemment accessible sur la suite du mot (position définie à l'écoute du stimulus graduellement dévoilé par auto présentation). Il en est de même pour localiser les positions P<sub>2</sub> et P<sub>3</sub>.

- ▷ cible : « clic » (bruit blanc, durée 10 ms.)
- ▷ sonorité distractive : « bip » (son périodique, durée 10 ms.)

Les facteurs lexicaux déterminent 4 types de mots expérimentaux ( $M_1.M_2$ ,  $M_1.nM_2$ ,  $nM_1.M_2$ ,  $nM_1.nM_2$ ). Les facteurs techniques déterminent 3 situations expérimentales ( $P_1.clic$ ,  $P_2.clic$ ,  $P_3.clic$ ). L'expérience comporte donc  $4 * 3 = 12$  conditions expérimentales (et 12 conditions distractrices, 1 condition anti-stratégique, 3 conditions contrôle).

#### 4.2.8.2.2 Hypothèses opérationnelles

- ▷ Hypothèse  $M_1$  : effet d'une superposition lexicale réelle au début d'un mot - 1 sonde

Les temps de détection d'une cible insérée en position  $P_1$  sont plus longs lorsque la première syllabe est un mot que ceux obtenus pour la détection d'une cible insérée au même endroit, dans un mot bisyllabique dont la première syllabe n'est pas un mot.

- ▷ Hypothèse  $M_2$  : effet d'une superposition lexicale réelle à la fin d'un mot - 2 sondes

∩ Sonde  $P_2$  - mot porteur non reconnu :

Si le mot porteur n'a pas encore été reconnu, les temps de détection d'une cible acoustique non linguistique insérée dans la deuxième syllabe d'un mot bisyllabique sont plus longs lorsque la deuxième syllabe est un mot que lorsqu'elle n'est pas un mot.

∩ Sonde  $P_3$  - mot porteur reconnu :

Si le mot porteur a été reconnu, les temps de détection d'une cible acoustique non linguistique insérée dans la deuxième syllabe d'un mot bisyllabique sont équivalents, que la deuxième syllabe soit un mot ou pas.

- ▷ Hypothèse  $M_1 / M_2$  : différence d'effet entre une superposition lexicale réelle au début et à la fin d'un mot

Les temps de détection de la sonde insérée en position  $P_1$  sont plus longs lorsque la première syllabe du mot porteur est un mot, que les temps de détection de la sonde lorsqu'elle est insérée en position  $P_2$  dans une deuxième syllabe lexicale.

#### 4.2.8.2.3 Matériel expérimental utilisé

Il s'agit des stimuli employés dans l'Expérience 7 (stimuli en voix naturelle. Mots expérimentaux bisyllabiques ; Distracteurs : pseudo-mots bisyllabiques (mots devenant pseudo-mot par le dernier phonème). La liste des items est donnée en Annexe 19 page 508). Les fichiers son de l'Expérience 7 ont été utilisés pour pouvoir faire une comparaison directe entre les résultats des Expériences 7 et 10.

Les stimuli ont subi un traitement numérique pour l'insertion des sonorités cible et distractive aux positions définies (décrit à la Section 3.4.1.3.2 page 144).

- ▷ Items expérimentaux :

- ∩ (20 mots expérimentaux) \* (4 types de mots) \* (3 positions de la sonorité cible) = 240
- ∩ 3 items vides renfermant chacun la sonorité cible à une position moyenne (P<sub>1</sub> moyenne, P<sub>2</sub> moyenne, P<sub>3</sub> moyenne)<sup>115</sup>
- ▷ Items distracteurs :
  - ∩ (20 mots expérimentaux) \* (4 types de mots) \* (3 positions de la sonorité distractrice) = 240
  - ∩ (20 pseudo-mots) \* (4 types de mots) \* (3 positions) \* (2 types de sonorité) = 480
  - ∩ 3 items vides renfermant chacun la sonorité distractrice à une position moyenne<sup>115</sup>
  - ∩ (12 mots P<sub>4</sub>) \* (4 types de mots) \* (2 types de sonorité) = 96
  - ∩ (12 pseudo-mots P<sub>4</sub>) \* (4 types de pseudo-mots) \* (2 types de sonorité) = 96

L'expérience comporte donc un total de 1158 stimuli.

#### 4.2.8.2.4 Organisation de l'expérience

Chaque participant est testé avec tous les mots expérimentaux (N = 80), chacun lui étant présenté une seule fois. Il est testé avec toutes les situations, au nombre de 70 (4 (structure d'item) \* 2 (lexicalité de l'item) \* 4 (position de cible) \* 2 (type de sonorité) + 6 (contrôle) = 70).

Cependant, un participant n'est pas testé avec tous les items : lorsqu'il est testé avec un mot, c'est pour l'une des 8 situations (P<sub>7</sub> \* Sonorité). Deux de ces 8 situations ont un statut particulier (les situations P<sub>4</sub> - distraction anti-stratégique) : elles ont été introduites dans l'expérience afin d'éviter que les auditeurs ne répondent stratégiquement sur la position P<sub>3</sub>. Les mesures en P<sub>4</sub> ne sont donc pas utiles ni analysées (de plus, le recueil des données en P<sub>4</sub> ne concerne que 12 des 20 mots expérimentaux par type de mot, et la proportion d'items P<sub>4</sub> dans les listes est de 1/3 par rapport aux items de l'une des 3 autres positions (plan non équilibré)). Pour que tous les mots soient testés dans toutes les situations, 6 groupes de participants (et non 8) sont utilisés.

Chaque participant a donc été testé avec 208 stimuli (1 fois chacun) :

- ▷ 80 mots expérimentaux [(3 items \* 3 positions \* 4 types de mots + 4 distracteur P<sub>4</sub>) \* 2 types de sonorités]
- ▷ 48 contrôle (8 items \* 3 positions \* 2 types de sonorité)

---

<sup>115</sup> Calcul à partir des valeurs P<sub>1</sub>, P<sub>2</sub>, P<sub>3</sub> sur tous les mots expérimentaux.



- ▷ 80 pseudo-mots [(3 items \* 3 positions \* 4 types de mots + 4 distracteur P<sub>4</sub>) \* 2 types de sonorités]

Seules 60 réponses sont utiles :

- ▷ 36 mesures expérimentales (3 items \* 3 positions \* 4 types de mots)
- ▷ 24 mesures contrôles (8 items \* 3 positions)

Le recueil est équilibré sur l'ensemble des listes.

#### 4.2.8.2.5 Population testée

Les données de 81 participants (N = 13 - Groupe 1 ; N = 14 - Groupes 2 et 3 ; N = 15 - Groupe 4 ; N = 13 - Groupe 5 ; N = 12 - Groupe 6), francophones monolingues, sans troubles auditifs, cognitifs et moteurs, étudiant à l'université de Genève, ont été utilisées pour les analyses. Le recueil a été effectué en janvier 2004.

#### 4.2.8.2.6 Tâche

Tâche de détection de cible. Modalité auditive.

#### 4.2.8.2.7 Procédure de recueil des données

Expérience réalisée à l'université de Genève. Détails dans la Partie III, Chapitre 4.

60 \* 81 = 4860 mesures effectuées.

La phase test de l'expérience se déroule sans pause, et dure une quinzaine de minutes.

#### 4.2.8.2.8 Méthode d'analyse

- ▷ Détail des procédures pour le tri des données et les analyses dans la Partie III, Chapitre 3.
- ▷ Analyse de régression multiniveaux, modèle à 2 niveaux (sujets, items. Corrélation intra classe : 27.15%).
- ▷ 84 valeurs manquantes au total (condition contrôle compris : 1.73% de non-réponses et de valeurs filtrées<sup>116</sup>), dont 36 valeurs manquantes pour les positions test (1.23% de non-réponses et de valeurs filtrées (détail par condition : M<sub>1</sub>.M<sub>2</sub> - P<sub>1</sub> 2.5% ; P<sub>2</sub> 1.2% ; P<sub>3</sub> 1.6% / M<sub>1</sub>.nM<sub>2</sub> - P<sub>1</sub> 0% ; P<sub>2</sub> 1.2% ; P<sub>3</sub> 0% / nM<sub>1</sub>.M<sub>2</sub> - P<sub>1</sub> 0.8% ; P<sub>2</sub> 1.6% ; P<sub>3</sub> 0.4% / nM<sub>1</sub>.nM<sub>2</sub> - P<sub>1</sub> 2.1% ; P<sub>2</sub> 1.6% ; P<sub>3</sub> 1.6%)).

---

<sup>116</sup> L'Expérience 9 comportait 13.26% de valeurs manquantes. Le très faible score observé ici indique probablement une meilleure qualité des mesures dans l'Expérience 10, peut-être liée au gabarit bisyllabique des items, peut-être liée à une audibilité meilleure de la différence acoustique entre sonorités cible et distractrice.

### 4.2.8.2.9 Paramètres contrôlés

#### À la construction du matériel et à la structuration de l'expérience

- ▷ Facteurs contrôlés lors de la création du matériel pour l'Expérience 7 (Sections 4.2.4.2.9, page 251, et 4.2.5.2.9, page 260).
- ▷ Type de sonorité (cible / distractrice), de manière à ce qu'elles soient suffisamment voisines pour obliger les auditeurs à écouter attentivement le signal (et donc traiter le mot), et de manière à ce qu'elles soient suffisamment distinctes pour ne pas handicaper l'apparition d'effets.
- ▷ Localisation manuelle des sondes sur des critères psychologiques (relatifs aux étapes du traitement du mot) : fin de la 1<sup>ère</sup> syllabe, 2<sup>ème</sup> syllabe avant la reconnaissance du mot porteur, 2<sup>ème</sup> syllabe à la reconnaissance du mot porteur.
- ▷ Localisation prenant en compte la coarticulation entre phonèmes afin de limiter le parasitage des traitements par l'information droite.
- ▷ Présentation aléatoire des items lors de la passation

#### Lors des analyses statistiques

- ▷ L'effet de liste n'a pas été directement contrôlé (se reporter à la note de bas de page n°105, page 201)
- ▷ Latéralité et sexe des participants
- ▷ Intégration au modèle de la position acoustique des cibles (en ms.) et de la durée des mots (en ms. – se reporter à l'explication fournie pour l'Expérience 9 dans la Section 4.2.7.2.9, page 277)
- ▷ Les facteurs linguistiques habituellement contrôlés (nombre de phonèmes, PU, PR, début PU, fréquence formelle phonologique, AOA, familiarité) n'ont pas été ajoutés au modèle, suite à l'absence d'effet observée pour ces covariables dans l'analyse de l'Expérience 9.
- ▷ La détectabilité acoustique des cibles n'a pas été analysée, étant donné que ce facteur s'était avéré non significatif dans l'Expérience 9.

### 4.2.8.3 Résultats obtenus

#### 4.2.8.3.1 Rapport des analyses statistiques

L'équation du modèle (Équation 47) et les tests de significativité relatifs aux covariables peuvent être consultés à l'Annexe 25 page 520.

Tableau 75 : Tests de significativité pour les facteurs expérimentaux et leurs interactions

	Coefficient	Chi <sup>2</sup> <sub>(1, N=4776)</sub>	p<
<b>Test</b>	-31.302	15.517	<b>9<sup>E-05</sup></b>
<b>M<sub>1</sub></b>	17.737	5.765	<b>.017</b>
<b>M<sub>2</sub></b>	-0.409	0.003	.96
<b>P<sub>2</sub></b>	1.452	0.041	.84
<b>P<sub>3</sub></b>	8.494	0.459	.50
<b>Test.P<sub>2</sub></b>	-9.251	0.608	.44
<b>Test.P<sub>3</sub></b>	-14.779	0.617	.44

	Coefficient	Chi <sup>2</sup> <sub>(1, N=4776)</sub>	p<
<b>M<sub>1</sub>.M<sub>2</sub></b>	-14.162	1.898	.17
<b>M<sub>1</sub>.P<sub>2</sub></b>	-23.199	5.076	<b>.025</b>
<b>M<sub>1</sub>.P<sub>3</sub></b>	-30.454	8.811	<b>.003</b>
<b>M<sub>2</sub>.P<sub>2</sub></b>	-2.557	0.062	.81
<b>M<sub>2</sub>.P<sub>3</sub></b>	-2.871	0.077	.79
<b>M<sub>1</sub>.M<sub>2</sub>.P<sub>2</sub></b>	19.462	1.802	.18
<b>M<sub>1</sub>.M<sub>2</sub>.P<sub>3</sub></b>	27.806	3.702	<b>.055</b>

Tableau 76 : Tests de significativité des contrastes entre conditions

		P <sub>1</sub> (fin S <sub>1</sub> )		P <sub>2</sub> (dans S <sub>2</sub> )		P <sub>3</sub> (P.R. mot)	
		χ <sup>2</sup> <sub>(1, N=4776)</sub>	p<	χ <sup>2</sup> <sub>(1, N=4776)</sub>	p<	χ <sup>2</sup> <sub>(1, N=4776)</sub>	p<
<b>M<sub>1</sub>.M<sub>2</sub></b>	<b>M<sub>1</sub>.nM<sub>2</sub></b>	4.005	<b>.046</b>	0.104	.75	2.022	.16
	<b>nM<sub>1</sub>.M<sub>2</sub></b>	0.235	.63	0	1	0.016	.90
	<b>nM<sub>1</sub>.nM<sub>2</sub></b>	0.188	.67	0.185	.67	0.101	.76
<b>M<sub>1</sub>.nM<sub>2</sub></b>	<b>nM<sub>1</sub>.M<sub>2</sub></b>	5.871	<b>.016</b>	0.117	.74	1.66	.20
	<b>nM<sub>1</sub>.nM<sub>2</sub></b>	5.765	<b>.017</b>	0.568	.46	3.091	<b>.079</b>
<b>nM<sub>1</sub>.M<sub>2</sub></b>	<b>nM<sub>1</sub>.nM<sub>2</sub></b>	0.003	.96	0.165	.69	0.195	.66

Ces analyses révèlent :

- ▷ un effet de la superposition lexicale initiale sur les temps de détection des cibles en position P<sub>1</sub> (augmentation moyenne estimée des temps de détection de 18 ms.)
- ▷ un effet de la superposition lexicale initiale sur les temps de détection des cibles en position P<sub>2</sub> (diminution moyenne estimée des temps de détection de 23 ms.)
- ▷ un effet de la superposition lexicale initiale sur les temps de détection des cibles en position P<sub>3</sub> (diminution moyenne estimée des temps de détection de 30 ms.)
- ▷ pas d'effet de superposition lexicale finale sur les temps de détection des cibles (positions P<sub>1</sub>, P<sub>2</sub> ou P<sub>3</sub>)
- ▷ Une interaction de la lexicalité de la 1<sup>ère</sup> et de la 2<sup>ème</sup> syllabe sur les temps de détection des cibles en position P<sub>3</sub> (augmentation moyenne estimée des temps de détection de 28 ms.).

#### 4.2.8.3.2 Présentation des résultats

Par rapport aux hypothèses de l'expérience, les résultats des analyses (Tableau 75, page 293) permettent de répondre que :

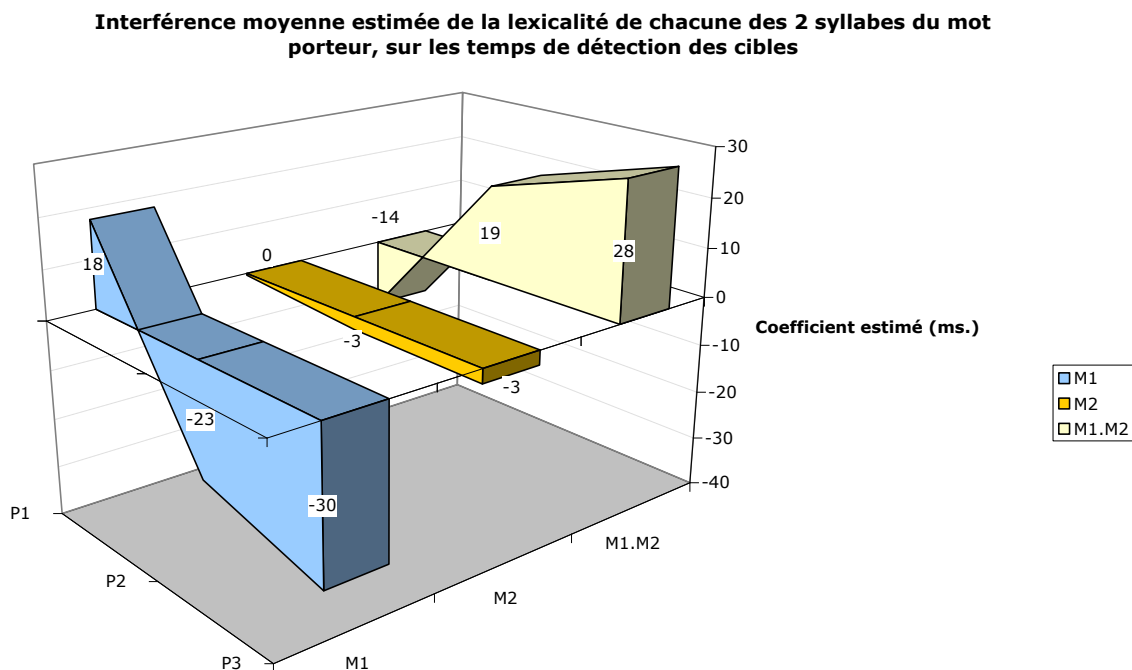
- ▷ L'hypothèse M<sub>1</sub> est validée (augmentation des temps de détection d'une cible insérée en position P<sub>1</sub> dans un mot présentant une superposition lexicale initiale, par rapport à un mot n'en présentant pas).
- ▷ L'hypothèse M<sub>2</sub> est invalidée pour ce qui concerne le sondage en position P<sub>2</sub> (l'augmentation prédite des temps de détection n'est pas observée pour un mot présentant une superposition lexicale finale, par rapport à un mot n'en présentant pas). L'hypothèse M<sub>2</sub> est validée pour ce qui concerne le sondage en position P<sub>3</sub> (aucune différence liée à la présence d'une superposition lexicale finale, lorsque le mot porteur est reconnu).

- ▷ L'hypothèse  $M_1 / M_2$  est validée (temps de détection d'une cible insérée en position  $P_1$  dans un mot présentant une superposition lexicale initiale plus longs que les temps de détection d'une cible insérée en  $P_2$  dans un mot présentant une superposition lexicale finale)

Concernant l'évolution de l'activité lexicale au cours du traitement du mot (voir Figure 65 ci-dessous), les analyses (Tableau 77 et Tableau 78 ci-après) révèlent que :

- ▷ les temps de détection des cibles sont augmentés à la fin de la première syllabe sous l'influence d'une superposition lexicale initiale (étant donné l'absence de superposition lexicale finale). On montre donc avec cette expérience ce que les Expériences 4 à 8 ne pouvaient pas mettre en évidence.
- ▷ les temps de détection des cibles diminuent durant la 2<sup>ème</sup> syllabe et jusqu'à la fin du mot sous l'influence d'une superposition initiale (étant donné l'absence de superposition finale). On retrouve donc le phénomène de facilitation mis en évidence dans les Expériences 4 à 8.
- ▷ la présence d'une superposition lexicale finale (étant donné l'absence de superposition lexicale initiale) sur la 2<sup>ème</sup> syllabe n'est responsable d'aucun effet sur les temps de traitement des cibles.
- ▷ une interférence plus importante existe lorsque la 1<sup>ère</sup> et la 2<sup>ème</sup> syllabe sont lexicales (interaction des 2 facteurs), se traduisant par une augmentation des temps de détection des cibles au cours de la 2<sup>ème</sup> syllabe et jusqu'à la fin du mot. Cet effet tendanciel (en position  $P_3$ , non significatif aux positions  $P_1$  et  $P_2$ ) suggère qu'un phénomène de segmentation lexicale est alors à l'œuvre.

Figure 65 : Coefficients estimés relatifs à la lexicalité des syllabes du mot stimulus et à la position de la cible pour chaque facteur expérimental et leur interaction



La forte symétrie entre  $M_1$  et  $M_1.M_2$ , avec un facteur  $M_2$  quasiment sans effet, indique que 3 des 4 conditions doivent avoir un profil de réponse voisin, mais qu'une des 4 conditions doit fortement diverger sous l'influence de  $M_1$  (voir Figure 66 ci-après).

Tableau 77 : Tests de significativité des contrastes entre les modalités des 2 facteurs expérimentaux ou d'un des deux facteurs et de leur interaction, pour chaque position de la cible

	$P_1$		$P_2$		$P_3$	
	$\text{Chi}^2_{(1, N=4776)}$	$p <$	$\text{Chi}^2_{(1, N=4776)}$	$p <$	$\text{Chi}^2_{(1, N=4776)}$	$p <$
$M_1 / M_2$	6.008	<b>.015</b>	4.271	<b>.039</b>	7.821	<b>.006</b>
$M_1 / M_1 * M_2$	4.333	<b>.038</b>	3.663	<b>.056</b>	6.663	<b>.01</b>
$M_2 / M_1 * M_2$	0.906	.35	0.973	.33	1.702	.20

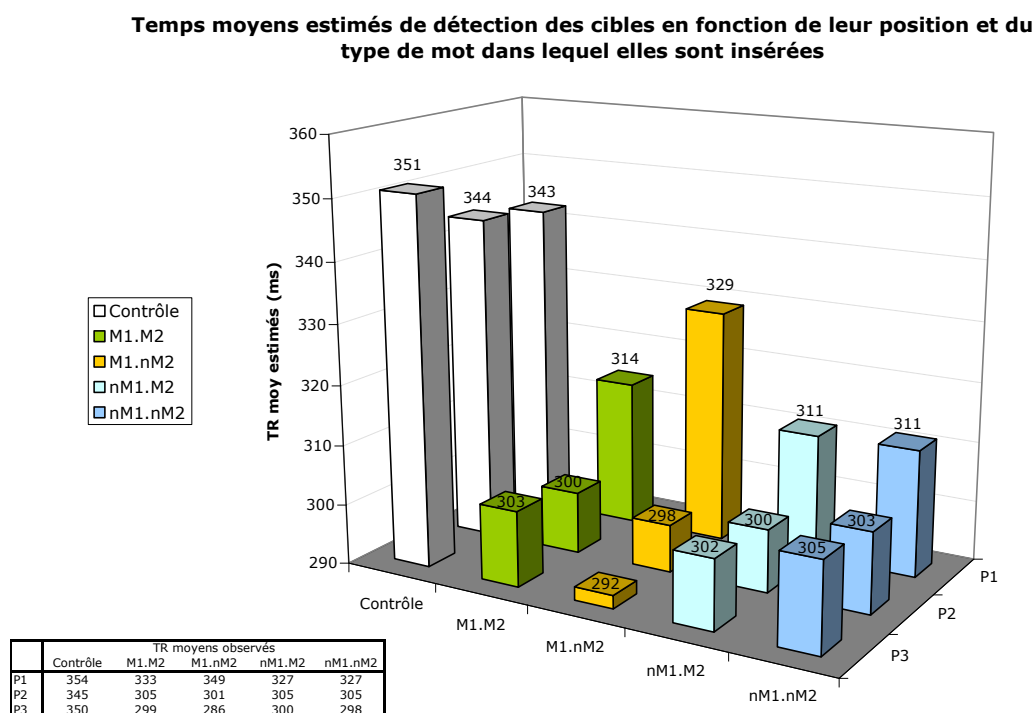
Tableau 78 : Tests de significativité des contrastes entre positions, pour chaque facteur expérimental et pour leur interaction

	$M_1$		$M_2$		$M_1 * M_2$	
	$\text{Chi}^2_{(1, N=4776)}$	$p <$	$\text{Chi}^2_{(1, N=4776)}$	$p <$	$\text{Chi}^2_{(1, N=4776)}$	$p <$
$P_1/P_2$	6.69	<b>.01</b>	0.026	.88	2.392	.13
$P_1/P_3$	9.193	<b>.003</b>	0.014	.91	3.59	<b>.059</b>
$P_2/P_3$	0.492	.49	0.004	.95	0.301	.59

Concernant l'analyse des contrastes entre conditions (voir Figure 66 ci-après et Tableau 76, page 294), on remarque que, par position de sondage, les temps moyens estimés de détection des cibles varient peu d'une condition à l'autre. Toutefois, il apparaît que les mots  $M_1.nM_2$  sont traités différemment des 3 autres types de mots :

- ▷ les cibles insérées en position  $P_1$  dans ces items sont détectées plus lentement que les cibles insérées à la même position dans les 3 autres catégories de mots.
- ▷ les cibles insérées en position  $P_3$  dans ces items semblent être détectées plus rapidement que les cibles insérées en position  $P_3$  dans les 3 autres catégories de mots (contrastes non significatifs. On remarque toutefois une différence importante dans l'ordre de grandeur des niveaux de probabilité pour ces contrastes, par rapport aux autres contrastes).

Figure 66 : Temps moyens estimés de détection des cibles en fonction de leur position et du type de mot dans lequel elles sont insérées



Cet effet suggère apparemment une difficulté accrue du traitement liée à la présence d'une superposition lexicale initiale. Cependant, comme l'allongement n'apparaît pas avec la condition  $M_1.M_2$ , qui présente elle aussi une superposition lexicale initiale, la difficulté provient d'une autre cause.

Un processus de segmentation peut être à l'œuvre - dans la condition  $M_1.nM_2$  aussi bien que dans la condition  $M_1.M_2$  - puisque la première syllabe est un mot, ce qui déclenche l'hypothèse que la syllabe qui suit est peut-être un mot ou le début d'un mot.

En position  $P_1$ , le phénomène n'est révélé que dans la condition  $M_1.nM_2$  parce qu'il est rapidement mis en échec par la configuration particulière des stimuli. En effet, la 2<sup>ème</sup> syllabe n'est ni un mot, ni le début d'un mot, et donc, le système se retrouve face à une impasse. Deux phénomènes se produisent probablement, entraînant un surcoût au niveau des traitements :

- ▷ le système lance une exploration sur toutes les possibilités lexicales de manière à ne pas omettre une hypothèse importante
- ▷ il réorganise ensuite ses hypothèses : la séquence  $M_1.nM_2$  ne pouvant pas être segmentée, l'hypothèse  $M_1$  doit être rejetée et l'hypothèse correcte est donc  $M_1.nM_2$ .

Dans le cas des mots  $M_1.M_2$ , la 2<sup>ème</sup> syllabe étant un mot, la segmentation de la séquence  $M_1.M_2$  en unités lexicales  $M_1$  et  $M_2$  est possible – donc le processus peut se poursuivre (exploration en parallèle des hypothèses  $M_1$ ,  $M_2$ , et  $M_1.M_2$ ) jusqu'à ce que l'information acoustique permette au système de procéder à un choix, éventuellement passif sous l'effet de l'information acoustique.

Un tel phénomène ne peut pas avoir lieu avec les deux conditions  $nM_1^*$ , puisque  $nM_1$  n'est pas un mot. Dans la condition  $nM_1.nM_2$  où les syllabes  $nM_2$  ne sont pas, tout comme dans la condition  $M_1.nM_2$ , des initiales de mot, le système ne s'engage donc pas dans un processus voué à l'échec.

En position  $P_3$ , le temps de détection des cibles tend à diminuer avec la condition  $M_1.nM_2$ , alors qu'il est en moyenne identique pour les 3 autres types de mots. Ce phénomène pourrait être l'indice que les hypothèses  $nM_2^*$  sont abandonnées au point de reconnaissance du mot bisyllabique, libérant soudainement des ressources, qui étaient mobilisées et deviennent disponibles pour la réalisation de la tâche<sup>117</sup>.

---

<sup>117</sup> Il faut remarquer que l'Expérience 9 a révélé un profil de comportement similaire (diminution des temps de détection des cibles insérées en fin de mot) avec une condition de structure comparable pour les facteurs expérimentaux (E.nI). Dans les deux expériences, il s'agit à chaque fois d'une situation où le degré de certitude pour la solution lexicale adoptée est maximal.

#### 4.2.8.4 Conclusions de l'Expérience 10

Des 3 hypothèses de l'Expérience 10, la 1<sup>ère</sup> et la 3<sup>ème</sup> sont validées par les résultats, la 2<sup>ème</sup> est partiellement validée.

L'Expérience 10 a permis de montrer que :

- ▷ L'interférence entre le traitement lexical et la tâche de détection est plus importante en initiale de mot (position P<sub>1</sub>), lorsque la 1<sup>ère</sup> syllabe est lexicale que lorsqu'elle ne l'est pas. L'Expérience 10 précise donc ce que les expériences antérieures n'avaient pu montrer : il existe bien une difficulté supplémentaire en début de mot, liée à la présence d'un mot enchâssé initial.
- ▷ cette difficulté n'est que passagère : elle laisse rapidement place à un phénomène de facilitation, non expliqué, qui se manifeste dès la 2<sup>ème</sup> syllabe, et semble devenir de plus en plus important à l'approche de la fin du mot. Ce résultat est conforme à ceux des Expériences 4 à 8, qui sondaient l'activité lexicale à la fin de l'item.
- ▷ la lexicalité de la 2<sup>ème</sup> syllabe est sans effet en tant que facteur principal.
- ▷ elle est cependant en interaction avec la lexicalité de la 1<sup>ère</sup> syllabe. Cette interaction se traduit par une augmentation des temps de réponse, quasi-significative en fin de mot, reflet d'un conflit probablement lié à une procédure de segmentation lexicale. L'Expérience 10 introduit donc l'idée d'une procédure de segmentation lexicale basée sur la syllabation de la séquence acoustique.



### **4.2.9 Expérience 11 : Détection de syllabes non lexicales du Français I – Exploration du rôle de la fréquence positionnelle**

Les Expériences 4 à 8 ont mis en évidence un phénomène de facilitation de la reconnaissance des mots en cas de superposition lexicale initiale. L'effet est paradoxal, puisque les théories de la reconnaissance des mots en psycholinguistique proposent que de la compétition lexicale existe avec ce type de configuration, donc que la reconnaissance du mot stimulus soit rendue plus difficile au lieu d'être facilitée.

Par ailleurs, l'utilisation de mots expérimentaux bisyllabiques dans les Expériences 7 et 10 a révélé un phénomène qui ressemble à une stratégie d'exploration du lexique mental (segmentation lexicale - M1.nM2), retrouvé dans l'Expérience 9 (E.nI), et qui semble être basé sur l'unité syllabique.

Les effets de facilitation sont en général considérés comme étant de nature infra-lexicale (fréquence de cooccurrence de phonèmes, etc.). L'existence d'un niveau de représentation syllabique pourrait donc peut-être fournir un début d'explication au paradoxe observé.

L'Expérience 11, première d'une série de 4 expériences relatives à l'exploration de l'hypothèse d'un niveau infra-lexical de représentation syllabique, cherche à montrer l'existence d'un biais statistique lié à la répartition positionnelle de syllabes non lexicales dans le lexique français. Ce biais serait à même d'influencer à la fois la reconnaissance d'un mot monosyllabique et le processus de segmentation du flux de parole en mots, en induisant une granularité perceptive de la taille de la syllabe dans le processus.

#### **4.2.9.1 Aspects théoriques**

Les syllabes forment des unités articulatoires fondamentales et naturelles. Elles correspondent à des « bouchées de parole » (un noyau vocalique éventuellement enrobé de consonnes), dont la cohérence interne est élevée (coarticulation plus importante entre 2 phonèmes successifs d'une même syllabe qu'entre 2 phonèmes successifs de 2 syllabes). Il semble peu logique que de telles unités existent dans le signal et ne soient pas utilisées comme unités perceptives. Le système de traitement perceptif a tout à gagner en exploitant cette information supplémentaire, et les hypothèses proposées par certains chercheurs vont généralement dans ce sens (exemple : SOSH).

Étant donné la coarticulation des phonèmes, qui rend impossible leur identification hors contexte et conduit à envisager l'existence d'une analyse compensatoire dans le

système (correction à partir du contexte immédiat – processus non précisés), un niveau infra-lexical d'unités plus élaborées que le phonème peut représenter une source de stabilité et de fiabilité de la reconnaissance, et correspondre à l'un des mécanismes de compensation de la coarticulation (phénomènes d'attraction par des unités plus stables).

Pour mettre en évidence la présence de telles unités, la fréquence d'occurrence des syllabes dans le lexique peut fournir des indices précieux. Afin de décorréler le niveau lexical du niveau syllabique, des syllabes non lexicales ont été testées dans un premier temps (contrairement au matériel utilisé dans les Expériences 4, 5 et 8) :

- ▷ si le niveau syllabique n'est pas représenté dans le système d'accès au lexique, qu'une syllabe soit fréquente ou non ne devrait pas changer quoique ce soit dans les temps nécessaires à sa détection.
- ▷ par contre, si une syllabe est représentée en mémoire à long terme sous quelle que forme que ce soit (y compris sous forme de cooccurrence de phonèmes), plus elle est fréquente dans la langue, plus elle devrait être détectée rapidement.

Si une syllabe est représentée en mémoire, l'information de fréquence ne doit pas être la seule représentée. La manière dont la syllabe s'intègre au sein du lexique doit également l'être (position de la syllabe dans les mots), notamment grâce aux connexions entre les éléments du niveau syllabique et les éléments du niveau lexical. La fréquence d'occurrence des syllabes dans le lexique peut alors se décliner en un nouveau facteur : la fréquence positionnelle d'occurrence des syllabes dans les mots

Ce facteur, relatif à la régularité statistique de la répartition des syllabes à travers les mots du lexique à une position donnée dans le mot, autorise l'existence de catégories de syllabes (exemple : syllabes fréquentes ou rares en initiale de mot).

Dans l'Expérience 11, 2 cas de figure sont distingués pour le facteur position : la position en initiale de mot, et la position en fin de mot. Ce choix se justifie par :

- ▷ l'homogénéité du facteur : la position des syllabes est dans tous les cas à l'une des 2 frontières d'un mot, quelle que soit la quantité de syllabes que celui-ci renferme.
- ▷ la fonctionnalité de la position : les syllabes en initiale de mot ont probablement un rôle important dans l'activation lexicale, que ne possèdent pas les syllabes de fin de mot (comparaison entre les deux types de frontières). On suppose qu'une syllabe initiale joue un rôle activateur du lexique plus fort qu'une syllabe non initiale, et donc qu'elle est reconnue d'autant plus facilement en tant qu'unité.

Si la représentation d'une syllabe prend à la fois en considération sa position dans les mots du lexique et sa fréquence d'occurrence à une position particulière, l'effet

syllabique devrait être plus important pour les syllabes fréquemment en initiale d'un mot que pour les syllabes fréquemment en fin de mot.

## 4.2.9.2 Aspects techniques

### 4.2.9.2.1 Facteurs expérimentaux et conditions

Deux facteurs à deux modalités chacun. Croisement complet (4 conditions) :

Tableau 79 : Facteurs expérimentaux

		Fréquence de la 2 <sup>ème</sup> syllabe en position finale	
		Fréquente (FrqF)	Rare (rF)
Fréquence de la 2 <sup>ème</sup> syllabe en position initiale	Fréquente (FrqI)	FrqI.FrqF ( $S_1$ -/naʁ/)	FrqI.rF ( $S_1$ -/spe/)
	Rare (rI)	rI.FrqF ( $S_1$ -/zɛt/)	rI.rF ( $S_1$ -/zõp/)

Tableau 80 : Fréquence moyenne d'occurrence de chaque type de syllabe expérimentale en initiale, milieu, fin de mot

	Dans N mots à cette position			Fréquence positionnelle/syllabe		
	Initiale	Milieu	Fin	Initiale	Milieu	Fin
<b>FrqI.FrqF</b>	28.5	118.1	46.9	47254.0	67547.8	58319.0
<b>FrqI.rF</b>	56.2	29.7	0.0	66206.9	15107.9	0.0
<b>rI.FrqF</b>	0.0	5.5	21.7	0.0	2102.4	29523.6
<b>rI.rF</b>	0.0	3.0	0.0	0.0	2073.7	0.0

En italique : donné pour complément d'information.

### 4.2.9.2.2 Hypothèses opérationnelles

- ▷ Hypothèse sur la fréquence initiale :  
plus une syllabe non lexicale est fréquente en initiale de mot dans le lexique, plus les temps de détection de cette syllabe seront courts.
- ▷ Hypothèse sur la fréquence finale :  
plus une syllabe non lexicale est fréquente en fin de mot dans le lexique, plus les temps de détection de cette syllabe seront courts.
- ▷ Hypothèse sur la différence d'effet entre fréquence positionnelle initiale et finale :  
une syllabe non lexicale fréquente en initiale de mot sera détectée plus rapidement qu'une syllabe non lexicale fréquente en fin de mot.

### 4.2.9.2.3 Matériel expérimental utilisé

Stimuli en voix naturelle, de gabarit bisyllabique pour mettre l'auditeur dans une situation où une segmentation peut survenir. La liste des items est donnée en Annexe 26 page 521.

La première syllabe des items (/dzyk/) est illégale en Français selon BRULEX. Elle est sélectionnée à partir d'un ensemble de « syllabes » illégales candidates, choisies de manière à :

- ▷ Se terminer par une consonne pour faciliter la syllabation.
- ▷ Ne pas évoquer trop de mots ou trop fortement un mot particulier (test auprès de 5 étudiants (« Spontanément, à quel mot cette syllabe vous fait penser ? »)).

La deuxième syllabe des items correspond aux cibles expérimentales à détecter (syllabes non lexicales du Français) ou aux distracteurs (mots monosyllabiques).

La première syllabe des items a initialement été choisie identique pour tous les items, afin de permettre un contrôle de la situation (influence de  $S_1$  toujours identique). Toutefois, cette solution s'est avérée pouvoir introduire un biais dans le matériel, car la syllabe /dzyk/ est impossible à concaténer avec certaines syllabes expérimentales (la phonation exige que la séquence « /dzyk/-syllabe cible » ne soit pas trop difficile à produire<sup>118</sup>, ni coarticulée<sup>119</sup>). Une concaténation artificielle n'aurait pas représenté une meilleure solution, car certains enchaînements auraient sonné trop artificiel (effet de surprise parasitant les TR, marquage acoustique de la frontière syllabique rendant la détection de  $S_2$  plus aisée, etc.). Il a donc fallu employer 3 variantes de la première syllabe en changeant le dernier phonème de manière à ce que l'enchaînement  $S_1S_2$  soit énonçable et vraisemblable. Le dernier phonème pouvait donc prendre 3 autres modalités : /p/, /b/ ou /g/. Ces 3 phonèmes de substitution sont employés uniquement lorsque la séquence /dzyk/ n'est pas utilisable<sup>(118 et 119)</sup>. Ils permettent d'obtenir une séquence plus adéquate<sup>120</sup>, par une modification minimale de la syllabe sur 1 trait articulatoire de son dernier phonème (ajustement du voisement pour contrôler la coarticulation du dernier phonème de la première syllabe avec le premier segment de la syllabe suivante (paire /k/-/g/ ou paire /p/-/b/); ajustement du lieu d'articulation

---

<sup>118</sup> Si la séquence  $S_1S_2$  est trop difficile à articuler, le locuteur produira une frontière acoustique artificiellement accentuée entre les deux syllabes (exemple : /dzykglo/).

<sup>119</sup> La coarticulation modifie parfois le premier phonème de la syllabe cible (exemple : l'association des 2 syllabes dans la séquence /dzykzif/ provoque une modification du /k/ qui se voise et tend vers un /g/, et du /z/ qui se dévoise et tend à devenir /s/. /sif/ n'est pas la syllabe souhaitée ni celle qui a été sélectionnée pour former un item expérimental).

<sup>120</sup> Par exemple, la séquence /dzykglo/ posant un problème d'articulation du groupe /kg/ devrait être /dzygglo/ afin que le premier phonème de la syllabe cible soit plus « prototypique » d'un /g/, mais cette séquence est aussi trop difficile à produire sans marquer acoustiquement la frontière syllabique. Un autre phonème voisé doit donc être choisi pour réduire l'effort articulatoire tout en préservant le voisement sur le premier phonème de la deuxième syllabe (/b/ : /dzybglo/).

(dorso-vélaire ou bilabial) pour contrôler l'aisance de la production et le naturel de l'item (paire /k/-/p/ ou paire /g/-/b/).

Ces modulations de la première syllabe posent cependant un problème : la répartition des 4 syllabes  $S_1$  dans les 4 conditions expérimentales est hétérogène (Tableau 81 et Tableau 82 ci-dessous), et joue peut-être une influence sur l'aisance acoustique ou phonologique à segmenter une séquence de phonèmes en unités de niveau supérieur.

Tableau 81 : Répartition des 4 syllabes  $S_1$  à travers les 4 conditions expérimentales

	Ri		Fi		Somme
	Rf	Ff	Rf	Ff	
/k/	1	1	8	4	14
/g/	11	11	1	7	30
/p/	0	0	3	2	5
/b/	1	1	1	0	3

Tableau 82 : Synthèse du tableau précédent (somme par modalité de chaque facteur, du nombre de syllabes  $S_1$  terminant avec l'un des 4 phonèmes)

	Ri	Fi	Rf	Ff		Ri	Fi	Rf	Ff
/k/	2	12	9	5	/p/	0	5	3	2
/g/	22	8	12	18	/b/	2	1	2	1

La répartition des variantes de  $S_1$  à travers les conditions expérimentales n'a pas pu être contrôlée malgré trois solutions considérées :

- ▷ Choisir les syllabes expérimentales en conséquence (trop de contraintes lexicales autres à respecter, limitant le nombre d'items de chaque condition)
- ▷ Concaténer une syllabe  $S_1$  unique à des syllabes  $S_2$  produites dans un contexte bisyllabique duquel elles sont extraites, pour contourner les contraintes de production (meilleur contrôle au niveau physique, mais aucune connaissance ni évaluation possible des conséquences que ces stimuli artificiels entraîneraient sur la perception).
- ▷ Contrôler le facteur aux analyses statistiques (répartition trop inégale entre conditions. Le modèle n'est pas capable de faire une estimation correcte).

Par conséquent, un matériel imparfait mais naturel est employé. Ce choix risque apparemment d'amputer la discussion d'une partie des conclusions qui auraient pu être tirées à partir d'une étude correctement contrôlée.

Cependant, un raisonnement logique sur la répartition des variantes de la syllabe  $S_1$  par rapport aux 2 facteurs expérimentaux permet de découvrir que l'influence de chaque type de syllabe  $S_1$  sur les temps de détection des syllabes  $S_2$ , ne peut pas expliquer les résultats obtenus. Ri.Rf renferme chacune des 4 syllabes  $S_1$  en proportions exactement identiques que Ri.Ff (Tableau 81), ce qui n'empêche pas les temps de

détection des syllabes  $S_2$  d'être très différents entre ces deux conditions. Par ailleurs, les proportions de chacun des 4 types de syllabes  $S_1$  pour les deux conditions  $Fi.*$  sont très différentes, alors que les temps moyens de détection des syllabes  $S_2$  sont très voisins (mais cette équivalence peut être due à une moyenne d'effets contraires). Il serait possible d'opposer les deux conditions  $Ri.*$  aux deux conditions  $Fi.*$  pour la proportion de syllabes  $S_1$  avec dernier phonème voisé / non voisé (respectivement : 24 / 2 – 17 / 9), mais la différence est faible et peut donc difficilement expliquer les différences comportementales. D'après le Tableau 82, page 304, et d'après le résultat selon lequel l'influence facilitante de la modalité  $Fi$  est supérieure (et significative) à celle de la modalité  $Ff$  (non significative), une syllabe  $S_1$  terminant par le phonème :

- ▷ /k/ faciliterait la segmentation de la séquence bisyllabique (7 de plus pour les conditions  $Fi.*$  que pour les conditions  $Ff.*$ ). Donc les syllabes  $S_2$  des conditions  $*.Rf$  devraient être plus faciles à détecter (plus de syllabes  $S_1$  terminant par /k/) que celles des conditions  $*.Ff$ .
- ▷ /g/ rendrait plus difficile la segmentation de la séquence bisyllabique (10 de plus pour les conditions  $Ff.*$  que pour les conditions  $Fi.*$ ). Donc les syllabes  $S_2$  des conditions  $*.Rf$  devraient être plus faciles à détecter (moins de syllabes  $S_1$  terminant par /g/) que celles des conditions  $*.Ff$ .
- ▷ /p/ faciliterait la segmentation de la séquence bisyllabique (3 de plus pour les conditions  $Fi.*$  que pour les conditions  $Ff.*$ ). Donc les syllabes  $S_2$  des conditions  $*.Rf$  devraient être plus faciles à détecter (plus de syllabes  $S_1$  terminant par /p/) que celles des conditions  $*.Ff$ .
- ▷ /b/ serait sans effet (même quantité).

Si ces relations entre syllabe  $S_1$  et temps de détection des syllabes  $S_2$  étaient vraies, la logique démontre que les syllabes expérimentales des conditions  $*.Rf$  devraient présenter des temps de réponse plus rapides que celles des conditions  $*.Ff$ , mais il n'en est rien. Il faut donc relativiser l'importance accordée à cette idée d'un biais lié à l'hétérogénéité de la répartition des types de syllabes  $S_1$  à travers les 4 conditions expérimentales.

Les caractéristiques du matériel sont donc les suivantes :

- ▷ Items expérimentaux :  
52 bisyllabes non lexicaux. Leur première syllabe est illégale en Français, et son dernier phonème prend 4 modalités (/k/, /g/, /p/, /b/). Leur deuxième syllabe est une syllabe non lexicale du Français, fréquente/rare en initiale/finale (4 conditions,  $N = 13$  par condition). La répartition fréquentielle dans le lexique de ces deuxièmes syllabes en tant que syllabes, et non en tant que séquences de phonèmes, est obtenue à partir d'une analyse de la base de données BRULEX (C. Lachaud ; J. Goslin).
- ▷ Items distracteurs :  
52 bisyllabes non lexicaux, construits selon la même structure que les items expérimentaux, mais dont la deuxième syllabe est un mot.

L'expérience renferme donc 104 items.

#### **4.2.9.2.4 Organisation de l'expérience**

Un participant est testé avec tous les items (N = 104), chacun lui étant présenté une seule fois.

#### **4.2.9.2.5 Population testée**

Les données de 16 individus francophones monolingues, sans troubles auditifs, cognitifs et moteurs, étudiant à l'université de Genève, ont été utilisées pour les analyses. Le recueil a été effectué en janvier 2003.

#### **4.2.9.2.6 Tâche**

Détection de pseudo-mot sur la deuxième syllabe du stimulus (tâche de décision lexicale en « go, no go »). Le participant, pour fournir une réponse, doit effectuer au préalable une analyse du stimulus afin d'isoler la cible. Modalité auditive.

#### **4.2.9.2.7 Procédure de recueil des données**

Expérience réalisée à l'université de Genève. Détails dans la Partie III, Chapitre 4.

16 \* 52 = 832 mesures utilisées.

La phase test de l'expérience se déroule sans pause, et dure moins de 10 minutes.

#### **4.2.9.2.8 Méthode d'analyse**

- ▷ Détail des procédures pour le tri des données et les analyses dans la Partie III, Chapitre 3.
- ▷ Analyse de régression multiniveaux, modèle à 2 niveaux (sujets, items. Corrélation intra classe : 29.45%).
- ▷ 107 valeurs manquantes (12.86% de non réponses et de valeurs filtrées (détail par condition : FrqI.FrqF 11.5% ; FrqI.rF 9.1% ; rI.FrqF 19.7% ; rI.rF 11.1%)).

#### **4.2.9.2.9 Paramètres contrôlés**

À la construction du matériel et à la structuration de l'expérience

- ▷ Les syllabes, et non les séquences de phonèmes, sont considérées pour la sélection du matériel.

- ▷ La première syllabe de la séquence est choisie illégale en Français pour ne pas induire de phénomènes d'interférence avec le niveau lexical et avec le traitement de la deuxième syllabe de l'item.
- ▷ La première syllabe n'évoque pas de mot particulier, et est choisie comme en évoquant un minimum après enquête auprès de 5 étudiants.
- ▷ « Prononçabilité » de la séquence  $S_1S_2$ .
- ▷ Contrôle de la coarticulation à la jonction syllabique. Pas de modification du premier phonème de  $S_2$  à cause des caractéristiques du dernier phonème de  $S_1$ .
- ▷ Les syllabes cible qualifiées de rares à une position donnée sont en réalité absentes de cette position selon BRULEX, en tant que syllabes. Elles peuvent cependant éventuellement exister en tant que séquence de phonèmes (non vérifié).
- ▷ Le nombre de phonèmes moyen a été à peu près équilibré à travers les conditions.
- ▷ Présentation aléatoire des items lors de la passation.

#### Lors des analyses statistiques

- ▷ Utilisation des temps de détection mesurés depuis le début de la deuxième syllabe (pour limiter la variance liée à la durée de l'onde acoustique qui précède la deuxième syllabe) .
- ▷ Latéralité manuelle des participants.
- ▷ Sexe des participants.
- ▷ Ordre de présentation des items.
- ▷ Nombre de phonèmes composant la syllabe cible.
- ▷ Nombre de mots évoqués par la syllabe<sup>121</sup>.
- ▷ Durée de la syllabe cible (ms.).
- ▷ Fréquence fondamentale moyenne de la syllabe cible (Hz).
- ▷ Intensité moyenne de la syllabe cible (dB).

Les facteurs prosodiques  $f_0$  et dB ont été intégrés aux analyses (en plus du facteur durée, qui l'est pour des raisons liées au contrôle de la temporalité des items avec l'utilisation d'une technique de chronométrie) puisqu'il a été montré qu'ils variaient significativement en rapport avec les facteurs expérimentaux (détail des analyses présenté dans l'Annexe 27 page 523). Les interactions qu'ils font avec les facteurs expérimentaux figurent aussi dans le modèle d'analyse.

---

<sup>121</sup> Les données sont obtenues par enquête sur Internet, selon le même principe que celle qui a permis de recueillir les informations d'âge d'acquisition estimé et de familiarité conceptuelle. Pour cela, les participants ont utilisé un questionnaire grâce auquel ils ont pu fournir au moins 1, au plus 3, mots qui leur sont venus spontanément à l'esprit lorsque le pseudo-mot était présenté.



### 4.2.9.3 Résultats obtenus

#### 4.2.9.3.1 Rapport des analyses statistiques

L'équation du modèle (Équation 49) et les tests de significativité relatifs aux covariables peuvent être consultés à l'Annexe 28 page 526.

Tableau 83 : Tests de significativité des facteurs étudiés et de leur interaction

	Coefficient	Chi <sup>2</sup> <sub>(1, N=725)</sub>	p<	% variance expliquée
<b>FrqI</b>	-138.847	6.558	<b>.011</b>	0.522
<b>FrqF</b>	-69.555	2.19	.14	0.438
<b>FrqI.FrqF</b>	82.9	1.355	.25	0.161

Tableau 84 : Tests de significativité des contrastes entre conditions

Contraste		Chi <sup>2</sup> <sub>(1, N=725)</sub>	p<
<b>rI.rF</b>	<b>rI.fF</b>	2.19	.14
	<b>fI.rF</b>	6.558	<b>.011</b>
	<b>fI.fF</b>	5.284	<b>.022</b>
<b>rI.fF</b>	<b>fI.rF</b>	1.353	.25
	<b>fI.fF</b>	1.551	.22
<b>fI.rF</b>	<b>fI.fF</b>	0.04	.85

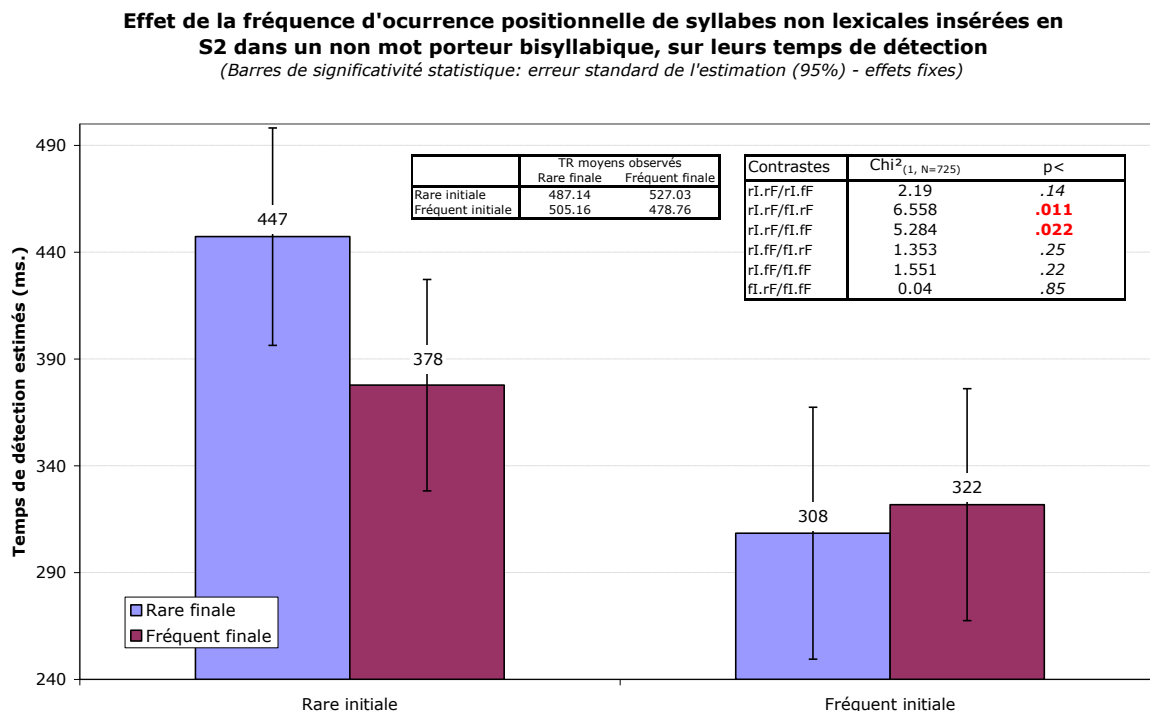
Ces analyses mettent en évidence un effet principal pour la fréquence d'occurrence d'une syllabe non lexicale en initiale de mot.

La fréquence d'occurrence d'une syllabe non lexicale en position non initiale n'a pas d'influence significative sur les temps de détection, de même que l'interaction des deux facteurs.

L'analyse des contrastes montre que les syllabes à la fois rares en initiale et en finale de mot sont traitées différemment des syllabes fréquentes en initiale de mot, qu'elles soient aussi rares ou fréquentes en fin de mot.

### 4.2.9.3.2 Présentation des résultats

Figure 67 : Temps moyens de détection des syllabes cible en fonction de leur fréquence positionnelle



▷ Hypothèse sur la fréquence initiale :

Le Tableau 83, page 308, révèle un effet principal de la fréquence d'une syllabe non lexicale initiale (traitée 139 ms. plus rapidement (estimation) lorsqu'elle est fréquente à cette position que lorsqu'elle y est rare). L'hypothèse sur la fréquence d'occurrence en position initiale est donc confirmée.

▷ Hypothèse sur la fréquence finale :

L'hypothèse sur la fréquence d'occurrence en position finale n'est pas confirmée : l'effet n'atteint pas le seuil de significativité (Tableau 83, page 308), bien qu'il semble également y avoir un phénomène de facilitation lié à la fréquence (syllabes fréquentes traitées 70 ms. plus rapidement (estimation) que les syllabes rares en position finale).

▷ Hypothèse sur la différence d'effet entre fréquence positionnelle initiale et finale :

L'hypothèse relative à la différence de traitement des syllabes initiales et finales n'a pas pu être confirmée par les données (comparaison entre les conditions rI.fF et fI.rF – Tableau 84, page 308).

#### 4.2.9.4 Conclusions de l'Expérience 11

Des 3 hypothèses de l'Expérience 11, seule la 1<sup>ère</sup> a pu être validée par les données. La fréquence positionnelle initiale des syllabes non lexicales influence leurs temps de traitement (effet significatif seulement pour les syllabes fréquentes en initiale de mot par rapport aux syllabes rares à cette position, mais pas de différence significative entre les syllabes fréquentes en position initiale (condition Fi.Rf) et celles fréquentes en position finale (condition Ri.Ff)).

Le fait qu'un effet significatif n'apparaisse que pour les syllabes en position initiale dans le lexique peut provenir :

- ▷ d'un traitement spécifique au début des mots (alignement et appariement)
- ▷ d'un biais expérimental lié à la tâche (syllabation et focalisation du système sur les syllabes de début de mot à cause de la segmentation implicitement demandée)

Il est impossible de départager ces deux suppositions. Il faut quand même remarquer que si l'effet obtenu provient d'un biais lié à la tâche, cela implique que le système est capable de distinguer les syllabes initiales de mot des syllabes de fin de mot.

Un point est néanmoins acquis grâce à l'Expérience 11 : les résultats obtenus indiquent qu'une syllabe est bien représentée en mémoire à long terme dans le système de traitement de la parole, quelle que soit sa forme de représentation, puisqu'un effet général de la fréquence est observé (facilitation).

### 4.2.10 Expérience 12 : Détection de syllabes non lexicales du Français II – Poursuite de l'exploration du rôle de la fréquence positionnelle

L'Expérience 11 a montré que le système disposait de représentations syllabiques (effet de fréquence), et que les syllabes initiales de mots disposaient d'un statut particulier (effet significatif) par rapport aux syllabes de fin de mot (effet non significatif). Son échec à mettre en évidence un effet de fréquence pour les syllabes non initiales s'explique peut-être par la tâche expérimentale, qui focalise le système sur les initiales de mots à cause du processus de segmentation implicitement demandé. Afin de retester l'hypothèse d'un niveau de représentation syllabique dans le système, tout en limitant l'impact du niveau lexical sur les résultats :

- ▷ des syllabes non lexicales sans rapport avec les frontières d'un mot sont employées (elles n'existent pas (selon BRULEX) en position initiale, ni en position finale). Les syllabes sont opposées selon leur fréquence d'occurrence (rares ou fréquentes).
- ▷ une tâche ne nécessitant pas de réaliser une procédure de segmentation est choisie. La tâche demandée dans l'Expérience 11 était relativement artificielle, car une analyse consciente des stimuli (syllabation) était nécessaire. L'Expérience 12 résout ce problème en utilisant une tâche chronométrique plus simple (tâche de décision lexicale en « go, no go »).

#### 4.2.10.1 Aspects théoriques

Le traitement de la parole commence par une phase d'analyse qui s'effectue dans le temps (donc, une syllabe initiale de mot possède peut-être un statut de code d'accès au lexique), implique à un moment donné la localisation des frontières lexicales dans le signal (donc les syllabes initiale et finale peuvent servir à baliser l'onde acoustique pour sa segmentation en mots), et conduit à l'évocation des représentations mentales lexicales (appariement entre unités infra-lexicales et unités lexicales).

Dans un mot polysyllabique composé d'au moins 3 syllabes (une syllabe initiale, une syllabe médiale et une syllabe finale), chacune de ces syllabes peut être impliquée dans l'une ou plusieurs de ces fonctions (code d'accès, balisage, appariement), dont la combinaison permet de les distinguer. Le Tableau 85 ci-dessous résume ces caractéristiques fonctionnelles possibles.

Tableau 85 : Caractéristiques fonctionnelles de 3 types de syllabes

	Syllabe initiale	Syllabe médiale	Syllabe finale
<b>Code d'accès</b>	1	0	0
<b>Balisage</b>	1	0	1
<b>Appariement</b>	1	1	1

Les syllabes médiales, si elles ne sont pas initiales ou finales dans un autre mot du lexique, doivent permettre de tester de façon mieux ciblée l'hypothèse sur l'existence de représentations mentales syllabiques, puisque les fonctions relatives à la localisation des unités lexicales dans le signal acoustique (code d'accès et balisage) sont évacuées. Ainsi :

- ▷ Si les syllabes sont représentées dans le système, elles doivent être reconnues d'autant plus rapidement qu'elles sont fréquemment utilisées en tant qu'unité.
- ▷ Si elles ne sont pas représentées dans le système, leur fréquence d'usage en tant qu'unité ne doit pas produire d'effet.

#### 4.2.10.2 Aspects techniques

##### 4.2.10.2.1 Facteur expérimental et conditions

Fréquence d'occurrence de la syllabe dans le lexique français (deux modalités : rare / initial). L'expérience comporte 2 conditions.

Tableau 86 : Conditions expérimentales

	Syllabe rare	Syllabe fréquente
<b>Exemples</b>	/nɔp/, /vɛal/	/lɥa/, /zɔp/

##### 4.2.10.2.2 Hypothèse opérationnelle

Les temps de détection de syllabes médiales fréquentes dans le lexique français seront plus courts que les temps de détection de syllabes médiales rares.

##### 4.2.10.2.3 Matériel expérimental utilisé

Stimuli en voix naturelle. La liste des items est donnée en Annexe 29 page 528.

- ▷ Items expérimentaux :
  - ∞ 18 syllabes non lexicales du Français, non initiales et non finales selon BRULEX. 9 sont rares (fréquence moyenne = 161) et 9 sont fréquentes (fréquence moyenne = 6057). Elles sont appariées entre conditions pour leur structure syllabique (4 CVC ; 4 CCV ; 1 CCVC).

Tableau 87 : Indices caractérisant les syllabes utilisées (moyennes)

	Commence N mots	Finit N mots	Syllabe dans N mots	Fréquence syllabe	Séquence dans N mots	Fréquence séquence
<b>Syllabes rares</b>	0	0	2	161	6	3659
<b>Syllabes fréquentes</b>	0	0	9	6057	15	6506

- ▷ Items distracteurs :

- ∩ 36 mots monosyllabiques
- ∩ 18 « syllabes » illégales en Français<sup>122</sup> selon BRULEX. La séquence de phonèmes qui les compose n'existe pas en Français, et cette séquence sans le dernier phonème se retrouve en moyenne dans 21.72 mots.

#### 4.2.10.2.4 Organisation de l'expérience

Chaque participant est testé avec tous les items. Il traite chaque stimulus 1 fois, et fournit 18 réponses utiles.

#### 4.2.10.2.5 Population testée

Les données de 11 francophones monolingues, sans troubles auditifs, cognitifs et moteurs, étudiant à l'université de Genève, ont été utilisées pour les analyses. Le recueil a été effectué en janvier 2003.

#### 4.2.10.2.6 Tâche

Tâche de détection (tâche de décision lexicale en « go, no go » - détection de pseudo-mots). Modalité auditive.

#### 4.2.10.2.7 Procédure de recueil des données

Expérience réalisée à l'université de Genève. Détails dans la Partie III, Chapitre 4.

11 \* 18 = 198 mesures utilisées.

La phase test de l'expérience dure 3 ou 4 minutes.

#### 4.2.10.2.8 Méthode d'analyse

- ▷ Détail des procédures pour le tri des données et les analyses dans la Partie III, Chapitre 3.
- ▷ Analyse de régression multiniveaux, modèle à 2 niveaux (sujets, items. Corrélation intra classe : 6.23%).
- ▷ 30 valeurs manquantes (15.15% de non réponses et de valeurs filtrées (détail par condition : Syllabes rares 19.2% ; Syllabes fréquentes 11.1%)).

---

<sup>122</sup> Initialement, des « syllabes » illégales étaient prévues pour représenter l'une des conditions expérimentales (tester l'absence d'appariement comme condition contrôle). Cependant, la distinction entre « syllabes » illégales et syllabes non lexicales du Français pouvant se faire sur la base d'indices phonotactiques, les réponses obtenues auraient présenté une ambiguïté quant au phénomène mesuré. Elles figurent donc comme distracteurs dans le compte rendu de cette expérience, et ne sont pas intégrées dans les analyses.

### 4.2.10.2.9 Paramètres contrôlés

#### À la construction du matériel et à la structuration de l'expérience

- ▷ Les syllabes médiales n'existent pas en position initiale ni en position finale dans le lexique français, en tant que syllabe (selon BRULEX)
- ▷ Les syllabes médiales rares sont appariées aux syllabes médiales fréquentes pour ce qui est leur nombre de phonèmes et leur structure phonétique
- ▷ Les syllabes ne renferment pas de mot enchâssé initial
- ▷ Présentation aléatoire des items lors de la passation

#### Lors des analyses statistiques

- ▷ Les facteurs prosodiques f0 et dB, étant donné leur corrélation avec la V.I. (détail des analyses acoustiques présentés dans l'Annexe 30 page 529).
- ▷ La durée des syllabes
- ▷ Le nombre de mots évoqués par la syllabe (obtenus par enquête via Internet)  
note 121, page 307

Le nombre de phonèmes composant le stimulus n'a pas été pris en considération car les deux conditions sont strictement appariées pour ce facteur.

L'ordre de présentation des items n'a pas été utilisé car il y a trop peu de stimuli analysés (et peu de stimuli dans l'expérience).

La latéralité manuelle et le sexe des participants de même : l'analyse porte sur trop peu de données pour intégrer beaucoup de variables au modèle.

### 4.2.10.3 Résultats obtenus

#### 4.2.10.3.1 Rapport des analyses statistiques

L'équation du modèle (Équation 51) et les tests de significativité relatifs aux covariables peuvent être consultés à l'Annexe 31 page 532.

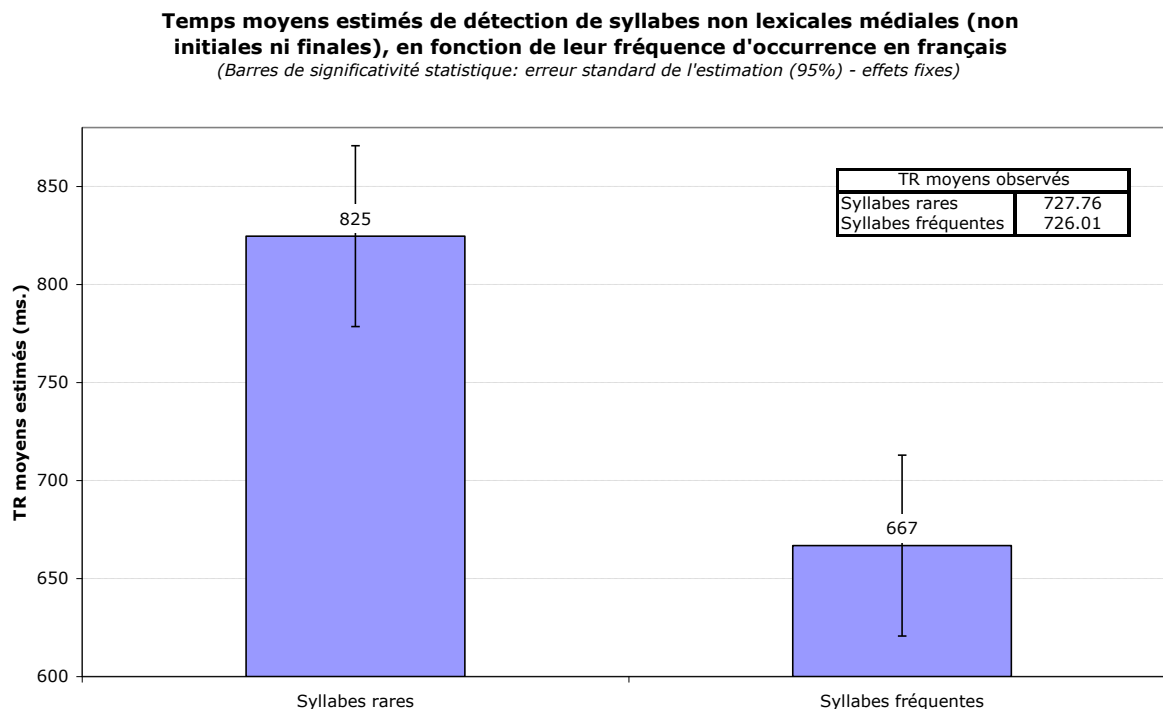
Tableau 88 : Test de significativité du facteur expérimental

	Coefficient	Chi <sup>2</sup> <sub>(1, N=168)</sub>	p<	% variance expliquée
<b>F</b>	-157.802	11.244	<b>.001</b>	5.23

Les syllabes fréquentes sont détectées 158 ms. plus rapidement que les syllabes rares (Tableau 88 ci-dessus).

### 4.2.10.3.2 Présentation des résultats

Figure 68 : Temps de détection moyens estimés des syllabes en fonction de leur fréquence d'occurrence dans le lexique



L'hypothèse de l'Expérience 12 est validée par les résultats, et l'hypothèse sur le niveau de représentation syllabique paraît confirmée. Cependant, des résultats de cet ordre pourraient aussi être obtenus dans le cas où le système est sensible à la fréquence d'occurrence d'une séquence de phonèmes (phénomène de nature phonotactique). Quatre analyses supplémentaires ont testé sur les temps de détection des items (Annexe 32 page 533) :

- ▷ l'effet du nombre de mots renfermant la séquence de phonèmes composant la syllabe expérimentale (la séquence n'est donc pas forcément une syllabe dans ces mots)
- ▷ l'effet du nombre de mots renfermant la syllabe
- ▷ l'effet de la fréquence d'occurrence de la séquence de phonèmes
- ▷ l'effet de la fréquence d'occurrence de la syllabe

Les résultats obtenus impliquent des effets plutôt liés à une forme de représentation spécifiquement syllabique, puisque :

- ▷ Un modèle d'analyse est plus explicatif des données comportementales lorsque les V.I. sont relatives à la syllabe (nombre de mots renfermant la syllabe, fréquence d'occurrence de la syllabe) et non à la séquence de phonèmes (AIC = 10.969 <sup>note 123 page 332</sup>).



- ▷ Le nombre de mots renfermant la syllabe explique mieux les données comportementales ( $\chi^2_{(1, N=168)} = 2.5$  ;  $p < .12$ ) que le nombre de mots renfermant la séquence de phonèmes ( $\chi^2_{(1, N=168)} = 0.7$  ;  $p < .39$ ), même si dans les deux cas aucun effet n'est significatif.
- ▷ Les fréquences d'occurrence de la syllabe et de la séquence de phonèmes ont une influence assez voisine et non significative (respectivement :  $\chi^2_{(1, N=168)} = 0.1$  ;  $p < .74$  -  $\chi^2_{(1, N=168)} = 0.2$  ;  $p < .67$ )
- ▷ L'interaction fréquence \* nombre de mots renfermant l'élément considéré (séquence de phonèmes ou syllabe) est fortement significative dans le cas des syllabes ( $\chi^2_{(1, N=168)} = 7.3$  ;  $p < .007$ ), alors qu'elle ne l'est pas dans le cas des séquences de phonèmes ( $\chi^2_{(1, N=168)} = 2.8$  ;  $p < .1$ ).

Le recensement des mots qui renferment une séquence de phonèmes donnée comptabilise les cas où la séquence correspond à une syllabe plus d'autres cas. Le résultat des analyses qui testent le rôle de la fréquence d'occurrence de la séquence de phonèmes et du nombre de mots dans lesquels cette séquence apparaît est donc relatif à l'influence syllabique plus autre chose. Si l'occurrence statistique d'une séquence de phonèmes était le facteur influençant le comportement des participants, l'effet devrait apparaître aussi avec des syllabes. Par contre, si la syllabe était le facteur influent, l'effet devrait être moins fort ou devrait disparaître avec la V.I. « séquences de phonèmes », car celle-ci introduit du bruit dans l'analyse. C'est ce dernier cas qui semble être mis en évidence par les analyses supplémentaires.

#### **4.2.10.4 Conclusions de l'Expérience 12**

L'hypothèse de l'Expérience 12 est validée par les résultats obtenus. Le traitement des syllabes non lexicales qui ne sont jamais en initiale ni en fin de mot dans le lexique français, est sensible à la fréquence d'usage des syllabes dans le lexique. Ce facteur opposant grossièrement syllabes rares et syllabes fréquentes, est relatif à deux aspects distincts : le nombre de mots qui renferment l'élément considéré et le cumul des fréquences d'occurrence de ces mots. Il est donc nécessaire de départager ces facteurs à un moment donné.

Les analyses supplémentaires effectuées montrent que l'effet syllabique serait plutôt de nature structurale que purement fréquentielle (influence plus importante du nombre de mots renfermant la syllabe que de la fréquence d'occurrence de la syllabe dans le lexique).

L'Expérience 12 a donc permis de confirmer les effets de facilitation qui étaient obtenus avec l'Expérience 11, liés à la fréquence d'occurrence d'une syllabe non lexicale, sur les temps de détection de celle-ci. Elle a précisé que cet effet était bien de nature syllabique, et non de nature phonotactique (telle la fréquence d'occurrence d'une séquence de phonèmes dans le lexique).

#### **4.2.11 Expérience 13 : Détection de syllabes lexicales I – Exploration du rôle de la fréquence positionnelle en initiale de mots**

Les deux expériences précédentes (Expériences 11 et 12) ont montré qu'il existe une forme de représentation des syllabes dans le système. L'Expérience 13 et la suivante (Expérience 14) testent l'effet de fréquence positionnelle pour des syllabes lexicales (mots monosyllabiques existant aussi comme syllabes de mots polysyllabiques). À la différence de l'Expérience 4, qui testait elle aussi l'influence de l'enchâssabilité à l'initiale, mais sans tenir compte de la distribution positionnelle des items dans le lexique, les Expériences 13 et 14 distinguent le cas des syllabes de début de mot (rares ou fréquentes en initiale de mot, dans cette expérience) de celui des syllabes de fin de mot (rares ou fréquentes en fin de mot - Expérience 14).

Par l'introduction de la lexicalité des syllabes dans l'expérience, une situation comparable à celle de l'Expérience 4 (facteur : superposition lexicale initiale virtuelle) est recréée dans le but d'expliquer les effets de facilitation dus aux superpositions lexicales par d'éventuels effets syllabiques.

Toutefois, ces deux expériences (Expérience 13 et Expérience 14) ne suffisent pas pour affirmer que l'origine de l'effet de fréquence positionnelle est syllabique, si un effet apparaît (il peut aussi être lié à la fréquence positionnelle d'occurrence de la séquence de phonèmes). Une série d'analyses supplémentaires des données comportementales obtenues dans ces deux expériences viendra donc compléter la prospection (Analyse exploratoire, Section 4.2.13, page 330).

##### **4.2.11.1 Aspects théoriques**

Si une syllabe non lexicale qui apparaît fréquemment en position initiale dans les mots est reconnue plus rapidement qu'une syllabe non lexicale peu fréquente à cette position (Expérience 11), alors, un mot monosyllabique fréquemment enchâssé à l'initiale d'autres mots devrait également être reconnu plus rapidement qu'un mot monosyllabique peu fréquemment enchâssé à cette position dans d'autres mots.

## 4.2.11.2 Aspects techniques

### 4.2.11.2.1 Facteur expérimental et conditions

Fréquence d'enchâssement du mot monosyllabique comme syllabe initiale dans le lexique français. Deux modalités : rarement en position initiale, fréquemment en position initiale. L'expérience comporte 2 conditions.

### 4.2.11.2.2 Hypothèse opérationnelle

Un mot monosyllabique fréquemment enchâssé comme première syllabe de mot polysyllabique est reconnu plus rapidement qu'un mot monosyllabique rarement enchâssé à cette position en tant que syllabe.

### 4.2.11.2.3 Matériel expérimental utilisé

Stimuli en voix naturelle. La liste des items est donnée en Annexe 33 page 535.

- ▷ Items expérimentaux :
  - ∩ 24 mots monosyllabiques dont la fréquence d'enchâssement en tant que première syllabe de mots polysyllabiques est élevée.
  - ∩ 24 mots monosyllabiques dont la fréquence d'enchâssement en tant que première syllabe de mots polysyllabiques est faible.

Le Tableau 89 ci-dessous fournit une description par condition pour plusieurs dimensions relatives à la fréquence d'enchâssement de la syllabe et du mot.

Tableau 89 : Description des caractéristiques linguistiques formelles moyennes des syllabes par condition

	Facteur expérimental	Facteur corrélé (non contrôlé)	Facteurs corrélés contrôlés		
	Fréquence d'enchâsst. de la syllabe en position initiale	Nombre de mots où la syllabe est à l'initiale	Fréquence Lexicale	Nombre de mots où la syllabe est en position non initiale (milieu + fin)	Fréquence d'enchâsst. de la syllabe en position non initiale
<b>Mot monosyll. fréquent en initiale</b>	33549.3	32.5	26043.8	6.6	7684.3
<b>Mot monosyll. rare en initiale</b>	3878.3	4.5	25449.1	6.7	7804.1

- ▷ Items distracteurs :
  - ∩ 24 syllabes non lexicales fréquentes en tant que première syllabe dans le lexique français (12 sont aussi fréquentes en position finale, 12 sont aussi rares en position finale).

- ∩ 24 syllabes non lexicales rares en tant que première syllabe dans le lexique français (12 sont aussi fréquentes en position initiale, 12 sont aussi rares en position initiale).

L'expérience comporte donc 48 items expérimentaux, 96 items au total.

#### **4.2.11.2.4 Organisation de l'expérience**

Un participant est testé avec les 96 items, une seule fois par item. Il fournit donc 48 réponses utiles.

#### **4.2.11.2.5 Population testée**

Les données de 15 francophones monolingues, sans troubles auditifs, cognitifs et moteurs, étudiant à l'université de Genève, ont été utilisées pour les analyses. Le recueil s'est déroulé en janvier 2003.

#### **4.2.11.2.6 Tâche**

Tâche de détection (tâche de décision lexicale en « go, no go » - détection de mots). Modalité auditive.

#### **4.2.11.2.7 Procédure de recueil des données**

Expérience réalisée à l'université de Genève. Détails dans la Partie III, Chapitre 4.

15 \* 48 = 720 mesures utilisées.

La phase test de l'expérience dure environ 5 minutes.

#### **4.2.11.2.8 Méthode d'analyse**

- ▷ Détail des procédures pour le tri des données et les analyses dans la Partie III, Chapitre 3.
- ▷ Analyse de régression multiniveaux, modèle à 2 niveaux (sujets, items. Corrélation intra classe : 20.33%).
- ▷ 43 valeurs manquantes (5.97% de non-réponses et de valeurs filtrées (détail par condition : syllabe rare en initiale 3.6% ; syllabe fréquente en initiale 8.3%)).

#### **4.2.11.2.9 Paramètres contrôlés**

À la construction du matériel et à la structuration de l'expérience

- ▷ Les syllabes doivent préférentiellement ne pas renfermer de mot enchâssé (ce facteur n'a cependant pas pu être totalement neutralisé, il a fait l'objet d'un

équilibre entre les conditions (6 items expérimentaux renferment un mot enchâssé initial par condition))

- ▷ Le nombre moyen de phonèmes composant les items est voisin dans les deux conditions (2.96 et 3.04)
- ▷ La structure phonologique des syllabes est appariée entre les conditions pour 40 des 48 mots (83.3% - Détail (fréquent / rare) : CV : 3 / 2 ; CVC : 8 / 11 ; CCV : 11 / 8 ; CCVC : 2 / 2 ; CCCV : 1 / 0)
- ▷ La fréquence lexicale est en moyenne voisine pour les deux conditions (Tableau 89, page 319)
- ▷ Le nombre de mots où la syllabe apparaît en position non initiale, de même que la fréquence d'enchâssement de la syllabe en position non initiale, est en moyenne voisine pour les deux conditions (Tableau 89, page 319)
- ▷ Présentation aléatoire des items lors de la passation

#### Lors des analyses statistiques

- ▷ Fréquence fondamentale et intensité moyennes des syllabes (ces facteurs sont corrélés à la variable indépendante. Les résultats de l'analyse sont présentés dans l'Annexe 34 page 536)
- ▷ Durée des syllabes
- ▷ Position acoustique du point de reconnaissance du mot (préféré au début du point d'unicité phonologique car meilleur prédicteur de la variance dans les données)
- ▷ Fréquence lexicale (dispersion autour de la moyenne de la fréquence lexicale 3 fois plus importante pour les syllabes de la condition « fréquente » que pour les syllabes de la condition « rare ». le facteur est donc introduit dans le modèle)
- ▷ Fréquence d'occurrence de la syllabe en position non initiale (mêmes raisons que précédemment)
- ▷ Âge d'acquisition estimé du mot (moyenne par forme phonologique)
- ▷ Familiarité conceptuelle du mot (moyenne par forme phonologique)
- ▷ Latéralité et sexe des participants
- ▷ Ordre de présentation des items

### **4.2.11.3 Résultats obtenus**

#### **4.2.11.3.1 Rapport des analyses statistiques**

L'équation du modèle (Équation 55) et les tests de significativité relatifs aux covariables peuvent être consultés à l'Annexe 35 page 539.

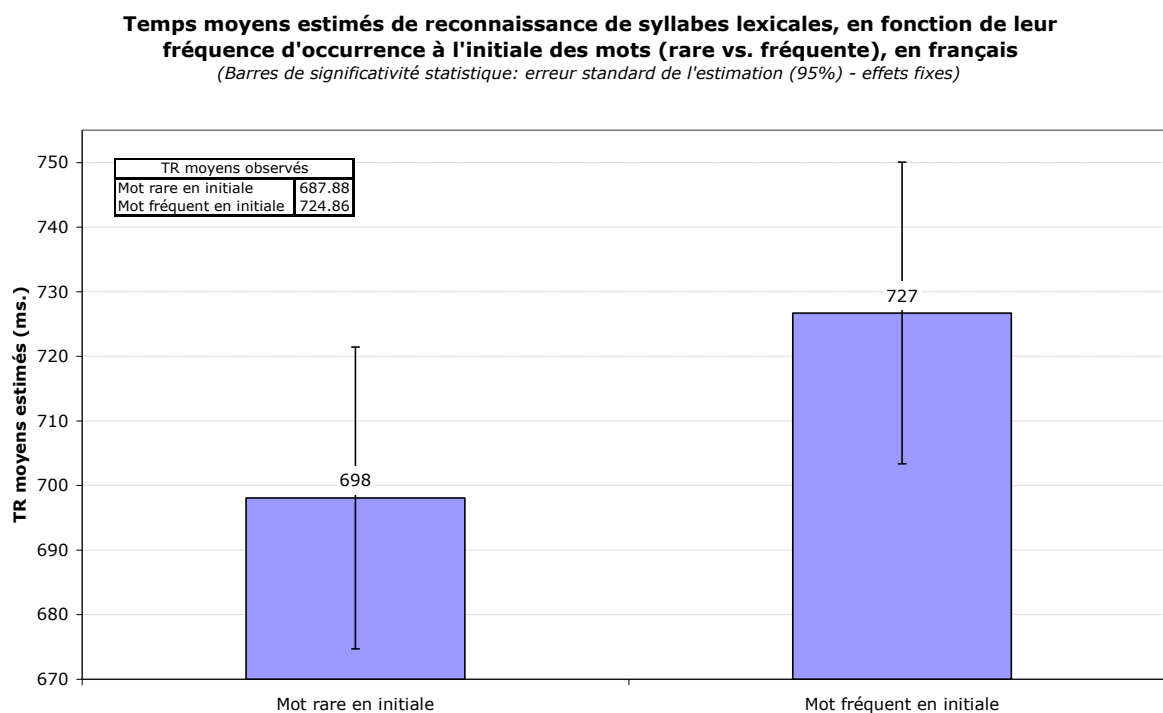
Tableau 90 : Test de significativité du facteur expérimental

	<b>Coefficient</b>	<b>Chi<sup>2</sup><sub>(1, N=677)</sub></b>	<b>p&lt;</b>	<b>% de variance expliquée</b>
<b>FrqI</b>	28.629	1.442	.23	0.125

Les mots monosyllabiques fréquemment enchâssés en position initiale ne sont pas traités à une vitesse différent statistiquement des mots rarement enchâssés à cette position.

#### 4.2.11.3.2 Présentation des résultats

Figure 69 : Temps de reconnaissance des mots en fonction de leur fréquence d'occurrence à l'initiale d'autres mots du lexique



Un mot fréquemment enchâssé en initiale de mot semblerait être plus difficilement reconnu qu'un mot rarement enchâssé en position initiale. Cependant, cet effet n'est pas significatif (Tableau 90, page 321). L'hypothèse n'est donc pas validée par ces résultats.

#### **4.2.11.4 Conclusions de l'Expérience 13**

L'hypothèse de l'Expérience 13 n'a pas été validée par les résultats. Un mot fréquemment enchâssé en tant que syllabe initiale n'est pas reconnu plus vite qu'un mot rarement enchâssé en tant que syllabe à cette position. Le caractère lexical du stimulus pourrait être la cause de la disparition de l'effet de facilitation dû à la fréquence d'occurrence syllabique, qui était observé avec des syllabes non lexicales dans les Expériences 11 et 12. Un phénomène de compétition lexicale tel que supposé par TRACE pourrait aussi avoir masqué un traitement facilitateur de nature infra-lexicale (comme semblent le montrer les résultats de l'Expérience 9 (augmentation de la difficulté à détecter les cibles en initiale de mot en cas de superposition lexicale initiale virtuelle)), ou encore, l'item étant lexical, les traitements de l'unité lexicale auraient pu masquer ceux de l'unité infra-lexicale.

Toutefois, cette absence d'effet apparaît curieuse si on se réfère aux observations fournies par l'Expérience 4, qui manipulait aussi le facteur enchâssabilité pour des mots monosyllabiques et révélait une facilitation de la reconnaissance. La différence de comportement pourrait cependant s'expliquer par la différence de facteurs : l'enchâssabilité selon l'Expérience 13 est relative à une syllabe, tandis qu'elle était relative à une séquence de phonèmes dans l'Expérience 4.



### **4.2.12 Expérience 14 : Détection de syllabes lexicales II – Exploration du rôle de la fréquence positionnelle en non initiale de mots**

L'Expérience 13 a échoué à révéler des effets syllabiques liés à la fréquence d'enchâssement d'un mot en tant que syllabe initiale dans le lexique. Il était cependant prévu de tester aussi l'incidence de la fréquence d'enchâssement d'un mot en tant que syllabe non initiale dans le lexique, afin de comparer systématiquement le rôle que la position d'enchâssement joue dans le traitement, pour des stimuli à la fois syllabiques et lexicaux. L'Expérience 14 est donc quand même effectuée.

#### **4.2.12.1 Aspects théoriques**

Si une syllabe non lexicale non initiale est plus rapidement reconnue lorsqu'elle est fréquente que lorsqu'elle est rare (Expérience 12), il pourrait en être de même pour une syllabe lexicale fréquemment enchâssée en position non initiale.

#### **4.2.12.2 Aspects techniques**

##### **4.2.12.2.1 Facteur expérimental et conditions**

Fréquence d'enchâssement du mot monosyllabique en tant que syllabe non initiale (milieu ou fin de mot) dans le lexique français. Deux modalités : rare en position non initiale, fréquent en position non initiale. L'expérience comporte 2 conditions.

##### **4.2.12.2.2 Hypothèse opérationnelle**

Un mot monosyllabique fréquemment enchâssé en tant que syllabe non initiale dans un mot polysyllabique est reconnu plus rapidement qu'un mot monosyllabique rarement enchâssé en position non initiale en tant que syllabe.

##### **4.2.12.2.3 Matériel expérimental utilisé**

Stimuli en voix naturelle. La liste des items est donnée en Annexe 36 page 541.

▷ Items expérimentaux :

- ∩ 24 mots monosyllabiques fréquemment enchâssés en tant que syllabe non initiale dans des mots polysyllabiques.
- ∩ 24 mots monosyllabiques rarement enchâssés en tant que syllabe non initiale dans des mots polysyllabiques.

Le Tableau 91 ci-dessous fournit une description par condition pour plusieurs dimensions relatives à la fréquence d'occurrence de la syllabe et du mot.

Tableau 91 : Description des caractéristiques moyennes des syllabes par condition

	Facteur expérimental	Facteur corrélé (non contrôlé)	Facteurs corrélés et contrôlés		
	Fréquence d'enchâsst. de la syllabe en position non initiale	Nombre de mots où la syllabe est en position non initiale	Fréquence Lexicale	Nombre de mots où la syllabe est en position initiale	Fréquence d'enchâsst. de la syllabe en position initiale
<b>Mot monosyll. fréquent en position non initiale</b>	19926.4	20.3	19691.1	3.8	4513.6
<b>Mot monosyll. rare en position non initiale</b>	1341.3	2.8	19497.7	3.8	4513.4

▷ Items distracteurs :

Les mêmes items que ceux utilisés dans l'Expérience 13 :

- ∩ 24 syllabes non lexicales fréquentes en tant que première syllabe dans le lexique français (12 sont aussi fréquentes en position finale, 12 sont aussi rares en position finale).
- ∩ 24 syllabes non lexicales rares en tant que première syllabe dans le lexique français (12 sont aussi fréquentes en position initiale, 12 sont aussi rares en position initiale).

L'expérience comporte donc 48 items expérimentaux, 96 items au total.

#### 4.2.12.2.4 Organisation de l'expérience

Un participant est testé avec les 96 items, une seule fois par item. Il fournit 48 réponses utiles.

#### 4.2.12.2.5 Population testée

Les données de 17 francophones monolingues, sans troubles auditifs, cognitifs et moteurs, étudiant à l'université de Genève, ont été utilisées pour les analyses. Le recueil s'est déroulé en janvier 2003.

#### 4.2.12.2.6 Tâche

Tâche de détection (tâche de décision lexicale en « go, no go » - détection de mots). Modalité auditive.

#### 4.2.12.2.7 Procédure de recueil des données

Expérience réalisée à l'université de Genève. Détails dans la Partie III, Chapitre 4.

17 \* 48 = 816 mesures utilisées.

La phase test de l'expérience dure environ 5 minutes.

#### 4.2.12.2.8 Méthode d'analyse

- ▷ Détail des procédures pour le tri des données et les analyses dans la Partie III, Chapitre 3.
- ▷ Analyse de régression multiniveaux, modèle à 2 niveaux (sujets, items. Corrélation intra classe : 3.34%).
- ▷ 85 valeurs manquantes (10.42% de non-réponses et de valeurs filtrées (détail par condition : syllabe rare en fin 8.1% ; syllabe fréquente en fin 12.8%)).

#### 4.2.12.2.9 Paramètres contrôlés

##### À la construction du matériel et à la structuration de l'expérience

- ▷ Les syllabes doivent préférentiellement ne pas renfermer de mot enchâssé (ce facteur n'a cependant pas pu être totalement contrôlé : 8 items expérimentaux renferment un mot enchâssé initial dans la condition « Fréquente », contre 11 dans la condition « Rare », soit une différence de 3 items entre les 2 conditions)
- ▷ Le nombre moyen de phonèmes composant les items est voisin dans les deux conditions (2.79 et 3.21)
- ▷ La structure phonologique des syllabes est appariée entre les deux conditions pour 36 des 48 mots (54.2% - Détail (fréquente / rare) : CV : 7 / 0 ; CVC : 10 / 13 ; CCV : 5 / 6 ; CCVC : 2 / 3 ; CCCV : 0 / 2)
- ▷ La fréquence lexicale est en moyenne voisine pour les deux conditions (Tableau 91, page 325)
- ▷ Le nombre de mots où la syllabe apparaît en position initiale, de même que la fréquence d'occurrence de la syllabe en position initiale, est en moyenne voisine pour les deux conditions (Tableau 91, page 325)
- ▷ Présentation aléatoire des items lors de la passation

##### Lors des analyses statistiques

- ▷ Fréquence fondamentale et intensité moyennes des syllabes (ces facteurs sont corrélés à la V.I. Détail des analyses acoustiques présenté dans l'Annexe 37 page 542).
- ▷ Durée des syllabes
- ▷ Position acoustique du point de reconnaissance du mot (préférée au début du point d'unicité phonologique car meilleur prédicteur de la variance dans les données)

- ▷ Fréquence lexicale (la dispersion de la fréquence lexicale autour de la moyenne est 4 fois plus importante pour les syllabes de la condition « fréquente » que pour les syllabes de la condition « rare »)
- ▷ Fréquence d'occurrence de la syllabe en position initiale (mêmes raisons que précédemment)
- ▷ Âge d'acquisition estimé du mot (moyenne par forme phonologique)
- ▷ Familiarité conceptuelle du mot (moyenne par forme phonologique)
- ▷ Ordre de présentation des items

Remarque : la latéralité manuelle des participants ne figure pas dans le modèle car tous sont droitiers.

### 4.2.12.3 Résultats obtenus

#### 4.2.12.3.1 Rapport des analyses statistiques

L'équation du modèle (Équation 57) et les tests de significativité relatifs aux covariables peuvent être consultés à l'Annexe 38 page 545.

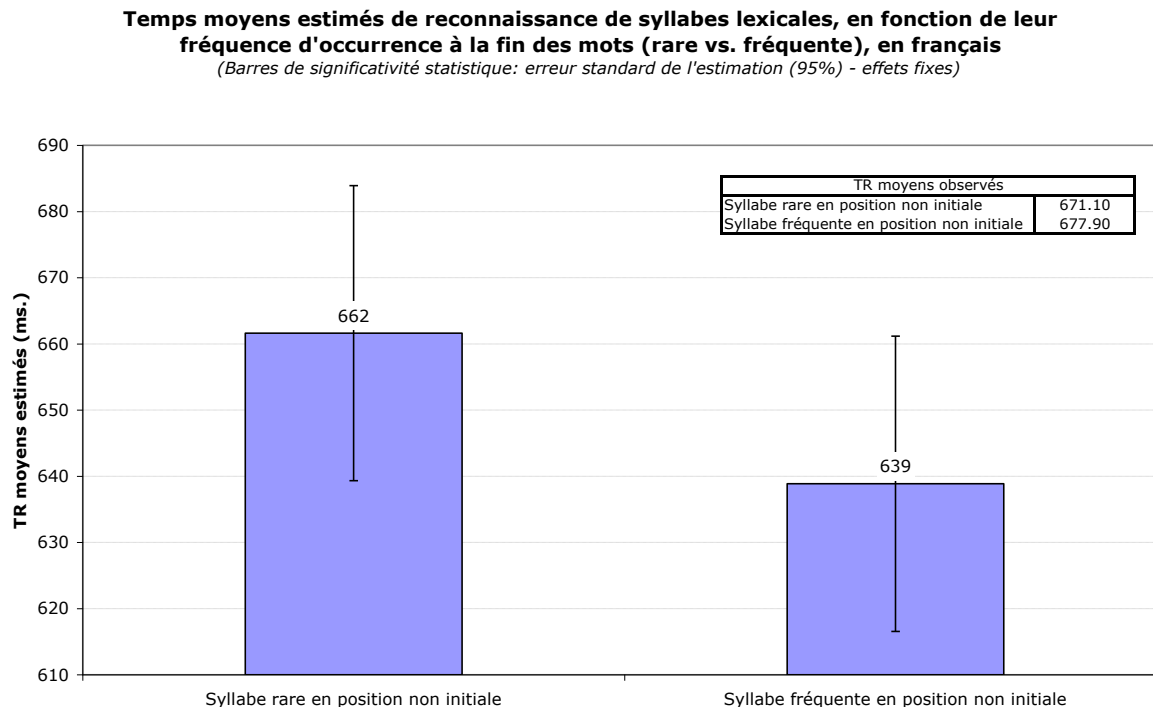
Tableau 92 : Test de significativité du facteur expérimental

	Coefficient	Chi <sup>2</sup> <sub>(1, N=731)</sub>	p<	% variance expliquée
<b>Fréquence d'enchâssement du mot monosyllabique en position non initiale</b>	-22.757	1.263	.27	0.202

La fréquence d'enchâssement en position non initiale d'un mot monosyllabique en tant que syllabe d'autres mots du lexique, n'a pas d'influence significative sur les temps de reconnaissance des items.

### 4.2.12.3.2 Présentation des résultats

Figure 70 : Temps de reconnaissance des mots en fonction de leur fréquence d'occurrence en position non initiale dans le lexique



Un mot monosyllabique fréquemment enchâssé en position non initiale à travers le lexique ne serait pas plus facilement reconnu qu'un mot monosyllabique rarement enchâssé en position non initiale. L'hypothèse n'est donc pas validée par ces résultats.

#### **4.2.12.4 Conclusions de l'Expérience 14**

L'hypothèse de l'Expérience 14 n'a pas été validée par les résultats obtenus. Cette expérience échoue à montrer un effet de la fréquence d'enchaînement d'un mot monosyllabique en tant que syllabe non initiale dans d'autres mots du lexique.

Comme avec l'Expérience 13, les items, du fait de leur lexicalité, pourraient avoir été traités en tant que mots isolés, et non en tant qu'unités syllabiques. Le statut lexical des items pourrait donc avoir masqué les effets infra-lexicaux qui étaient obtenus avec des syllabes non lexicales dans les Expériences 11 & 12.

Une réanalyse des données comportementales des Expériences 13 & 14 à l'aide de nouvelles variables explicatives est proposée afin d'explorer de manière plus approfondie l'hypothèse syllabique, avant de conclure définitivement. L'analyse est présentée à la Section suivante.

### 4.2.13 Analyse exploratoire

Le but de cette Analyse exploratoire, ou réanalyse des données comportementales obtenues dans les Expériences 13 et 14, est de déterminer plus précisément lequel, parmi 4 facteurs, rend le mieux compte de la variance dans les temps de réponse. Ces 4 facteurs sont :

- ▷ La fréquence d'occurrence de la syllabe à une position donnée dans les mots du lexique
- ▷ Le nombre de mots du lexique renfermant la syllabe à une position donnée
- ▷ La fréquence d'occurrence de la séquence de phonèmes à une position donnée dans les mots du lexique
- ▷ Le nombre de mots du lexique renfermant la séquence de phonèmes à une position donnée

#### 4.2.13.1 Aspects techniques

La fréquence d'occurrence des syllabes lexicales a été considérée comme une variable catégorielle (rare / fréquent) dans les deux expériences précédentes. Le problème qui découle d'un tel format est lié à une perte d'information, et donc de précision dans l'évaluation de l'effet. Pour remédier à cet inconvénient, il conviendrait d'utiliser la fréquence d'occurrence dans un format numérique, continu. Toutefois, 3 précautions doivent encore être prises :

1. la fréquence d'occurrence d'une syllabe résulte de la combinaison de deux facteurs :
  - ∩ Un facteur structural : le nombre de mots où la syllabe survient à une position donnée
  - ∩ Un facteur fonctionnel : le nombre de fois où un mot renfermant la syllabe à une position donnée est utilisé

Si l'on souhaite déterminer lequel de ces deux aspects de la fréquence est impliqué dans l'effet de fréquence, il conviendrait de les distinguer.
2. Une séquence de phonèmes ne correspond pas forcément à une syllabe (elle peut être une portion de syllabe ou inclure plus d'une syllabe).
3. Une syllabe peut exister à d'autres positions que la position spécifiquement testée dans chacune des deux expériences. La fréquence d'occurrence à ces autres positions (non initiale dans l'Expérience 13 ; initiale dans l'Expérience 14) peut avoir une incidence sur les temps de reconnaissance du mot monosyllabique. Bien que le matériel ait été équilibré en moyenne par condition pour la position non étudiée (fréquence plus faible que pour la position d'intérêt, rendant l'effet du facteur moins visible, et en moyenne identique dans les 2 conditions), de la variance demeure à travers les items, qui doit être contrôlée.

En tenant compte de tous ces points, les analyses supplémentaires sont organisées de la manière suivante :

- ▷ Pour chaque expérience, séparément, 4 modèles d'analyse de régression multiniveaux sont étudiés. Tous ont la même architecture afin d'en faciliter la comparaison.
- ▷ Chaque modèle considère à la fois l'occurrence (structurale ou fonctionnelle) d'un élément (séquence de phonèmes ou syllabe) en initiale et en non initiale de mot.
- ▷ Chaque modèle examine séparément l'influence sur le comportement (TR), de l'occurrence d'un élément à une position donnée. Ceci aboutit, par croisement complet du type d'occurrence avec le type d'élément considéré, aux 4 modèles d'analyse par expérience, soit un total de 8 modèles pour les deux expériences. Pour chaque expérience, les V.I. suivantes sont donc tour à tour considérées :
  - ∩ La fréquence d'occurrence de la syllabe à une position donnée (structure + fonction)
  - ∩ Le nombre de mots renfermant la syllabe à une position donnée (structure seule)
  - ∩ La fréquence d'occurrence de la séquence de phonèmes à une position donnée (structure + fonction)
  - ∩ Le nombre de mots renfermant la séquence de phonèmes à une position donnée (structure seule)

Le détail des résultats d'analyse est présenté à l'Annexe 39 page 547.

#### 4.2.13.2 Rapport des analyses statistiques

Les 2 tableaux ci-après (Tableau 93 et Tableau 94) récapitulent les effets obtenus pour chaque V.I. examinée pour la position d'intérêt. Ces 2 tableaux fournissent 3 types d'information :

- ▷ Le degré d'adéquation du modèle aux données (IGLS Deviance)
- ▷ Le coefficient estimé pour le facteur (influence sur les temps de réponse, en ms.)
- ▷ La significativité de chaque facteur (probabilité testée par  $\text{Chi}^2_{(1, N_{\text{Exp}13}=677 ; N_{\text{Exp}14}=731)}$ )

Tableau 93 : Résultats fournis par les 4 modèles d'analyse, pour l'Expérience 13

	Syllabe		Séquence de phonèmes	
	Fréquence d'occurrence en initiale	Dans N mots en initiale	Fréquence d'occurrence en initiale	Dans N mots en initiale
« IGLS Deviance » ou degré d'adéquation du modèle aux données	8376.725	8389.884	8396.868	8390.977
Coefficient (p<)	<b>45.1 (.05)</b>	-0.7 (.77)	-2.4 (.95)	<b>-1.3 (.05)</b>



Tableau 94 : Résultats fournis par les 4 modèles d'analyse, pour l'Expérience 14

	Syllabe		Séquence de phonèmes	
	Fréquence d'occurrence en position non initiale	Dans N mots en position non initiale	Fréquence d'occurrence en position non initiale	Dans N mots en position non initiale
« IGLS Deviance » ou degré d'adéquation du modèle aux données	8684.334	8708.209	8707.822	8680.584
Coefficient (p<)	<b>41.2 (.002)</b>	<b>2.9 (.0002)</b>	<b>22 (.033)</b>	<b>2 (.001)</b>

### 4.2.13.3 Présentation des résultats

#### 4.2.13.3.1 Le degré d'adéquation du modèle aux données

L'adéquation des modèles aux données comportementales est en moyenne meilleure dans l'Expérience 13 (8388.61) que dans l'Expérience 14 (8695.24). Cet écart peut provenir d'une variation de précision de la variable indépendante étudiée entre les 2 expériences, puisque dans l'Expérience 13, la variation de fréquence porte sur une seule position (position initiale), tandis que dans l'Expérience 14, la variation de fréquence porte sur plusieurs positions (toutes les autres positions que la première).

Les deux tableaux ci-dessous récapitulent l'information obtenue (« IGLS Deviance » - surligné : meilleure adéquation du modèle aux données).

Tableau 95 : Degrés d'adéquation des modèles aux données pour l'Expérience 13, en fonction du type d'unité considéré et du type d'occurrence dans le lexique

	Fréquence	Nombre de mots
<b>Séquence de phonèmes</b>	8396.868	8390.977
<b>Syllabe</b>	8376.725	8389.884

Tableau 96 : Degrés d'adéquation des modèles aux données pour l'Expérience 14, en fonction du type d'unité considéré et du type d'occurrence dans le lexique

	Fréquence	Nombre de mots
<b>Séquence de phonèmes</b>	8707.822	8680.584
<b>Syllabe</b>	8684.334	8708.209

Techniquement, il n'est pas possible d'évaluer statistiquement la différence entre 2 modèles non emboîtés tels que ceux étudiés. Dans ce type de situation, un critère de comparaison peut néanmoins être calculé, le Critère d'Information de Akaike<sup>123</sup> (AIC). Il

<sup>123</sup> “The Akaike Information Criterion (AIC - Hirotugu Akaike, 1970), is a statistical model fit measure. It quantifies the relative goodness-of-fit of various previously derived statistical models, given a sample of data. The driving idea behind the AIC is to examine the complexity of the model together with goodness of its fit to

permet d'évaluer la différence d'adéquation des modèles aux données, sans toutefois correspondre à un test statistique. Dans le cas présent, «  $p$  » étant identique dans chacun des 4 modèles testés puisqu'ils ont la même structure, la comparaison porte directement sur la valeur de «  $l$  » (degré d'adéquation du modèle aux données) rapportée dans les Tableau 95 et Tableau 96.

On remarque qu'en fonction du type d'unité considéré (Séquence de phonèmes / Syllabe) et du type d'occurrence dans le lexique (Fréquence / Nombre de mots), certains modèles expliquent mieux les données comportementales. Ce sont ceux basés sur la fréquence des syllabes (initiale et non initiale), ce qui plaide en faveur d'un niveau syllabique de représentation, indépendant du niveau lexical (si tel n'était pas le cas, les modèles qui ont utilisé le nombre de mots renfermant la syllabe seraient les plus explicatifs).

Dans le cas des données de l'Expérience 14, le facteur nombre de mots renfermant la syllabe en position non initiale ressort comme l'un des plus significatifs. Il y a donc un phénomène lexical qui se manifeste spécifiquement pour les positions non initiales, qui n'apparaît pas dans le cas des positions initiales (Expérience 13). Ce phénomène peut être en lien avec un processus d'exploration du lexique, relatif à la recherche des débuts de mots et à la segmentation lexicale, qui ne se produit donc pas avec des syllabes alignées.

#### **4.2.13.3.2 L'effet des facteurs**

##### **4.2.13.3.2.1 La syllabe :**

Lorsque la fréquence d'occurrence d'une syllabe lexicale augmente, les temps de reconnaissance augmentent significativement, indépendamment de la position d'occurrence de la syllabe dans les mots du lexique (+45.1 ms. pour la position initiale, +41.2 ms. pour les positions non initiales - Tableau 93 et Tableau 94, pages 331 et 332). Plus une syllabe est fréquente, plus elle doit être prégnante à la perception, donc plus elle doit être facile à reconnaître. Toutefois, puisque les indices de fréquence utilisés renferment à la fois une information relative à la structure du lexique (le nombre de mots qui contiennent la syllabe) et une information relative à l'utilisation du lexique (nombre de fois où un mot est utilisé), plus une syllabe est fréquente et plus il est probable qu'elle

---

the sample data, and to produce a measure which balances between the two." Information compilée à partir du site web "thefreedictionary.com".

$AIC = l + 2p$  ;  $l$  = écart d'IGLS Deviance entre les deux modèles ;  $p$  = différence du nombre de paramètres entre les deux modèles.

évoque un nombre élevé de lexèmes, donc qu'elle induise une compétition lexicale importante (effet de compétition ou effet d'exploration du lexique ?). L'analyse basée sur le nombre de mots renfermant la syllabe ne montre cependant aucun effet de ce facteur pour l'initiale des mots ( $p < .77$  - Tableau 93, page 331), alors qu'un effet existe pour la position non initiale ( $p < .0002$  - Tableau 94, page 332). Celui-ci va également dans le sens d'un allongement des temps de reconnaissance (+2.9 ms. / mot supplémentaire), et comme il est plus significatif que dans le cas de la fréquence de la syllabe (respectivement  $p < .0002$  et  $p < .002$ ), il est possible que le facteur fréquence renferme du bruit. L'absence d'effet du nombre de mots renfermant la syllabe dans le cas des positions initiales laisse une nouvelle fois penser que les débuts de mots sont traités différemment des non débuts. Un phénomène compensant la compétition lexicale a alors peut-être lieu, qui masquerait l'effet observé avec les syllabes lexicales non initiales.

#### **4.2.13.3.2.2 La séquence de phonèmes :**

L'effet de fréquence qui existait avec la syllabe pour la position initiale ( $p < .05$ , IGLS = 8377 - Tableau 93, page 331) n'est plus significatif avec la séquence de phonèmes ( $p < .95$ , IGLS = 8397 - Tableau 93, page 331), tandis qu'il se réduit pour la position non initiale (respectivement  $p < .002$  (IGLS = 8684) /  $p < .033$  (IGLS = 8708) - Tableau 94, page 332). Toute séquence de phonème n'est donc pas traitée de manière équivalente, celles n'étant pas une syllabe introduisant du bruit lorsqu'elles sont considérées dans l'analyse. Ce résultat relatif à la fréquence d'usage appuie donc encore l'hypothèse d'une forme de représentation de la syllabe dans le système.

Concernant le nombre de mots dans lequel la séquence de phonèmes apparaît, on retrouve une différence de comportement en fonction de la position d'enchaînement considérée (initiale vs. non initiale). Une séquence de phonèmes enchâssée à l'initiale d'un nombre élevé de mots est plus facile à traiter (l'effet de l'Expérience 4 est ainsi retrouvé), alors qu'une séquence de phonèmes enchâssée en position non initiale d'un nombre élevé de mots est plus difficile à traiter. On retrouve donc encore une fois une distinction de traitement entre initiales de mots et non initiales.

Il faut remarquer que dans le premier cas, l'adéquation du modèle aux données par rapport à un modèle basé sur le nombre de mots renfermant à l'initiale la séquence en tant que syllabe, n'est pas meilleure (respectivement, IGLS = 8390 et 8391), alors que dans le deuxième cas, elle l'est (respectivement, IGLS = 8708 et 8681). Il peut s'agir une nouvelle fois d'un indice en faveur de l'existence d'un processus d'exploration du lexique, relatif à la recherche des débuts de mots et à la segmentation lexicale, capté par l'utilisation du facteur séquence de phonèmes.

#### 4.2.13.4 Conclusions de l'Analyse exploratoire

Les motifs de résultat fournis par ces analyses sont complexes et difficiles à synthétiser. Trois points importants peuvent néanmoins être dégagés :

- ▷ La syllabe est le facteur qui explique le mieux les données, comparativement à la séquence de phonèmes (effet de la fréquence supérieur pour la syllabe (position initiale et non initiale), « forte » adéquation du modèle aux données pour la fréquence de la syllabe (position initiale et non initiale), effet du nombre de mots renfermant la syllabe fortement significatif pour la position non initiale).
- ▷ Les effets obtenus pour 3 des 4 conditions (Syllabe \* N mots, Séquence \* Fréquence, Séquence \* N mots) varient selon la position d'occurrence considérée (initiale vs. non initiale). L'origine de cette différence semble être liée à un traitement distinct des initiales et des non initiales de mots, comme supposé dans la Section 4.2.13.3.1 (page 332). Bien qu'elle puisse aussi provenir d'une différence de précision des facteurs (pour la position initiale, les syllabes sont toujours des syllabes initiales ; pour la position non initiale, les syllabes peuvent être médiales ou finales), il est fortement probable que tel ne soit pas le cas (effet différent d'une superposition lexicale initiale et non initiale, obtenu dans les Expériences 6 et 7, par exemple). Les unités lexicales pourraient donc présenter une polarité interne utile à leur reconnaissance à partir du début du mot, ce qui adresse question à un système basé sur un format localiste de représentation.
- ▷ L'effet structural du nombre de mots renfermant la séquence de phonèmes, facilitateur dans le cas de la position initiale, inhibiteur dans le cas des positions non initiales, est un indice supplémentaire pour appuyer l'idée d'un traitement particulier des initiales de mots relativement au reste du mot. Ce phénomène suggère l'existence de relations complexes des lexèmes entre eux à travers le lexique mental.
- ▷ Enfin, l'absence d'effet significatif du nombre de mots renfermant la syllabe en position initiale ( $p < .77$ ), bien qu'un effet significatif de compétition existe pour le nombre de mots renfermant la syllabe en position non initiale ( $p < .0002$  - augmentation moyenne des TR de 3 ms. par mot supplémentaire), pourrait indiquer qu'avec les syllabes initiales, l'effet de compétition lexicale est compensé par un phénomène de facilitation, de même nature que celui observé dans les Expériences 4 à 10.

#### 4.2.14 Conclusions des Expériences 11 à 14, et des Analyses supplémentaires et exploratoire

Le but des Expériences 11 à 14, des Analyses supplémentaires et de l'Analyse exploratoire, était de mieux cerner l'organisation du système de traitement, afin de disposer d'indices supplémentaires pour essayer de déterminer la cause des effets de facilitation liés aux superpositions lexicales, qui ont été révélés par les Expériences 4 à 8.

Le Tableau 97 récapitule les effets qui ont été obtenus.

Tableau 97 : Récapitulatif des effets obtenus dans les Expériences 11 à 14, les Analyses supplémentaires et l'Analyse exploratoire

Syllabe non lexicale	Syllabe lexicale
<u>Expérience 11</u> Syllabe fréquente en position initiale : TR $\searrow$ ( $p < .02$ ) Syllabe fréquente en position finale : n.s.	<u>Expérience 13</u> Syllabe fréquente en position initiale : n.s.
<u>Expérience 12</u> Syllabe fréquente en position médiale : TR $\searrow$ ( $p < .0008$ )	<u>Expérience 14</u> Syllabe fréquente en position non initiale : n.s.
<u>Analyses supplémentaires</u> Effet syllabe > effet séquence de phonèmes	<u>Analyse exploratoire</u> Effet syllabe > effet séquence de phonèmes Effet de la fréquence : TR $\nearrow$ Effet du nombre de mots (séquence de phonèmes) : TR $\searrow$ (position initiale), TR $\nearrow$ (position non initiale)

Les résultats des Expériences 11 & 12, ainsi que des Analyses supplémentaires, militent en faveur d'un niveau de représentation syllabique ayant une influence facilitatrice sur les temps de réponse.

Cependant, lorsque les stimuli ont un statut lexical, l'effet de facilitation lié à la syllabe disparaît (Expériences 13 & 14).

L'Analyse exploratoire montre alors que cet effet syllabique est remplacé par un autre phénomène, interprétable comme de la compétition lexicale (les temps de réponse augmentent avec la fréquence de la syllabe). Lorsque les mots monosyllabiques sont enchâssés à l'initiale d'un nombre élevé de mots, ce phénomène de compétition disparaît, peut-être compensé par un autre phénomène, facilitant (Analyse exploratoire). Cette hypothèse est appuyée par l'existence d'un effet de facilitation du nombre de mots renfermant la séquence de phonèmes en initiale dans l'Analyse exploratoire.

Le niveau syllabique serait en interaction forte avec le niveau lexical, puisque les effets observés présentent une variation liée à la position d'occurrence des syllabes dans les mots à travers le lexique, et à la lexicalité des syllabes :

- ▷ pour les syllabes non lexicales, l'effet facilitant est moindre aux positions non initiales qu'en position initiale (Expérience 11)
- ▷ pour les syllabes lexicales, l'effet facilitant est masqué par ce qui pourrait être interprété comme de la compétition lexicale (Expériences 13 et 14, Analyse exploratoire), elle-même masquée par un troisième phénomène, facilitant et de nature lexicale, dans le cas de la position initiale (Analyse exploratoire).

Le niveau phonémique serait peut-être aussi en interaction avec le niveau lexical, puisque les effets d'une séquence de phonèmes varient en fonction de sa position d'occurrence à travers les mots du lexique (Analyse exploratoire). Toutefois, il n'est pas possible de se prononcer sur ce point avec certitude, étant donné qu'une séquence de phonèmes peut aussi être une syllabe.

L'analyse exploratoire révèle donc des subtilités de fonctionnement du système de reconnaissance des mots :

- ▷ Le système distingue le début des mots comme une position particulière (polarisation des représentations lexicales ?)
- ▷ Le système effectue des opérations exploratoires supplémentaires sur les positions non initiales (mise en relation des mots les uns avec les autres ?).

En conclusion, un niveau syllabique d'analyse semble exister lors du processus de reconnaissance des mots parlés. Il est logiquement considéré comme un niveau infra-lexical de représentation. Toutefois, il pourrait correspondre à un processus de traitement basé sur les relations des mots entre eux dans le lexique, plutôt qu'à un niveau de représentation.

Le but de ces 4 expériences, des Analyses supplémentaires, et de l'Analyse exploratoire, était de fournir des éléments de réponse quant à l'origine infra-lexicale de l'effet de facilitation obtenu dans les Expériences 4 à 8 (effet d'une superposition lexicale initiale). La réponse à donner à partir de ces recherches est que si un effet de facilitation infra-lexicale (syllabique) semble bien exister, l'effet de facilitation obtenu dans les Expériences 4 à 8 est peu vraisemblablement d'origine exclusivement infra-lexicale. Une autre piste est proposée lors de la discussion générale.

### 4.2.15 Conclusions relatives aux études présentées dans le Chapitre 2

Tableau 98 : Récapitulatif des résultats obtenus

Expérience n° (modalité)	Type d'item	V.I. (superposition lexicale)	Moment du sondage	V.D. (effet comportemental – TR)		
<b>4 (auditive)</b>	Mot monosyll.	Réelle initiale	Terminal	↘		
<b>5 (auditive)</b>		Virtuelle initiale		↘		
<b>6 (auditive)</b>	Mot bisyll.	Réelle initiale		↘		
<b>7 (auditive)</b>		Réelle finale		n.s.		
<b>8 (visuelle)</b>	Mot monosyll.	Réelle initiale		↘		
<b>9 (auditive)</b>		- Réelle initiale		En cours de traitement		
<b>10 (auditive)</b>		- Virtuelle initiale				
<b>11 (auditive)</b>	- Réelle initiale	- Pas de conflit visible entre mot porteur et mot enchâssé lors de la superposition, facilitation en fin de mot - Conflit dès le début du porteur, résolu avant la fin du stimulus - ↗ en début de mot, ↘ ensuite - n.s. conflit quand superposition initiale + finale = processus de segmentation lexicale				
<b>12 (auditive)</b>	- Réelle finale					
<b>13 (auditive)</b>	- Réelle finale					
<b>11 (auditive)</b>	Pseudo-mot monosyll.	Virtuelle initiale	Terminal	↘		
<b>12 (auditive)</b>		Virtuelle finale		n.s. (↘ tendancielle lointaine)		
<b>Analyse supplmtair.</b>	Mot monosyll.	Virtuelle médiale		↘		
<b>13 (auditive)</b>		Virtuelle initiale		Effets liés à la syllabe, et non à la séquence de phonèmes		
<b>14 (auditive)</b>	Mot monosyll.	Virtuelle non initiale		n.s.	Effet masqué par le caractère lexical des items	
<b>Analyse exploratoire</b>		Mot + pseudo-mot monosyll.		[Fréquence vs. N mots] * [Syllabe vs. Séquence de phonèmes]		
				- Effet syllabique - Différences de traitement entre initiales et non initiales de mots - Traitement particulier de la séquence de phonèmes de début de mot (↘) - Recherche des initiales de mot par séquençage phonémique		

Les études présentées dans le Chapitre 2 permettent de proposer le bilan suivant :

▷ Niveau syllabique de représentation

Un niveau infra-lexical de représentation syllabique semble exister dans le système, à en juger d'après les observations des Expériences 11 et 12, des Analyses complémentaires et de l'Analyse exploratoire (diminution des temps de détection de syllabes fréquentes par rapport aux syllabes rares, effets plus importants lorsqu'une séquence de phonèmes est une syllabe que lorsqu'elle n'est pas une syllabe).

▷ La syllabe comme unité de segmentation lexicale

L'existence d'une superposition lexicale initiale réelle n'entraîne aucune surcharge en calculs dans le système, lors du traitement de syllabes lexicales (Expérience 9). Par contre, l'existence d'une superposition lexicale initiale virtuelle (mots monosyllabiques) produit une surcharge transitoire en calculs dans le système (Expérience 9). Il en va de même lorsque la première syllabe d'un mot bisyllabique est lexicale (surcharge en calculs transitoirement observée en fin de première syllabe - Expérience 10). Par ailleurs, l'effet facilitant de la fréquence d'occurrence de syllabes non lexicales est moins important aux positions non initiales qu'en position initiale (Expérience 11).

Ces observations laissent penser que le niveau syllabique est fortement en interaction avec le lexique. Si la séquence n'est pas segmentable en syllabes (mot monosyllabique - Expérience 9), la présence d'un mot enchâssé initial n'entraîne pas de difficulté supplémentaire, mais si la séquence est segmentable en syllabes (mot plurisyllabique - Expérience 10) ou entre en interaction avec le lexique (mot monosyllabique enchâssable - Expérience 9, syllabe non lexicale enchâssable en position non initiale - Expérience 11), le traitement apparaît alors comme plus difficile. Ce phénomène semble être confirmé par le fait que lorsque les items ont une structure qui conduit le système à une impasse de segmentation (première syllabe lexicale, deuxième syllabe non lexicale), une augmentation importante des temps de réponse se produit (Expérience 10).

▷ Compétition lexicale discrète et secondaire

Les tentatives de ce travail pour mettre en évidence un effet de compétition lexicale sur la reconnaissance d'un mot stimulus, tel que ce phénomène est décrit par le modèle TRACE en cas de superposition lexicale initiale réelle, sont un échec. Toutes les expériences testant ce type de configuration (Expériences 4, 5, 6, 7, 8, 9 et 10) montrent une facilitation, et non une difficulté accrue, de la reconnaissance du mot porteur. Le phénomène est récurrent et robuste, révélant que l'effet de compétition lexicale est finalement plutôt très discret. L'Expérience 9 (détection de cibles non linguistiques) montre qu'il n'y a aucune difficulté supplémentaire du traitement au moment où survient la superposition lexicale, et que la facilitation se produit en fin d'item, conformément aux résultats obtenus avec la tâche de décision lexicale (l'augmentation transitoire de difficulté / complexité du traitement en fin de première syllabe (Expérience 10) semble plutôt être relative au processus de segmentation syllabique / lexicale qu'à un processus de compétition lexicale). L'Expérience 9 révèle toutefois un phénomène qui peut être assimilable à de la compétition lexicale en cas de superposition lexicale initiale virtuelle (cependant, il peut aussi correspondre à un autre traitement que de la compétition lexicale, par exemple l'exploration du lexique). Ce phénomène survient au début du mot (difficulté accrue pour détecter une cible), et disparaît rapidement, bien avant la fin du mot, et sans induire de facilitation comme les superpositions lexicales initiales réelles. Les Expériences 13 et 14 qui testent aussi ce facteur présentent une absence d'effet, ce qui confirme les observations de l'Expérience 9 mais va à l'encontre de ceux de l'Expérience 4 (toutefois, dans cette expérience, la situation est différente (traitement de l'enchâssement sous forme de séquence de phonèmes et non sous forme de syllabe)). On aurait par contre la trace possible d'un phénomène de compétition lexicale (annulation de la facilitation infra-lexicale), retrouvée dans l'Analyse exploratoire pour le facteur « fréquence positionnelle des syllabes » (effet significatif de compétition lexicale). Par conséquent, la compétition lexicale apparaît être un phénomène plutôt relatif et discret, dont l'impact réel sur la reconnaissance d'un mot est difficile à mettre en évidence.

▷ Asymétrie lexicale ou polarisation des représentations lexicales



Le traitement des initiales de mots est favorisé par rapport à celui des non initiales. L'Expérience 11 révèle que les effets de fréquence sont plus importants pour des syllabes non lexicales fréquentes en initiale de mots que pour des syllabes non lexicales fréquentes en position non initiale. L'Analyse exploratoire montre que la reconnaissance d'une séquence de phonèmes est privilégiée proportionnellement au nombre de mots qui la contiennent en initiale, alors qu'elle est défavorisée proportionnellement au nombre de mots qui la renferment en position non initiale. Le lexique est donc structuré de manière à favoriser, et non inhiber, la reconnaissance de mots dont le début est identique (groupes). Le système semble par ailleurs disposer d'une protection contre les segmentations intra lexicales, puisqu'une séquence de phonèmes correspondant à un début de mot fréquemment en position non initiale est plus difficile à détecter (conflits entre groupes de mots ?). Tous ces éléments indiquent que le système réagit différemment aux initiales de mot et aux non initiales de mots, ce qui peut représenter un élément essentiel dans la procédure de segmentation du flux sonore en mots.

▷ Stratégie de recherche des débuts de mots

L'asymétrie lexicale précédemment décrite, couplée à une activation syllabique du lexique, conduit à une favorisation du traitement des initiales de mots souvent en initiale dans les mots du lexique. Toutefois, les initiales qui sont aussi non initiales dans un nombre élevé de mots étant plus difficilement reconnues (facilitation contrecarrée par un processus limitant les segmentations intra lexicales, qui pourrait être de type compétition lexicale), les mots qui commencent par cette séquence de phonèmes doivent donc être plus difficiles à reconnaître sur la base d'une stratégie de détection de début de mot (????: non vérifié).

### **4.3 Chapitre 3 - Exploration de l'influence combinée des aspects supra-segmentaux du signal et architecturaux du lexique mental sur le traitement lexical**

Le premier chapitre de la partie expérimentale était consacré à la recherche d'un code prososyllabique en rapport avec les caractéristiques morphosyntaxiques et la topologie des syllabes, code supposé servir à guider de manière ascendante le traitement de segmentation du flux de parole en mots.

Le deuxième chapitre de la partie théorique était consacré à la recherche de corrélats entre structure du lexique et reconnaissance des mots. Il a essentiellement permis de révéler une contradiction entre théorie et réalité, en montrant que la reconnaissance des mots est facilitée par la présence de superpositions lexicales initiales, et que cette facilitation lexicale est peu probablement d'origine infra-lexicale.

Le troisième chapitre de la partie expérimentale se concentre sur l'aspect perceptif et cognitif de la segmentation du flux de parole en mots, en explorant l'influence combinée des facteurs prosodiques et des aspects structuraux du lexique.

La question centrale abordée dans ce troisième chapitre est de savoir si le guidage ascendant et le guidage descendant de la segmentation du flux de parole en mots agissent linéairement comme deux phénomènes distincts, ou agissent en se potentialisant mutuellement, pour orienter l'élaboration du percept lexical.

Dans ce but, deux expériences sont réalisées (Expériences 15 et 16). Elles testent la cohésion perceptive de stimuli dans lesquels sont manipulés simultanément la prosodie (existence ou non d'une frontière lexicale dans l'item) et la structure superposante de l'item (présence et position des superpositions lexicales).

Le chapitre se termine par une récapitulation succincte des effets obtenus dans ces deux expériences et une brève conclusion.

### 4.3.1 Expérience 15 : Exploration de l'influence de la cohésion acoustique et psychologique d'une chaîne parlée sur sa perception I

Lorsqu'une chaîne de syllabes constitue une séquence formellement ambiguë, c'est à dire pouvant être segmentée en mots de plus d'une façon, la prosodie doit logiquement permettre de lever l'ambiguïté en guidant le système vers la solution, puisqu'elle offre un ensemble d'indices servant à baliser la séquence sonore (Chapitre 1).

L'Expérience 15 teste l'effet d'une modulation de la cohésion acoustique et/ou lexicale de mots bisyllabiques sur l'élaboration du percept lexical.

#### 4.3.1.1 Aspects théoriques

Si une séquence de deux syllabes renferme un marquage prosodique de frontière lexicale entre ses deux syllabes ( $] [$ ), sa cohésion acoustique est inférieure à celle de la même séquence de syllabes dépourvue d'une telle démarcation ( $[ ]$ ).

De la même manière, un mot bisyllabique non segmentable en deux mots monosyllabiques ( $M_1 S_2$ ) présente une cohésion lexicale supérieure à celle d'un mot bisyllabique segmentable en deux mots monosyllabiques ( $M_1 M_2$ ).

Si un mot bisyllabique non segmentable en deux mots présente une cohésion acoustique faible ( $M_1 ] [ S_2$ ), son traitement sera simplement ralenti par rapport à celui du même mot à forte cohésion acoustique ( $[ M_1 S_2 ]$ ). Cet effet sera de nature purement psychophysique.

Par contre, si un mot bisyllabique segmentable en deux mots monosyllabiques présente une cohésion acoustique faible ( $M_1 ] [ M_2$ ), son traitement sera ralenti par rapport à celui du même mot à forte cohésion acoustique ( $[ M_1 M_2 ]$ ), pour des raisons psychophysiques et cognitives. En effet, la marque acoustique de frontière lexicale pourra alors conduire le système à activer les deux mots enchâssés dans le stimulus.

Si on se place dans le cadre d'une hypothèse proposant les effets de compétition comme ayant lieu entre deux alternatives de segmentation, on obtiendra un conflit entre l'alternative  $[ M_1 ] + [ M_2 ]$  et l'alternative  $[ M_1 M_2 ]$ .

Si on se place dans le cadre d'une hypothèse proposant les effets de compétition comme ayant lieu entre alternatives lexicales, on obtiendra un conflit entre  $[ M_1 ]$ ,  $[ M_2 ]$  et  $[ M_1 M_2 ]$ .

Ce degré de précision des hypothèses n'est toutefois pas testable dans le cadre de l'Expérience 15.

Deux autres hypothèses peuvent par contre être envisagées quant à l'impact de la combinaison prosodie / structure lexicale sur la segmentation du flux de parole en mots :

- ▷ Soit la prosodie et la structure lexicale agissent séparément mais additivement (Hypothèse additive)
- ▷ Soit la prosodie est un révélateur de la structure lexicale, cette dernière n'ayant pas d'effet si elle n'est pas potentialisée par des marqueurs suprasegmentaux (Hypothèse potentialisation)

On s'attend à obtenir un schéma de difficulté du traitement variable en fonction de l'hypothèse considérée :

- ▷ « Hypothèse additive » :  $[M_1S_2]$  (pas de ralentissement) <  $M_1][S_2$  (ralentissement psychophysique) =  $[M_1M_2]$  (ralentissement cognitif) <  $M_1][M_2$  (ralentissement psychophysique et cognitif)
- ▷ « Hypothèse potentialisation » :  $[M_1S_2]$  (pas de ralentissement) =  $[M_1M_2]$  (pas de ralentissement cognitif car pas de démarcation prosodique) <  $M_1][S_2$  (ralentissement psychophysique) <  $M_1][M_2$  (ralentissement cognitif rendu possible par la démarcation prosodique)

L'Expérience 15, à travers la manipulation de la cohésion acoustique et/ou lexicale des items, teste donc 3 effets (effet de la cohésion acoustique d'un mot, effet de la cohésion lexicale d'un mot, interaction des deux facteurs) afin de mieux comprendre la dynamique de la segmentation du flux de parole en mots.

### 4.3.1.2 Aspects techniques

#### 4.3.1.2.1 Facteurs expérimentaux et conditions

Deux facteurs à deux modalités chacun :

- ▷ Lexicalité de la deuxième syllabe de l'item (mot ( $M_2$ ) vs. pseudo-mot ( $S_2$ ))
- ▷ Prosodie de l'item (prosodie de mot bisyllabique  $[]$  vs. prosodie de frontière lexicale entre les deux syllabes de l'item  $][]$ )

L'expérience comporte 4 conditions résumées dans le Tableau 99 ci-dessous.

Tableau 99 : Conditions expérimentales

	$[]$	$][]$
$M_2$	$[M_1.M_2]$ /bagaʁ/	$M_1][M_2$ /ba/-/gaʁ/
$S_2$	$[M_1.S_2]$ /banan/	$M_1][S_2$ /ba/-/nan/

### 4.3.1.2.2 Hypothèses opérationnelles

- ▷ Cohésion acoustique :  
lorsque la prosodie du mot stimulus est une prosodie de mot bisyllabique (cohésion acoustique élevée du stimulus - [ ]), la reconnaissance du mot stimulus est plus rapide que lorsque la prosodie marque une frontière lexicale entre les deux syllabes du mot stimulus (cohésion acoustique faible du stimulus - ][).
- ▷ Cohésion lexicale :  
lorsque la deuxième syllabe du mot stimulus est un mot (cohésion lexicale faible du stimulus), les temps de reconnaissance de l'item sont plus longs que lorsque la deuxième syllabe est non lexicale (cohésion lexicale plus élevée du stimulus).
- ▷ Interaction des deux facteurs :  
dans le cas où la cohésion acoustique et la cohésion lexicale agissent séparément sur l'élaboration du percept lexical, on observera une influence additive ou simple de chacun des deux facteurs.  
Dans le cas où la cohésion lexicale est potentialisée par la cohésion acoustique, la condition  $M_1][M_2$  présentera les TR les plus longs, tandis que la condition  $[M_1M_2]$  conduira à des TR du même ordre de grandeur que la condition  $[M_1S_2]$ .

### 4.3.1.2.3 Matériel expérimental utilisé

Stimuli en voix naturelle. La liste des items est donnée en Annexe 40 page 561.

La première syllabe des stimuli est toujours un mot monosyllabique ( $M_1$ ), et leur deuxième syllabe peut être un mot ( $M_2$ ) ou non ( $S_2$ ). En majorité, les syllabes  $S_2$  ne sont pas des syllabes initiales de mots (analyse syllabique de BRULEX, J. Goslin : 31 sur 40, soient 77.5%). Elles peuvent toutefois exister à l'initiale d'un mot en tant que séquence de phonèmes selon BRULEX.

- ▷ 160 items expérimentaux :
  - ∩ 40 mots bisyllabiques dont la première et la deuxième syllabe sont lexicales
  - ∩ 40 mots bisyllabiques dont seule la première syllabe est lexicale
 Ces 80 items sont enregistrés sous 2 formes prosodiques :
  - ∩ Prosodie de mot bisyllabique (unité)
  - ∩ Prosodie de fin de mot bisyllabique et de début de mot bisyllabique (frontière)

La prosodie spécifique est créée naturellement à la production des stimuli, par l'utilisation de séquences de 2 items bisyllabiques. Ces séquences sont toujours produites avec la même prosodie. Dans le cas d'une prosodie de mot bisyllabique, le premier item de la séquence correspond au mot stimulus, le deuxième à un pseudo-mot. Dans le cas d'une prosodie de frontière de mots, 2 pseudo-mots bisyllabiques sont produits, et la séquence [2<sup>ème</sup> syllabe du premier pseudo-mot + 1<sup>ère</sup> syllabe du deuxième pseudo-mot] correspond à l'item.

- ▷ 160 items distracteurs :
    - ∩ 80 pseudo-mots bisyllabiques composés :
      - d'une première syllabe lexicale autre que celles composant les mots expérimentaux (1<sup>ère</sup> ou 2<sup>ème</sup>)
      - de la deuxième syllabe des items expérimentaux
- Ces 80 items sont également enregistrés sous 2 formats prosodiques, selon le calibre décrit pour les items expérimentaux.

#### 4.3.1.2.4 Organisation de l'expérience

Les stimuli sont répartis en 4 listes de 40 items expérimentaux (et 40 items distracteurs), chaque liste comportant 10 mots de chaque condition.

Les stimuli d'une même liste sont appariés par la première syllabe selon que leur deuxième syllabe soit lexicale ou non (exemple : /bagaʁ/-/banan/). Cette structure a été choisie afin de pouvoir faire une comparaison directe du rôle de la deuxième syllabe. Toutefois, un amorçage formel par la première syllabe est alors possible. Une vérification sur l'ensemble des réponses des participants confirme que ce phénomène n'a pas lieu plus fréquemment dans une condition que dans l'autre<sup>124</sup>.

Les 40 items distracteurs renferment les 40 deuxièmes syllabes des items expérimentaux (exemple : /gaʁ/ dans « *bagarre* » et /bjɛgaʁ/). Cette structure avait été initialement choisie pour pouvoir faire une comparaison directe de l'effet de la lexicalité de l'item porteur en vue de la tâche de détection de mot sur la deuxième syllabe (Expérience 16). Un amorçage par la deuxième syllabe des mots expérimentaux par les pseudo-mots distracteurs a donc dû également influencer les réponses des participants. Cependant, ce biais se répartit également de façon équivalente à travers les conditions expérimentales.

Un participant est testé avec 80 items (40 mots – 40 pseudo-mots), une fois par item. Il fournit 40 réponses utiles.

#### 4.3.1.2.5 Population testée

Les données de 56 francophones monolingues, sans troubles auditifs, cognitifs et moteurs, étudiant à l'université de Genève, ont été utilisées pour les analyses. Le recueil s'est déroulé en janvier 2003.

---

<sup>124</sup> Théoriquement, les mots dont la deuxième syllabe est un mot doivent être présentés avant ceux dont la deuxième syllabe n'est pas lexicale dans autant de cas que après, puisque la présentation des items est aléatoire. Dans les faits, 52.86% des items M<sub>1</sub>M<sub>2</sub> sont présentés avant les items M<sub>1</sub>S<sub>2</sub>, ce qui est proche du taux théorique.

#### 4.3.1.2.6 Tâche

Tâche de détection (tâche de décision lexicale en « go, no go » - détection de mots bisyllabiques). Modalité auditive.

#### 4.3.1.2.7 Procédure de recueil des données

Expérience réalisée à l'université de Genève. Détails dans la Partie III, Chapitre 4.

56 \* 40 = 2240 mesures utilisées.

La phase test de l'expérience dure moins de 10 minutes.

#### 4.3.1.2.8 Méthode d'analyse

- ▷ Détail des procédures pour le tri des données et les analyses dans la Partie III, Chapitre 3.
- ▷ Analyse de régression multiniveaux, modèle à 2 niveaux (sujets, items. Corrélation intra classe : 12.72%).
- ▷ 405 valeurs manquantes (18.08% de non-réponses et de valeurs filtrées (détail par condition : [ ]M<sub>2</sub> 10.7% ; [ ]S<sub>2</sub> 17.9% ; ][M<sub>2</sub> 17.7% ; ][S<sub>2</sub> 26.1%)).

#### 4.3.1.2.9 Paramètres contrôlés

À la construction du matériel et à la structuration de l'expérience

- ▷ La majorité des syllabes S<sub>2</sub> n'est pas un début de mot selon BRULEX (77.5%)
- ▷ Format bisyllabique uniquement
- ▷ Mots singuliers, noms, sans e muet, sans gémiation, sans homophones, sans homographes, non morphologiques (pas de préfixation), commençant par une consonne, dont la deuxième syllabe commence par une consonne, dont la séquence S<sub>1</sub>S<sub>2</sub> peut être syllabée sans ambiguïté en S<sub>1</sub> et S<sub>2</sub>
- ▷ Chaque syllabe S<sub>2</sub> ou M<sub>2</sub> est unique dans les items expérimentaux
- ▷ La séquence M<sub>1</sub>M<sub>2</sub> est non signifiante (contre-exemple : dauphin (dos fin), douleur (doux leurre))
- ▷ Même nombre d'items M<sub>1</sub>M<sub>2</sub> avec M<sub>2</sub> ayant des homophones qu'avec M<sub>2</sub> sans
- ▷ Même nombre d'items M<sub>1</sub>M<sub>2</sub> dont la position du point d'unicité (en nombre de phonèmes) est aussi la position du point d'unicité du mot M<sub>2</sub>, que d'items dont la position du point d'unicité est différente de celle du mot M<sub>2</sub>
- ▷ Position moyenne du point d'unicité (en nombre de phonèmes) identique entre M<sub>1</sub>M<sub>2</sub> (5.45) et M<sub>1</sub>S<sub>2</sub> (5.23)
- ▷ Nombre moyen de phonèmes identique entre M<sub>1</sub>M<sub>2</sub> (5.5) et M<sub>1</sub>S<sub>2</sub> (5.53)
- ▷ Nombre moyen de phonèmes identiques entre M<sub>2</sub> (3.025) et S<sub>2</sub> (3.05)
- ▷ Structure des deuxièmes syllabes équilibrée entre conditions M<sub>1</sub>M<sub>2</sub> et M<sub>1</sub>S<sub>2</sub> (respectivement : CV 3 / 3 ; CVC 33 / 32 ; CCVC 4 / 4 ; CVCC 0 / 1)
- ▷ Présentation aléatoire des items lors de la passation

Lors des analyses statistiques

- ▷ Facteurs prosodiques pour les syllabes  $S_1$  et  $S_2$  (fréquence fondamentale moyenne (Hz), intensité moyenne de la syllabe (dB), durée de la syllabe (ms.))
- ▷ Durée du mot stimulus
- ▷ Position acoustique du point de reconnaissance du mot stimulus
- ▷ Position acoustique du point de reconnaissance de la deuxième syllabe
- ▷ Latéralité manuelle des participants
- ▷ Ordre de présentation des items
- ▷ Interactions entre facteurs, en particulier entre facteurs prosodiques et V.I.
- ▷ Dans un premier modèle, la fréquence lexicale des stimuli avait été contrôlée puisqu'une différence moyenne existe entre les mots  $M_1M_2$  (1289.68) et les mots  $M_1S_2$  (528.8). L'effet est non significatif et le facteur n'a donc pas été intégré dans le modèle final à cause de la quantité élevée de termes qu'il renferme.
- ▷ Le facteur liste n'est pas directement contrôlé (se reporter à la note de bas de page n°105, page 201).

**4.3.1.3 Résultats obtenus****4.3.1.3.1 Rapport des analyses statistiques**

L'équation du modèle de régression multiniveaux (Équation 66) et les tests de significativité relatifs aux covariables peuvent être consultés à l'Annexe 41 page 563.

Tableau 100 : Tests de significativité pour les facteurs expérimentaux

	Coefficient	Chi <sup>2</sup> <sub>(1, N=1835)</sub>	p<	% variance expliquée
<b>Absence de frontière</b>	-115.659	7.907	<b>.005</b>	0.514
<b>S<sub>2</sub> lexicale</b>	114.691	3.608	<b>.058</b>	0.028
<b>Frontière * S<sub>2</sub></b>	-129.96	2.528	.12	0.047

Tableau 101 : Tests de significativité des contrastes entre conditions

Contraste		Chi <sup>2</sup> <sub>(1, N=1835)</sub>	p<
[S <sub>1</sub> S <sub>2</sub> ]	[S <sub>1</sub> M <sub>2</sub> ]	0.076	.79
	S <sub>1</sub> ][S <sub>2</sub>	7.907	<b>.005</b>
	S <sub>1</sub> ][M <sub>2</sub>	13.442	<b>.0003</b>
[S <sub>1</sub> M <sub>2</sub> ]	S <sub>1</sub> ][S <sub>2</sub>	6.984	<b>.0083</b>
	S <sub>1</sub> ][M <sub>2</sub>	12.86	<b>.0004</b>
S <sub>1</sub> ][S <sub>2</sub>	S <sub>1</sub> ][M <sub>2</sub>	3.608	<b>.058</b>

Les analyses révèlent :

- ▷ un effet significatif de la cohésion acoustique sur les temps de reconnaissance des mots stimuli. Les facteurs prosodiques ayant été contrôlés aux analyses, en particulier la durée des syllabes et des items, les effets obtenus ici ne sont

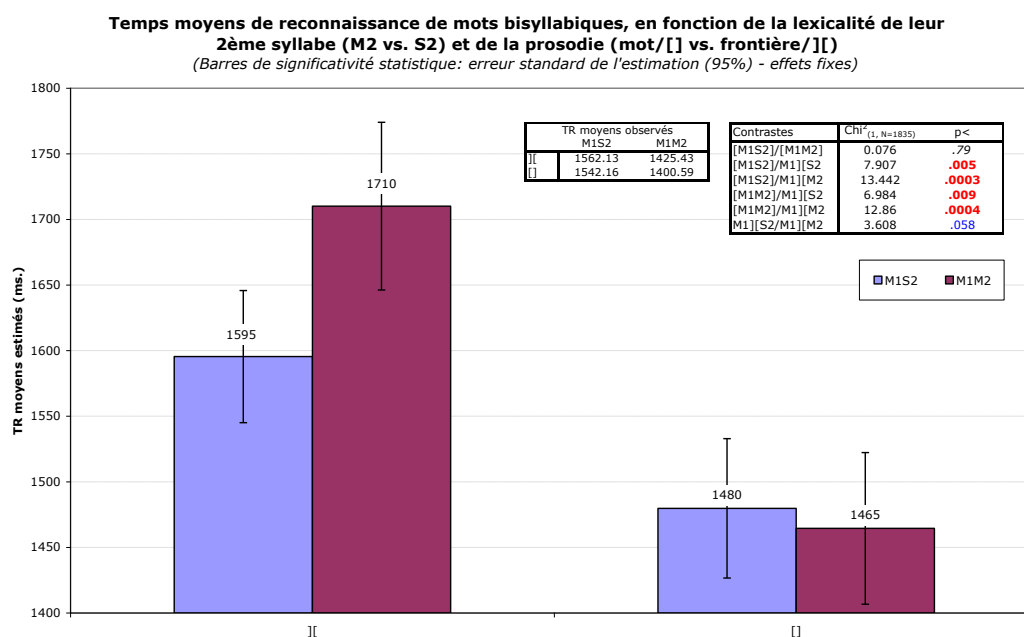


pas relatifs à des variations de durée des stimuli, mais à des variations dans le traitement.

- ▷ une tendance de la cohésion lexicale sur les temps de reconnaissance des mots stimuli
- ▷ pas d'interaction significative entre les deux facteurs, mais une probabilité presque tendancielle

#### 4.3.1.3.2 Présentation des résultats

Figure 71 : Temps de reconnaissance des mots en fonction de la présence d'une frontière lexicale



Les auditeurs sont plus lents pour fournir une réponse de lexicalité avec des items à faible cohésion acoustique, relativement à des items ayant une forte cohésion acoustique (différence moyenne estimée de 116 ms.). L'effet est systématiquement retrouvé entre les conditions M<sub>1</sub>][ ? et les conditions [M<sub>1</sub> ?] (contrastes dans le Tableau 101 page 347). L'hypothèse relative à la cohésion acoustique des items est donc validée par ces résultats.

Les auditeurs ne sont pas systématiquement plus lents pour fournir une réponse de lexicalité avec des items à faible cohésion lexicale. Au contraire, seuls les items marqués d'une frontière prosodique entre les deux syllabes conduisent à un allongement des TR (1710 ms. vs. 1465 ms.), indiquant que la cohésion lexicale des items est révélée par la prosodie dans la dynamique du traitement, mais n'est pas un facteur primordial sinon. L'hypothèse relative à la cohésion lexicale des items n'est donc pas entièrement validée, mais l'« Hypothèse potentialisation » apparaît comme étant correcte.

#### **4.3.1.4 Conclusions de l'Expérience 15**

L'Expérience 15 montre clairement que le traitement de la cohésion lexicale d'un mot ne produit pas systématiquement d'impact sur l'élaboration du percept lexical, comme cela pourrait être supposé à partir d'un modèle comme TRACE. Cela signifie que toute séquence de phonème n'est pas découpée en toute séquence possible de mots.

La cohésion lexicale d'un mot apparaît au contraire comme dépendre fortement de la cohésion acoustique, qui potentialise ou non les possibilités de la structure lexicale. Cela signifie que l'élaboration du percept lexical est fortement contrainte par les facteurs acoustiques, avant de l'être par des facteurs lexicaux de type structure superposante des mots.

L'Expérience 15 fournit donc une preuve supplémentaire en faveur du caractère essentiel du guidage de la segmentation du flux de parole en mots par la prosodie. Elle démontre que l'activation systématique de tous les lexèmes possibles par une séquence de phonèmes, si elle a vraiment lieu, a peu d'impact sur la reconnaissance d'un mot parlé. Il semble au contraire, au vu de tels résultats, que tous les lexèmes possibles ne soient pas activés, mais seulement ceux qui correspondent aux motifs définis par la prosodie.

### **4.3.2 Expérience 16 : Exploration de l'influence de la cohésion acoustique et psychologique d'une chaîne parlée sur sa perception II**

L'Expérience 15 montre qu'un mot stimulus bisyllabique est plus difficile à reconnaître si sa cohésion acoustique est réduite ou si sa cohésion acoustique et sa cohésion lexicale sont réduites. L'Expérience 16 s'attache à montrer la contrepartie de ce phénomène : il est plus facile de segmenter un mot bisyllabique composé de 2 syllabes lexicales lorsque la cohésion acoustique de l'item est moindre. Contrairement à l'Expérience 15, l'Expérience 16 n'explore donc que les effets de la prosodie sur la segmentation du flux de parole en mots.

#### **4.3.2.1 Aspects théoriques**

Lorsqu'un mot bisyllabique dont les deux syllabes sont lexicales est prononcé avec une prosodie de frontière lexicale, il est segmenté en deux mots. La détection du mot  $M_2$  est donc favorisée dans une telle situation, par rapport à une condition où le stimulus est prononcé comme une unité bisyllabique.

#### **4.3.2.2 Aspects techniques**

##### **4.3.2.2.1 Facteurs expérimentaux et conditions**

Un facteur à deux modalités : prosodie de l'item (prosodie de mot bisyllabique [ ] vs. prosodie de frontière lexicale entre les deux syllabes de l'item [ ])

L'expérience comporte donc 2 conditions.

##### **4.3.2.2.2 Hypothèse opérationnelle**

Les temps de reconnaissance de la deuxième syllabe lexicale d'un mot bisyllabique sont plus longs lorsque l'item présente une prosodie de mot bisyllabique, que lorsqu'il présente une prosodie de frontière entre ses deux syllabes.

##### **4.3.2.2.3 Matériel expérimental utilisé**

Stimuli en voix naturelle.

Le même matériel que celui de l'Expérience 15 est utilisé (Annexe 42 page 566).

#### 4.3.2.2.4 Organisation de l'expérience

Les items sont répartis en listes selon la même organisation que dans l'Expérience 15.

Un participant est testé avec tous les items (une fois chaque), mais il ne fournit que 20 réponses utiles.

#### 4.3.2.2.5 Population testée

Les données de 58 francophones monolingues (Groupe 1 : N = 13 ; Groupe 2 : N = 14 ; Groupe 3 : N = 15 ; Groupe 4 : N = 16), sans troubles auditifs, cognitifs et moteurs, étudiant à l'université de Genève, ont été utilisées pour les analyses. Le recueil a été effectué en mai 2002.

#### 4.3.2.2.6 Tâche

Tâche de détection (tâche de décision lexicale en « go, no go » - détection de mots en position de deuxième syllabe dans l'item bisyllabique, impliquant au préalable une analyse par syllabation des items). Modalité auditive.

#### 4.3.2.2.7 Procédure de recueil des données

Expérience réalisée à l'université de Genève. Détails dans la Partie III, Chapitre 4.

$58 * 20 = 1160$  mesures utilisées.

La phase test de l'expérience dure moins de 10 minutes.

#### 4.3.2.2.8 Méthode d'analyse

- ▷ Détail des procédures pour le tri des données et les analyses dans la Partie III, Chapitre 3.
- ▷ Analyse de régression multiniveaux, modèle à 2 niveaux (sujets, items. Corrélation intra classe : 20.49%).
- ▷ 224 valeurs manquantes (19.31% de non-réponses et de valeurs filtrées (détail par condition : ][ 19.1% ; [] 19.5%)).

#### 4.3.2.2.9 Paramètres contrôlés

Il s'agit des mêmes paramètres que dans l'Expérience 15.

### 4.3.2.3 Résultats obtenus

#### 4.3.2.3.1 Rapport des analyses statistiques

Le modèle de régression multiniveaux étudié (Équation 67, page 568 (Annexe 43)) ne renferme pas d'autres facteurs que la V.I., les facteurs acoustiques, et les facteurs influençant le moment auquel le mot est reconnu (durée du mot et position du point de reconnaissance. Les tests de significativité relatifs aux covariables peuvent être consultés à l'Annexe 43 page 568). Les 2 conditions qui sont opposées concernent en effet les mêmes mots.

Tableau 102 : Test de significativité du facteur expérimental

	Coefficient	Chi <sup>2</sup> <sub>(1, N=936)</sub>	p<	% variance expliquée
<b>Frontière prosodique S<sub>1</sub>][S<sub>2</sub></b>	-376.224	9.129	<b>.0026</b>	0.835

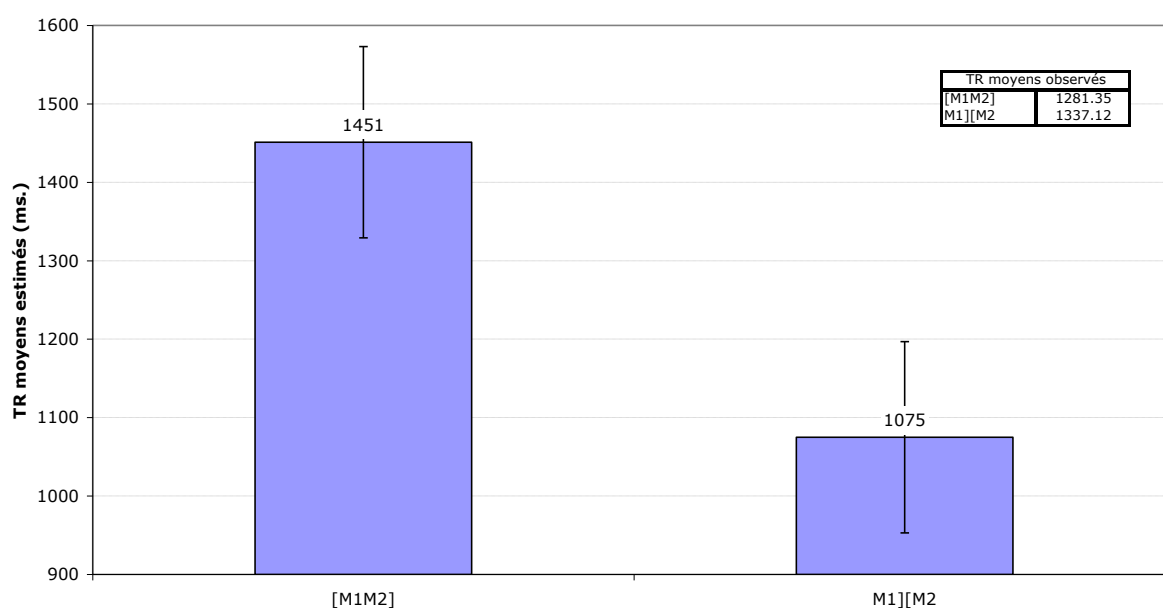
L'analyse révèle un fort effet du facteur « Frontière prosodique » sur le traitement.

#### 4.3.2.3.2 Présentation des résultats

Figure 72 : Temps de détection des mots en position de 2<sup>ème</sup> syllabe dans le stimulus, en fonction de la présence d'une frontière lexicale prosodique entre les syllabes de l'item

**Temps moyens de détection des mots insérés (2ème syllabe), en fonction de la présence d'une marque prosodique de frontière lexicale dans l'item (][)**

(Barres de significativité statistique: erreur standard de l'estimation (95%) - effets fixes)



Les temps de détection des mots enchâssés en position de deuxième syllabe dans des mots bisyllabiques sont en moyenne 376 ms. plus rapides lorsque la prosodie marque une frontière lexicale entre les deux syllabes, que lorsque la prosodie correspond à une unité lexicale bisyllabique. L'hypothèse de l'expérience est donc validée par ces résultats.

#### **4.3.2.4 Conclusions de l'Expérience 16**

L'Expérience 16 montre que la prosodie, en modulant la cohésion acoustique de la séquence sonore, dirige fortement l'accès au lexique. La prosodie favorise ou handicape donc la segmentation du flux sonore en mots.

Ce résultat, associé aux résultats de l'Expérience 15, confirme le rôle de balisage et de guidage des traitements cognitifs par la prosodie.

### **4.3.3 Conclusions relatives aux études présentées dans le Chapitre 3, Partie IV**

La cohésion perceptive est un concept défini comme le résultat de l'influence combinée de plusieurs facteurs sur la perception d'un objet par le sujet psychologique. Dans les 2 expériences présentées au Chapitre 3, la cohésion perceptive dépend avant tout de la cohésion acoustique des stimuli (prosodie) et de leur cohésion lexicale (superpositions lexicales).

Les deux expériences présentées montrent l'importance capitale de la prosodie sur la dynamique du traitement de la parole. Une séquence phonémique à faible cohésion lexicale n'est pas systématiquement segmentée en plusieurs mots. Elle ne le serait que si la prosodie en donne le signal.

Ceci a des implications très importantes pour les modèles de la reconnaissance des mots parlés. Soit tous les mots possibles à partir d'une séquence de phonèmes ne sont pas activés, ce qui permettrait de réduire le coût de la phase de sélection. Soit tous les mots possibles sont activés à partir de la séquence de phonèmes, mais l'impact de cette activation multiple sur la reconnaissance du mot stimulus est peu important, car la prosodie sert à guider le système pour effectuer la sélection.

Dans les deux cas, un mécanisme de sélection basé sur de la compétition lexicale devient inutile, ce qui expliquerait la difficulté que nous avons eu pour voir un tel phénomène apparaître dans les expériences du Chapitre 2, alors la plupart du temps attribuable à la résolution de « garden path » lexicaux.

## 4.4 Chapitre 4 – Exploration des données issues des Chapitres 2 et 3 : relation entre caractéristiques linguistiques et prosodiques des items, et entre prosodie et temps de réponse

Les résultats apportés par les expériences du troisième chapitre, mais aussi ceux apportés par le travail présenté au premier chapitre, impliquent que le système locuteur encode dans la prosodie une information relative à la manière dont il manipule mentalement le langage, information qui permet de guider le système auditeur pour effectuer le décodage du message.

### 4.4.1 Présentation du contenu et de l'organisation du Chapitre 4

À partir de ces observations, il était intéressant de reprendre les analyses des données expérimentales obtenues dans les Chapitres 2 et 3, déjà à disposition dans les annexes, afin de connaître plus en détail la relation entre facteurs lexicaux (superpositions lexicales) et prosodie, et entre prosodie et comportement.

Ces données n'ayant pas été pleinement exploitées, le Chapitre 4 présente un travail avant tout complémentaire et exploratoire.

Plus spécifiquement, les buts du Chapitre 4 sont :

- ▷ de dresser une synthèse des résultats fournis par les analyses acoustiques du matériel utilisé dans les Expériences 4, 6, 7, 11, 12, 13 et 14<sup>125</sup>, pour déterminer si un rapport constant existe entre représentation des mots dans le lexique et codage acoustique de la prosodie, en fonction de l'existence de superpositions lexicales et de la fréquence statistique d'occurrence des syllabes à une position donnée dans les mots du lexique.
- ▷ de dresser un inventaire des motifs de réactivité des participants face à ces variations prosodiques (variations dans les temps de réponse, en rapport avec les variations prosodiques).
- ▷ de proposer des pistes de recherche éventuelles.

Ce court chapitre est organisé comme suit :

---

<sup>125</sup> Expériences 1, 2, et 3 : non utilisés dans ces analyses ; Expérience 5 : aucun effet de la prosodie ; Expérience 8 : modalité visuelle ; Expériences 9 et 10 : même matériel que les Expériences 4 et 7 respectivement.



- ▷ Présentation des effets sous une forme synthétique :
  - ∩ récapitulation des relations entre paramètres prosodiques ( $f_0$ , dB et durée) et facteurs lexicaux (Section 4.4.2.1 : superposition lexicale initiale (réelle ou virtuelle) ou non initiale ; Section 4.4.2.2 : fréquence positionnelle d'occurrence des syllabes). Les tableaux de ces sections ne présentent que les modulations liées aux facteurs lexicaux sur les paramètres prosodiques, dans la dimension de chacun de ces derniers (coefficients estimés, fournis par le modèle). Les valeurs de probabilité ne sont pas fournies afin d'alléger la présentation (se reporter aux annexes pour les détails : Annexe 12 page 494, Annexe 17 page 502, Annexe 20 page 510, Annexe 27 page 523, Annexe 30 page 529, Annexe 34 page 536, et Annexe 37 page 542), mais une indication sur la significativité de l'effet figure grâce à la nomenclature utilisée dans ce manuscrit pour la notification des résultats statistiques (gras&rouge = significatif à  $p \leq .05$ , normal&bleu = tendanciel ( $.05 < p \leq .1$ ), italique&noir = non significatif ( $p > .1$ )).
  - ∩ présentation de la relation entre temps de réponse des participants et  $f_0$  moyenne, intensité moyenne, et durée de la syllabe. Il s'agit de mettre en parallèle les modulations de la prosodie (en lien avec les facteurs lexicaux manipulés), avec une éventuelle modification de la réactivité des participants.
- ▷ Discussion des résultats.
- ▷ Conclusion

## 4.4.2 Présentation des effets

### 4.4.2.1 Effets liés à l'implication des mots stimulés dans des superpositions lexicales

#### 4.4.2.1.1 Expérience 4

##### 4.4.2.1.1.1 Rapport entre facteurs lexicaux et prosodie

Tableau 103 : Effet d'une superposition lexicale initiale réelle (E) et/ou virtuelle (I) sur les paramètres prosodiques, dans le cas de mots monosyllabiques

	<b>E</b>	<b>I</b>
<b>f0 (Hz)</b>	<b>8.9</b>	<b>-3.3</b>
<b>Intensité (dB)</b>	<b>-0.3</b>	<b>-4.9</b>
<b>Durée (ms.)</b>	<b>-79</b>	<b>100</b>

Une superposition lexicale initiale réelle se traduit par une augmentation de la f0 moyenne du mot stimulus, et par une augmentation du débit d'élocution.

Une superposition lexicale initiale virtuelle se traduit par un motif prosodique inverse, avec en plus, une diminution de l'intensité moyenne.

##### 4.4.2.1.1.2 Rapport entre prosodie et réactivité des participants

La f0 est en relation avec les temps de reconnaissance des mots stimulés (allongement des temps de reconnaissance de 7.3 ms. par Hz. supplémentaire –  $p < .0001$ ), alors que l'intensité est sans incidence sur la réactivité des auditeurs ( $p < .95$ ).

Une superposition lexicale initiale réelle est donc responsable d'une augmentation des temps de réponse des participants à cause de l'augmentation associée de f0 (de 65 ms.), alors qu'une superposition lexicale initiale virtuelle conduit à une diminution des temps de réponse des participants à cause de la diminution associée de f0 (de 24 ms.).

L'influence de la durée sur les temps de réponse des auditeurs est marginale (allongement de 0.15 ms. par ms. supplémentaire de durée –  $p < .08$ ).

Une superposition lexicale initiale réelle est donc responsable d'une diminution statistiquement marginale des temps de réponse des participants (12 ms.) à cause de l'augmentation associée du débit d'élocution, alors qu'une superposition lexicale initiale virtuelle est responsable d'une augmentation statistiquement marginale des temps de réponse des participants (15 ms.) à cause d'un ralentissement du débit du locuteur.

En cas de superposition lexicale initiale réelle, on a donc une augmentation moyenne estimée des temps de réponse de 53 ms. à cause des caractéristiques de la prosodie.

En cas de superposition lexicale initiale virtuelle, on a donc une augmentation moyenne estimée des temps de réponse de 9 ms. à cause des caractéristiques de la prosodie.

#### **4.4.2.1.1.3 Conclusions pour l'Expérience 4**

En présence d'une superposition lexicale initiale réelle ou virtuelle, le locuteur induit un ralentissement du traitement perceptif chez l'auditeur à cause des caractéristiques prosodiques qu'il encode :

- ▷ par le biais de la mélodie :
  - ∩ il ralentit le traitement des mots stimuli impliqués dans une superposition lexicale initiale réelle.
  - ∩ il augmente la vitesse de traitement des mots stimuli impliqués dans une superposition lexicale initiale virtuelle.
- ▷ par le biais de la durée de ses productions :
  - ∩ il donne plus rapidement l'information décrivant le mot stimulus au système récepteur dans le cas d'une superposition lexicale initiale réelle (où le risque d'activer le mot enchâssé avant le mot porteur est élevé). En agissant ainsi, le locuteur fournit finalement une information acoustique ayant une résolution moindre au système récepteur. Cette stratégie permettrait de diminuer le risque d'activer le mot enchâssé initial à la place du mot porteur (réduction de la durée d'activation transitoire du mot enchâssé, ce dernier ne parvenant donc peut-être pas à un niveau d'activation suffisant pour être reconnu, ou pour interférer fortement avec le mot porteur). Cette stratégie pourrait aussi révéler un degré de difficulté de traitement moindre avec ce type de configuration (si le traitement est plus facile du fait de la présence d'un enchâssement initial, l'information peut être délivrée plus rapidement).
  - ∩ Il donne moins rapidement l'information décrivant le mot stimulus au système récepteur dans le cas d'une superposition lexicale initiale virtuelle, afin que le système récepteur accorde plus de poids dans l'espace perceptif au mot stimulus, et départage ce dernier des mots concurrents (ceux qui commencent par le mot stimulus). En procédant ainsi, le système dispose de plus de temps pour résoudre d'éventuels conflits dans le lexique. Il travaille par ailleurs à partir d'une information acoustique ayant une résolution meilleure.

Le locuteur tend donc à minimiser la prégnance perceptive du mot enchâssé et à maximiser celle du mot enchâssable dans l'esprit de l'auditeur, grâce au débit de ses productions, de manière à ce que le système dispose d'information pour lever l'ambiguïté lors du traitement de la séquence sonore.

#### 4.4.2.1.2 Expérience 5

Aucun effet de la présence d'une superposition lexicale initiale réelle n'est observé sur les facteurs prosodiques, pour des mots monosyllabiques phonologiquement appariés entre conditions. Ils sont donc aussi acoustiquement appariés, fait certainement lié à la méthode d'enregistrement des stimuli (production successive des stimuli appariés, par souci du locuteur de produire des items acoustiquement similaires), qui a pu gommer le naturel de la prononciation.

#### 4.4.2.1.3 Expériences 6 et 7

##### 4.4.2.1.3.1 Rapport entre facteurs lexicaux et prosodie

Tableau 104 : Effet de la lexicalité de la première ( $M_1$ ) et/ou de la deuxième syllabe ( $M_2$ ) d'un mot bisyllabique, sur les paramètres prosodiques caractérisant chacune des 2 syllabes du mot ( $S_1$  et  $S_2$ ).

		$M_1$		$M_2$	
		$S_1$	$S_2$	$S_1$	$S_2$
<b>f0 (Hz)</b>	<b>Exp. 6</b>	<b>-2.2</b>	<b>-0.4</b>	<b>-2.6</b>	<b>0.8</b>
	<b>Exp. 7</b>	<b>-3</b>	<b>-2.4</b>	<b>-1.8</b>	<b>-0.3</b>
<b>Intensité (dB)</b>	<b>Exp. 6</b>	<b>3.8</b>	<b>2.7</b>	<i>-0.01</i>	<b>1.9</b>
	<b>Exp. 7</b>	<b>3.8</b>	<b>5</b>	<b>-1.7</b>	<b>2.3</b>
<b>Durée (ms.)</b>	<b>Exp. 6</b>	<b>-69</b>	<b>47</b>	<i>0.7</i>	<b>-43</b>
	<b>Exp. 7</b>	<b>-46</b>	<b>41</b>	<b>9</b>	<b>-35</b>

En cas de superposition lexicale initiale réelle ( $S_1$  est un mot), les locuteurs :

- ▷ diminuent la f0 de la première syllabe
- ▷ augmentent son intensité
- ▷ diminuent sa durée

En cas de superposition lexicale finale réelle ( $S_2$  est un mot), les locuteurs :

- ▷ augmentent ou diminuent la f0. Le motif est donc inconsistant.
- ▷ augmentent l'intensité de la deuxième syllabe
- ▷ diminuent sa durée

On remarque donc que lorsqu'une syllabe est lexicale, le locuteur la prononce plus rapidement et plus fort que lorsqu'elle n'est pas lexicale.

##### 4.4.2.1.3.2 Rapport entre prosodie et réactivité des participants

- ▷ Première syllabe :

- ∩  $f_0$  : sans incidence sur les temps de réaction des auditeurs ( $p < .61$ ) dans l'Expérience 6, mais son augmentation de 1 Hz. produit une augmentation moyenne de 4.2 ms. dans les temps de réponse, dans l'Expérience 7 ( $p < .0001$ ). Lorsque la première syllabe est un mot, les temps de réponse des participants diminuent donc en moyenne de 13 ms. à cause de la  $f_0$ .
- ∩ intensité : son augmentation de 1 dB. produit une augmentation statistiquement tendancielle ( $p < .08$ ) des temps de réponse des auditeurs (2.1 ms.) dans l'Expérience 6, et une augmentation significative dans l'Expérience 7 (3 ms. -  $p < .0004$ ). À cause de l'augmentation de l'intensité de la première syllabe lorsqu'elle est lexicale, les auditeurs présentent des temps de réponse plus longs de 11 ms.
- ∩ durée : sa diminution de 1 ms. produit une diminution des temps de réponse des auditeurs dans les deux expériences (respectivement de 0.7 et 1.8 ms., soit une diminution totale respective de 48 ms. et 83 ms. à cause de la durée, lorsque la première syllabe est lexicale -  $p < .0001$ ).

Globalement, lorsque la première syllabe est lexicale, les temps de réponse diminuent à cause de l'influence des paramètres prosodiques (48 ms. dans l'Expérience 6, 85 ms. dans l'Expérience 7).

▷ Deuxième syllabe :

- ∩  $f_0$  : son augmentation de 1 Hz. est responsable d'une augmentation moyenne de 5.8 ms. des temps de réponse dans l'Expérience 6 ( $p < .0007$ . Les temps de réponse augmentent donc en moyenne de 5 ms. à cause de l'augmentation de  $f_0$  lorsque la 2<sup>ème</sup> syllabe est lexicale) et d'une augmentation tendancielle dans l'Expérience 7 (1.3 ms. -  $p < .07$ . Les temps de réponse ne varient donc pratiquement pas lorsque la 2<sup>ème</sup> syllabe est lexicale, puisque la variation moyenne de  $f_0$  est minimale).
- ∩ intensité : son augmentation de 1 dB. est responsable d'un allongement moyen des temps de réponse de 3.2 ms. dans l'Expérience 6 ( $p < .05$ . Les temps de réponse augmentent en moyenne de 6 ms. à cause de la modulation d'intensité lorsque la 2<sup>ème</sup> syllabe est lexicale), mais est sans effet dans l'Expérience 7 ( $p < .84$ ).
- ∩ durée : sa diminution de 1 ms. s'accompagne d'une diminution moyenne des temps de réponse de 0.3 ms. dans l'Expérience 6 ( $p < .0002$ . Les temps de réponse diminuent donc en moyenne de 13 ms. à cause de la durée lorsque le 2<sup>ème</sup> syllabe est lexicale) mais est sans effet dans l'Expérience 7 ( $p < .9$ ).

À cause de la prosodie du mot stimulus lorsque sa 2<sup>ème</sup> syllabe est lexicale, les temps de réponse des participants diminuent donc en moyenne de 2 ms. dans l'Expérience 6, mais ne varient pas dans l'Expérience 7.

Les motifs de réactivité des participants à chaque paramètre prosodique apparaissent à première vue comme variables d'une expérience à l'autre, alors que les motifs de production sont assez similaires entre les deux expériences. Pourtant, on remarque des phénomènes perceptifs constants :

- ▷ augmentation des temps de réponse liés à l'augmentation de la  $f_0$  de la syllabe
- ▷ augmentation des temps de réponse liés à l'augmentation de l'intensité de la syllabe

- ▷ diminution des temps de réponse liés à l'augmentation du débit de production de la syllabe.

#### **4.4.2.1.3.3 Conclusions pour les Expériences 6 et 7**

Avec les paramètres durée et intensité, le système locuteur semble spécifiquement marquer le caractère lexical des syllabes.

Comme dans le cas de mots monosyllabiques (Expérience 4), le système délivre plus rapidement l'information acoustique lorsqu'une syllabe du mot porteur bisyllabique est un mot que lorsqu'elle n'est pas un mot (en d'autres termes, il donne moins d'information acoustique pour décrire les portions du mot stimulus qui peuvent causer une ambiguïté dans le traitement). Parallèlement à cela, il alerte le système en augmentant l'intensité des syllabes lexicales, peut-être afin de signifier une difficulté éventuelle. Ces phénomènes permettraient de moduler la dynamique du traitement en minimisant l'importance perceptive des mots enchâssés, tel que cela a été décrit dans la Section 4.4.2.1.1.3, page 358 (Conclusions pour l'Expérience 4). Le fait que les modulations de durée de la première syllabe aient des conséquences supérieures à celles de la deuxième syllabe sur la vitesse du traitement, indique que pour le traitement de mots isolés, le guidage prosodique sur l'initiale du mot est important (il se peut qu'en situation de phrase, le guidage prosodique sur la fin du mot soit important également). Cette importance du début des mots peut être retrouvée au niveau acoustique, puisque les modulations mélodiques et intensitaires sont moins importantes pour la deuxième syllabe du mot que pour la première. Le système pourrait donc marquer plus fortement les initiales de mot, qui représentent des régions cruciales pour le traitement perceptif.

On retrouve cela dans la réaction des auditeurs : leurs temps de réponse diminuent lorsque la première syllabe est lexicale, à cause des modulations prosodiques, alors que ces modulations prosodiques n'entraînent aucune modification des temps de réponse lorsque la deuxième syllabe est lexicale.

## 4.4.2.2 Effets liés à la position statistique d'occurrence des syllabes dans les mots du lexique

### 4.4.2.2.1 Expérience 11

#### 4.4.2.2.1.1 Rapport entre facteurs lexicaux et prosodie

Tableau 105 : Effet de la fréquence d'occurrence dans les mots du lexique, de syllabes non lexicales en position initiale (FqI) et/ou en position finale (FqF), sur les paramètres prosodiques

	FqI	FqF
<b>f0 (Hz)</b>	<b>-1.4</b>	<b>1.8</b>
<b>Intensité (dB)</b>	<b>-0.4</b>	<b>0.4</b>
<b>Durée (ms.)</b>	<b>-69</b>	<b>-22</b>

Les syllabes non lexicales fréquentes sont prononcées avec :

- ▷ une f0 et une intensité moins élevée pour la position d'occurrence en initiale de mot
- ▷ une f0 et une intensité plus élevée pour la position non initiale
- ▷ un débit plus rapide que pour les syllabes rares. Les syllabes fréquentes en position initiale sont prononcées plus rapidement que les syllabes fréquentes en position finale.

#### 4.4.2.2.1.2 Rapport entre prosodie et réactivité des participants

Une augmentation de la f0 moyenne des syllabes de 1 Hz. est responsable d'une augmentation moyenne des temps de réponse des auditeurs d'environ 46 ms. ( $p < .0003$ ).

L'augmentation d'intensité moyenne des syllabes ne produit pas de modification significative dans le comportement des auditeurs ( $p < .11$ ).

Le débit d'élocution est inversement proportionnel à la durée nécessaire au participant pour répondre (les temps de réponse diminuent de 3 ms. par ms. supplémentaire de durée des syllabes –  $p < .0007$ ).

On a donc une augmentation moyenne des temps de réponse des participants, à cause des modulations prosodiques, de 143 ms. pour les syllabes d'initiale de mot et de 145 ms. pour les syllabes de fin de mot, lorsque ces syllabes sont fréquentes à ces 2 positions respectives.

#### **4.4.2.2.1.3 Conclusions pour l'Expérience 11**

Le motif de f0 produit par le locuteur et relatif à la fréquence d'occurrence des syllabes est inversé en fonction de la localisation de la syllabe dans les mots du lexique, et la réactivité des auditeurs suit cette variation. On a donc une diminution des temps de réponse de 64 ms. à cause de la f0 pour les syllabes fréquentes en position initiale, et une augmentation des temps de réponse de 83 ms. à cause de la f0 pour les syllabes fréquentes en position non initiale.

Comme dans le cas des mots renfermant un enchâssement initial (Expérience 4) ou des syllabes lexicales (Expériences 6 et 7), une augmentation du débit d'élocution est retrouvée avec les syllabes fréquentes, ayant un effet sur les temps de traitement par les auditeurs. Ce phénomène pourrait indiquer que plus le traitement perceptif est facile, moins le système locuteur prend de peine pour fournir une information acoustique détaillée. Néanmoins, cette diminution de la durée semble rendre le traitement plus difficile, puisque les temps de réponse des auditeurs s'allongent proportionnellement à l'augmentation du débit des locuteurs (de 207 ms. dans le cas des syllabes fréquentes en position initiale, de 66 ms. dans le cas de syllabes fréquentes en position finale).

On voit donc se dessiner, sur l'ensemble des paramètres prosodiques, un motif complexe de gestion de la dynamique du traitement, avec des paramètres augmentant ou diminuant la vitesse du traitement, et d'autres agissant simultanément de façon contraire. Ceci laisse penser que le niveau auquel chaque paramètre a une influence, et les modalités de cette influence, sont spécifiques à chaque paramètre.

Le résultat comportemental global de l'ensemble des variations des 3 paramètres prosodiques se traduit cependant de la même manière, par un allongement des temps de réponse pour les syllabes non lexicales fréquentes.



#### 4.4.2.2.2 Expérience 12

##### 4.4.2.2.2.1 Rapport entre facteurs lexicaux et prosodie

Tableau 106 : Effet de la fréquence d'occurrence dans les mots du lexique, de syllabes non lexicales en position médiale, sur les paramètres prosodiques

	Syllabe fréquente (réf = rare)
<b>f0 (Hz)</b>	-0.4
<b>Intensité (dB)</b>	-0.6
<b>Durée (ms.)</b>	<b>17</b>

Lorsqu'une syllabe médiale non lexicale est fréquente, elle est produite avec une f0 légèrement plus basse que lorsqu'elle est rare, de même que pour l'intensité (tendances, respectivement :  $p < .08$  et  $p < .1$ ). Elle est par contre produite avec une durée plus importante (diminution du débit -  $p < .03$ ).

##### 4.4.2.2.2.2 Rapport entre prosodie et réactivité des participants

- ▷ Plus la f0 est haute, plus les temps de réponse sont allongés ( $p < .0002$ . Les temps de réponse diminuent donc de 80 ms. à cause de la f0 lorsque la syllabe est fréquente).
- ▷ La relation entre intensité et temps de réponse n'est pas manifeste ( $p < .11$ ).
- ▷ Plus le débit d'élocution diminue, plus les temps de réponse diminuent (pour 1 ms. supplémentaire, diminution moyenne de 2.5 ms. -  $p < .007$ . Les temps de réponse diminuent de 43 ms. à cause de la durée lorsque la syllabe est fréquente).

À cause de la prosodie, les temps de réponse des participants diminuent donc en moyenne de 123 ms. lorsque la syllabe est fréquente.

##### 4.4.2.2.2.3 Conclusions pour l'Expérience 12

Une syllabe médiale non lexicale est produite de manière à être reconnue plus rapidement lorsqu'elle est fréquente dans les mots du lexique (diminution de sa f0 et augmentation de sa durée), que lorsqu'elle est rare. Ce résultat n'est pas surprenant d'un point de vue logique : les syllabes fréquentes étant plus faciles d'accès, aussi bien en production qu'en perception, leur traitement peut être effectué plus aisément que les syllabes rares.

Ce n'était toutefois pas ce qui était observé dans les résultats de l'Expérience 11, qui considérait des syllabes de frontière lexicale. Il pourrait donc y avoir un traitement particulier sur les frontières des mots, qui ne se produit pas ici (réactivité spécifique à un motif prosodique de la syllabe spécifique ?).

### 4.4.2.2.3 Expériences 13 et 14

#### 4.4.2.2.3.1 Rapport entre facteurs lexicaux et prosodie

Tableau 107 : Effet de la fréquence d'occurrence dans les mots du lexique, de syllabes lexicales en position initiale (FqI, Expérience 13) ou en position non initiale (FqnI, Expérience 14), sur les paramètres prosodiques

	FqI (réf = rare I)	FqnI (réf = rare nI)
<b>f0 (Hz)</b>	<b>5</b>	<b>-1.3</b>
<b>Intensité (dB)</b>	<b>1.2</b>	<b>1.5</b>
<b>Durée (ms.)</b>	<b>-15</b>	<b>-42</b>

- ▷ f0 : ce paramètre permet de distinguer les syllabes lexicales en fonction de leur fréquence positionnelle d'occurrence<sup>126</sup> :
  - ∩ syllabes lexicales fréquentes en position initiale : f0 plus élevée (réf = syllabes lexicales rares en position initiale).
  - ∩ syllabes lexicales fréquentes en position non initiale : f0 moins élevée (réf = syllabes lexicales rares en position non initiale).
- ▷ intensité : une syllabe lexicale fréquente est produite avec une intensité plus élevée qu'une syllabe lexicale rare, quelle que soit sa position d'occurrence.
- ▷ durée : une syllabe lexicale fréquente est produite avec un débit plus rapide qu'une syllabe lexicale rare, quelle que soit sa position d'occurrence.

#### 4.4.2.2.3.2 Rapport entre prosodie et réactivité des participants

La relation entre prosodie et réactivité des participants est peu manifeste, car peu d'effets significatifs des variations prosodiques ont été obtenus sur les temps de reconnaissance :

- ▷ f0 : pas d'effet (Expérience 13 :  $p < .14$  ; Expérience 14 :  $p < .65$ ).
- ▷ intensité : pas d'effet pour la fréquence positionnelle initiale ( $p < .52$ ), augmentation moyenne des temps de reconnaissance de 9.5 ms. par dB. supplémentaire ( $p < 0006$ ) pour la fréquence positionnelle non initiale.
- ▷ durée : diminution moyenne des temps de reconnaissance de 0.6 ms. par ms. en moins de durée des syllabes dans l'Expérience 13 (fréquence positionnelle initiale -  $p < .0002$ ), aucun effet dans l'Expérience 14 (fréquence positionnelle non initiale -  $p < .53$ ).

Seules les modulation d'intensité et de durée des syllabes ont une incidence sur les temps de réponse des auditeurs, l'intensité produisant un effet dans le cas des syllabes lexicales non initiales (son augmentation conduit à une augmentation des temps de réponse de 14 ms.), la durée produisant un effet dans le cas des syllabes lexicales initiales (sa diminution produit une diminution des temps de reconnaissance de 25 ms.).

<sup>126</sup> La fréquence fondamentale moyenne des syllabes rares est équivalente d'une position à l'autre (146.5 Hz. en position initiale, 146.9 Hz. en position non initiale).

On retrouve une asymétrie des traitements en fonction de la position d'occurrence des items dans les mots du lexique.

#### **4.4.2.2.3.3 Conclusions pour les Expériences 13 et 14**

Il semble que le locuteur encode une distinction :

- ▷ entre syllabes lexicales rares et fréquentes. Les syllabes fréquentes sont produites avec une intensité et un débit plus rapides que les syllabes rares.
- ▷ entre syllabes lexicales fréquentes, en fonction de leur position d'occurrence (initiale vs. non initiale). Une syllabe lexicale fréquente en position initiale sera distinguée d'une syllabe lexicale fréquente en position non initiale par une augmentation (vs. une diminution) de la  $f_0$ , et par une modulation qui semble moins forte, de l'intensité et du débit.

L'auditeur n'apparaît pas très sensible aux modulations de  $f_0$  dans ces deux expériences, et ne semble pas réagir aux modulations d'intensité des syllabes de position préférentiellement initiale, ni aux modulations de débit des syllabes de position préférentiellement non initiale. La réactivité des participants aux caractéristiques prosodiques des items est donc différente en fonction de la position d'occurrence préférentielle des syllabes lexicales dans les mots du lexique (forte réaction en cas d'augmentation de l'intensité des syllabes fréquentes en position non initiale (augmentation de la vitesse de traitement de 14 ms.), et en cas d'accélération du débit des syllabes fréquentes en position initiale (diminution de la vitesse du traitement de 25 ms.)).

### 4.4.3 Discussion

Les motifs prosodiques exposés dans ce quatrième chapitre, relatifs aux facteurs lexicaux et à la réactivité des participants, sont complexes, mais ne paraissent pas être aléatoires. Ils sont synthétisés dans la section suivante.

#### 4.4.3.1 Encodage à la production

Lors de la production, des différences acoustiques (prosodie) sont encodées par le locuteur en fonction des facteurs lexicaux manipulés dans les expériences de l'étude.

Concernant la cause de ce phénomène :

- ▷ soit le locuteur (tout locuteur) donne inconsciemment des informations au système récepteur afin de guider le traitement de décodage ultérieur et d'optimiser la communication.
- ▷ soit le locuteur (le producteur des stimuli), connaissant le matériel et les V.I., a inconsciemment produit une différence.

Ce dernier argument doit être considéré, sans pour autant occulter la plausibilité du premier. Il ne faudrait pas ignorer, par exemple, l'ensemble des résultats qui ont été obtenus dans le Chapitre 1 (Partie IV) à partir d'un matériel produit par une locutrice naïve.

##### 4.4.3.1.1 f<sub>0</sub> et intensité

Les motifs qui s'observent dans les modulations de la f<sub>0</sub> et de l'intensité varient souvent pour des facteurs lexicaux identiques (exemple pour la f<sub>0</sub> : augmentation en cas de mot enchâssé initial dans l'Expérience 4 et diminution dans les Expériences 6 et 7). Cependant, aucune règle n'apparaît clairement, peut-être par manque d'exemples (toutes les situations expérimentales sont différentes. Dans l'exemple cité précédemment pour la f<sub>0</sub>, les mots sont monosyllabiques dans l'Expérience 4, bisyllabiques dans les Expériences 6 et 7. Cette variation de format lexical peut être la cause de la différence).

##### 4.4.3.1.2 durée

La prégnance perceptive des mots enchâssés (Expériences 4, 6 et 7) et des syllabes fréquentes (non lexicales (Expérience 11) ou lexicales (Expériences 13 et 14)) est réduite par une augmentation du débit d'élocution, comparativement à des mots sans enchâssement ou à des syllabes rares. Il semblerait que cette minimisation de l'importance perceptive d'une séquence de phonèmes vise à minimiser les risques

d'activation lexicale indésirable (augmentation de la cohésion perceptive de l'unité porteuse) et survienne lorsque les portions de parole sont d'un traitement plus facile (pas de nécessité pour le locuteur de faire un effort articulatoire). Il est à remarquer que la présence d'un mot enchâssé pourrait aussi correspondre à une zone du signal plus facile à traiter du fait de la superposition lexicale, d'où une accélération du débit.

#### 4.4.3.2 Réactivité lors de la perception

Des constantes dans la manière dont les auditeurs réagissent aux fluctuations de chacun des 3 paramètres prosodiques ont pu être remarquées. Elles sont récapitulées dans le tableau ci-dessous, et varient en fonction de la lexicalité des items.

Tableau 108 : Sens de variation des temps de réponse des auditeurs en fonction des variations des paramètres prosodiques

	Syllabes lexicales		Syllabes non lexicales	
	Sens de variation	Expérience	Sens de variation	Expérience
<b>f0</b>	Identique	4, 6, 7	Identique	11, 12
<b>Intensité</b>	Identique	6, 7, 14	Opposé	11, 12
<b>Durée</b>	Identique	4, 6, 7, 13	Opposé	11, 12

Avec des syllabes lexicales, le sens des variations de vitesse de traitement suit toujours celui des variations mélodiques, intensitaires et de durée.

Avec des syllabes non lexicales, le sens des variations de vitesse de traitement est le même que celui des variations mélodiques, mais est inversé pour l'intensité et la durée.

Cette observation indique que le lexique intervient probablement dans la réactivité des participants. Pour les syllabes non lexicales, traitées à un niveau infra-lexical ou acoustique, une augmentation d'intensité et de durée se traduit par une meilleure définition acoustique des segments (audibilité), et donc une facilitation de l'analyse de la parole. Dans le cas des syllabes lexicales, le traitement infra-lexical est probablement facilité de manière identique, mais cette facilitation permet, aux niveaux supérieurs, à des processus ralentissant le traitement de se produire. Il peut s'agir de phénomènes analytiques (mieux un segment est acoustiquement défini, plus il est à même d'induire une activation généralisée dans le lexique, et plus il faut de temps pour trouver la représentation lexicale correcte car moins elle est prégnante dans un ensemble plus vaste), auxquels le système préférera une reconnaissance de forme plus globale en cas de moindre audibilité (convergence par résolution de contraintes multiples).

L'augmentation de f0 produit curieusement une augmentation des temps de traitement. Un tel phénomène pourrait être lié à une meilleure définition acoustique des segments, avec pour conséquence une activation plus forte du lexique, et donc une difficulté plus importante pour reconnaître un candidat dans un ensemble (tel qu'expliqué précédemment). Cette augmentation des temps de réponse pourrait aussi être liée à une mobilisation de l'attention consciente vers la source du signal, ou vers le sens contenu dans le message, faisant diversion sur la réalisation consciente de la tâche.

Enfin, la combinaison « modulation prosodique à l'encodage et à la production + réactivité des participants », conduit, sur l'ensemble des paramètres prosodiques, à un motif global de comportement (temps de réponse) excessivement variable d'une expérience à l'autre. Le tableau ci-dessous récapitule pour chaque expérience les effets produits par les fluctuations de la prosodie sur la réaction des auditeurs.

Tableau 109 : Récapitulation des motifs de réponse des participants, obtenus dans les expériences en fonction de la lexicalité des stimuli et des facteurs manipulés

Expérience	Lexicalité des stimuli (1 = mot)	Facteurs manipulés	Sens de l'effet sur les TR	Amplitude de l'effet (ms.)
<b>4</b>	1	E	↗	53
		I	↗	9
<b>6 / 7</b>	1	M <sub>1</sub>	↘	48 / 85
		M <sub>2</sub>	-	-
<b>11</b>	0	FqI	↗	143
		FqF	↗	149
<b>12</b>	0	FqM	↘	123
<b>13 / 14</b>	1	FqI	↘	25
		FqnI	↗	14

*Avec : E = superposition lexicale initiale réelle ; I = superposition lexicale initiale virtuelle ; M1/2 = la première / deuxième syllabe du stimulus est un mot ; FqI/F = la syllabe est fréquemment rencontrée en position initiale / finale (première / dernière syllabe) dans les mots du lexique ; FqM = la syllabe est fréquemment rencontrée en position non initiale et non finale dans les mots du lexique ; FqnI = la syllabe est fréquemment rencontrée en position non initiale dans les mots du lexique.*

Dans le cas des mots, le motif de réactivité à la prosodie paraît varier en fonction du gabarit du mot (monosyllabique ou bisyllabique) et de la position considérée dans le mot (syllabe de début ou syllabe non initiale).

Lorsqu'un mot monosyllabique contient un enchâssement initial, la prosodie paraît être responsable d'une augmentation de la difficulté du traitement (le traitement du niveau infra-lexical pourrait être en cause).

Lorsqu'un mot bisyllabique contient une syllabe initiale lexicale, ou lorsqu'une syllabe lexicale est fréquente à l'initiale d'un mot plurisyllabique, la prosodie produit un effet facilitant. Lorsque cette configuration concerne la position non initiale d'un mot

plurisyllabique, l'effet serait très ténu et irait dans le sens d'une augmentation de la difficulté.

Il faut ici remarquer que cette différence en fonction du gabarit lexical est logique dans un système à segmentation syllabique. Les mots monosyllabiques ne peuvent en effet pas être segmentés en mots plus petits, tandis que la segmentation d'un mot bisyllabique en 2 mots monosyllabiques est possible.

Dans le cas de pseudo-mots, la prosodie des syllabes qui composent fréquemment la frontière initiale ou finale d'un mot est responsable d'une augmentation de la difficulté du traitement perceptif, alors que la prosodie des syllabes médiales est responsable d'une facilitation des traitements perceptifs. Ceci pourrait signifier qu'une prosodie de frontière existe, et que la difficulté remarquée provient d'une interférence avec le lexique, déclenchée par les caractéristiques de la prosodie.

#### 4.4.4 Conclusions relatives aux analyses présentées dans le Chapitre 4, Partie IV

Les quelques observations rassemblées dans le Chapitre 4 permettent de dégager les 3 pistes de recherche suivantes :

- ▷ D'un point de vue méthodologique, l'analyse acoustique des détails prosodiques de productions verbales représente une méthode comportementale directe et naturelle pour explorer la manière dont un locuteur pense et manipule inconsciemment le langage lorsqu'il parle. Cette technique, utilisée dans les Chapitres 1 et 4, a permis de montrer que les paramètres prosodiques ne sont pas modulés au hasard par le locuteur, mais relativement à des processus psychologiques liés à la cognition et à la communication. Tout un champ de recherche est donc à développer à partir de ce travail, puisque des études beaucoup plus systématiques que celles présentées ici pourraient être mises en chantier.
- ▷ La manière dont les paramètres prosodiques influencent la réactivité des participants, et donc, guident le processus de décodage de l'information acoustique et d'activation lexicale, n'est pas totalement clarifiée. Ce travail a cependant montré que la prosodie produit une influence comportementale mesurable sur les auditeurs lors de la reconnaissance de mots parlés. Si quelques hypothèses ont été proposées (fenêtrage temporel par la durée, modulation du niveau de la  $f_0$  et l'intensité, modulation de la résolution acoustique des segments par la durée et la  $f_0$ ), il reste à valider scientifiquement ces idées par un travail expérimental plus systématique.
- ▷ Il est possible de retrouver comment la prosodie a influencé le traitement perceptif (réactivité à la combinaison des 3 paramètres prosodiques), ce qui fournit des renseignements sur la dynamique cognitive du traitement. Les recherches en psycholinguistique, relatives à la perception de la parole, pourraient tirer avantage de cette méthode d'exploration.



## 4.5 Chapitre 5 – Récapitulation synthétique des résultats présentés dans la partie expérimentale

Ce cinquième et dernier chapitre de la partie expérimentale récapitule de manière synthétique les résultats obtenus dans les Chapitres 1 à 4 (Partie IV).

### Les études présentées dans le Chapitre 1 ont montré :

- ▷ que les modulations de la prosodie générée par le locuteur ( $f_0$ , durée) étaient en relation, au niveau syllabique, avec des facteurs morphosyntaxiques et topologiques, donc avec les traitements cognitifs du langage (Étude préalable).
- ▷ que ces modulations étaient bien perçues par l'auditeur, et donc qu'elles remplissaient probablement un rôle de code, utilisable pour la segmentation du flux de parole en mots (Expérience 1).
- ▷ que ces modulations permettaient à des auditeurs de détecter et de localiser les frontières lexicales entre mots de classe ouverte, mais pas entre mots de classe fermée et mots de classe ouverte. Dans ce cas, une unité prosodique (mot prosodique ?) est perçue avec une cohésion interne supérieure à celle d'un mot polysyllabique de classe ouverte (Expérience 2).
- ▷ que ces modulations prosodiques agissent sur le système récepteur parce qu'elles modulent la résolution acoustique de la ligne segmentale. Par conséquent, certaines unités infra-lexicales sont physiquement moins bien définies, alors que d'autres le sont de façon plus précise. Les mécanismes employés à cette fin seraient relatifs au débit d'élocution et à la fréquence fondamentale (Expérience 3).

### Les études présentées dans le Chapitre 2 ont montré :

- ▷ que la présence d'un enchâssement lexical initial ne cause pas une plus grande difficulté pour reconnaître un mot, mais produit au contraire une facilitation. Ce résultat est invraisemblable dans le contexte théorique actuel relatif à la reconnaissance des mots parlés en psycholinguistique.
- ▷ que la compétition lexicale, en tant que mécanisme de sélection tel que défini par des modèles théoriques comme TRACE, est un processus assez discret, et donc peut-être aussi assez secondaire pour la reconnaissance des mots parlés (Expériences 4, 5, 6, 7, 8, 9 et 10). Il pourrait ne survenir que dans des cas très spécifiques d'ambiguïté lexicale (« garden path » lexicaux), et être de nature sémantique plutôt que formelle (absence d'effet de compétition lié à la présence de mots enchâssés).
- ▷ qu'il existerait un niveau syllabique de représentation dans le système de reconnaissance des mots parlés (Expériences 10, 11, 12, Analyses supplémentaires (Annexe 32 page 533), Analyse exploratoire). Ce niveau infra-lexical pourrait servir à initier l'accès au lexique, ou à organiser la genèse d'un percept lexical, étant donné les interférences observées dans les Expériences 9 et 10.
- ▷ que les mots sont des unités perceptives polarisées, induisant une réactivité importante du système par leur initiale (Expérience 11, Analyse exploratoire).
- ▷ que le lexique serait structuré en familles de mots, selon une structure arborescente, les mots d'une famille étant ceux qui commencent par la même

syllabe. Plus une famille comporterait de membres, plus les membres de cette famille seraient rapidement reconnus (Expérience 4, Analyse exploratoire).

Les études présentées dans le Chapitre 3 ont montré :

- ▷ que la cohésion acoustique d'une séquence de syllabes permettait de moduler la cohésion perceptive de cette séquence (Expériences 15 et 16).
- ▷ que la cohésion lexicale d'une séquence de syllabes n'était pas une donnée traitée a priori lors de l'élaboration du percept lexical, mais qu'elle ne l'était que sous l'impulsion d'un signal véhiculé par la prosodie (Expérience 15). Il s'agit là d'une preuve en faveur de l'importance du guidage du traitement cognitif par la prosodie.
- ▷ que la prosodie rend inutile l'utilisation de mécanismes de compétition pour la segmentation du flux de parole en mots. Elle semble au contraire contraindre fortement les traitements, soit au niveau des processus d'activation du lexique, soit au niveau des processus de sélection.
- ▷ que la notion de cohésion perceptive, multidimensionnelle et relative à la dynamique des traitements sensoriels et cognitifs, permettait de concevoir le flux de parole comme une alternance de zones à forte et à faible cohésion perceptive, et que la segmentation du flux sonore en mots résultait probablement de ce type d'irrégularités (Expériences 15 et 16).

Les études présentées dans le Chapitre 4 ont montré :

- ▷ que le système locuteur module la prosodie ( $f_0$ , intensité, durée) en fonction des facteurs lexicaux manipulés dans les expériences, confirmant les observations du Chapitre 1.
- ▷ que ces modulations agissent sur l'audibilité des unités linguistiques produites (phonème, syllabe, mot), par l'intermédiaire de leur résolution acoustique et de leur intensité sonore.
- ▷ que le but de ces modulations semble être d'agir directement sur la dynamique du traitement du flux sonore puisque des répercussions au niveau lexical sont observées (la réactivité des participants aux paramètres prosodiques varie en fonction de la lexicalité des syllabes considérées). Les syllabes qui peuvent être traitées aisément (syllabes fréquentes), ou dont l'importance perceptive doit être minimisée pour éviter une ambiguïté lexicale (syllabes lexicales dans un mot polysyllabique) seraient produites avec une résolution acoustique moindre visant à optimiser le traitement du mot porteur et défavoriser le traitement du mot enchâssé. Ces résultats vont dans la même direction que ceux de l'Expérience 15.

## 5 Partie V - Discussion

Le travail de recherche présenté dans ce manuscrit a abordé le problème de la segmentation du flux de parole en mots à travers deux perspectives, l'une relative au guidage ascendant des traitements par la dimension suprasegmentale du signal, l'autre relatif au guidage dit descendant des traitements par l'organisation du lexique mental. Il a également cherché à comprendre comment la combinaison de ces dimensions, suprasegmentale et lexicale, pouvaient interagir pour orienter, voire contraindre, la reconnaissance des mots parlés.

La cinquième partie de ce manuscrit propose une synthèse des résultats obtenus dans ce travail. Les implications théoriques de ces résultats sont discutées, et un modèle mathématique de prégnance perceptive des mots est proposé. Ce modèle permet de rendre compte des phénomènes paradoxaux rapportés dans cette étude, quant au rôle des superpositions lexicales sur la reconnaissance des mots parlés.

La cinquième partie est organisée en quatre sections comme suit :

- ▷ Une première section (Section 5.1 page 375) dresse le bilan du travail relatif à l'exploration du rôle de la prosodie sur la reconnaissance des mots parlés, effectué dans les expériences des Chapitres 1, 3 et 4 (Partie IV).
- ▷ Une deuxième section (Section 5.2 page 377) dresse le bilan du travail relatif à l'exploration du rôle des superpositions lexicales sur la reconnaissance des mots parlés, effectué dans les expériences du Chapitre 2 (Partie IV).
- ▷ Une troisième section (Section 5.3 page 382) propose un modèle mathématique de prégnance perceptive des mots, le modèle SLEXS, qui permet de rendre compte des effets de facilitation obtenus avec les expériences du Chapitre 2 (Partie IV).
- ▷ Le modèle SLEXS est testé avec des données d'expériences, et les résultats de ces tests sont présentés dans la quatrième section (Section 5.4 page 395)

Afin de faciliter la lecture de cette cinquième partie, le lecteur peut avoir recours aux synthèses présentées aux :

- ▷ Section 2.4 page 77
- ▷ Section 3.5 page 155
- ▷ Section 4.5 page 372

## 5.1 L'exploration du rôle de la prosodie dans la reconnaissance des mots parlés

Le premier versant de ce travail, relatif à l'impact de la prosodie sur la démarcation lexicale, a permis de montrer des phénomènes prosodiques étayant l'hypothèse d'un guidage ascendant du traitement de segmentation en mots par l'information suprasegmentale.

Des indices ont pu être mis en évidence en faveur de l'existence d'un code prososyllabique, grâce aux analyses acoustiques et simulations (Partie IV, Chapitre 1, Étude préalable). Les caractéristiques prosodiques des syllabes apparaissent être déterminées en fonction de facteurs topologiques et morphosyntaxiques. Elles sont à même d'être utilisées pour le guidage de la segmentation en mots à un très bas niveau, le système disposant ainsi d'une information pour localiser la position des frontières lexicales et de la trame syntaxique.

Le test d'auditeurs naïfs a révélé leur capacité à distinguer perceptivement les catégories prososyllabiques ainsi définies, puisqu'ils parviennent à utiliser l'information suprasegmentale pour organiser inconsciemment le matériel expérimental (Partie IV, Chapitre 1, Expérience 1). Ils parviennent également à détecter et localiser consciemment la présence de frontières lexicales dans des items déségmentés (Partie IV, Chapitre 1, Expérience 2), indiquant par là qu'ils disposent aussi d'une représentation du motif prosodique qui doit être retrouvé aux frontières entre certains types de mots.

Toutefois, si des indices surgissent çà et là pour appuyer l'hypothèse de représentations mentales prosodiques impliquées dans les opérations de démarcation lexicale, ce travail s'est surtout focalisé sur la recherche d'une information acoustique à même de guider le système de manière ascendante dans le décodage de la parole. Il n'apporte donc pas de conclusion sur l'existence de représentations mentales prosodiques. Cependant, il fournit suffisamment d'indices pour inciter une recherche dans cette direction.

Il a été proposé que l'efficacité de l'information prosodique à remplir une fonction démarcative était basée sur l'exploitation de propriétés psychophysiques fondamentales, telles que la résolution acoustique des éléments de la ligne segmentale (phonèmes, syllabes). L'Expérience 3 a alors montré que ce rôle peut être accompli aussi bien par le biais de la durée que de la  $f_0$ . Par ailleurs, le système locuteur agit sur son débit de manière à moduler le fenêtrage temporel de la ligne segmentale (quantité d'information

phonologique délivrée par unité de temps) en fonction de facteurs lexicaux, comme cela a pu être montré dans l'Analyse exploratoire (Partie IV, Chapitre 2) ou dans le Chapitre 4 (Partie IV). Ce procédé permettrait d'adapter l'information entrante à la dynamique du traitement, aussi bien pour éviter que certains phénomènes ne se produisent au mauvais moment (exemple : une segmentation lexicale dans un mot), ou pour en favoriser d'autres (exemple : reconnaissance d'un mot court libre). Bien que non étudié expérimentalement, il a également été proposé que la prosodie soit impliquée dans une modulation du niveau vigile du locuteur par le biais des paramètres  $f_0$  et intensité, et remplisse ainsi un rôle de guidage attentionnel (au sens psychophysologique). Le but de ce guidage attentionnel serait essentiellement de potentialiser le système sur les zones de traitement qui présentent une plus grande difficulté cognitive (ambiguïté lexicale, syntaxique ou sémantique, information sémantique nouvelle, etc.). Les analyses statistiques présentées au Chapitre 4 (Partie IV) visaient aussi à donner des indices relativement à cette question, mais malheureusement, aucune certitude n'a pu être fournie.

Enfin, les Expériences 15 et 16 ont confirmé que les caractéristiques prosodiques de la parole sont impliquées dans un guidage des processus lexicaux qui semble être relativement contraignant. L'Expérience 15, en particulier, démontre que les traitements lexicaux sont subordonnés aux indices véhiculés par la prosodie.

Au terme de ce travail, on dispose donc d'un ensemble de résultats qui indiquent que la prosodie joue un rôle important dans la démarcation ascendante des unités lexicales en Français. L'hypothèse d'un code prososyllabique relatif, exploitant des phénomènes contrastifs de proche en proche, permet avantageusement de rendre compte de ce guidage ascendant, et ce en dépit de l'absence de marques acoustiques systématiques et caractéristiques de frontières de mots. Les résultats (Simulation 2 et Expérience 1) montrent d'ailleurs une manière originale d'exploiter le code prososyllabique en reconnaissance automatique de la parole, ce système présentant des performances supérieures aux techniques traditionnelles de classification telles que celles utilisées par Caillaud (1996) et Munteanu (1996).

Aussi, les schémas théoriques de la reconnaissance des mots parlés exclusivement basés sur le traitement d'une information représentationnelle de nature phonétique ou phonologique apparaissent comme insuffisants pour rendre pleinement compte de la segmentation perceptive du flux sonore en mots. L'information suprasegmentale est essentielle pour limiter la combinatoire des possibilités lexicales. En dirigeant l'activation du lexique ou en organisant les procédures de sélection, la prosodie rend tout à fait

secondaire les procédures de segmentation du flux de parole en mots basées sur de la compétition lexicale. Il semblerait plutôt que la compétition lexicale soit un processus tardif de type correctif, mis en œuvre en cas d'ambiguïté dans la segmentation lexicale (« garden path » lexicaux), donc qu'elle n'intervienne pas directement dans la reconnaissance d'un mot.

## **5.2 L'exploration du rôle des superpositions lexicales dans la reconnaissance des mots parlés**

Le deuxième versant de ce travail a permis de mettre en évidence un phénomène paradoxal dans le cadre théorique actuel de la reconnaissance des mots parlés en psycholinguistique : la facilitation de la reconnaissance d'un mot renfermant une superposition lexicale initiale.

Robuste, cet effet a pu être répliqué dans plusieurs expériences de ce travail utilisant une tâche de décision lexicale, aussi bien en modalité auditive (Expériences 4 à 7) que visuelle (Expérience 8). Le phénomène a aussi pu être retrouvé avec une tâche de détection de cibles non linguistiques insérées dans les mots stimuli (Expériences 9 et 10). Il est par ailleurs décrit dans la littérature, quoique très rarement rapporté.

Étant donné son caractère absolument contraire au schéma théorique général du domaine, la première explication à considérer est celle du biais, qui a pu subsister en dépit d'efforts importants pour maîtriser la situation expérimentale et les différentes étapes du traitement des données.

Le principal biais expérimental pouvant être à l'origine de ce phénomène est lié à l'utilisation de la tâche de décision lexicale. La facilitation de la reconnaissance d'un mot en cas de superposition lexicale initiale est peut-être provoquée par un phénomène d'amorçage de la réponse de lexicalité par le mot enchâssé initial. Selon cette hypothèse, la réponse « mot » doit logiquement être donnée plus rapidement pour le mot stimulus porteur d'un enchâssement initial que pour un mot stimulus ne renfermant pas de mot enchâssé initial. Par contre, ce phénomène ne devrait pas être observé avec une superposition lexicale en position finale, ou l'être avec une faible amplitude, si on considère qu'alors, l'amorçage de la réponse « mot » n'a pas le temps de produire son effet ou un effet mesurable. C'est généralement ce qui a été observé (Expériences 6 et 7), confortant l'hypothèse du biais lié à la tâche expérimentale. Toutefois, si cette logique est suivie, on ne devrait pas obtenir d'effet de facilitation en cas de superposition lexicale initiale virtuelle en modalité auditive (Expérience 4), puisque la reconnaissance porte sur l'amorce, et non sur une cible renfermant une amorce. Or, une facilitation a bien lieu. De

même, en cas de superposition lexicale initiale réelle lors du traitement d'un mot en modalité visuelle (Expérience 8), l'effet d'amorçage de la réponse « mot » par l'enchâssement initial ne devrait pas avoir lieu, puisque l'amorce de la réponse « mot » et la cible de la tâche apparaissent au même moment. Or, un effet de facilitation est encore obtenu. On peut toutefois encore proposer, pour expliquer les effets de facilitation dans le cas des Expériences 4 et 8, que la réponse « mot » ait bénéficié de l'activation dans le lexique d'une cohorte plus importante que dans le cas de mots non impliqués dans une superposition lexicale initiale virtuelle ou réelle. Cette plus forte activation se traduirait par des temps de réponse plus rapides pour fournir la réponse « mot ». Selon cette explication, l'effet de facilitation serait essentiellement de nature stratégique.

Toutefois, il faut remarquer que ce phénomène de facilitation a pu être retrouvé avec l'utilisation d'autres tâches expérimentales que la TDL. Ceci suggère que l'effet pourrait être indépendant de la tâche utilisée, donc être relatif au traitement lexical.

Ainsi, Luce et Lyons (1999) l'ont observé avec une tâche de « shadowing », employée sur un matériel expérimental très similaire à celui qui a été utilisé dans les Expériences 6 et 7. Ces auteurs contrôlaient par ailleurs la facilité de récupération phonologique du matériel expérimental (taille de la cohorte initiale pondérée par la fréquence).

Dans la présente recherche, une facilitation de la détection de cibles non linguistiques insérées dans des mots a pu être observée en fin de mot porteur, en présence de superpositions lexicales initiales réelles (Expériences 9 et 10). Dans l'Expérience 9, la diminution des temps de détection des cibles apparaît alors en fin de mot, bien après que n'ait eu lieu la superposition lexicale initiale, tandis qu'aucune augmentation des temps de détection des cibles n'est mesurée lorsqu'elles sont insérées au moment où se produit la superposition lexicale. Dans l'Expérience 10, un phénomène comparable est obtenu pour la superposition lexicale initiale, aucun effet n'apparaissant pour une superposition lexicale non initiale.

Le biais d'amorçage stratégique lié à la TDL apparaît donc être une hypothèse peu plausible pour expliquer les effets de facilitation lexicale obtenus dans toutes les expériences de ce travail. En effet, ce biais ne permet pas d'expliquer comment l'effet peut subsister avec des tâches qui n'induisent pas d'amorçage de la réponse (par exemple pour les Expériences 9 et 10, aucun lien n'existe entre les caractéristiques de la cible non linguistique, sur laquelle porte la tâche, et les caractéristiques lexicales des items dans lesquels elle est insérée).

Si cet effet de facilitation n'est pas un biais, on peut donc s'interroger sur son origine, lexicale ou infra-lexicale. A priori, et d'après le contexte théorique contemporain sur la problématique de la reconnaissance des mots parlés, il semble peu vraisemblable que l'effet soit de nature lexicale. En effet, d'après la grande majorité des modèles, la reconnaissance d'un mot passe par une phase transitoire d'activation de candidats multiples, tout particulièrement dans les situations de superpositions lexicales étudiées dans le présent travail. Or, lorsque plusieurs mots sont activés simultanément, le système doit mettre en œuvre des procédures de sélection (compétition lexicale), qui varient d'une théorie à l'autre, mais qui sont toutes, au pire coûteuses en énergie et/ou en temps, au mieux neutres ou discrètes. Par contre, aucune de ces procédures ne permet d'envisager que l'activation multiple de candidats puisse aboutir à une facilitation de la reconnaissance lexicale. Les effets de facilitation en cas de superposition lexicale initiale auraient donc plus vraisemblablement une origine infra-lexicale, impliquant des processus de traitement différents, par opposition aux effets de compétition, de nature lexicale (formelle ou sémantique).

L'hypothèse d'une distinction du sens des effets en fonction du niveau de traitement (infra-lexical vs. lexical) est défendue par les observations de certains chercheurs, comme par exemple Slowiaczek et Hamburger (1992) ou Vitevitch et Luce (1998). Les effets de facilitation seraient de nature infra-lexicale<sup>127</sup>, tandis que les effets de compétition seraient de nature lexicale et surviendraient plus tardivement dans le traitement<sup>128</sup>.

Partant de ces observations, il était logique de supposer que les effets facilitateurs aient peut-être une origine infra-lexicale. De plus, les items testés dans les situations de superposition lexicale impliquant toujours le niveau syllabique (mots monosyllabiques, mots bisyllabiques pour lesquels la superposition lexicale coïncide avec la syllabe), la question pouvait porter plus spécifiquement sur le niveau syllabique. Ce point a été exploré dans les Expériences 11 à 14 (enchâssabilité des syllabes), les Analyses supplémentaires, et l'Analyse exploratoire. Bien que des effets facilitateurs des

---

<sup>127</sup> Selon Slowiaczek et collègues (2000), la force des effets de facilitation en amorçage phonologique est proportionnelle à la quantité de phonèmes en commun entre amorce et cible, donc à la distance formelle entre items, ou encore selon la terminologie utilisée dans ce travail, au pourcentage de superposition de l'information entre items.

<sup>128</sup> Les effets de compétition lexicale disparaissent lorsque la durée entre l'amorce et la cible se réduit (Vroomen & de Gelder, 1995), indice que la compétition lexicale est un phénomène tardif. Étant donné que la compétition lexicale disparaît également lorsque les items à traiter sont des pseudo-mots, pour laisser place à des effets facilitateurs de nature phonotactique, c'est qu'il s'agit vraisemblablement d'un phénomène de nature lexicale (Vitevitch & Luce, 1998).



superpositions syllabiques virtuelles aient été obtenus pour des syllabes non lexicales (Expériences 11 et 12), confirmant l'existence d'un niveau syllabique de représentation (par ailleurs retrouvé avec les Analyses supplémentaires), une absence paradoxale d'effet est observée lorsque les syllabes sont lexicales (Expériences 13 et 14). Ce fait est d'autant plus curieux que l'Expérience 4, analogue à l'Expérience 13 (tâche et facteur), révélait des effets de facilitation dus au facteur superposition lexicale initiale virtuelle. Toutefois, l'Expérience 4 opérationnalisait la superposition lexicale initiale virtuelle par la séquence de phonèmes, et non par la syllabe (Expérience 13). La différence dans les profils de réponse des deux expériences s'expliquerait ainsi, ce que semble confirmer l'Analyse exploratoire (effet facilitant du nombre de mots renfermant le mot stimulus en position initiale selon un comptage de la séquence de phonèmes, pas d'effet selon un comptage de la syllabe).

Les Expériences 11 à 14, les Analyses supplémentaires et l'Analyse exploratoire, dont le but était d'aider à cerner l'origine des effets de facilitation obtenus dans les Expériences 4 à 10, ne permettent pas d'affirmer que la facilitation de la reconnaissance des mots sous l'influence de superpositions lexicales initiales soit d'origine infra-lexicale.

L'Analyse exploratoire ayant montré que des phénomènes de compétition lexicale ont lieu (Section 4.2.13 page 330), il faut en déduire qu'ils ont aussi probablement eu lieu lors de la production des données par les auditeurs dans les expériences 4 à 8. Si les effets de compétition lexicale correspondent à des phénomènes tardifs, cela peut expliquer pourquoi aucune augmentation des temps de détection des cibles non linguistiques n'a pu être obtenue dans les Expériences 9 et 10, lorsque les cibles sont insérées dans le mot au moment où se produit une superposition lexicale réelle, donc au moment où un conflit lexical est susceptible d'avoir lieu. Toutefois, le conflit devrait devenir visible ultérieurement, mais il n'apparaît toujours pas lorsque ces mêmes cibles sont insérées en milieu ou en fin de mot, parfois plusieurs centaines de millisecondes après la superposition lexicale initiale. C'est au contraire une facilitation de la détection des cibles qui est observée. De même, un allongement des temps de réponse devrait être obtenu dans les Expériences 4 à 7, au lieu d'une diminution. Toutefois, cet allongement, s'il avait lieu sous l'influence d'un phénomène de compétition lexicale, pourrait aussi être compensé par des phénomènes infra-lexicaux facilitants (exemple : Vroomen et de Gelder (1995), Expériences 11 et 12). Cependant, pour que des phénomènes infra-lexicaux facilitants contrecarrent des phénomènes tardifs de compétition au point de les faire totalement disparaître, il faudrait que ces effets infra-lexicaux aient une amplitude supérieure à celle des effets liés aux mécanismes de compétition, ou que ces derniers soient si tardifs que leur impact sur la reconnaissance puisse être questionné, donc que

leur influence soit relativement secondaire dans le processus général de la reconnaissance d'un mot.

Puisque les effets de facilitation obtenus dans les Expériences 4 à 10 ne semblent pas être liés à un biais, et qu'ils ne s'expliquent pas franchement par l'organisation infra-lexicale du système, une hypothèse relative à la dynamique du traitement lexical est proposée : l'hypothèse d'un amorçage lexical.

La technique de l'amorçage phonologique permet de montrer qu'une amorce correspondant à un mot enchâssé initial (exemple : /ku/) facilite la reconnaissance d'un mot-cible porteur de ce mot enchâssé en position initiale (exemple : /kulis/ - Dufour et Peereman (2004)). Ce phénomène est cependant supposé être de nature phonologique, puisqu'il se produit aussi lorsque l'amorce correspondant au début du mot cible n'est pas lexicale (exemple : Slowiaczek et Hamburger (1992)). Pour expliquer une différence de traitement en fonction de la lexicalité du début du stimulus, il faut donc envisager un autre type d'amorçage, qualifié ici de lexical. Étant donné que le signal acoustique décrivant un mot se déploie dans le temps à partir d'un moment initial correspondant au début du mot, un auditeur entendra le mot enchâssé à l'initiale du mot stimulus avant d'entendre ce dernier. Lors de la présentation du mot « *garage* », par exemple, le mot « *gars* » sera d'abord entendu, puis le mot « *gare* », puis le mot « *gara* », et enfin le mot « *garage* ». Selon l'hypothèse de l'amorçage lexical, un mot comme « *garage* » bénéficierait donc d'un amorçage par le lexème /gaxa/, le mot enchâssé initial « *gara* » ayant lui-même été amorcé par le lexème /gax/ au cours du processus, le mot enchâssé initial « *gare* » lui-même amorcé par le lexème /ga/. Toutefois, l'hypothèse de l'amorçage lexical expliquée par l'arrivée progressive d'information dans le système ne permet pas de rendre compte d'un amorçage lexical avec des stimuli visuels, pourtant obtenu dans l'Expérience 8. Pour maintenir une telle hypothèse, il faut donc supposer que l'amorçage lexical est un phénomène amodal qui se déroule en haut niveau, dans le lexique mental ou en mémoire de travail, une fois que les éléments nécessaires à l'élaboration du percept lexical sont disponibles ou activés. Proposer une telle hypothèse implique donc de supposer que le percept lexical soit progressivement élaboré par niveaux de complexité, du moins au plus complexe, selon un principe d'agrégation graduelle d'éléments (de quelle nature ?), plutôt que par un simple processus d'activation d'unités lexicales non composites. Par conséquent, les phénomènes de compétition lexicale supposés par les théories de la psycholinguistique pourraient être rejetés à un niveau antérieur du traitement (sélection d'éléments stockés en mémoire à long terme),

ou postérieur (procédures de correction en cas d'hypothèse de segmentation fallacieuse<sup>129</sup>), d'où leur faible visibilité dans les expériences de ce travail de thèse.

Partant de ces observations, réflexions et propositions, un modèle mathématique de prégnance perceptive des mots parlés est présenté pour montrer comment l'activation séquentielle de lexèmes alignés peut faciliter, et non inhiber, la reconnaissance d'un mot.

### **5.3 Description du modèle de reconnaissance des mots SLEXS1**

Le modèle SLEXS1 (« èslèksès » - Structure LEXicale Superposante, version 1) est un modèle mathématique permettant de calculer un score exprimant le degré de prégnance perceptive d'un mot, en fonction de l'énergie qu'il reçoit depuis le niveau infra-lexical, et en fonction des échanges d'énergie entre lexèmes lors de leur intégration perceptive en mémoire de travail sous l'effet d'une représentation mentale des relations entre mots dans le lexique.

SLEXS1 est défini à partir d'un ensemble de principes logiques de fonctionnement, tirés de la littérature dans le domaine et de propositions originales faites à partir des résultats expérimentaux obtenus dans ce travail de thèse.

Le modèle n'est pour l'heure implémenté que sous forme mathématique, mais a priori, sa traduction en programme informatique est techniquement aisée à effectuer à partir des équations proposées plus bas. La manière dont les traces mémorielles sont incarnées dans un substrat physique ne sera donc pas abordée en détails ici.

Cette section commence par décrire les principes logiques à l'œuvre dans SLEXS1 (Section 5.3.1) avant d'expliquer comment le modèle fonctionne théoriquement (Sections 5.3.2 et 5.3.3). la formalisation mathématique du modèle est ensuite proposée à la Section 5.3.4.

---

<sup>129</sup> Exemple de l'Expérience 10 avec la condition M1.nM2 : le système a pu considérer que, M1 étant un mot, il fallait segmenter le signal à la fin de la première syllabe et engager un nouvel accès lexical sur la suivante. Cependant, nM2, qui n'est ni un mot ni le début d'un mot, resterait non appariée. Il faut donc que le système revienne à une alternative plus élaborée (M1.nM2). Dans ce cas, on peut supposer que le traitement pour annuler le premier schéma de segmentation lexicale pourrait être la cause de l'interférence observée.

### 5.3.1 Principes logiques à l'œuvre dans SLEXS1

#### 5.3.1.1 Principes généraux de la reconnaissance d'un mot

Tout percept lexical est élaboré en mémoire de travail à partir d'éléments infra-lexicaux et lexicaux retrouvés en mémoire à long terme (détecteurs de phonèmes et détecteurs de lexèmes<sup>130</sup>), et d'un ensemble de forces contradictoires, facilitantes et interférentes. Ces forces correspondent essentiellement à la manière dont les lexèmes sont liés entre eux dans le lexique, liens qui forment une représentation mentale appelée matrice lexicale ou matrice perceptive.

Les forces facilitantes sont dues à l'existence de lexèmes initialement superposés dans la matrice lexicale (Expériences 4, 5, 6, 7, 8, 9, 10). Les forces interférentes sont dues à l'existence de lexèmes non alignés dans la matrice lexicale.

L'effet de ces forces contradictoires sur la perception est synthétisé dans un score de prégnance lexicale, supposé traduire la facilité avec laquelle un mot est reconnu.

#### 5.3.1.2 Exploration du lexique

Toute unité infra-lexicale est potentiellement une initiale de mot, comme cela est proposé dans TRACE. L'accès au lexique est donc exhaustif, puisqu'une séquence de phonèmes potentialise tous les lexèmes possibles en mémoire à long terme. Toutefois, il est vraisemblable que le système dispose ensuite de moyens pour localiser les initiales de mot, qui peuvent correspondre à ceux classiquement envisagés (prosodie, SOSH<sup>131</sup> – non modélisé), ou dépendre de la mécanique de fonctionnement du système (le début d'un mot est l'entrée initiale d'une matrice lexicale, certaines matrices étant rejetées car aboutissant à un score de prégnance perceptive trop faible – voir plus loin). Il s'agirait d'un système d'alignement syllabique des matrices lexicales, qui n'exclue pas la possibilité d'ajuster le point d'alignement sur le phonème suivant (exemple des resyllabations). Il faudrait donc distinguer deux niveaux d'exploration du lexique. Un premier, exhaustif, correspondant à l'accès lexical (activation des lexèmes), et un second, sélectif, correspondant à l'intégration perceptive des lexèmes sous l'influence d'une représentation mentale des liens formels entre lexèmes.

---

<sup>130</sup> Le système ne modélise pas le niveau syllabique dans sa version 1.

<sup>131</sup> Il s'agit peut-être des effets syllabiques mis en évidence dans les Expériences 11 et 12, et dans l'Analyse exploratoire, Partie IV, Chapitre 2.

### **5.3.1.3 Accès au lexique**

Sous l'effet d'une stimulation acoustique, un lexème est potentialisé par l'information qu'il reçoit depuis le niveau des détecteurs phonémiques. Cette information dépend de 2 aspects : le niveau d'énergie atteint par chaque détecteur phonémique, et l'organisation des relations entre détecteurs phonémiques et détecteurs lexicaux (lexèmes). L'énergie d'un détecteur phonémique est une valeur supérieure ou égale à zéro, qui dépend essentiellement de la réactivité propre du détecteur et des caractéristiques de la stimulation acoustique, comme la prosodie ou la coarticulation<sup>132</sup>. Cet aspect n'est pour l'heure pas modélisé dans SLEXS1. La transmission de cette énergie infra-lexicale vers le niveau lexical correspond à un mécanisme classique d'intégration (fonctionnement de type réseau de neurones). Elle s'effectue de telle manière que tout lexème maximalelement potentialisé présente un niveau d'énergie identique à tout autre lexème qui est dans le même état, et est donc indépendant du nombre de phonèmes desquels ce lexème a reçu de l'information.

### **5.3.1.4 Intégration perceptive des lexèmes en mémoire de travail : gradient d'information dans une matrice lexicale**

SLEXS1 propose que l'information véhiculée par chaque phonème dans un mot soit polarisée suivant un gradient décroissant depuis le début du mot jusqu'à sa fin ou jusqu'à son point d'unicité. L'information du début des mots prime donc sur celle de leur fin (Analyse exploratoire). Dans le modèle, cet aspect n'est pas pris en compte au niveau des détecteurs (mémoire à long terme) pour des raisons de facilité, mais au niveau de la représentation d'un mot dans le lexique (matrice lexicale). Ceci est traduit par une pondération de la valence des lexèmes lors de leur intégration perceptive par la matrice lexicale. Cette pondération dépend donc de la matrice lexicale considérée, ou encore du point d'alignement choisi par le système.

L'hypothèse d'un gradient d'information au sein d'un mot a été validée lors du test du modèle SLEXS1 sur des données comportementales issues d'expériences de laboratoire relatives à la reconnaissance des mots. SLEXS1 est plus à même de rendre compte de ces données lorsqu'un tel gradient est modélisé que lorsqu'il ne l'est pas.

---

<sup>132</sup> Un détecteur est plus fortement activé si la forme sonore auquel il répond est voisine du prototype mémorisé par ce détecteur. Toutes les variations acoustiques conduisant à un meilleur ciblage du prototype mémorisé aboutissent à une plus forte activation du détecteur.

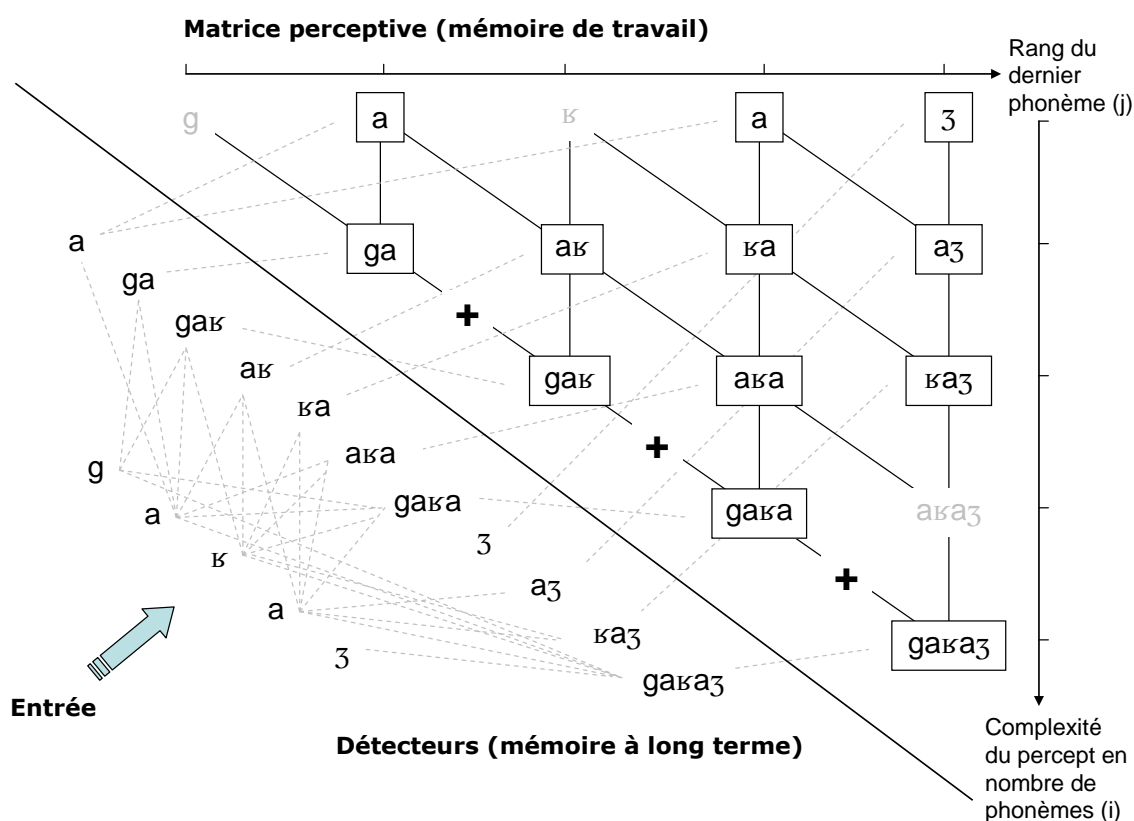
### 5.3.1.5 Dynamique du traitement

L'élaboration du percept lexical en mémoire de travail s'effectue par l'intégration progressive et non linéaire de lexèmes potentialisés. Ce phénomène est supposé être indépendant de la modalité sensorielle.

### 5.3.2 Fonctionnement théorique général

D'après SLEXS1, l'intégration perceptive d'un mot stimulus s'effectue selon l'organisation présentée dans la Figure 73.

Figure 73 : Organisation du traitement du mot « garage » dans SLEXS1



SLEXS1 comporte un ensemble de détecteurs de nature phonologique (« mémoire à long terme »), correspondant aux phonèmes et aux lexèmes. Ils sont chargés de fournir une estimation statistique de la probabilité d'existence d'un événement dans le signal acoustique (« Entrée »). L'information qu'ils émettent est ensuite progressivement intégrée en mémoire de travail, à l'aide d'une « matrice perceptive » qui organise les rapports entre détecteurs en fonction de la complexité formelle (nombre de phonèmes  $i$ ) des entités auxquels ils renvoient (« Complexité du percept en nombre de phonèmes  $i$  ») et du rang du détecteur dans la séquence traitée relatif au point d'alignement de la matrice (« Rang du dernier phonème  $j$  »). Une matrice perceptive ou matrice lexicale

est une représentation mentale des relations formelles qui existent entre lexèmes au sein du lexique mental. Elle sert probablement aussi à intégrer les aspects abstraits ou contenu d'un mot (non modélisé).

Dans une matrice lexicale, l'information de complexité  $c-1$  est intégrée à l'information de complexité  $c$  selon une procédure récursive qui explore exhaustivement un chemin lexical aligné avec le début du mot stimulus, ou chemin candidat, et des alternatives, ou chemins concurrents, non alignés avec le début du mot stimulus. Dans la Figure 73, les chemins sont symbolisés par les diagonales qui relient les lexèmes dans la matrice. Les traits verticaux dans la matrice symbolisent les points possibles d'interférence des lexèmes entre eux (d'autres relations interférentes existent mais ne sont pas représentées. Exemple : interférence entre /aʁ/ et /gaʁa/).

Selon l'exemple de la Figure 73, le chemin candidat de cette matrice commence par le phonème /g/ au rang  $r_1$  ; à  $r_2$  le percept du mot « gars » peut éventuellement<sup>133</sup> émerger sous l'influence des détecteurs /g/ + /a/ + « a » + « gars » ; à  $r_3$ , le mot « gare » peut être perçu (/g/ + /a/ + /ʁ/ + « a » + « gars » + « gare ») ; etc., le processus d'élaboration progressant ainsi de proche en proche. Le percept élaboré sur le chemin candidat bénéficie de la potentialisation des lexèmes alignés, qui sont utilisés comme amorce des percepts alignés à venir. Les chemins concurrents sont exhaustivement explorés dans la matrice à partir des points d'alignement alternatifs (tout phonème). Selon cet exemple, un premier point d'alignement alternatif commence à  $r_2$ , à partir duquel le chemin [/a/ + « a »], [/a/ + /ʁ/ + « a » + « art »], [/a/ + /ʁ/ + /a/ + « a » + « art » + « ara »], etc., peut être suivi. À  $r_3$ , un deuxième point d'alignement alternatif permet de suivre chemin concurrent [/ʁ/], [/ʁ/ + /a/ + « rat »], [/ʁ/ + /a/ + /ʒ/ + « rat » + « rage »], etc. À chaque pas de l'élaboration du percept lexical du mot stimulus, un chemin concurrent peut perturber le traitement (exemple : /aʁ/ dans /gaʁ/). Cependant, le chemin concurrent le plus perturbant est le chemin immédiatement adjacent au chemin candidat ( $r_2$ ). En effet, plus un chemin concurrent est éloigné du chemin candidat, plus le nombre de détecteurs qu'il implique diminue, d'une part, et d'autre part, plus la valence informationnelle de la séquence est faible sous l'effet du gradient d'information dans la matrice. En termes psychologiques, cela signifie que plus le décalage entre le mot candidat et une alternative lexicale est important, moins cette alternative interfère sur la reconnaissance du mot candidat (exemple : « orage »

---

<sup>133</sup> Pour cela, le système a besoin de temps. Si le débit est trop rapide, l'information qui arrive après /a/ est agrégée et un autre percept est élaboré.

interfère fortement sur la reconnaissance de « *forage* », alors que « *rage* » interfère moins sur la reconnaissance de « *forage* »).

Enfin, il faut remarquer que l'interférence entre chemins est unidirectionnelle au sein d'une matrice, depuis les chemins concurrents qui surviennent plus tardivement sur les chemins qui surviennent avant eux. Le chemin candidat n'induit donc pas d'interférence sur les chemins concurrents.

### 5.3.3 Fonctionnement théorique spécifique

#### 5.3.3.1 Segmentation du flux de parole en mots

Le fonctionnement de proche en proche du modèle permet de limiter la quantité d'alternatives lexicales possibles à des chemins lexicaux, et non à l'ensemble des lexèmes potentialisés par une séquence infra-lexicale. L'explosion combinatoire proposée par des modèles comme TRACE est donc ici fortement réduite grâce à la procédure d'élaboration du percept par agrégation progressive.

Pour une séquence sonore donnée, une matrice est générée à partir des points d'alignement probable (début d'un mot, début d'une syllabe), et se développe sous l'effet de contraintes acoustiques (prosodie), phonétiques (coarticulation) et lexicales (la séquence ne peut plus constituer un mot). Le système n'est pas supposé générer exhaustivement des matrices perceptives sous l'influence de chaque détecteur phonémique. Les chevauchements de matrices sont néanmoins possibles, en théorie, mais leur impact sur la perception est limité, du fait que certaines alternatives aboutissent à des scores de prégnance perceptive faibles (voir par exemple les scores présentés plus bas dans le Tableau 110). Ces matrices à faible score sont donc passivement éliminées si une matrice alternative génère un score de prégnance perceptive élevé. Dans le cas où deux matrices se chevauchant partiellement aboutiraient au même score de prégnance perceptive, on peut supposer que la première bénéficierait de son antériorité sur la seconde pour la reconnaissance du mot. Cette discussion ne concerne que le cas où les matrices lexicales ne prennent en compte que des aspects formels des mots. Dans le cas inverse, les ambiguïtés discutées ici n'existeraient pas.

Un tel mode de fonctionnement permettrait d'expliquer les « garden path » lexicaux, du type « *marchandage* » / « *marchand d'art* ». L'augmentation probable de durée pour la reconnaissance du mot « *art* » dans la séquence /*maʁʃãdaʁ*/ par rapport à un contrôle (exemple : « *vendeur d'art* »), s'expliquerait par la complexité supérieure du traitement à effectuer pour reprendre une hypothèse lexicale alternative abandonnée 2



phonèmes plus tôt (nécessité d'inhiber l'hypothèse lexicale « *marchandage* », et de réactiver l'hypothèse lexicale « *marchand* »).

Bien que le but de SLEXS1 soit de modéliser la prégnance perceptive d'un mot, et non d'expliquer comment s'effectue la segmentation du signal en mots, le modèle a des implications théoriques pour la segmentation du flux de parole en mots, d'une part, et d'autre part, il est possible de conjecturer son comportement de segmentation lexicale à l'aide d'exemples numériques (nécessite toutefois plus de développement). Quatre variantes du modèle (se reporter plus loin à la section « Formalisation mathématique de SLEXS1 » pour le détail) ont été utilisées pour générer les scores de prégnance perceptive du mot « *garage* », et des mots « *gars* » et « *rage* » sous leur forme enchâssée dans « *garage* » et sous leur forme libre. Les scores sont présentés dans le Tableau 110.

Tableau 110 : Scores de prégnance perceptive des mots libres « *garage* », « *gars* », « *rage* » et des mots « *gars* » et « *rage* » enchâssés dans « *garage* »

	$f_{4-}$	$f_{4+}$	$f_{5-}$	$f_{5+}$
« <i>garage</i> » (porteur)	10.27	6.83	11.15	5.95
/ga/- (enchâssé)	0.7	0.5	0.7	0.5
-/RaZ/ (enchâssé)	-0.11	0.11	-0.28	0.28
« <i>gars</i> » (libre)	1.75	1.25	1.75	1.25
« <i>rage</i> » (libre)	3.56	2.11	3.61	2.06

Il apparaît que les mots sont moins prégnants lorsqu'ils sont traités comme étant sous une forme enchâssée que lorsqu'ils sont traités comme étant sous une forme libre, indépendamment des différences prosodiques qui distinguent ces 2 configurations, d'une part, et d'autre part, que les mots courts ont une prégnance perceptive moindre que les mots plus longs, essentiellement à cause de la présence de mots enchâssés initiaux dans les mots longs. Par conséquent, la reconnaissance des mots longs tendra à prendre le pas sur celle des mots plus courts en l'absence de toute pondération par une information supplémentaire. C'est le comportement qui est généralement observé avec le système naturel : pour que la séquence /gɑʁɑʒ/, qui peut former un mot, soit segmentée en /ga/ + /ʁɑʒ/, le locuteur doit démarquer les deux mots de manière d'autant plus insistante que la séquence des 2 mots n'a pas de sens alors que le mot bisyllabique en a un. En règle générale, le locuteur aide l'auditeur à distinguer un mot court libre d'une initiale de mot (version enchâssée) en lui donnant une étoffe importante au niveau de la durée (analyses acoustiques présentées dans la Partie IV, Chapitre 1). Il semble par ailleurs utiliser la durée pour limiter l'impact perturbateur qu'une superposition lexicale peut avoir sur le traitement, et éviter ainsi qu'elle n'induisse une segmentation lexicale et ne génère de « *garden path* » lexical. Ceci est attesté par les analyses présentées dans la

Partie IV, Chapitre 4, qui montrent une accélération du débit spécifiquement localisée sur la syllabe correspondant à une superposition lexicale (Expériences 6 et 7)<sup>134</sup>. Ces modulations de durée représenteraient donc un système utilisé à dessein par le système locuteur pour éviter des erreurs de décodage par le système auditeur sur les zones à risque.

Comme pour le système naturel, SLEXS1 devrait donc utiliser la prosodie pour compenser le biais qui rend plus prégnant un mot long que les mots plus courts composés d'une même séquence de phonèmes. Ces procédures pourraient être intégrées aisément en ajoutant des termes de pondération au modèle.

### 5.3.3.2 Compétition lexicale

Le fonctionnement de SLEXS1 ne précise pas sur quel mécanisme se fonde la perturbation qu'il modélise entre lexèmes non alignés au sein d'une même matrice perceptive. La compétition lexicale se limite à une interférence unidirectionnelle depuis tout chemin initialisé postérieurement à un chemin donné au sein d'une matrice. Il n'existe pas de compétition entre détecteurs. Il se peut qu'une interférence existe entre matrices, bien qu'un mot soit reconnu parce qu'il présente le score de prégnance perceptive le plus élevé (et donc que ce type d'interférence ne paraisse pas a priori nécessaire).

### 5.3.3.3 Amorçage lexical

Le modèle SLEXS1 est conçu pour rendre compte du phénomène d'amorçage lexical par le biais des lexèmes correspondant à des mots enchâssés initiaux. Ceux-ci augmentent la prégnance perceptive du mot stimulus par un effet de type cohorte. Par contre, l'inverse n'est pas vrai : l'activité du lexème du mot stimulus est sans effet sur l'activité des lexèmes des mots enchâssés.

Rien n'empêche que ce phénomène, actuellement limité à l'intérieur d'une matrice lexicale, ne soit plus général au sein du lexique, notamment si les matrices perceptives sont organisées en une méta-structure arborescente qui autorise l'existence de familles de mots. Par exemple, un mot monosyllabique de 3 phonèmes comme « *gare* » forme l'initiale de 193 mots (LEXIQUE). Il existe donc 193 matrices qui commencent toutes au

---

<sup>134</sup> L'impact de cette minimisation de l'information acoustique initiale sur la reconnaissance serait compensé par au moins deux phénomènes, la polarisation de l'information selon un gradient décroissant entre début et fin de mot, et l'organisation générale du lexique avec concentration importante des superpositions lexicales sur les initiales de mot, au rôle facilitant sur la reconnaissance du mot porteur.

même « tronc ». Il est donc possible qu'au début du traitement, il n'y ait en réalité qu'une seule matrice qui soit générée, et qu'elle se développe ensuite sur un modèle de croissance végétale (la super-matrice potentialisée par « *gare* » se ramifie en 15 branches ou 15 sous-matrices au 4<sup>ème</sup> phonème, qui se ramifient à leur tour au phonème suivant, etc.). Lors du traitement, la perception suivrait donc un chemin au sein d'une matrice, mais aussi un chemin matriciel au sein du lexique. Par conséquent, en cas de présentation du « tronc » au système, tout un embranchement du lexique serait potentialisé, et un phénomène massif de facilitation lexicale devrait être obtenu sous l'effet du phénomène d'amorçage lexical précédemment évoqué. Les résultats de l'Expérience 4 obtenus avec des superpositions lexicales initiales virtuelles semblent appuyer une telle hypothèse.

### **5.3.4 Formalisation mathématique du modèle SLEXS1**

La formalisation mathématique développée dans cette section est une traduction des principes logiques de fonctionnement énoncés dans les sections précédentes. Selon SLEXS1, la reconnaissance d'un mot s'effectue en 2 étapes : la potentialisation de détecteurs (phonèmes et lexèmes) et l'intégration des lexèmes dans l'élaboration d'un percept lexical.

La potentialisation des lexèmes est calculée par la fonction  $E_i$ , qui exprime la quantité d'énergie infra-lexicale que reçoivent les lexèmes. L'intégration perceptive des lexèmes est calculée par deux fonctions. La première fonction, dite de pondération des lexèmes, ou  $P(x_{ij})$ , dépend de la position des lexèmes dans une matrice lexicale. La deuxième fonction, dite de prégnance perceptive du mot stimulus, ou  $f(x_{ij})$ , correspond à la résultante du processus.

#### **5.3.4.1 Modélisation de la potentialisation d'un lexème : la fonction $E_i$**

Tout lexème  $x_{ij}$  reçoit une quantité  $e_i$  d'énergie infra-lexicale de chaque phonème  $p$  qui entre dans la constitution du mot auquel ce lexème correspond. La quantité  $e_i$  est choisie comme égale à la valeur  $L$  qu'un lexème prend lorsqu'il est activé, supposée a priori identique pour tous les lexèmes (elle peut donc par exemple être fixée à 1), et divisée par le nombre  $i$  de phonèmes constituant le mot correspondant à ce lexème. La fonction  $E_i$  permettant de calculer l'énergie infra-lexicale totale reçue par un lexème est donnée dans l'Équation 6.

Équation 6 : Fonction permettant de calculer l'énergie infra-lexicale  $E_i$

$$E_i = \sum_1^i e_i$$

Avec  $e_i = L/i$ ,  $L$  pour la valeur d'activation du lexème lorsqu'il est activé et  $i$  pour le nombre de phonèmes constituant le mot stimulus.

La quantité d'énergie infra-lexicale transmise par phonème est par exemple de 0.25 pour un mot de 4 phonèmes, et de 0.2 pour un mot de 5 phonèmes.

### 5.3.4.2 Modélisation de la pondération d'un lexème dans la matrice lexicale : la fonction $P(x_{ij})$

La pondération des lexèmes est obtenue par la fonction  $P(x_{ij})$ , qui intègre deux paramètres : la position de chaque lexème dans la matrice perceptive  $ij$  et l'alignement d'un enchâssement lexical dans cette matrice.

Le premier paramètre, relatif à la position des lexèmes dans le mot stimulus, dépend de deux facteurs : l'existence d'un gradient décroissant d'information sur l'axe  $j$  dans la matrice lexicale (soit entre le début et la fin du mot), et la complexité du mot correspondant à un lexème donné (dépendant du nombre de phonèmes composant ce mot), représentée sur l'axe  $i$  de la matrice.

Ce paramètre est calculé par la fonction  $V_{ij}$ , correspondant à la somme des valences informationnelles  $v_i$  de chaque phonème selon la matrice  $ij$ , pour un lexème dont le dernier phonème occupe la position  $j$  dans la matrice lexicale. La valence informationnelle  $v_i$  d'un phonème est égale à  $v_u$ , la valence informationnelle unitaire d'un phonème libre, posée a priori égale à 1 pour tout phonème, divisée par son rang  $r_i$  dans un mot stimulus de  $i$  phonèmes. La fonction  $V_{ij}$  est donnée dans l'Équation 7.

Équation 7 : Fonction  $V_{ij}$  permettant de calculer la valence informationnelle d'un lexème  $x_{ij}$  dans la matrice lexicale

$$V_{ij} = \sum_1^i v_i$$

Avec  $v_i = v_u/r_i$ ,  $v_u$  la valeur unitaire du phonème et  $r_i$  son rang dans le mot.

Par exemple, la valence informationnelle du 2<sup>ème</sup> phonème d'un mot est de 0.5, celle du 6<sup>ème</sup> phonème de 0.167. La valence informationnelle d'un lexème aligné de 4 phonèmes dans une matrice lexicale de 5 phonèmes est donc de 0.417, et celle d'un lexème non aligné de 4 phonèmes (matrice lexicale de 5 phonèmes) est de 0.257.

Le deuxième paramètre (noté  $S_p$ ), relatif à l'alignement d'un enchâssement lexical dans la matrice, correspond simplement à l'attribution d'un signe en fonction de l'alignement de l'enchâssement. Le signe est positif dans le cas du lexème d'un mot aligné avec le mot stimulus. Il est négatif sinon. Le modèle est testé avec ou sans  $S_p$ .

La pondération d'un lexème  $x_{ij}$  de  $i$  phonèmes est calculée selon l'Équation 8 à partir de ces deux paramètres.

Équation 8 : Fonction de pondération  $P(x_{ij})$  d'un lexème  $x_{ij}$  dans la matrice lexicale  $ij$

$$P(x_{ij}) = S_p E_i V_{ij}$$

Avec  $S_p$  le signe de la pondération,  $E_i$  la quantité d'énergie infra-lexicale reçue par le lexème  $x_{ij}$  d'un mot de  $i$  phonèmes,  $V_{ij}$  la valence informationnelle d'un lexème  $x_{ij}$  dans une matrice  $ij$ .

### **5.3.4.3 Modélisation de la prégnance perceptive d'un mot stimulus : la fonction $f(x_{ij})$**

La prégnance perceptive d'un mot stimulus dépend de deux aspects. Le premier aspect est relatif à la valence informationnelle des lexèmes dans la matrice lexicale, à la fois influencée par la structure phonémique d'un mot et l'organisation ou contenu du lexique. Le deuxième aspect est la dynamique de l'élaboration du percept, à savoir le procédé par lequel les lexèmes sont utilisés pour composer un mot dans l'esprit du sujet.

Le premier aspect (valence informationnelle des lexèmes) est défini dans le modèle SLEXS1 par la fonction  $P(x_{ij})$  de pondération d'un lexème, déjà présentée à la section précédente. Le deuxième aspect (dynamique de l'élaboration du percept) est défini par la fonction de prégnance perceptive  $f(x_{ij})$  elle-même.

La dynamique de l'élaboration du percept lexical peut être conçue de plusieurs manières. Cinq procédures ont été considérées et testées sur des corpus de mots et sur des données comportementales afin de déterminer quelle conception était la plus représentative de la réalité. Elles sont décrites ci-dessous.

#### Dynamique 1

La première procédure envisagée est non récursive. Elle ne considère que l'impact facilitant des enchâssements initiaux afin de tenter de modéliser l'effet des facteurs manipulés dans les Expériences 4 à 10. Elle suppose que les enchâssements non initiaux sont sans influence sur la reconnaissance d'un mot stimulus. Cette procédure est décrite mathématiquement par la fonction  $f_1(x_{ij})$  dans l'Équation 9.

Équation 9 : Fonction  $f_1(x_{ij})$  de prégnance perceptive d'un mot stimulus

$$f_1(x_{ij}) = P(x_{ij}) + \sum_{q=1}^{q=i-1} P(x_{(i-q)(j-q)})$$

Avec  $i$  pour le nombre de phonèmes composant un mot,  $j$  le rang du dernier phonème du mot considéré dans la matrice  $ij$ , et  $q$  une variable prenant des valeurs entières comprises entre 1 et  $(i-1)$ .

Cette procédure s'avère cependant fournir des scores de prégnance perceptive qui ne sont pas prédictifs des temps de réponse. Elle n'a donc pas été conservée.

Dynamique 2

La deuxième procédure envisagée est également non récursive. Elle considère que toute superposition lexicale alignée de  $i-q$  phonèmes (avec  $1 < q < i-1$ ) facilite la reconnaissance d'un mot stimulus de  $i$  phonèmes, tandis que toute superposition lexicale non alignée de  $i-q$  phonèmes perturbe la reconnaissance du mot stimulus. Les phénomènes lexicaux lors de la reconnaissance d'un mot sont donc additifs, pour la facilitation comme pour la compétition. Cette procédure est décrite mathématiquement par la fonction  $f_2(x_{ij})$  dans l'Équation 10.

Équation 10 : Fonction  $f_2(x_{ij})$  de prégnance perceptive d'un mot stimulus

$$f_2(x_{ij}) = P(x_{ij}) + \sum_{q=1}^{q=i-1} P(x_{(i-q)(j-q)}) - \sum_{k=0}^{k=i-2} \left\{ \sum_{q=1}^{q=i-2} P(x_{(i-(q+k))(j-q+1)}) \right\}$$

Avec  $k$  une variable prenant des valeurs entières.

La procédure de calcul  $f_2$  n'a pas non plus été conservée, car elle ne permet pas d'obtenir un modèle expliquant correctement les données comportementales.

Dynamique 3

La troisième procédure envisagée est partiellement récursive. Elle considère que toute superposition lexicale alignée de  $i-1$  phonèmes facilite la reconnaissance d'un mot de  $i$  phonèmes, et que toute superposition lexicale non alignée de  $i-1$  phonèmes perturbe la reconnaissance d'un mot de  $i$  phonèmes. Le processus est récursif de proche en proche, ce qui signifie qu'un mot de  $i-1$  phonèmes est facilité par un mot enchâssé à son initiale de  $i-2$  phonèmes, et perturbé par un mot enchâssé non aligné de  $i-2$  phonèmes, etc. La prégnance perceptive d'un mot enchâssé de  $i$  phonèmes sera donc augmentée s'il renferme un mot enchâssé initial, et son interférence avec le mot  $i+1$  phonèmes sera donc plus importante. Par contre, elle sera diminuée s'il renferme un enchâssement non initial. Cette procédure est décrite mathématiquement par la fonction  $f_3(x_{ij})$  dans l'Équation 11.

Équation 11 : Fonction  $f_3(x_{ij})$  de prégnance perceptive d'un mot stimulus

$$f_3(x_{ij}) = P(x_{ij}) + f_3(x_{(i-1)(j-1)}) - f_3(x_{(i-1)j})$$

Les tests du modèle avec  $f_3$  ont montré que la procédure ne permet pas non plus prédire correctement les données comportementales mesurées. Elle a donc été rejetée.

Dynamique 4

La quatrième procédure envisagée est totalement récursive. Elle suit un schéma voisin de celui de la procédure précédente, en considérant que toute superposition lexicale alignée de  $i-q$  phonèmes (avec  $1 \leq q \leq i-1$ ) facilite la reconnaissance du mot-porteur de  $i$  phonèmes, et que toute superposition lexicale non alignée de  $i-q$  phonèmes perturbe la reconnaissance du mot-porteur de  $i$  phonèmes. Cette procédure est décrite mathématiquement par la fonction  $f_4(x_{ij})$  dans l'Équation 12.

Équation 12 : Fonction  $f_4(x_{ij})$  de prégnance perceptive d'un mot stimulus

$$f_4(x_{ij}) = [P(x_{ij})] + \left[ \sum_{q=1}^{q=i-1} f_4(x_{(i-q)(j-q)}) \right] - \left[ \sum_{k=0}^{k=i-2} \left\{ \sum_{q=1}^{q=i-1} f_{4na}(x_{(i-(q+k))(j-q+1)}) \right\} \right]$$

Avec  $f_{4na}(x_{ij})$  la fonction de prégnance perceptive d'un mot non aligné définie dans l'Équation 13.

Équation 13 : Fonction  $f_{4na}(x_{ij})$  de prégnance perceptive d'un mot non aligné

$$f_{4na}(x_{ij}) = [P(x_{ij})] - \left[ \sum_{q=1}^{q=i-1} f_{4na}(x_{(i-q)(j-(q-1))}) \right]$$

Cette quatrième procédure permet une bonne prédiction des données comportementales, à condition que  $S_p$  soit appliqué à  $P(x_{ij})$  (que les lexèmes soient pondérés positivement s'ils sont alignés avec le mot stimulus, et négativement s'ils ne sont pas alignés avec le mot stimulus). Elle a donc été conservée et utilisée dans un comparatif des alternatives possibles du modèle SLEXS.

Dynamique 5

La cinquième procédure est la procédure la plus élaborée des 5 procédures testées. Elle est très similaire à la quatrième pour ce qui est du plan général de la fonction. Elle en diffère cependant par la manière de calculer le score de prégnance perceptive des mots enchâssés non initiaux selon une procédure partiellement récursive. Si toute superposition lexicale non alignée est susceptible d'avoir un impact indirect sur la prégnance perceptive d'un mot non aligné de  $i$  phonèmes par le biais de la récursivité partielle de la fonction  $f_{na}$ , seulement 3 mots enchâssés peuvent avoir un impact direct

sur le score de prégnance perceptive du mot enchâssé considéré. Il s'agit de la superposition lexicale non alignée de  $i-1$  phonèmes et des superpositions lexicales alignées de  $i-1$  et  $i-2$  phonèmes. Un tel fonctionnement se base sur l'idée que la prégnance perceptive d'un mot non aligné de  $i-q$  phonèmes est diminuée par la présence d'un mot enchâssé aligné de  $i-(q+1)$  phonèmes, comme par la présence d'un mot enchâssé non aligné de  $i-(q+1)$ . Cependant, l'application récursive de ce principe de base fait apparaître un effet résiduel de la compétition entre mots non alignés sur  $i-q$ , qui se traduit par la réduction de l'interférence provoquée par le mot enchâssé de  $i-(q+1)$  aligné sur le mot  $i-q$ . La fonction  $f_{5na}(x_{ij})$  est donnée dans l'Équation 14.

Équation 14 : Fonction  $f_{5na}(x_{ij})$  de prégnance perceptive d'un mot non aligné

$$f_{5na}(x_{ij}) = P(x_{ij}) - P(x_{(i-1)j}) - P(x_{(i-1)(j-1)}) + P(x_{(i-2)(j-1)})$$

La fonction  $f_5(x_{ij})$  de prégnance perceptive d'un mot stimulus est donnée dans l'Équation 15.

Équation 15 : Fonction  $f_5(x_{ij})$  de prégnance perceptive d'un mot stimulus

$$f_5(x_{ij}) = [P(x_{ij})] + \left[ \sum_{q=1}^{q=i-1} f_5(x_{(i-q)(j-q)}) \right] - \left[ \sum_{k=0}^{k=i-2} \left\{ \sum_{q=1}^{q=i-1} f_{5na}(x_{(i-(q+k))(j-q+1)}) \right\} \right]$$

Ce modèle fournit des scores de prégnance perceptive qui permettent une bonne prédiction des données comportementales.

SLEXS1 comporte donc à ce stade 2 variantes,  $f_4$  et  $f_5$ , testées sur des corpus de mots et sur des données comportementales. Les résultats de ces tests sont présentés dans la section suivante.

## 5.4 Test du modèle SLEXS1

Le modèle SLEXS1 a été testé sur les corpus et les données comportementales des Expériences 4, 7 et 8, ainsi que sur le corpus et les données comportementales d'une expérience d'amorçage phonologique réalisée par Sophie Dufour et Ronald Peereman (2004)<sup>135</sup>, gracieusement mises à disposition par le premier auteur.

---

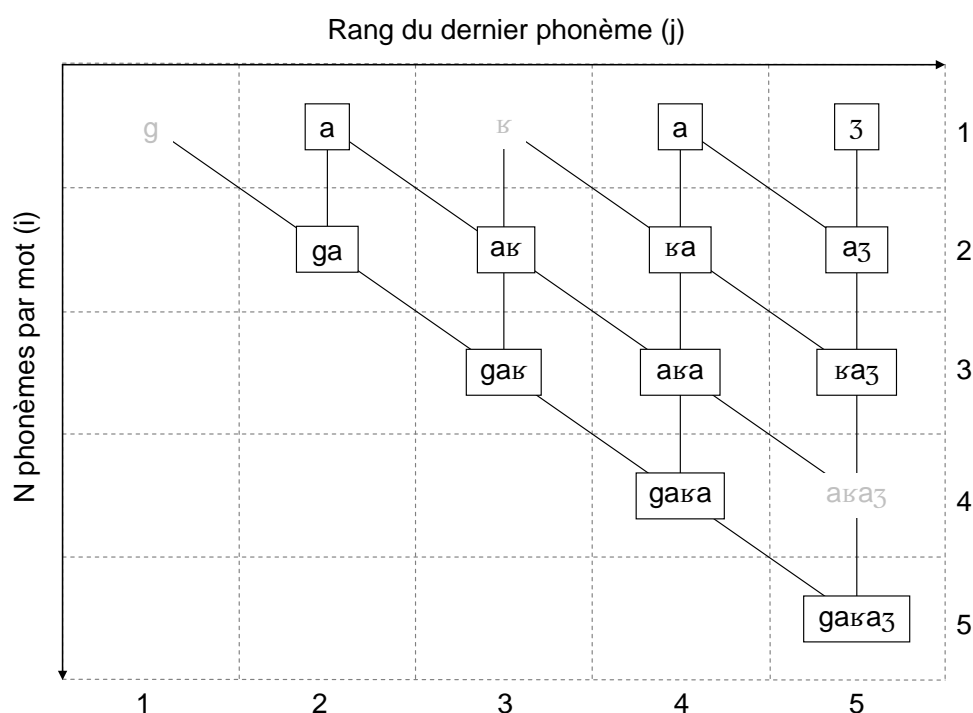
<sup>135</sup> Expérience 1, amorçage phonologique de mots cible bisyllabiques par les mots monosyllabiques qu'ils présentent enchâssés à leur initiale.

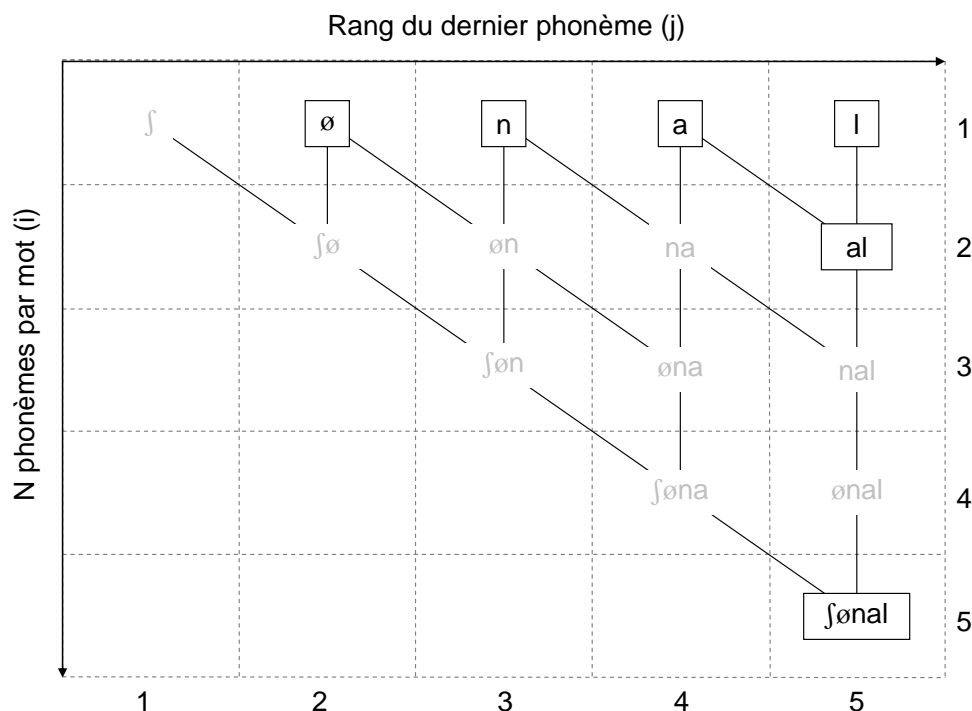


Les 2 variantes  $f_4$  et  $f_5$  du modèle, ainsi que la nécessité d'une pondération positive ou négative des lexèmes en fonction de l'alignement des mots enchâssés avec la séquence stimulus ( $S_p+$  en l'absence de pondération : tous les lexèmes ont une valence positive.  $S_p-$  en cas de pondération : les lexèmes alignés ont une valence positive, les lexèmes non alignés ont une valence négative), sont les 2 facteurs ayant servi pour organiser le test. Le test suit un plan complet (modèle ( $f_4$  vs.  $f_5$ ) \* Pondération ( $S_p+$  vs.  $S_p-$ )), qui détermine 4 manières différentes de calculer le score de prégnance perceptive ( $f_{4+}$ ,  $f_{5+}$ ,  $f_{4-}$ ,  $f_{5-}$ ).

Les scores de prégnance perceptive  $f_{4+}$ ,  $f_{5+}$ ,  $f_{4-}$ , et  $f_{5-}$  ont été calculés pour chaque mot stimulus à partir d'une analyse matricielle de leur structure superposante, telle que celles présentées dans les diagrammes de la Figure 74.

Figure 74 : Diagrammes de l'organisation superposante des mots « *garage* » et « *chenal* »





Avec en encadré, les lexèmes correspondant aux superpositions lexicales.

Pour l'expérience de Dufour et Peereman, les scores de prégnance perceptive de l'amorce ont également été calculés pour tester la capacité du modèle à rendre compte de données d'amorçage ayant testé spécifiquement le type de configuration lexicale modélisé par SLEXS1.

Chacun des 4 scores ( $f_{4+}$ ,  $f_{5+}$ ,  $f_{4-}$ ,  $f_{5-}$ ) est utilisé dans des analyses statistiques, comme variable dépendant des facteurs manipulés dans l'Expérience 4, dans l'Expérience 7, dans l'Expérience 8, et comme prédicteur des temps de réponse des participants, séparément pour chacune des 4 expériences.

Les analyses statistiques sont réalisées grâce à des modèles de régression linéaire à un niveau (test de la relation entre scores et caractéristiques lexicales des items) ou à deux niveaux (test de la relation entre scores et comportement).

### 5.4.1 Résultats des tests

#### 5.4.1.1 Expérience 4 : Test du modèle pour la reconnaissance de mots monosyllabiques traités en modalité auditive

Les scores de prégnance perceptive moyens par condition, fournis par chacune des 4 variantes de SLEXS1, sont donnés dans le Tableau 111, ainsi que les temps de réaction moyens par condition, estimés par le modèle utilisé pour l'analyse de l'Expérience 4 (Section 4.2.2.3 page 239).

Tableau 111 : Scores moyens de prégnance perceptive par condition et par modèle, et TR moyens estimés par condition

Condition expérimentale	$f_{4+}$	$f_{4-}$	$f_{5+}$	$f_{5-}$	TR
<b>E.I</b>	3.69	2.33	3.60	2.41	422
<b>E.nI</b>	3.96	2.86	3.93	2.89	665
<b>nE.I</b>	2.60	1.72	2.64	1.68	593
<b>nE.nI</b>	2.55	1.92	2.62	1.85	711

Les 4 modes de calcul des scores de prégnance perceptive génèrent des valeurs qui, en moyenne, apparaissent être cohérentes entre les modèles. Elles ne le sont par contre pas avec les TR moyens estimés, puisqu'on attendrait que les conditions soient ordonnées de la manière suivante selon des valeurs moyennes de TR de la plus petite à la plus élevée : E.nI, E.I, nE.I, nE.nI (plus un score est élevé, plus les temps de réponse sont bas). Or on a : E.I, nE.I, E.nI, nE.nI. Ceci s'explique probablement par le fait que SLEXS1 ne modélise pas l'effet de I, un effet de cohorte relatif à l'organisation des matrices lexicales entre elles (le facteur E étant relatif à l'organisation des lexèmes entre eux dans une matrice lexicale). S'il est fait abstraction de ce facteur I, le modèle semble capturer correctement l'influence du facteur E (scores moyennés sur E.I + E.nI plus élevés que les scores moyennés sur nE.I + nE.nI, et TR moyennés sur E.I + E.nI plus bas que les scores moyennés sur nE.I + nE.nI).

Le Tableau 112 fournit les résultats des analyses statistiques concernant la relation entre facteurs manipulés dans le matériel (présence d'un mot enchâssé, enchâssabilité du mot) et scores de prégnance perceptive.

Tableau 112 : Résultats des analyses de régression entre facteurs expérimentaux et scores de prégnance perceptive

		Coefficient	p<	Chi <sup>2</sup> <sub>(1, N=60)</sub>	%variance expliquée
<b>f4-</b>	<b>E</b>	1.401	<b>2<sup>E-09</sup></b>	37.02	38.1
	<b>I</b>	0.049	.84	0.05	0
	<b>E.I</b>	-0.319	.33	0.96	0.9
<b>f4+</b>	<b>E</b>	0.932	<b>3<sup>E-06</sup></b>	22.59	27.2
	<b>I</b>	-0.205	.30	1.09	1.3
	<b>E.I</b>	-0.325	.25	1.38	1.5
<b>f5-</b>	<b>E</b>	1.302	<b>4<sup>E-09</sup></b>	34.95	36.8
	<b>I</b>	0.022	.93	0.01	0
	<b>E.I</b>	-0.349	.27	1.25	1.4
<b>f5+</b>	<b>E</b>	1.031	<b>2<sup>E-07</sup></b>	27.52	31.4
	<b>I</b>	-0.177	.37	0.81	0.9
	<b>E.I</b>	-0.296	.29	1.13	1.2

Si les 4 scores utilisés permettent tous de modéliser le facteur E manipulé dans l'Expérience 4, seules les deux variantes  $f_{4-}$  et  $f_{5-}$  du modèle SLEXS1 offrent la meilleure congruence avec le facteur E (respectivement, 38 et 37% de la variance des scores de

prégnance perceptive est expliquée par E). Ce pourcentage n'est pas plus élevé, probablement à cause de la différence de précision entre les scores de SLEXS1, relatifs à la structure superposante complète de chaque mot (précision élevée), et les facteurs expérimentaux, relatifs à un sous-ensemble restreint de tous les enchâssements présents dans le mot stimulus (précision plus faible).

Aucun des 4 scores ne permet de rendre compte du facteur I, ni de l'interaction E.I. Cela apparaît comme étant tout à fait normal puisque SLEXS1 ne modélise pas dans sa version actuelle l'organisation des matrices dans le lexique mental.

Afin de connaître si les scores de SLEXS1 sont tout de même explicatifs des TR mesurés dans l'Expérience 4, l'analyse suivante a été effectuée (résultats présentés dans le Tableau 113).

Tableau 113 : Résultats de l'analyse de régression multiniveaux Scores/Comportement

	Coefficient	Chi <sup>2</sup> (1, N=2100)	p<	% variance expliquée
$f_{4-}$	-18.44	34.7	$4^{E-09}$	1.42
$f_{4+}$	9.266	6.1	.014	0.23
$f_{5-}$	-18.669	31.8	$2^{E-08}$	1.27
$f_{5+}$	6.569	3.2	.076	0.11

Deux des 4 variantes du modèle ( $f_{4-}$  et  $f_{5-}$ ) fournissent des scores significativement corrélés avec les TR mesurés ( $p < .0001$ ) et expliquent respectivement 1.4 et 1.3% de la variance comportementale. Le modèle capture donc un phénomène qui peut être en relation avec le facteur E (facteur I non modélisé par SLEXS1).

#### 5.4.1.2 Expérience 7 : Test du modèle pour la reconnaissance de mots bisyllabiques traités en modalité auditive

Les scores de prégnance perceptive moyens par condition, fournis par chacune des 4 variantes de SLEXS1, sont donnés dans le Tableau 114, ainsi que les temps de réaction moyens par condition estimés par le modèle utilisé pour l'analyse de l'Expérience 7 (Section 4.2.7.3 page 279).

Tableau 114 : Scores moyens de prégnance perceptive par condition et par modèle, et TR moyens estimés par condition

Condition expérimentale	$f_{4+}$	$f_{4-}$	$f_{5+}$	$f_{5-}$	TR
$M_1.M_2$	5.22	9.16	6.29	9.31	712
$M_1.nM_2$	6.30	9.92	7.32	10.01	704
$nM_1.M_2$	2.16	5.06	2.57	5.31	782
$nM_1.nM_2$	2.73	5.70	3.28	5.95	765

Les 4 modes de calcul des scores de prégnance perceptive génèrent des valeurs qui semblent être, en moyenne, en rapport avec les caractéristiques lexicales des items. En effet, en présence d'un mot enchâssé initial, le score de prégnance perceptive est plus élevé qu'en présence d'un mot enchâssé non initial. Par ailleurs, les scores apparaissent être en corrélation inverse avec le comportement des auditeurs (plus un score est élevé, plus les temps de réponse sont bas). Donc, les scores générés par SLEXS1 semblent en moyenne présenter une adéquation remarquable aux données comportementales.

Le Tableau 115 fournit les résultats des analyses statistiques concernant la relation entre facteurs manipulés dans le matériel (présence d'un mot enchâssé sur la première syllabe, présence d'un mot enchâssé sur la deuxième syllabe) et scores de prégnance perceptive.

Tableau 115 : Résultats des analyses de régression entre facteurs expérimentaux et scores de prégnance perceptive

		Coefficient	p<	Chi <sup>2</sup> <sub>(1, N=80)</sub>	%variance expliquée
f <sub>4</sub> -	M <sub>1</sub>	4.651	<b>.0001</b>	15.14	15.91
	M <sub>2</sub>	-0.659	.59	0.30	0.38
	M <sub>1</sub> .M <sub>2</sub>	0.073	.97	0.00	0.00
f <sub>4</sub> +	M <sub>1</sub>	3.573	<b>6<sup>E-09</sup></b>	34.18	29.94
	M <sub>2</sub>	-0.572	.35	0.88	1.09
	M <sub>1</sub> .M <sub>2</sub>	-0.508	.56	0.35	0.45
f <sub>5</sub> -	M <sub>1</sub>	4.06	<b>2<sup>E-05</sup></b>	19.01	19.20
	M <sub>2</sub>	-0.64	.50	0.47	0.58
	M <sub>1</sub> .M <sub>2</sub>	-0.062	.97	0.00	0.00
f <sub>5</sub> +	M <sub>1</sub>	4.043	<b>2<sup>E-09</sup></b>	36.35	31.25
	M <sub>2</sub>	-0.712	.29	1.13	1.38
	M <sub>1</sub> .M <sub>2</sub>	-0.322	.74	0.12	0.16

Comme pour l'Expérience 4, les 4 scores utilisés permettent tous de modéliser le facteur M<sub>1</sub><sup>136</sup> manipulé dans l'Expérience 7, mais les deux variantes f<sub>4</sub>+ et f<sub>5</sub>+ du modèle SLEXS1 offrent la meilleure performance, avec respectivement 30 et 32% de la variance des scores de prégnance perceptive expliquée par M<sub>1</sub>. Ce pourcentage n'est pas plus élevé, probablement à cause de la différence de précision entre les scores de SLEXS1, relatifs à la structure superposante complète de chaque mot, et les facteurs expérimentaux, relatifs à un sous-ensemble restreint de tous les enchâssements présents dans le mot stimulus.

---

<sup>136</sup> Facteur E dans l'Expérience 4.

Aucun des 4 scores ne permet de rendre compte du facteur  $M_2$ , ni de l'interaction  $M_1.M_2$ . Cela peut être dû à la différence d'information entre scores et facteurs expérimentaux et/ou au gradient de pondération de l'information dans un mot.

Cependant, le motif de résultats qui est obtenu est totalement concordant avec celui des résultats comportementaux de l'Expérience 7, dans laquelle un fort effet de la présence d'un enchâssement lexical initial était obtenu sur les temps de reconnaissance des mots, tandis que l'impact de la présence d'un enchâssement lexical non initial sur les temps de reconnaissance des mots restait caché. En d'autres termes, les scores de prégnance perceptive calculés sur le corpus de l'Expérience 7, utilisés comme variable dépendante à la place des temps de réponse, permettent d'obtenir un profil de résultats similaire à celui obtenu lors du test perceptif. Ce point est confirmé par les analyses suivantes, mettant en relation les scores de prégnance perceptive générés par SLEXS1 avec les temps de reconnaissance des mots mesurés chez l'Homme (Tableau 116).

Tableau 116 : Résultats de l'analyse de régression multiniveaux Scores/Comportement

	Coefficient	Chi <sup>2</sup> (1, N=5296)	p<	% variance expliquée
$f_{4-}$	-2.067	23	$2^{E-06}$	0.363
$f_{4+}$	-0.226	0.14	.72	0.004
$f_{5-}$	-1.404	6.65	<b>0.01</b>	0.110
$f_{5+}$	-2.155	22.98	$2^{E-06}$	0.365

Ces résultats montrent que 3 des 4 variantes du modèle ( $f_{4-}$ ,  $f_{5-}$  et  $f_{5+}$ ) permettent de générer des scores de prégnance perceptive prédisant le comportement des auditeurs, dont deux plus fortement ( $f_{4-}$  et  $f_{5+}$ ). Ainsi, plus un mot est perceptivement prégnant, plus la durée nécessaire pour le reconnaître diminue (en moyenne de 2 ms. par point supplémentaire de score de prégnance perceptive). L'effet est très fortement significatif ( $p < .0001$ ), bien que la quantité de variance totale expliquée soit assez faible (0.36%).

Les tests du modèle sur le corpus et les données de l'Expérience 7 apportent donc une validation des principes de SLEXS1.

#### 5.4.1.3 Expérience 8 : Test du modèle pour la reconnaissance de mots monosyllabiques traités en modalité visuelle

Les scores de prégnance perceptive moyens par condition, fournis par chacune des 4 variantes de SLEXS1, sont donnés dans le Tableau 117, ainsi que les temps de réaction moyens par condition estimés par le modèle utilisé pour l'analyse de l'Expérience 8 (Section 4.2.6.3 page 270).

Tableau 117 : Scores moyens de prégnance perceptive par condition et par modèle, et TR moyens estimés par condition

	$f_{4+}$	$f_{4-}$	$f_{5+}$	$f_{5-}$	TR
<b>Sans Enchâss<sup>t</sup>.</b>	4.92	3.74	5.11	3.56	451
<b>Enchâssement</b>	6.73	5.15	6.6	5.36	420

On retrouve ce qui était observé avec les données des Expériences 4 et 7 : les scores de prégnance perceptive sont plus élevés en cas d'enchâssement initial que dans un mot sans enchâssement initial, et ils semblent être en corrélation inverse avec les temps de reconnaissance des mots.

Le Tableau 118 présente les résultats des analyses de la relation entre le facteur expérimental (enchâssement) et les scores de prégnance perceptive.

Tableau 118 : Résultats des analyses de régression entre enchâssement et scores de prégnance perceptive

	Coefficient	Chi <sup>2</sup> (1, N=36)	p<	% variance expliquée
$f_{4-}$	1.952	7.44	.007	17.13
$f_{4+}$	1.409	2.41	.13	6.26
$f_{5-}$	1.564	4.42	.036	15.48
$f_{5+}$	1.796	5.27	.022	12.76

Comme pour les Expériences 4 et 7, SLEXS1 permet de modéliser la présence d'un enchâssement initial, puisque les modèles  $f_{4-}$  et  $f_{5+}$  conduisent à l'effet le plus significatif du facteur Enchâssement initial sur les scores de prégnance perceptive ( $.007 < p < .022$ , explication respective de 17% et 13 % de la variance des scores par E).

Le Tableau 119 présente les résultats de l'analyse de régression multiniveaux testant la relation entre scores de prégnance perceptive et comportement.

Tableau 119 : Résultats des analyses de régression multiniveaux Scores/Comportement

	Coefficient	Chi <sup>2</sup> (1, N=408)	p<	% variance expliquée
$f_{4-}$	-2.396	1.53	.22	0.40
$f_{4+}$	-1.777	1.19	.28	0.31
$f_{5-}$	-1.385	0.53	.47	0.15
$f_{5+}$	-3.135	2.91	.088	0.74

Contrairement à ce qui était obtenu avec les Expériences 4 et 7, SLEXS1 échoue à modéliser de manière significative l'effet comportemental obtenu dans l'Expérience 8. Une tendance existe cependant pour la variante  $f_{5+}$  du modèle, qui de plus explique le pourcentage le plus élevé de variance des temps de réponse.

Cette carence du modèle peut être due à la modélisation du gradient d'information lexicale selon un mode propre à l'oral, qui ne se justifierait pas pour des mots monosyllabiques présentés en modalité visuelle (totalité de l'information présente au temps  $t$ ).

De nouveaux scores ont donc été calculés en retirant le gradient de la fonction de pondération. Le Tableau 120 présente les résultats de l'analyse de régression multiniveaux testant la relation entre ces nouveaux scores de prégnance perceptive et le comportement des auditeurs.

Tableau 120 : Résultats des analyses de régression multiniveaux Scores sans gradient/Comportement

	Coefficient	Chi <sup>2</sup> (1, N=408)	p<	% variance expliquée
$f_{4bis-}$	-0.858	0.62	.44	0.18
$f_{4bis+}$	-0.424	0.62	.26	0.08
$f_{5bis-}$	0.244	0.05	.83	0.001
$f_{5bis+}$	-1.79	2.74	.099	0.74

Les résultats obtenus avec ces scores diffèrent peu ou sont plus mauvais que ceux obtenus avec des scores modélisant le gradient d'information dans un mot.

Par conséquent, le modèle SLEXS1 apparaît mal adapté pour rendre compte des phénomènes d'enchâssement avec des mots monosyllabiques traités en modalité visuelle. Toutefois, une tendance est retrouvée avec la variante  $f_{5+}$ , indiquant que les principes modélisés dans SLEXS1 ne sont pas totalement invalides dans cette situation.

#### 5.4.1.4 Expérience d'amorçage : Test du modèle sur les données de l'expérience d'amorçage lexical de Dufour et Peereman (2004)

Les scores d'activation des cibles bisyllabiques et des amorces monosyllabiques correspondant aux mots enchâssés ont été utilisés comme prédicteurs des temps de réponse pour reconnaître les mots cible. Les résultats des analyses statistiques sont présentés dans le Tableau 121.

Tableau 121 : Résultats des analyses de régression multiniveaux Scores/Comportement

Condition	Prédicteur	Coefficient	Chi <sup>2</sup> (1, N=1749)	p<	% variance expliquée
Amorçage	A	-12.404	10.167	<b>.0015</b>	0.31
	$f_{4-}$	-1.45	17.537	<b>3<sup>E-05</sup></b>	0.49
	$f_{4+}$	-4.157	35.989	<b>2<sup>E-09</sup></b>	0.98
	$f_{5-}$	-2.36	29.085	<b>7<sup>E-08</sup></b>	0.83
	$f_{5+}$	-2.589	21.835	<b>3<sup>E-06</sup></b>	0.57



<b>Non relié</b>	<b>A</b>	-1.065	0.087	.77	0	
	<b>f<sub>4</sub>-</b>	-0.799	6.065	<b>.014</b>	0.21	
	<b>f<sub>4</sub>+</b>	-3.136	23.459	<b>2<sup>E-06</sup></b>	0.71	
	<b>f<sub>5</sub>-</b>	-1.463	12.718	<b>.0004</b>	0.40	
	<b>f<sub>5</sub>+</b>	-1.718	10.918	<b>.001</b>	0.36	
<b>Amorçage</b>	<b>Score amorce</b>	<b>f<sub>4</sub>-</b>	-5.062	5.504	<b>.019</b>	0.17
		<b>f<sub>4</sub>+</b>	-11.852	9.869	<b>.0017</b>	0.33
		<b>f<sub>5</sub>-</b>	-5.733	6.329	<b>.012</b>	0.20
		<b>f<sub>5</sub>+</b>	-9.832	8.021	<b>.005</b>	0.26

Avec A pour le facteur amorçage manipulé dans l'expérience de Dufour et Peereman.

Ces résultats montrent que le modèle SLEXS1, dans ses versions  $f_4+$  et  $f_5+$ , capture correctement le rôle de l'amorce sur le traitement de la cible,  $f_4+$  permettant d'obtenir un effet quasi identique à celui du facteur manipulé dans l'expérience de Dufour et Peereman.

Ces résultats montrent de plus que le modèle explique de manière significative la structure superposante des mots cible, le meilleur candidat sur ces données étant  $f_4+$ .

### 5.4.2 Conclusions concernant le modèle SLEXS1

Le modèle SLEXS1 apporte une confirmation mathématique de la proposition faite dans cette étude, selon laquelle les superpositions lexicales initiales favorisent le traitement lexical au lieu de le perturber. L'organisation du lexique telle qu'elle est présentée dans la Section 2.3 page 61 (répartition des superpositions lexicales en quantité plus importante à l'initiale des mots qu'en position non initiale) correspondrait bien à un indice de la réalité fonctionnelle facilitante que remplissent les superpositions lexicales initiales lors de la reconnaissance des mots parlés.

De plus, le modèle SLEXS1 semble montrer mathématiquement que les superpositions lexicales non initiales ont effectivement un rôle perturbateur sur la reconnaissance du mot stimulus, puisque les variantes du modèle qui n'intègrent pas une composante compétitive pour les mots enchâssés non alignés, génèrent des scores qui semblent être moins bien en rapport avec les données comportementales. SLEXS1 confirme ainsi les propositions avancées dans la littérature par des modèles comme TRACE.

SLEXS1 met en avant l'insuffisance des modèles de la psycholinguistique, qui ne différencient pas l'impact des superpositions lexicales sur la reconnaissance du mot stimulus en fonction de leur position (initiale facilitante, non initiale perturbante). La manière dont les représentations mentales lexicales entrent mutuellement en relation, ainsi que la structure interne de ces représentations, pourrait donc être plus élaborée que cela n'a été supposé jusqu'ici. En particulier, le processus de reconnaissance d'un mot pourrait être plus élaboré qu'un simple phénomène d'activation d'unités stockées en mémoire à long terme. Avec SLEXS, on s'achemine vers un phénomène de construction perceptive intégratif, qui dépend certes de l'activation d'éléments en mémoire à long terme, mais ne s'y limite pas.

Néanmoins, malgré ces premiers résultats encourageants, la performance du modèle est assez fluctuante, indiquant que celui-ci doit être retravaillé, complété, et testé systématiquement dans toutes ses variantes sur un lot plus important de données afin de poser un diagnostic précis. En particulier, le fait que les performances varient selon que des mots monosyllabiques ou bisyllabiques sont utilisés laisse penser que des phénomènes de cohorte non modélisés interviennent (mots monosyllabiques), et que des phénomènes de segmentation lexicale non modélisés ont lieu en plus de phénomènes de cohorte dans le cas des mots bisyllabiques. Ces points ne remettent cependant pas en question les bases jetées par SLEXS1, puisque le modèle s'avère à même de capturer le

phénomène d'amorçage lexical qu'il est supposé représenter (Expériences 4 et 7, expérience de Dufour et Peereman). À titre d'illustration (Tableau 122), la performance du modèle SLEXS1 (% de variance des TR expliqué par le modèle) est comparée au pourcentage de variance comportementale expliqué par les facteurs expérimentaux utilisés dans les expériences.

Tableau 122 : Comparaison du pourcentage de variance comportementale expliquée par le modèle SLEXS1 et par les facteurs expérimentaux E ou M<sub>1</sub>

% de variance expliquée		SLEXS1	Facteur expérimental E ou M <sub>1</sub>
Expérience 4	f <sub>4</sub> -	1.30	0.94
	f <sub>5</sub> -	1.40	
Expérience 7	f <sub>4</sub> -	0.36	0.13
	f <sub>5</sub> +	0.36	
Expérience 8	f <sub>4</sub> -	0.40	2.09
	f <sub>5</sub> +	0.74	
Expérience Dufour & Peereman	f <sub>4</sub> +	0.33 (amorçage)	0.31
		0.98 (cible)	-

Hormis pour l'Expérience 8, SLEXS1 permet d'expliquer plus précisément la variance comportementale que les facteurs utilisés.

Par ailleurs, le modèle SLEXS pourrait aisément intégrer l'information prosodique afin de moduler les scores de prégnance perceptive des mots en fonction de cette dimension organisatrice.

Enfin, SLEXS apporte une solution au problème de la segmentation du flux de parole en mots, et une méthode pour tester cette hypothèse (calcul de scores de prégnance perceptive). La localisation des frontières lexicales ne serait pas un écueil théorique et pratique, car elle résulterait naturellement de la prégnance perceptive des mots, les mots produits ayant une prégnance perceptive supérieure aux mots possibles, les mots à prégnance perceptive supérieure étant les mots reconnus.

## 6 Conclusion générale - Les apports de ce travail

Ce travail de thèse de doctorat s'est penché sur la problématique de la reconnaissance des mots parlés en Français et son corolaire, le problème de la segmentation du flux de parole en mots.

Il a exploré le phénomène sous un double aspect, en cherchant d'une part à préciser par quels moyens le système perceptif est susceptible de localiser les frontières lexicales à l'aide de la prosodie, et d'autre part, en cherchant à comprendre comment il exploite l'organisation superposante du lexique pour biaiser positivement la prégnance perceptive des mots avec, comme conséquence, une facilitation de la segmentation du flux de parole en mots. Il a également cherché à comprendre comment ces deux dimensions, la prosodie et le lexique, pouvaient interagir pour aboutir à la production des percepts lexicaux.

Les points essentiels de ce travail sont que :

- ▷ Il existe un code prososyllabique dans le signal de parole, exprimant directement la manière dont le locuteur manipule mentalement le langage lorsqu'il le produit, et à même de guider efficacement le système auditeur pour la localisation des frontières lexicales, et donc pour la segmentation du flux sonore en mots.
- ▷ La prosodie contraint de manière relativement forte le traitement lexical.
- ▷ Il existe une organisation particulière du lexique, telle que la concentration des superpositions lexicales initiales soit près de deux fois plus importante que sur une autre position, et que la variété des syllabes initiales soit moindre que celle des syllabes non initiales.
- ▷ Cette asymétrie existe à dessein, parce qu'elle permet au système d'exploiter le rôle facilitant des superpositions lexicales initiales sur la reconnaissance des mots parlés.
- ▷ Il existe une asymétrie lexicale qui se manifeste par une polarité ou gradient décroissant d'information dans une représentation lexicale. Le début des mots est donc primordial pour le traitement.
- ▷ Le comportement de facilitation observé lors de la reconnaissance de mots renfermant une superposition lexicale initiale, quoique paradoxal dans le cadre théorique de la psycholinguistique contemporaine, peut s'expliquer par un phénomène d'amorçage lexical. Il est mathématiquement modélisé par SLEXS1 (Structure LEXicale Superposante, version 1).
- ▷ Le modèle SLEXS propose d'expliquer la segmentation du flux de parole en mots comme résultat de la prégnance perceptive des mots. Mots prégnants = mots reconnus = mots produits ; mots possibles = mots non prégnants = mots non reconnus.

Ce travail de recherche s'achève sur un important projet de développement et de test du modèle SLEXS :

- ▷ L'élaboration du modèle devra être poursuivie, certains points restant à formuler, notamment le rapport que les matrices lexicales entretiennent entre elles dans le lexique, seul à même de rendre compte des effets des superpositions lexicales virtuelles.
- ▷ Une batterie de tests des différentes variantes du modèle devra être effectuée systématiquement sur un ensemble plus vaste de données comportementales. Les causes de la variation des performances du modèle sur des jeux de données provenant de différentes expériences devront également être précisées.
- ▷ Le modèle devra être implémenté dans un dispositif permettant de l'utiliser pour effectuer des simulations.
- ▷ Des prédictions devront être générées par le modèle, puis être testées expérimentalement auprès de sujets humains.

## Bibliographie et logigraphie

- Alario, F. X., & Ferrand, L. (1999). A set of 400 pictures standardized for French: norms for name agreement, image agreement, familiarity, visual complexity, image variability, and age of acquisition. *Behavior Research Methods, Instruments and Computers*, 31(3), 531-552.
- Amit, D. J. (1995). The Hebbian paradigm reintegrated: Local reverberations as internal representations. *Behavioral and Brain Sciences*, 18(4), 617-657.
- Anderson, J. L., Morgan, J. L., & White, K. S. (2003). A Statistical Basis for Speech Sound Discrimination. *Language and speech*, 46(2-3), 155-182.
- Andrews, S. (1986). Morphological influences on lexical access: Lexical or nonlexical effects? *Journal of Memory and Language*, 25(6), 726-740.
- Aptech Systems. (1998). Gauss. Maple Valley, WA, USA.
- Baayen, H., Pipenbrook, R., & Guilikers, L. (1995). *The Celex lexical database (v2.5)*
- Bagou, O., Fougeron, C., & Frauenfelder, U. (2002, 11-13 avril 2002). *Contribution of prosody to the segmentation and storage of "words" in the acquisition of a new mini-language*. Paper presented at the Speech prosody, Aix-en-Provence.
- Barry, C., & Gerhand, S. (2003). Both concreteness and age-of-acquisition affect reading accuracy but only concreteness affects comprehension in a deep dyslexic patient. *Brain and Language*, 84(1), 84-104.
- Barry, C., Hirsh, K., Johnston, R., & Williams, C. (2001). Age of acquisition, word frequency, and the locus of repetition priming of picture naming. *Journal of memory and language*, 44(3), 350-375.
- Barry, C., Morrison, C. M., & Ellis, A. W. (1997). Naming the Snodgrass and Vanderwart pictures: Effects of age of acquisition, frequency and name agreement. *The Quarterly Journal of Experimental Psychology: Human Experimental Psychology*, 50A(3), 560-585.
- Baumeister, A. A. (1985). Age of acquisition and meaningfulness as predictors of word availability. *Journal of General Psychology*, 112(1), 109-112.
- Bird, H., Franklin, S., & Howard, D. (2001). Age of acquisition and imageability ratings for a large set of words, including verbs and function words. *Behavior Research Methods, Instruments and Computers*, 33(1), 73-79.
- Boersma, P., & Weenink, D. (2000). Praat (Version 3.9 & +). Amsterdam, The Netherlands: Institute of Phonetic Sciences, University of Amsterdam.

- Bonin, P., Chalard, M., Meot, A., & Fayol, M. (2002). The determinants of spoken and written picture naming latencies. *British Journal of Psychology*, 93(Pt 1), 89-114.
- Bonin, P., Fayol, M., & Chalard, M. (2001). Age of acquisition and word frequency in written picture naming. *The Quarterly Journal of Experimental Psychology A*, 54(2), 469-489.
- Bonin, P., Peereman, R., Malardier, N., Méot, A., & Chalard, M. (2003). A new set of 299 pictures for psycholinguistic studies: French norms for name agreement, image agreement, conceptual familiarity, visual complexity, image variability, age of acquisition, and naming latencies. *Behavior Research Methods, Instruments and Computers*, 35(1), 158-167.
- Booij, G., Nespors, M., & Vogel, I. (1988). Prosodic Phonology (Studies in Generative Grammar). *Journal of Linguistics*, 24(2), 515-525.
- Bowers, J. S., Colin, J. D., & Derek, A. H. (2005). Automatic activation of embedded words: is there a "hat" in "that"? *Journal of Memory and Language*, 52(1), 131-143.
- Brown, G. D., & Watson, F. L. (1987). First in, first out: word learning age and spoken word frequency as predictors of word familiarity and word naming latency. *Memory and Cognition*, 15(3), 208-216.
- Brysbaert, M. (1996). Word frequency affects naming latency in Dutch when age of acquisition is controlled. *European Journal of Cognitive Psychology*, 8(2), 185-193.
- Brysbaert, M., Lange, M., & Van Wijnendaele, I. (2000). The effects of age-of-acquisition and frequency-of-occurrence in visual word recognition: Further evidence from the Dutch language. *European Journal of Cognitive Psychology*, 12(1), 65-85.
- Brysbaert, M., Van Wijnendaele, I., & De Deyne, S. (2000). Age-of-acquisition effects in semantic processing tasks. *Acta Psychologica*, 104(2), 215-226.
- Caelen-Haumont, G. (1996). *Prosody and disambiguation*. Paper presented at the Post-Coling seminar on Interactive Disambiguation, MIDDIM-96, Col-de-Porte.
- Caelen-Haumont, G., & Keller, E. (1997). La prosodie, de la parole à la synthèse: l'apport de la sémantique et de la pragmatique. *Etudes de lettres, Revue de la Faculté de lettres de l'Université de Lausanne*, 103-130.
- Caillaud, B. (1996). *Apprentissage de connaissances prosodiques pour la reconnaissance automatique de la parole*. Grenoble.
- Caillaud, B., Antoine, J.-Y., Caelen, J., & Caelen-Haumont, G. (1993). *MICRO : un système multi-agents pour la compréhension de la parole*. Paper presented at the Reconnaissance des Formes Intelligence Artificielle, RIFA'93, Paris, France.
- Caillaud, B., Munteanu, P., Serignat, J.-F., & Caelen, J. (1998). Prosodic Knowledge Acquisition for Lexical Access Improvement. *The Journal for the Integrated Study of*

*Artificial Intelligence, Cognitive Science and Applied Epistemology (CC AI)*, 15(3), 255-278.

Calliope. (1989). *La parole et son traitement automatique*: Masson.

Chalard, M., Bonin, P., Méot, A., Boyer, B., & Fayol, M. (2003). Objective age-of-acquisition (A0A) norms for a set of 230 object names in French: Relationships with psycholinguistic variables, the English data from Morrison et al. (1997), and naming latencies. *European Journal of Cognitive Psychology*, 15(2), 209-245.

Clark, H. H. (1973). The language-as-fixed-effect fallacy: a critique of language statistics in psychological research. *Journal of Verbal Learning and Verbal Behavior*, 12, 335-359.

Clark, H. H. (1976). Discussion of Wike and Church's comments. *Journal of Verbal Learning and Verbal Behavior*, 15, 257-266.

Cochran, W. G. (1947). Some consequences when the assumptions for the analysis of variance are not satisfied. *Biometrics*, 3, 22-38.

Cole, R., & Jakimik, J. (1980). A model of speech perception. In R. Cole (Ed.), *Perception and production of fluent speech* (pp. 133-163). Hillsdale, NJ: Erlbaum.

Connine, C. M., Mullennix, J., Shernoff, E., & Yelen, J. (1990). Word familiarity and frequency in visual and auditory word recognition. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 16(6), 1084-1096.

Content, A., Kearns, R. K., & Frauenfelder, U. H. (2001). Boundaries versus onsets in syllabic segmentation. *Journal of Memory and Language*, 45(2), 177-199.

Content, A., Mousty, P., & Radeau, M. (1990). BRULEX. Une base de données lexicales informatisée pour le français écrit et parlé. *L'année Psychologique*, 90(4), 551-566.

Costa, A., Cutler, A., & Sebastian-Galles, N. (1998). Effects of phoneme repertoire on phoneme decision. *Perception and Psychophysics*, 60(6), 1022-1031.

Crowder, R. G. (1978). Mechanisms of auditory backward masking in the stimulus suffix effect. *Psychological Review*, 85, 502-524.

Crowder, R. G. (1981). The role of auditory memory in speech perception and discrimination. In T. Myers, J. Laver & J. Anderson (Eds.), *The cognitive representation of speech* (pp. 167-179). New York.

Culter, A. (1986). Forbear is a homophone: lexical prosody does not constrain lexical access. *Language and speech*, 29(3), 201-220.

Cutler, A. (2000). Real words, phantom words and impossible words. In D. Burnham, S. Luksaneeyanawin, C. Davis & M. Lafourcade (Eds.), *Approaches to Language*



*Processing: The International Conference on Human and Machine Processing of Language and Speech* (pp. 32-42). Bangkok: NECTEC.

Cutler, A., & Carter, D. (1987). The predominance of strong initial syllables in the English vocabulary. *Computer speech and language*, 2(3-4), 133-142.

Cutler, A., Dahan, D., & van Donselaar, W. (1997). Prosody in the comprehension of spoken language: a literature review. *Language and speech*, 40 ( Pt 2), 141-201.

Cutler, A., Kearns, R. K., Norris, D., & Scott, D. R. (1993). Problems with click detection: Insights from cross-linguistic comparisons. *Speech Communication*, 13(3-4), 401-410.

Cutler, A., McQueen, J. M., Jansonius, M., & Bayerl, S. (2002). *The lexical statistics of competitor activation in spoken-word recognition*. Paper presented at the 9th Australian International Conference on Speech Science and Technology, Melbourne.

Cutler, A., McQueen, J. M., Norris, D., & Somejuan, A. (2001). The roll of the silly ball. In E. Dupoux (Ed.), *Language, brain, and cognitive development: Essays in honor of Jacques Mehler* (pp. 181-194).

Cutler, A., & Norris, D. (1988). The role of strong syllables in segmentation for lexical access. *Journal of Experimental Psychology: Human Perception and Performance*, 14(1), 113-121.

Cutler, A., & van Donselaar, W. (2001). Voornaam is not (really) a homophone: Lexical prosody and lexical access in Dutch. *Language and speech*, 44(Pt 2), 171-195.

Dashorst, M. (1997). *Designing a finetuning algorithm for RAW*. University of Twente, Twente.

Davidson, R., & MacKinnon, J. G. (1993). *Estimation and inference in econometrics*. New-York: Oxford University Press.

Davis, M. H. (2000). *Lexical segmentation in spoken word recognition*. University of London, London.

Davis, M. H., Gaskell, M. G., & Marslen-Wilson, W. D. (1997a). *Learning to recognise embedded words in connected speech*. Paper presented at the First Computational Psycholinguistics workshop, Berkeley, CA.

Davis, M. H., Gaskell, M. G., & Marslen-Wilson, W. D. (1997b). *Recognising embedded words in connected speech: context and competition*. Paper presented at the 4th Neural Computation in Psychology Workshop, Centre for Speech and language, London.

Davis, M. H., Marslen-Wilson, W. D., & Gaskell, M. G. (2000). *Lexical segmentation and ambiguity: investigating the recognition of onset-embedded words*. Paper presented at the Spoken Word Access Processes (SWAP).

- Davis, M. H., Marslen-Wilson, W. D., & Gaskell, M. G. (2002). Leading up the lexical garden path: Segmentation and ambiguity in spoken word recognition. *Journal of Experimental Psychology: Human Perception and Performance*, 28(1), 218-244.
- Demierre, P., & Serena, F. (2002). Formulaire de recueil des inscriptions aux activités facultatives. Genève, GE, Switzerland: Division Informatique, Développement et Maintenance, Université de Genève.
- Dempster, A. P., Laird, N. M., & Rubin, D. B. (1977). Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society*, 39(B), 1-38.
- Descout, R., Serignat, J. F., Cervantes, O., & Carré, R. (1986). *BDSONS: une base de données des sons du français*. Paper presented at the 12th International congress on acoustics, Toronto, Canada.
- Desrochers, A., & Bergeron, M. (2000). Valeurs de fréquence subjective et d'imagerie pour un échantillon de 1916 substantifs de la langue française. *Revue Canadienne de Psychologie Expérimentale*, 54(4), 274-325.
- Dreyfus, G., Martinez, J.-M., Samuelides, M., Gordon, M. B., Badran, F., Thiria, S., et al. (2002). *Réseaux de neurones*. Paris: Eyrolles.
- Dufour, S., & Peereman, R. (2004). Phonological priming in auditory word recognition: initial overlap facilitation effect varies as a function of target word frequency. *Current psychology letters: Behaviour, brain and cognition*, 14(3).
- Dumay, N., Frauenfelder, U. H., & Content, A. (2002). The role of the syllable in lexical segmentation in French: word-spotting data. *Brain and language*, 81(1-3), 144-161.
- Dutoit, T., Pagel, V., Pierret, N., Bataille, F., & Van der Vreken, O. (1996). *The MBROLA project: towards a set of high-quality speech synthesizers free of use for non-commercial purposes*. Paper presented at the ICSLP'96, Philadelphia, PA, USA.
- Efron, R. (1970). The minimum duration of a perception. *Neuropsychologia*, 8(1), 57-63.
- Eimas, P. D., Cooper, W. E., & Corbit, J. D. (1973). Some properties of linguistic feature detectors. *Perception and Psychophysics*, 13(2), 247-252.
- Ellis, A. W., & Morrison, C. M. (1998). Real age-of-acquisition effects in lexical retrieval. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 24(2), 515-523.
- Elman, J. L., & McClelland, J. L. (1988). Cognitive penetration of the mechanisms of perception: compensation for coarticulation of lexically restored phonemes. *Journal of Memory and Language*, 27(2), 143-165.
- Entropic Research Laboratory. (1998). ESPS (Version 5.3.1). Washington, DC, USA.

- Erman, L. D., & Lesser, U. R. (1980). The HEARSAY-II speech understanding system: a tutorial. In W. A. Lea (Ed.), *Trends in speech recognition* (pp. 361-381). Englewood Cliffs, NJ: Prentice-Hall.
- Espesser, R. (1994). MES Signaux. Aix-en-Provence: Laboratoire Parole et Langage, CNRS et Université de Provence, France.
- Ferrand, L., Grainger, J., & New, B. (2003). Normes d'âge d'acquisition pour 400 mots monosyllabiques. *L'année Psychologique*, *104*, 445-468.
- Forster, K. I. (1976). Accessing the mental lexicon. In R. J. W. E. Walker (Ed.), *New approaches to language mechanisms* (pp. 257-287). Amsterdam: North Holland.
- Forster, K. I., & Dickinson, R. G. (1976). More on the Language-as-fixed-effect fallacy: Monte Carlo estimates of error rates for F1, F2, F' and min F'. *Journal of Verbal Learning and Verbal Behavior*, *15*, 135-142.
- Frauenfelder, U. H., & Peeters, G. (1990). On lexical segmentation in TRACE: An exercise in simulation. In G. T. Altmann (Ed.), *Cognitive models of speech processing: psycholinguistic and computational perspectives*: MIT Press.
- Garlock, V. M., & Walley, A. C. (2001). Age-of-acquisition, word frequency, and neighborhood density effects on spoken word recognition by children and adults. *Journal of memory and language*, *45*, 468-492.
- Gaskell, M. G., & Marslen-Wilson, W. D. (1997). Integrating form and meaning: A distributed model of speech perception. *Language and Cognitive Processes*, *12*(5-6), 613-656.
- Gaskell, M. G., & Marslen-Wilson, W. D. (2001). Lexical ambiguity resolution and spoken word recognition: Bridging the gap. *Journal of Memory and Language*, *44*(3), 325-349.
- Gerhand, S., & Barry, C. (1998). Word frequency effects in oral reading are not merely age-of-acquisition effects in disguise. *Journal of experimental psychology: learning, memory and cognition*, *24*(2), 267-283.
- Gerhand, S., & Barry, C. (1999a). Age of acquisition, word frequency, and the role of phonology in the lexical decision task. *Memory and Cognition*, *27*(4), 592-602.
- Gerhand, S., & Barry, C. (1999b). Age-of-acquisition and frequency effects in speeded word naming. *Cognition*, *73*(2), B27-36.
- Gilhooly, K. J., & Gilhooly, M. L. (1980). The validity of age-of-acquisition ratings. *British Journal of Psychology*, *71*(1), 105-110.
- Goldstein, H. (2003). *Multilevel statistical models* (3 ed.). London: Hodder Arnold, New York, Wiley.

- Grosjean, F. (1985). The recognition of words after their acoustic offset: evidence and implications. *Perception and Psychophysics*, 38(4), 299-310.
- Harris, Z. (1955). From Phoneme to Morpheme. *Language*, 31(2), 190-222.
- Haykin, S. (1999). *Neural networks. A comprehensive foundation*. Upper Saddle River, New Jersey: Prentice Hall International, Inc.
- Hirsh, K., & Funnell, E. (1995). Those old, familiar things: Age of acquisition, familiarity and lexical access in progressive aphasia. *Journal of neurolinguistics*, 9(1), 23-32.
- Hox, J. (2002). *Multilevel analysis: techniques and applications*. London: Lawrence Erlbaum Associates, Publishers.
- Imbs, P. (1971). *Etudes statistiques sur le vocabulaire français. Dictionnaire de fréquences. Vocabulaire littéraire des XIX' et XX' siècles*. Paris-Nancy: Librairie Marcel-Didier.
- Isel, F., & Bacri, N. (1999). Spoken word recognition: the access to embedded words. *Brain and language*, 68(1-2), 61-67.
- Izura, C., & Ellis, A. W. (2002). Age of acquisition effects in word recognition and production in first and second languages. *Psicologica*, 23(2), 245-281.
- Jakobson, R., Fant, G., & Halle, M. (1952). *Preliminaries to speech analysis*. Cambridge: MIT Press.
- Jones, M. R. (1976). Time, our lost dimension: toward a new theory of perception, attention and memory. *Psychological Review*, 83, 323-355.
- Jones, M. R., Kidd, G., & Wetzel, R. (1981). Evidence for rhythmic attention. *Journal of Experimental Psychology: Human Perception and Performance*, 7, 1059-1073.
- Jorm, A. F. (1991). The validity of word age-of-acquisition ratings: A longitudinal study of a child's word knowledge. *British journal of developmental psychology*, 9(3), 437-444.
- Kirk, R. E. (1968). *Experimental design: procedures for the behavioral sciences* (1968 ed.). Belmont, California, USA: Wadsworth Publishing Company, Inc.
- Klatt, D. H. (1979). Speech perception: a model of acoustic-phonetic analysis and lexical access. *Journal of Phonetics*, 7, 279-312.
- Klatt, D. H. (1981). Lexical representations for speech production and perception. In T. Myers, J. Laver & J. Anderson (Eds.), *The cognitive representation of speech*. Amsterdam: North-Holland Publishing Company.

- Klatt, D. H. (1989). Review of selected models of speech perception. In W. Marslen-Wilson (Ed.), *Lexical representation and process* (pp. 169-226). Cambridge, MA: MIT Press.
- Kluender, K. R. (2000). Contributions of nonhuman animal models to understanding human speech perception. *The Journal of the Acoustical Society of America*, 107(5), 2835.
- Kluender, K. R., Lotto, A. J., & Holt, L. L. (sous presse). Contributions of nonhuman animal models to understanding human speech perception. In S. Greenberg & W. Ainsworth (Eds.), *Listening to Speech: An Auditory Perspective*: Lawrence Erlbaum Associates.
- Kolinsky, R., & Morais, J. (1996). Migrations in speech recognition. *Language and Cognitive Processes*, 11(6), 611-619.
- Kolinsky, R., Morais, J., & Cluytens, M. (1995). Intermediate representations in spoken word recognition: Evidence from word illusions. *Journal of Memory and Language*, 34(1), 19-40.
- Kremin, H., Hamerel, M., Dordain, M., De Wilde, M., & Perrier, D. (2000). Age of acquisition and name agreement as predictors of mean response latencies in picture naming of French adults. *Brain and cognition*, 43(1-3), 286-291.
- Lachaud, C. M., Caelen-Haumont, G., Pynte, J., & Espesser, R. (1999). *The role of prosodic cues in ASR, expert knowledge and Human perception : a comparison of performance for french word recognition*. Paper presented at the Proceedings of ESCA Workshop on Dialogue and Prosody, De Koningshof, Veldhoven, The Netherlands.
- Lacheret-Dujour, A., & Beaugendre, F. (1999). *La prosodie du Français*. Paris: CNRS Editions.
- Lahiri, A., & Marslen-Wilson, W. D. (1991). The mental representation of lexical form: a phonological approach to the recognition lexicon. *Cognition*, 38(3), 245-294.
- Lahiri, A., & Reetz, H. (2002). Underspecified recognition. In C. G. N. Warner (Ed.), *Laboratory phonology VII* (pp. 637-675). Berlin/New York: Mouton de Gruyter.
- Laporte, E. (1993). Phonetic syllables in French: Combinations, structure, and formal definitions. *Acta linguistica Hungarica*, 41, 175-189.
- Lehiste, I. (1976). Influence of fundamental frequency patterns on the perception of duration. *Journal of Phonetics*, 4, 113-117.
- Lindfield, K. C., Wingfield, A., & Goodglass, H. (1999a). The contribution of prosody to spoken word recognition. *Applied psycholinguistics*, 20, 395-405.
- Lindfield, K. C., Wingfield, A., & Goodglass, H. (1999b). The role of prosody in the mental lexicon. *Brain and language*, 68(1-2), 312-317.

- Luce, P. A. (1986). A computational analysis of uniqueness points in auditory word recognition. *Perception and Psychophysics*, 39(3), 155-158.
- Luce, P. A., Goldinger, S. D., Auer, E. T., & Vitevitch, M. S. (2000). Phonetic priming, neighborhood activation, and PARSYN. *Perception and Psychophysics*, 62(3), 615-625.
- Luce, P. A., & Lyons, E. A. (1999). Processing lexically embedded spoken words. *Journal of Experimental Psychology: Human Perception and Performance*, 25(1), 174-183.
- Luce, P. A., & Pisoni, D. B. (1998). Recognizing spoken words: the neighborhood activation model. *Ear and Hearing*, 19(1), 1-36.
- Luce, R. D. (1986). *Response times: Their role inferring elementary mental organization*. New-York: Oxford University Press, Inc.
- Lyons, A. W., Teer, P., & Rubenstein, H. (1978). Age-at-acquisition and word recognition. *Journal of psychological research*, 7(3), 179-187.
- Marr, D. (1982). Understanding complex information-processing systems. In *Vision. A computational investigation into the human representation and processing of visual information*. San Francisco: Henry Holt & Company.
- Marslen-Wilson, W. D. (1987). Functional parallelism in spoken word-recognition. *Cognition*, 25(1-2), 71-102.
- Marslen-Wilson, W. D., & Warren, P. (1994). Levels of perceptual representation and process in lexical access: words, phonemes, and features. *Psychological Review*, 101(4), 653-675.
- Marslen-Wilson, W. D., & Welsh, A. (1978). Processing interactions and lexical access during word recognition in continuous speech. *Cognitive Psychology*, 10(1), 29-63.
- Massaro, D. W. (1989). Testing between the TRACE model and the fuzzy logical model of speech perception. *Cognitive Psychology*, 21(3), 398-421.
- Massaro, D. W., & Oden, G. C. (1980). Evaluation and integration of acoustic features in speech perception. *The Journal of the Acoustical Society of America*, 67(3), 996-1013.
- Mattys, S. L. (1997). The use of time during lexical processing and segmentation: a review. *Psychonomic Bulletin and Review*, 4(3), 310-329.
- Mattys, S. L. (2004). Stress Versus Coarticulation: Toward an Integrated Approach to Explicit Speech Segmentation. *Journal of Experimental Psychology: Human Perception and Performance*, 30(2), 397-408.
- Mattys, S. L., & Herslow, J. (2002). Lexical activity in speech processing: evidence from pause detection. *Journal of Memory and Language*, 47(3), 343-359.

- McClelland, J. L., & Elman, J. L. (1986). The TRACE model of speech perception. *Cognitive Psychology*, 18(1), 1-86.
- McClelland, J. L., Fiez, J. A., & McCandliss, B. D. (2002). Teaching the /r/-/l/ discrimination to Japanese adults: behavioral and neural aspects. *Physiology and behavior*, 77, 657-662.
- McClelland, J. L., & Rumelhart, D. E. (1981). An interactive activation model of context effects in letter perception: I. An account of basic findings. *Psychological Review*, 88(5), 375-407.
- McLeod, P., Plaut, D. C., & Shallice, T. (2001). Connectionist modelling of word recognition. *Synthese*, 129(2), 173-183.
- McLeod, P., Shallice, T., & Plaut, D. C. (2000). Attractor dynamics in word recognition: converging evidence from errors by normal subjects, dyslexic patients and a connectionist model. *Cognition*, 74(1), 91-114.
- McQueen, J. M. (1998). Segmentation of continuous speech using phonotactics. *Journal of Memory and Language*, 39(1), 21-46.
- McQueen, J. M., Culter, A., Briscoe, T., & Norris, D. (1995). Models of continuous speech recognition and the contents of the vocabulary. *Language and Cognitive Processes*, 10(3-4), 309-331.
- McQueen, J. M., Norris, D., & Cutler, A. (1994). Competition in spoken word recognition: Spotting words in other words. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 20(3), 621-638.
- Mehler, J., Dommergues, J.-Y., Frauenfelder, U. H., & Segui, J. (1981). The syllable's role in speech segmentation. *Journal of Verbal Learning and Verbal Behavior*, 20, 289-305.
- Meschyan, G., & Hernandez, A. (2002). Age of acquisition and word frequency: Determinants of object-naming speed and accuracy. *Memory and cognition*, 30(2), 262-269.
- Miller, L. M., Escabi, M. A., & Schreiner, C. E. (2001). Feature selectivity and interneuronal cooperation in the thalamocortical system. *The Journal of Neuroscience*, 21(20), 8136-8144.
- Monaghan, J., & Ellis, A. W. (2002). Age of acquisition and the completeness of phonological representations. *Reading and writing: an interdisciplinary journal*, 15, 759-788.
- Morrison, C. M., Chappell, T., & Ellis, A. W. (1997). Age of acquisition norms for a large set of object names and their relation to adult estimates and other variables. *The Quarterly Journal of Experimental Psychology: Human Experimental Psychology*, 50A(3), 528-559.

- Morrison, C. M., & Ellis, A. W. (1995). Roles of word frequency and age of acquisition in word naming and lexical decision. *Journal of experimental psychology: learning, memory and cognition*, 21(1), 116-133.
- Morrison, C. M., & Ellis, A. W. (2000). Real age of acquisition effects in word naming and lexical decision. *British Journal of Psychology*, 91 ( Pt 2), 167-180.
- Morrison, C. M., Hirsh, K., Chappell, T., & Ellis, A. W. (2002). Age and age of acquisition: An evaluation of the cumulative frequency hypothesis. *European Journal of Cognitive Psychology*, 14(4), 435-459.
- Morrison, C. M., Hirsh, K. W., & Duggan, G. B. (2003). Age of acquisition, ageing, and verb production: normative and experimental data. *The Quarterly Journal of Experimental Psychology A*, 56(4), 705-730.
- Morton, J. (1969). Interaction of information in word recognition. *Psychological Review*, 76(2), 165-178.
- Munteanu, P. (1996). *Extraction de connaissances dans les bases de données parole : apport de l'apprentissage symbolique*. Grenoble.
- Munteanu, P., Caillaud, B., Serignat, J. F., & Caelen-Haumont, G. (1997). *Automatic word demarcation based on prosody*. Paper presented at the 5th European Conference On Speech Communication and Technology (EUROSPEECH 97), Rhodes (Greece).
- New, B., Pallier, C., Ferrand, L., & Matos, R. (2001). Une base de données lexicales du français contemporain sur internet: LEXIQUE. *L'année Psychologique*, 101(3), 447-462.
- Nooteboom, S. G. (1981). Lexical retrieval from fragments of spoken words: beginnings vs endings. *Journal of Phonetics*, 9, 407-424.
- Norris, D. (1994). Shortlist: A connectionist model of continuous speech recognition. *Cognition*, 52(3), 189-234.
- Norris, D., McQueen, J. M., & Cutler, A. (1995). Competition and segmentation in spoken-word recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 21(5), 1209-1228.
- Norris, D., McQueen, J. M., & Cutler, A. (2000a). Feedback on feedback on feedback: It's feedforward. *The Behavioral and Brain Sciences*, 23(3), 352-370.
- Norris, D., McQueen, J. M., & Cutler, A. (2000b). Merging information in speech recognition: feedback is never necessary. *Behavioral and Brain Sciences*, 23(3), 299-325.
- Norris, D., McQueen, J. M., Cutler, A., & Butterfield, S. (1997). The possible-word constraint in the segmentation of continuous speech. *Cognitive Psychology*, 34(3), 191-243.



- Papousek, H., & Papousek, M. (1987). Intuitive parenting: a dialectic counterpart to the infant's integrative competence. In J. Osofsky (Ed.), *Handbook of infant development* (2 ed., pp. 669-720). New-York: Wiley & Sons.
- Papousek, M., & Papousek, H. (1981). Musical elements in the infant's vocalization: their significance for communication, cognition and creativity. In L. P. Lipsitt (Ed.), *Advances in infancy research* (Vol. 1, pp. 163-224). Norwood, NJ.
- Peereman, R., & Content, A. (1999). LEXOP: A lexical database providing orthography-phonology statistics for French monosyllabic words. *Behavior Research Methods, Instruments and Computers*, 31(2), 376-379.
- Perea, M., Rosa, E., & Consolacion, G. (2002). Is the go/no-go lexical decision task an alternative to the yes/no lexical decision task? *Memory and Cognition*, 30(1), 34-45.
- Psychology Software Tools. (2002). E-Prime (Version 1.0). Pittsburgh, PA, USA.
- Raaijmakers, J. G., Schrijnemakers, J. M., & Gremmen, F. (1999). How to deal with "the language-as-fixed-effect fallacy": common misconceptions and alternative solutions. *Journal of Memory and Language*, 41(3), 416-426.
- Radeau, M., Morais, J., Mousty, P., & Bertelson, P. (2000). The effect of speaking rate on the role of the uniqueness point in spoken word recognition. *Journal of Memory and Language*, 42(3), 406-422.
- Rasbash, J., Browne, W., Healy, M., Cameron, B., & Charlton, C. (2001). MLwiN (Version 1.1): Multilevel Models Project, Institute of Education, London, UK.
- Reddy, D. R., Erman, L. D., Fennell, R. D., & Neely, R. B. (1973). *The HEARSAY speech understanding system: an example of the recognition process*. Paper presented at the International conference on Artificial Intelligence, Stanford, CA.
- Reetz, H. (2000). *Underspecified phonological features for lexical access*. Paper presented at the Phonetics and Phonology in ASR. Parameters and Features, and their Implications, Institute of Phonetics in Saarbrücken.
- Rietveld, A. C. (1980). Word boundaries in the french language. *Language and speech*, 23(3), 289-296.
- Rumelhart, D. E., & McClelland, J. L. (1985). Levels indeed! A response to Broadbent. *Journal of Experimental Psychology: General*, 114(2), 193-197.
- Sendlmeier, W. F. (1995). Feature, phoneme, syllable or word: how is speech mentally represented? *Phonetica*, 52(3), 131-143.
- Shillcock, R. C. (1990). Lexical hypotheses in continuous speech. In G. T. Altmann (Ed.), *Cognitive models of speech processing: psycholinguistic and computational perspectives* (pp. 24-49). Cambridge, MA: MIT Press.

- Shtyrov, Y., Kujala, A., Palva, S. J., Ilmoniemi, R. J., & Naatanen, R. (2000). Discrimination of speech and complex nonspeech sounds of different temporal structure in the left and right hemispheres. *Neuroimage*, *12*, 657-663.
- Slowiaczek, L. M., & Hamburger, M. (1992). Prelexical facilitation and lexical interference in auditory word recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *18*(6), 1239-1250.
- Slowiaczek, L. M., McQueen, J. M., Soltano, E. G., & Lynch, M. (2000). Phonological representations in prelexical speech processing: Evidence from form-based priming. *Journal of Memory and Language*, *43*(3), 530-560.
- Snijders, T., & Bosker, R. (1999). *Multilevel analysis: an introduction to basic and advanced multilevel modeling*. London, UK: SAGE Publications Ltd.
- Sonic Foundry. (2000). Sound Forge. Madison, WI, USA.
- Spinelli, E., Segui, J., & Radeau, M. (2001). Phonological priming in spoken word recognition with bisyllabic targets. *Language and Cognitive Processes*, *16*(4), 367-392.
- Statsoft. (2001). Statistica (Version 6). Tulsa, OK, USA.
- Stern, D. N., Spieker, S., & MacKain, K. (1982). Intonation contours as signals in maternal speech to prelinguistic infants. *Developmental Psychology*, *18*, 727-735.
- Syntrillium Software Corporation. (2003). Cool Edit Pro (Version 2.1). Phoenix, AZ, USA.
- Tabossi, P., Collina, S., Mazzetti, M., & Zoppello, M. (2000). Syllables in the processing of spoken Italian. *Journal of Experimental Psychology: Human Perception and Performance*, *26*(2), 758-775.
- Taft, L. A. (1984). *Prosodic constraints and lexical parsing strategies*. University of Massachusetts, Amherst.
- Taft, M., & Forster, K. I. (1976). Lexical storage and retrieval of polymorphemic and polysyllabic words. *Journal of Verbal Learning and Verbal Behavior*, *15*(6), 607-620.
- Tartter, V. C., & Eimas, P. D. (1975). The role of auditory feature detectors in the perception of speech. *Perception and Psychophysics*, *18*(4), 293-298.
- van Kuijk, D., Wittenburg, P., & Dijkstra, T. (1996, October 3-6). *Raw: a real-speech model for human word recognition*. Paper presented at the International conference on spoken language processing (ICSLP 96), Philadelphia, PA, USA.
- Vitevitch, M. S., & Luce, P. A. (1998). When words compete: Levels of processing in perception of spoken words. *Psychological Science*, *9*(4), 325-329.

- Vitevitch, M. S., & Luce, P. A. (1999). Probabilistic phonotactics and neighborhood activation in spoken word recognition. *Journal of Memory and Language*, 40, 374-408.
- Vitevitch, M. S., Luce, P. A., Pisoni, D. B., & Auer, E. T. (1999). Phonotactics, neighborhood activation, and lexical access for spoken words. *Brain and language*, 68(1-2), 306-311.
- Vroomen, J., & de Gelder, B. (1995). Metrical Segmentation and Lexical Inhibition in Spoken Word Recognition. *Journal of Experimental Psychology: Human Perception and Performance*, 21(1), 98-108.
- Vroomen, J., & van Gelder, B. (1997). Activation of embedded words in spoken word recognition. *Journal of Experimental Psychology: Human Perception and Performance*, 23(3), 710-720.
- Walley, A. C., & Sloane, M. E. (2001). The perceptual magnet effect: a review of empirical findings and theoretical implications. In F. Columbus (Ed.), *Advances in Psychology research* (Vol. 4, pp. 65-92). Huntington, N.Y.: Nova Science.
- Warren, P., & Marslen-Wilson, W. D. (1987). Continuous uptake of acoustic cues in spoken word recognition. *Perception and Psychophysics*, 41(3), 262-275.
- Warren, R. M., Healy, E. W., & Chalikia, M. H. (1996). The vowel-sequence illusion: Intrasubject stability and intersubject agreement of syllabic forms. *Journal of the Acoustical Society of America*, 100(4), 2452-2461.
- Wike, E. L., & Church, D. J. (1976). Comments on Clark's "The language-as-fixed-effect fallacy". *Journal of Verbal Learning and Verbal Behavior*, 15, 249-255.
- Winer, B. J., Brown, D. R., & Michels, K. M. (1991). *Statistical principles in experimental design* (1991, 3ème ed.). New-York, USA: McGraw-Hill, Inc.
- Wingfield, A., Goodglass, H., & Lindfield, K. C. (1997). Word recognition from acoustic onsets and acoustic offsets: Effects of cohort size and syllabic stress. *Applied psycholinguistics*, 18(1), 85-100.
- Xanthos, A. (2004). *Combining utterance-boundary and predictability approaches to speech segmentation*. Paper presented at the COLING, Geneva, Switzerland.
- Yates, F. (1933). The analysis of replicated experiments when field results are incomplete. *Empire journal of experimental agriculture*(1), 129-142.
- Zell, A., Mamier, G., Vogt, M., Mache, N., Hübner, R., Döring, S., et al. (1989). SNNS (Stuttgart Neural Network Simulator) (Version 4.1). Stuttgart, Federal Republic of Germany: SNNS Group, Institute for parallel and distributed High Performance Systems (IPVR), University of Stuttgart.

## **Annexe 1**

### **Obtention des informations permettant le contrôle des facteurs psychologiques « âge d'acquisition » et « familiarité conceptuelle »**

L'âge d'acquisition et la familiarité conceptuelle d'un mot sont des facteurs décisifs pour l'accès au lexique, en production comme en perception. Leur impact sur les temps de réponse mesurés lors des expériences de laboratoire nécessitait un contrôle de leurs effets. Or, n'ayant aucune information sur ces 2 dimensions pour les items utilisés dans les expériences présentées dans ce travail, il fallait trouver un moyen permettant de contrôler le biais potentiel qu'ils représentent.

Un recueil de ces données a donc été effectué grâce à une enquête par Internet, méthode moderne et performante qui a permis d'acquérir rapidement une information psychologique de qualité sur un corpus relativement volumineux (1225 mots monosyllabiques et bisyllabiques du Français). Ces informations ont ensuite été utilisées pour contrôler a posteriori l'incidence de l'âge d'acquisition estimé et de la familiarité conceptuelle sur les TR mesurés dans les expériences.

Cette annexe présente le détail du travail qui a été effectué et quelques résultats.

## Introduction

En psychologie du langage, on admet généralement que la fréquence d'usage d'un mot dans une communauté linguistique (fréquence d'occurrence d'un mot dans l'usage de la langue) influence statistiquement sa vitesse de traitement par les membres de cette communauté, en production comme en perception. Les psycholinguistes contrôlent donc ce paramètre lors de l'élaboration de leur matériel expérimental pour éviter tout biais dans les résultats. Ce contrôle est rendu possible par l'existence de bases de données lexicales, comme BRULEX pour le Français (Content, Mousty et Radeau (1990)). De telles banques de données sont issues d'analyses automatisées portant sur des corpus écrits volumineux (littérature, médias, etc.), qui recensent exhaustivement, pour un grand nombre de mots d'une langue donnée (environ 36000 pour BRULEX), un ensemble de caractéristiques linguistiques formelles (nombre de lettres, de phonèmes, de syllabes, fréquence d'occurrence, genre, etc.).

Ces deux dernières décennies ont vu se multiplier des recherches rapportant l'effet robuste, remarquable et récurrent sur le traitement des mots, de facteurs psychologiques comme l'âge d'acquisition ou la familiarité conceptuelle. Ces résultats soulèvent d'une part la question de la relation entre âge d'acquisition, familiarité et fréquence d'occurrence, et d'autre part, la question du contrôle et de la neutralisation des effets liés à l'âge d'acquisition et à la familiarité dans les recherches en psychologie du langage. Pratiquer un tel contrôle soulèverait au moins 3 obstacles techniques. Le premier provient de la lourdeur excessive des études qui en résulteraient. Comme il n'existe pas à ce jour, en Français, de bases de données lexicales renfermant une quantité suffisante d'items pour des variables psychologiques telles que l'âge d'acquisition ou la familiarité, les chercheurs qui désirent contrôler ces dimensions doivent pourvoir à l'obtention des informations par sondage d'une quantité importante d'individus. Un tel recueil, excessivement coûteux en travail et en temps, est d'autant plus difficilement acceptable d'un point de vue stratégique et productif qu'il ne représente qu'un préliminaire à l'étude (contrôle du matériel expérimental). Le deuxième obstacle est relatif aux contraintes lexicales. Comme les psycholinguistes ne créent pas les mots qu'ils étudient, mais choisissent parmi ceux qui existent, la quantité disponible est inversement proportionnelle au nombre de facteurs contrôlés. Manipuler deux facteurs supplémentaires pourrait rendre impossible l'expérimentation en réduisant encore la quantité de mots utilisables. Le troisième obstacle est dû à la colinéarité des variables âge d'acquisition, familiarité et fréquence d'occurrence (on est plus familier avec un mot appris très jeune, on utilise plus un mot avec lequel on est familier, etc.). Leur contrôle

impliquerait d'avoir recours à des précautions statistiques supplémentaires pour éviter les pièges liés à la non-indépendance des prédicteurs.

Cette annexe présente la méthodologie mise en place pour contourner ces difficultés techniques. Après une brève revue de question théorique concernant l'influence de l'âge d'acquisition, de la fréquence d'usage et de la familiarité conceptuelle sur l'accès au lexique, la procédure qui a été suivie est décrite. Cette procédure se base sur deux éléments :

- ▷ Un outil permettant la constitution rapide d'une base de données lexicales pour des facteurs de nature psychologique (résolution de la première contrainte technique).
- ▷ Une méthode spécifique d'analyse statistique (modélisation multiniveaux – résolution des problèmes de contrôle ou de neutralisation liés aux contraintes lexicales et à la colinéarité des variables).

Les résultats de l'enquête (résultats bruts et méthode de traitement, analyses quantitatives et qualitatives de la base de données obtenue) sont ensuite présentés, suivis d'une analyse des relations statistiques entre facteurs étudiés (âge d'acquisition estimé, familiarité conceptuelle, fréquence d'usage). Enfin, les résultats d'une méta analyse sont présentés, explorant l'incidence de l'âge d'acquisition estimé, de la familiarité conceptuelle et de la fréquence d'usage sur les temps de reconnaissance de mots en modalité auditive. Les apports de cette étude sont enfin récapitulés.

### Revue de question théorique

#### Quelques faits sur l'influence de l'âge d'acquisition, de la fréquence d'usage et de la familiarité conceptuelle (résultats majoritairement obtenus pour l'Anglais)

- ▷ En perception :

L'âge d'acquisition semble avoir une incidence sur le traitement des mots, en modalité visuelle (Lyons, Teer et Rubenstein (1978)) comme en modalité auditive (Garlock et Walley (2001) ; Gerhand et Barry (1999a)). Un mot appris plus tôt au cours du développement serait reconnu plus rapidement qu'un mot acquis plus tard. Cet effet serait indépendant de celui de la fréquence d'occurrence d'un mot (montré en modalité visuelle par Bonin et col. (2001), Brysbaert et col. (2000), Morrison et Ellis (1995; 2000)). Selon Garlock et Walley (2001), les effets de fréquence ne seraient pas systématiquement retrouvés, ou seraient minimaux en comparaison à ceux de l'âge d'acquisition. La familiarité semble aussi jouer un rôle sur la perception des mots, selon Connine et col. (1990), qui trouvent, lors d'une tâche de décision lexicale, quelle que soit la modalité sensorielle testée, une vitesse de traitement plus élevée pour des mots à score de familiarité élevé, que pour des mots à score de familiarité faible.

- ▷ En production

Un même motif de résultats qu'en perception est retrouvé pour l'âge d'acquisition. Ce facteur influence fortement les temps de production écrite et orale des mots (Bonin et col. (2002), Bonin et col. (2001), Chalard et col.

(2003), Ellis et Morrison (1998), Garlock et Walley (2001), Kremin et col. (2000), Morrison et col. (2002), Morrison et col. (2003)). Plus un mot a été appris tôt durant le développement de l'individu, plus les temps nécessaires pour le produire sont courts. Cet effet de l'âge d'acquisition semble être indépendant de celui de la fréquence d'utilisation d'un mot, en tous cas dans les tâches de dénomination d'image (Barry et col. (2001), Bonin et col. (2001), Chalard et col. (2003)), où seul l'âge d'acquisition joue un rôle. D'autres chercheurs comme Barry et col. (1997), Brysbaert (1996), Gerhand et Barry (1998), ou Meschyan et Hernandez (2002), trouvent cependant des effets de fréquence, mais qui restent indépendants de ceux de l'âge d'acquisition. Selon Barry et col. (1997), les effets de fréquence ne seraient observables que dans le cas des mots acquis tardivement (tâches de dénomination d'image, exploration du rôle de la fréquence d'utilisation, de l'âge d'acquisition, et de l'interaction des 2 facteurs). À l'opposé, les analyses de régression produites dans l'étude de Bird et col. (2001) semblent appuyer l'idée qu'âge d'acquisition et fréquence d'occurrence d'un mot seraient en lien. Concernant le rôle de la familiarité en production, les travaux sont plutôt rares et portent sur des individus pathologiques. Kremin et col. (2000) échouent à trouver un effet de ce facteur dans une tâche de dénomination d'images (sauf pour 1 personne sur 16). Ce problème pourrait être en lien avec la pathologie des patients testés ou la méthode d'analyse utilisée (régression multiple avec 7 facteurs). Hirsh et Funnell (1995) trouvent pour leur part un effet de ce facteur sur un patient aphasique progressif (mots familiers produits avec une meilleure précision). La familiarité pourrait donc avoir une influence sur la production des mots.

### Mécanismes proposés pour interpréter le rôle de ces facteurs

Pour la familiarité et la fréquence d'occurrence, il semble que le mécanisme explicatif proposé soit essentiellement fonctionnel. Plus on utilise une structure plus on devient capable de l'utiliser (parce qu'elle se développe pour être plus performante, parce qu'on acquiert une expertise, ou éventuellement, une habitude). Par contre, des phénomènes structuraux seraient nécessaires pour rendre compte des effets de l'âge d'acquisition. Ci-dessous, un tour d'horizon de la littérature présente les hypothèses qui sont généralement proposées.

Selon Barry et Gerhand (2003), ainsi que Ellis et Morrison (1998) et Gerhand et Barry (1998), l'âge d'acquisition affecterait l'aisance à élaborer le code phonologique d'un mot lors de sa production, idée que semblent appuyer les effets de facilitation observés lors d'amorçage par répétition dans les tâches de dénomination d'image (Barry et col. (2001)), ou l'augmentation de la force de l'effet lors de tâches de lecture sous contrainte de temps (Gerhand et Barry (1999b)). Cependant, d'autres travaux comme ceux de Brysbaert et col. (2000) laissent penser que les effets de l'âge d'acquisition ne se limitent pas au système d'encodage phonologique, mais ont aussi une origine sémantique. Selon Baumeister (1985), l'âge d'acquisition n'influencerait pas directement la vitesse de production. Un mot acquis plus tôt serait lié de façon plus centrale aux autres mots dans le lexique. Le système linguistique d'un individu se structure en effet selon une

chronologie qui correspond à l'ordre dans lequel les mots sont appris, plutôt que selon une dichotomie entre période critique d'acquisition du vocabulaire (durant l'enfance) et acquisition au-delà de cette phase développementale (Izura et Ellis (2002)). Les mots acquis le plus tôt seraient donc les plus facilement accessibles. De plus, les mots acquis tardivement pourraient être mémorisés sous forme fragmentée (Brown et Watson (1987)), impliquant des traitements plus difficiles (« Phonological Completeness Hypothesis », cependant non validée lors d'une étude spécifique (Monaghan et Ellis (2002)). Dans le cas de mots d'une langue seconde, l'ordre d'acquisition des formes lexicales serait déterminant pour leur vitesse de traitement, et non l'ordre d'acquisition des correspondants dans la langue maternelle. Ceci indique que le niveau de l'âge d'acquisition, en tous cas pour une langue seconde, serait plutôt de nature formelle que conceptuelle (Izura et Ellis (2002)).

#### Quelle validité pour l'âge d'acquisition estimé ?

Les données d'âge d'acquisition employées dans la recherche proviennent, la plupart du temps, d'une estimation subjective fournie par des adultes, et non d'une mesure objective de l'âge d'acquisition réel. Un certain nombre d'objections a été émis quant à la validité de l'indice âge d'acquisition estimé. De telles questions ne se posent pas avec le facteur fréquence d'occurrence, directement calculé à partir d'une analyse de corpus littéraires, ni avec le facteur familiarité, par essence subjectif. Par exemple, Chalard et col. (2003) ont montré que l'âge d'acquisition objectif était un meilleur prédicteur des latences de production parlée et écrite que l'âge d'acquisition estimé. Cependant, les effets obtenus avec l'âge d'acquisition objectif se retrouvent avec l'âge d'acquisition estimé, les 2 facteurs étant fortement corrélés ( $r = .69$ , selon Chalard et col. (2003),  $r = .76$  selon Morrison et col. (1997)), si bien que l'âge estimé d'acquisition, même s'il est moins précis, semble pouvoir être considéré comme un indicateur valable de l'âge d'acquisition objectif (Gilhooly et Gilhooly (1980), Morrison et col. (1997)). Une étude longitudinale a pu démontrer la validité de l'âge d'acquisition estimé, en tant que reflet fidèle de l'âge d'acquisition objectif (Jorm (1991)). De plus, les échantillons de population testés à travers différentes études sont généralement consistants sur leurs estimations (Morrison et col. (1997)). Étant donné les difficultés évidentes pour obtenir un âge d'acquisition objectif (population d'enfants, difficile à obtenir et à enquêter, nécessité d'un échantillon très élevé dû à la démultiplication des tranches d'âge, etc.), l'usage d'un âge d'acquisition estimé semble donc représenter un compromis technique des plus acceptables. Il s'agit de la méthode généralement utilisée pour produire des normes de cette dimension (Morrison et col. (1997), Alario et col. (1999), Bonin et col. (2003)).



## Conclusion

Il ressort de la littérature que plus un mot est utilisé fréquemment, est appris tôt, ou est un mot avec lequel on est familier, plus il est produit ou perçu rapidement. Ceci implique que la fréquence, l'âge d'acquisition et la familiarité devraient faire l'objet d'un contrôle plus systématique dans les expériences psycholinguistiques. Disposer de l'information dans une base de données lexicale serait donc utile, d'où le présent article contribue à fournir aux chercheurs un outil de travail.

## Méthode

### Le corpus

Le corpus employé comporte l'ensemble originel des mots utilisés pour les expériences de ce travail de thèse. S'y ajoutent les mots enchâssés dans ces mots, ainsi que les duplicats des homophones de ces mots enchâssés, sélectionnés à partir du dictionnaire « Petit Larousse ». Le corpus comprends en majorité des noms, mais aussi d'autres catégories grammaticales (exemple : verbes conjugués).

### Information recueillie et opérationnalisation du recueil

L'enquête a consisté à recueillir 3 types d'information de nature psychologique : l'âge d'acquisition estimé d'un mot, le degré de certitude sur l'estimation de l'âge d'acquisition, et la familiarité des individus avec les mots présentés. Cette section décrit et justifie la manière dont ces 3 dimensions ont été opérationnalisées pour le recueil.

### Âge d'acquisition estimé

Concernant l'estimation de l'âge d'acquisition, il a été choisi de demander aux participants de l'évaluer selon la procédure habituellement employée pour recueillir ce type d'information (voir par exemple Ferrand et col. (2003), ainsi que la discussion présentée dans la section « Quelle validité pour l'âge d'acquisition estimé ? » pour une justification de cette procédure au lieu d'une mesure de l'âge réel d'acquisition), en attirant leur attention sur l'âge auquel ils pensaient avoir été mis en contact pour la première fois avec l'objet désigné par le mot présenté - objet physique ou concept. Plusieurs raisons motivent l'emploi d'une procédure focalisant l'attention du participant sur l'objet désigné par le mot plutôt que sur le mot, le but étant d'optimiser la précision du rappel. Premièrement, lorsque l'individu découvre un objet, il cherche spontanément, dès l'enfance, à acquérir la forme lexicale qui le désigne, afin de pouvoir communiquer à son sujet avec d'autres humains. La corrélation dans le temps entre l'expérience de l'objet et l'acquisition du signifiant relatif doit donc être très élevée. Deuxièmement, le participant devrait avoir une plus grande facilité pour localiser dans le temps son vécu à

un objet qu'à l'apprentissage d'une forme pour le désigner. L'expérience concrète liée au signifié peut donc représenter une aide essentielle, et rendre la tâche plus facile pour les participants. Enfin, focaliser les participants sur l'utilisation d'une stratégie particulière permet de réduire la variance dans les données en évitant que plusieurs stratégies ne soient mises en œuvre par les différents participants.

### Degré de certitude de l'estimation pour l'âge d'acquisition

Le degré de certitude sur l'estimation de l'âge d'acquisition est un indice secondaire ajouté au recueil dans le but de disposer d'une information supplémentaire pour évaluer la qualité de l'estimation des âges d'acquisition par les participants. Il apporte donc une aide pour valider l'estimation. Le degré de certitude de l'âge d'acquisition est simplement opérationnalisé en posant directement la question aux participants.

### Familiarité

La familiarité d'un individu avec un mot semble dépendre de 2 aspects : la fréquence avec laquelle un mot est rencontré / utilisé, et la complexité / le niveau d'abstraction du concept qu'il véhicule. Ces 2 aspects peuvent être évalués séparément grâce à la fréquence subjective pour le premier, et la concrétude pour le second (par exemple Bonin et col. (2003)). Cependant, comme la fréquence subjective et la fréquence objective sont fortement corrélées<sup>137</sup>, l'intérêt d'un recueil de la fréquence subjective est discutable parce qu'il risque de n'apporter qu'une quantité limitée d'information nouvelle, tandis que la concrétude d'un mot ne mesure pas directement la sensation de familiarité éprouvée par le sujet psychologique. C'est donc un indice composite qui a été utilisé pour mesurer directement la familiarité, selon une procédure comparable à celle suivie par exemple par Bonin et col. (2003) : les participants décident si l'item présenté est un mot qu'ils connaissent bien ou non. Cet indice présente plusieurs avantages. Premièrement, il est facile à produire par tout individu, qui peut dire s'il connaît un mot ou non (inversement, un indice comme la fréquence subjective nécessite une réflexion pour être estimé). Deuxièmement, il véhicule une information psychologique multidimensionnelle à même de servir directement de prédicteur de la difficulté du traitement d'un mot. En effet, l'impression de bien connaître un mot est probablement basée à la fois sur des aspects fréquentiels, sémantiques, et organisationnels dans le lexique. Ceci en fait un estimateur riche, qui peut ensuite être

---

<sup>137</sup>  $.41 \leq r \leq .79$  d'après les études citées par Desrochers et Bergeron (2000), impliquant que l'estimation de la fréquence subjective fournirait une information redondante jusqu'à près de 62% ( $r^2 = .62$ ) de celle déjà disponible dans les bases de données lexicales du Français.

soumis à une analyse statistique à l'aide d'indices objectifs pour en connaître les composantes et l'influence que chacune a sur la réponse fournie par les participants.

### L'outil de recueil

Il s'agit d'une interface électronique structurée autour de 2 questionnaires (inscription et enquête), utilisable à l'aide d'un navigateur Internet. Elle était accessible à partir du site du Laboratoire de Psycholinguistique Expérimentale de l'Université de Genève durant la période de recueil des données. Cette interface présente de l'information à l'utilisateur et récolte ses réponses. Ces dernières sont stockées au fur et à mesure qu'elles sont fournies, sur un serveur de l'Université de Genève, pour finalement constituer une base de données. Cette méthode de recueil, moderne et efficace, a permis d'échantillonner largement la population francophone en France et en Suisse Romande, et d'obtenir rapidement un corpus important et de qualité. Le système a été conçu pour obtenir des informations générales sur le participant, utilisées pour la validation des données brutes lors de leur traitement. Il a également été conçu pour être souple pour les utilisateurs. La quantité d'items à traiter étant importante, il était peu probable qu'une même personne les traite tous. Il était encore moins probable qu'elle les traite tous en une seule session de travail si elle choisissait de faire l'enquête dans sa totalité. Il fallait donc offrir la possibilité à quelqu'un ne disposant pas de beaucoup de temps, mais volontaire, de pouvoir s'impliquer de manière cumulée en reprenant à n'importe quel moment une enquête interrompue. C'est pour cette raison qu'un système d'identification anonyme a été développé. Enfin, le système devait autoriser l'analyse statistique des données récoltées. Les items ont donc été présentés aléatoirement à chaque participant, une seule fois chacun, que l'enquêté ait procédé en une ou plusieurs sessions de travail. Théoriquement, chaque item est en moyenne traité un nombre équivalent de fois lorsque le nombre de participants est élevé. De plus, les données sont déjà sous un format électronique, ne nécessitant pas une saisie fastidieuse avant leur traitement.

### Le questionnaire d'inscription

Le participant qui commence l'enquête doit d'abord créer son identité anonyme au sein du système. Celle-ci résulte de l'association entre un profil personnel et une clé d'accès (identifiant numérique). Le profil personnel comprend des informations sur la tranche d'âge, le sexe, le niveau d'éducation, l'origine linguistique et la compétence linguistique du participant (listes de choix à renseigner obligatoirement). Ces informations sont absolument anonymes, et ne permettent en aucun cas, directement ou par recoupement, de connaître la source d'émission des données, qu'il s'agisse

d'individus ou d'organisations. L'identifiant numérique correspond à une combinaison classique de type login + mot de passe, tous deux choisis par l'utilisateur. À cette identité anonyme est associée la liste des items restant à traiter spécifiquement par ce participant (soit, lors de la première connexion, les 1225 mots du corpus).

Le système présente une faille potentielle par son incapacité à contrôler le facteur humain. L'un des principaux points concerne la perte de la clé numérique personnelle (login, mot de passe, ou combinaison des deux). Un participant qui se trouve dans cette situation et souhaite poursuivre l'enquête doit à nouveau remplir le questionnaire d'inscription, puis créer une nouvelle identité dans le système (combinaison login + mot de passe obligatoirement nouvelle, le système refusant tout doublon à ce niveau). S'il fait cela, des items qu'il avait déjà traités sous son ancienne identité numérique peuvent à nouveau lui être présentés. Toutefois, ce cas de figure avait une très faible probabilité de survenir. Premièrement, il a été explicitement demandé aux participants, dans la page de présentation de l'étude, de ne plus participer s'ils venaient à se trouver dans ce cas de figure, en leur expliquant pourquoi. Deuxièmement, devoir remplir à nouveau le questionnaire d'inscription est fastidieux. L'enquête elle-même n'est pas un exercice des plus attrayants. Il est psychologiquement peu vraisemblable, compte tenu du coût, qu'un participant qui ne peut plus accéder au système, fasse une nouvelle fois la démarche. Troisièmement, les participants savaient qu'ils contribuaient à une étude scientifique, de laquelle ils n'allaient retirer aucune autre satisfaction que celle d'avoir aidé au développement de la Connaissance. Ils savaient également qu'ils auraient accès aux résultats de l'enquête dès que les conclusions seraient disponibles (via le site web), ce qui pouvait représenter une motivation. Enfin, vu la grande qualité des données obtenues, il est plutôt probable que les participants ont travaillé très sérieusement.

### Le questionnaire d'enquête

Le questionnaire d'enquête présente au participant un mot écrit dont le sens est précisé entre parenthèses par un synonyme (exemple : « *air (mélodie)* »). Le participant l'utilise pour fournir l'âge auquel il pense avoir appris ce mot et le degré de certitude de son estimation (1ère page du questionnaire), puis son degré de familiarité avec l'item (deuxième page du questionnaire). L'âge d'acquisition est donné librement (choix ouvert, réponse non optionnelle), avec un format numérique en années (nombre entier à 1 ou 2 chiffres), sans possibilité de répondre « ne sais pas ». La consigne écrite était : « Vers quel âge pensez-vous avoir appris ce mot ? Estimez l'âge auquel vous avez appris ce mot (âge auquel vous pensez avoir été mis pour la première fois en contact avec l'objet réel ou le concept correspondant à l'item) ». Le degré de certitude sur la réponse fournie pour l'âge d'acquisition estimé est recueilli à l'aide d'une liste de 6 choix (cases à cocher), du

moins certain (---) au plus certain (+++). La réponse est non optionnelle. La consigne écrite était : « Cochez la case correspondant à votre degré de certitude sur l'âge évalué », et le participant voyait sous la série de cases la légende suivante : « (moins certain) --- / -- / - / + / ++ / +++ (plus certain) ». Le degré de familiarité avec le mot est recueilli selon la même procédure que le degré de certitude (réponse non optionnelle dans une liste de 6 choix sur une échelle du même type que précédemment – consigne écrite : « Quel est votre degré de familiarité avec ce mot ? Estimez le degré de familiarité que vous avez avec ce mot (est-ce un mot que vous connaissez parfaitement ou est-ce un mot qui vous est inconnu ?) »).

Lorsqu'un item a été traité sur les 3 dimensions, le participant valide ses réponses en cliquant sur un bouton. Cette action a pour effet d'enregistrer les informations dans la base de données et de présenter l'item suivant. Le questionnaire présente aléatoirement les items un par un. La liste résiduelle (liste de laquelle l'item présenté a été tiré moins l'item qui vient d'être présenté) est ainsi parcourue, jusqu'à ce qu'elle soit vide.

#### Conception, développement et mise en production de l'outil de recueil

Les concepteurs sont M. F. Serena et M. P. Demierre, du Service de Développement à la Division Informatique de l'Université de Genève, et l'auteur du projet, M. C. Lachaud.

Le développement a entièrement été assuré par Messieurs Serena et Demierre au cours du mois de mars 2003.

#### L'enquête

L'enquête a été lancée à l'aide d'une campagne de communication incitative auprès de 53 universités de France métropolitaine et de Suisse romande, dont seulement une partie a répondu positivement en diffusant l'information auprès de son personnel et de ses étudiants. Elle n'a pas duré plus de 3 mois, ayant pris fin en juillet lorsque les participants ont cessé d'avoir une activité sur le site.

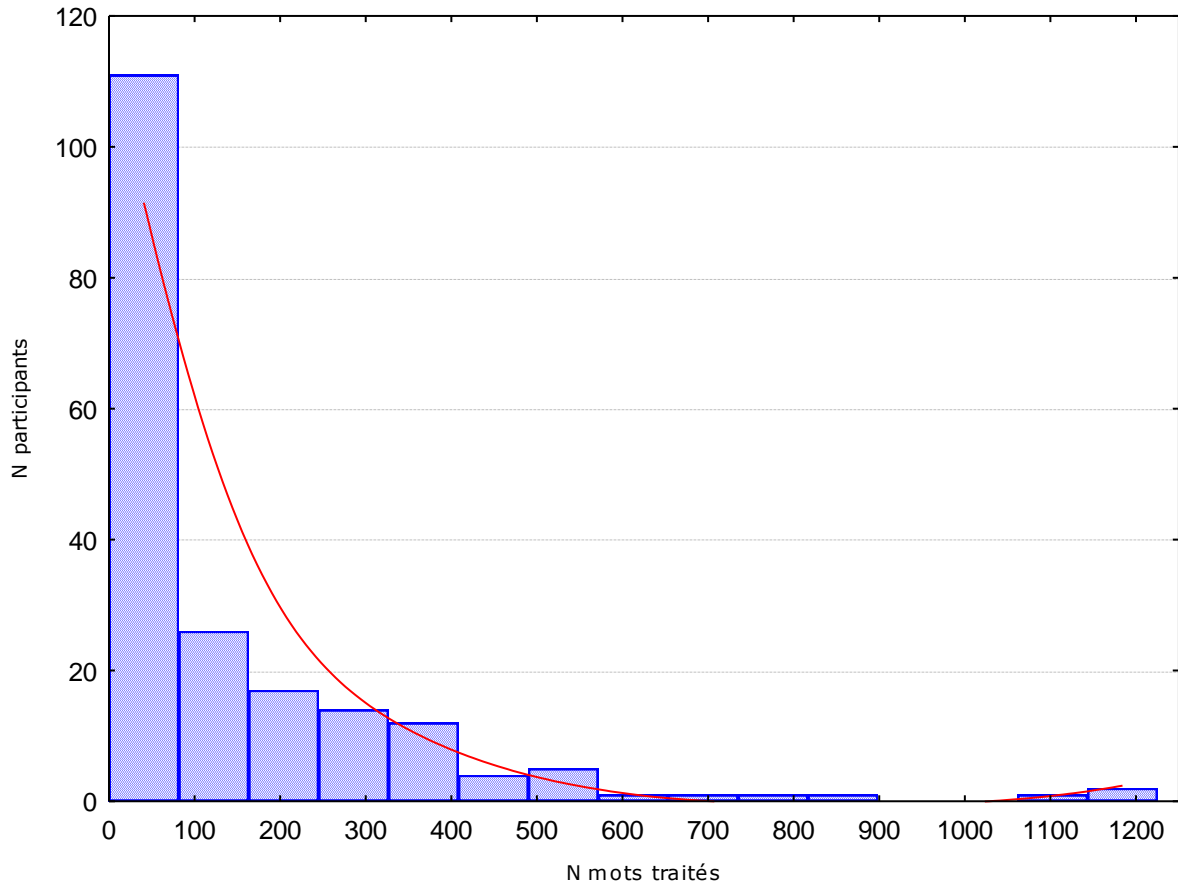
#### Présentation des données brutes et des procédures de traitement

##### Présentation des données brutes

Les données brutes ont été générées par 196 participants (121 femmes et 75 hommes), majoritairement âgés de 18 à 25 ans (N = 97 - répartition pour les autres tranches d'âge : 3 non majeurs, 50 âgés de 25 à 35 ans, 38 de 35 à 55 ans, et 8 de plus de 66 ans). Au total, 27778 réponses ont été recueillies. Chaque mot a été traité en

moyenne par 22,68 participants (détail de la distribution dans la Figure 75). Le participant moyen a traité environ 142 mots.

Figure 75. Distribution du nombre de mots traités par participant (courbe : distance pondérée des moindres carrés)



### Procédures de transformation des données brutes

#### Filtrage des participants

Dans un premier temps, les participants ont été filtrés sur la base de 2 critères. Le premier critère est linguistique. Ce filtre utilise les profils linguistiques renseignés à l'aide du premier questionnaire de l'enquête, pour ne conserver que les individus francophones unilingues, ou bilingues avec un excellent niveau de maîtrise de la langue française. Le deuxième critère est un critère statistique de normalité des réponses par mot. Ce filtre permet de détecter les individus « aberrants » (c'est-à-dire ayant un nombre de réponses anormales sur un nombre élevé d'items ou sur des items qui ne doivent pas causer de difficulté pour la tâche), qui seraient passés à travers le premier filtre (exemple : individus se caractérisant comme francophones unilingues d'après les informations renseignées par le questionnaire d'inscription, mais ayant une maîtrise douteuse de la langue française. On trouve ainsi dans leurs réponses des mots très

courants comme « œufs » estimés appris à 34 ans). L'application de ces 2 critères a conduit à éliminer 26 participants, à la fois non francophones (8 de langue maternelle allemande, 1 de langue maternelle anglaise, 4 de langue maternelle espagnole, 5 de langue maternelle italienne, et 8 de langue maternelle non répertoriée dans la liste à choix), et présentant des distributions curieuses. Les 170 participants restant sont 151 français unilingues, 11 bilingues par le milieu social (la langue maternelle est le Français, la langue seconde est apprise dès après 7 ans sous l'influence du contexte social ou scolaire : allemand, N = 5 ; anglais N = 4 ; autre, N = 2), et 8 bilingues par le milieu familial (les deux langues sont apprises dès avant 4 ans : français / [anglais, N = 2 ; espagnol, N = 1 ; italien, N = 1 ; portugais, N = 2 ; autre, N = 2]).

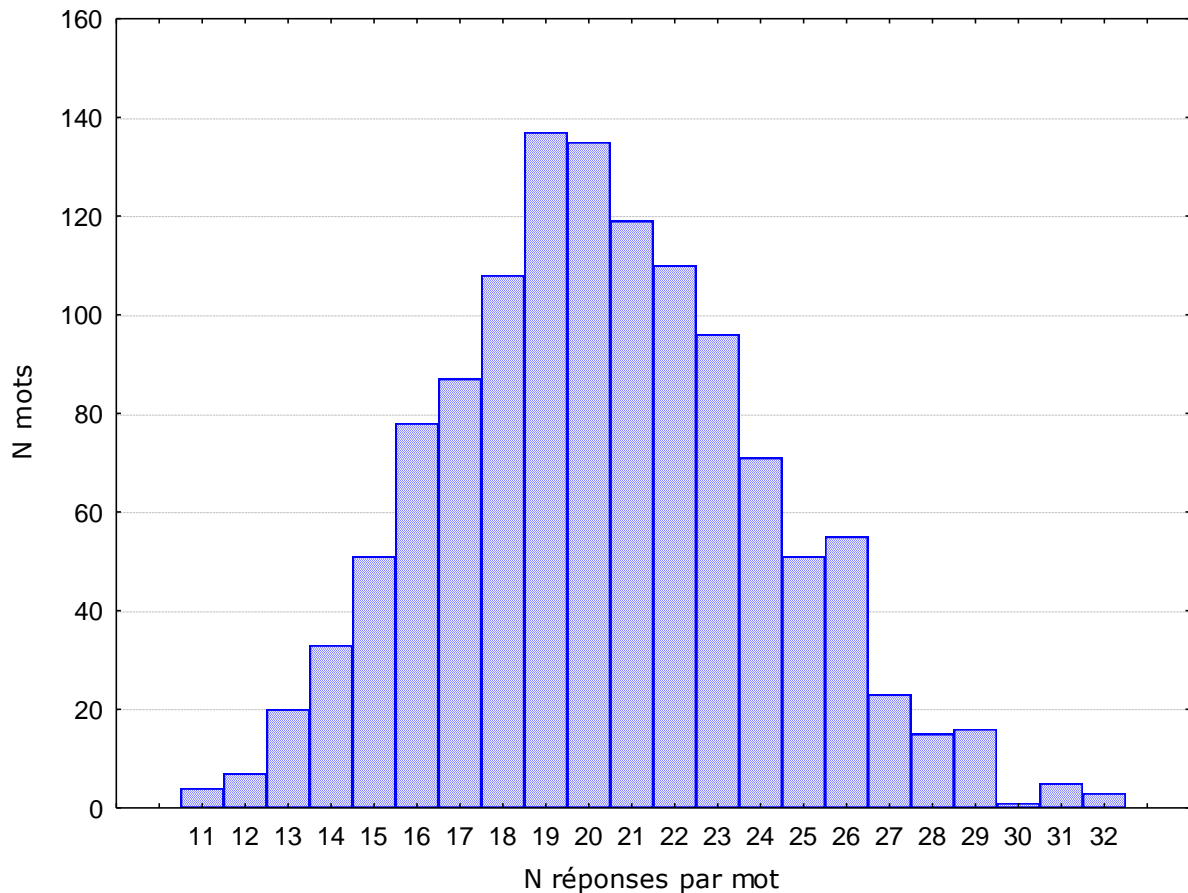
#### Filtrage par item des réponses des 170 participants restants

La distribution des données a été vérifiée pour chaque item de manière à filtrer toute réponse anormale. Quelques réponses ont encore été éliminées (exemples : réponses 0 an ou 99 ans, probablement pour « ne sais pas », réponses non normales et suspectes, comme « *trac (anxiété)* » appris à 2 ans (un enfant de 2 ans peut avoir peur, mais plus invraisemblablement avoir le trac), etc.). Au total, 2867 réponses ne satisfaisant pas aux critères statistiques et qualitatifs définis plus haut ont été éliminées. La base de données présentée ici a donc été constituée à partir de 24911 réponses (de 11 à 32 réponses par item ; moyenne = 20.34 ; écart-type = 3.73)<sup>138</sup>. La Figure 76 présente la distribution du nombre de mots en fonction du nombre de réponses par mot.

---

<sup>138</sup> Caractéristiques de la distribution brute avant filtrage : 27778 réponses, 12 à 34 réponses par item, moyenne = 22.68, écart-type = 3.89.

Figure 76. Distribution du nombre de mots en fonction du nombre de réponses par mot



Il est à noter que, si le nombre moyen de réponses par item est plus faible dans la présente base de données ( $N = 20$ ) que dans les bases d'Alario et col. ( $N = 26$  pour l'âge d'acquisition,  $N = 30$  pour la familiarité) ou de Ferrand et col. ( $N = 40$ ), la taille de l'échantillon reste acceptable si on s'en réfère à la base de Bonin et col., qui présente une information d'âge d'acquisition estimée à partir d'un échantillon de 20 personnes. Par ailleurs, la précision supérieure des échelles utilisées dans notre base (pas annuel pour l'estimation de l'âge d'acquisition, échelle en 6 points pour l'estimation de la familiarité) par rapport à celles utilisées dans les banques de données citées en référence (pas bi ou trisannuel pour l'estimation de l'âge d'acquisition, échelle en 5 points pour l'estimation de la familiarité), compense la perte éventuelle de précision de l'estimation à cause d'un échantillon plus faible. De plus, l'examen détaillé des données ne permet pas de relever, sur les items qui présentent un échantillon faible ( $N < 15$ ), d'écart moyen remarquable avec les mêmes items dans les bases de référence (écart avec la base de Bonin et col. (2003) :  $N = 10$  pour le nombre de mots communs entre bases, différence moyenne de 0.001 pour l'indice de familiarité, de 0.124 pour l'âge d'acquisition estimé ; écart avec la base de Ferrand et col. (2003) :  $N = 8$ , différence moyenne de 0.937 pour l'âge



d'acquisition estimé ; écart avec la base de Alario et col. (1999) :  $N = 5$ , différence moyenne de 0.243 pour l'indice de familiarité, de 0.548 pour l'âge d'acquisition estimé. Enfin, il faut préciser que la quantité d'items pour lesquels l'échantillon est faible (inférieur ou égal à 15) est peu importante (11 réponses par mot : 4 items (0.33%) ; 12 réponses : 7 items (0.57%) ; 13 réponses : 20 items (1.6%) ; 14 réponses : 33 items (2.7%) ; 15 items : 51 items (4.1%) – soit au total 115 items ou 9.4% de la base).

#### Codage des données filtrées et extraction de l'information

Les réponses fournies pour la certitude de l'âge d'acquisition estimé et pour la familiarité ne sont pas numériques, contrairement aux réponses données pour l'âge d'acquisition. Elles ont donc été codées numériquement pour permettre de réaliser les calculs de moyenne et d'écart type (--- = 1 ; -- = 2 ; - = 3 ; + = 4 ; ++ = 5 ; +++ = 6). Le nombre de réponses par item ( $N$ ) a été recensé. Pour chaque item et pour chaque dimension psychologique étudiée (âge d'acquisition estimé, degré de certitude pour l'âge d'acquisition fourni, familiarité), la moyenne ( $M$ ), la valeur minimum ( $Min$ ), la valeur maximum ( $Max$ ) et l'écart type ( $E.T.$ ) ont été extraits. Ces indices forment le corps de la base de données pour chacun des facteurs ( $M$ ,  $Min$ ,  $Max$ ,  $E.T.$ , en plus de  $N$ ).

#### Validation de l'information contenue dans la base de données

L'information contenue dans la banque de données a subi une double procédure de validation. La première concerne la vérification de la validité du recueil, la deuxième concerne la vérification de la conformité des données avec les normes existantes. Le ciblage de la valeur centrale a de plus été examiné.

#### Vérification de la validité du recueil

Si le nombre de réponses fournies pour chaque mot varie en fonction de l'âge d'acquisition estimé, de la familiarité, ou du degré de certitude sur l'âge estimé, c'est l'indice qu'il pourrait exister un biais dans la production ou le recueil des réponses. En effet, le système a été testé avant sa mise en production afin de s'assurer de la présentation aléatoire des items, et donc, le nombre de réponses par item ne doit pas varier en fonction d'un des 3 facteurs, sauf si le comportement des participants est influencé par l'une de ces 3 dimensions. Pour s'assurer que ce n'est pas le cas, une analyse statistique de type analyse de régression linéaire a été effectuée (variable dépendante = nombre de réponses recueillies par item ; prédicteurs = âge d'acquisition estimé, degré de certitude de l'estimation fournie pour l'âge d'acquisition, familiarité, interactions 2 à 2 entre ces 3 prédicteurs, interaction triple, et covariance entre prédicteurs). Le modèle mathématique de cette analyse est présenté dans l'Équation 16.

Équation 16 : Équation du modèle étudiant l'effet des facteurs âge d'acquisition, certitude sur l'estimation de l'âge, et familiarité, sur le nombre de réponses produites par les participants

$$n \text{ rép}_{It} \sim N(XB, \Omega)$$

$$n \text{ rép}_{It} = \beta_{0It} \text{Cons} + 0.065(0.053) \text{AoA}_{It} + 1.109(0.259) \text{Cert}_{It} + \\ 0.199(0.338) \text{Fam}_{It} + -0.154(0.103) \text{AoA.Cert}_{It} + \\ 0.035(0.039) \text{AoA.Fam}_{It} + -1.064(0.666) \text{Cert.Fam}_{It} + \\ 0.082(0.058) \text{AoA.Cert.Fam}_{It}$$

$$\beta_{0It} = 20.408(0.126) + e_{0It}$$

$$[e_{0It}] \sim N(0, \Omega_e) : \Omega_e = [13.624(0.550)]$$

$$-2 * \text{loglikelihood(IGLS Deviance)} = 6675.870(1225 \text{ of } 1225 \text{ cases in use})$$

Note. « n rép » le nombre de réponses par item (It), dont la distribution est supposée normale (N), de moyenne  $X\beta$  et de variance  $\Omega$ .  $\beta_0$  est le coefficient correspondant à la moyenne générale de la distribution (« intercept ») estimée par le modèle (20.408), accompagnée de l'erreur standard de l'estimation donnée entre parenthèses (0.126 – de même pour chaque facteur). « AoA » pour l'âge d'acquisition estimé, « Cert » pour le degré de certitude sur l'âge estimé, « Fam » pour le degré de familiarité (interactions : « AoA.Cert », « AoA.Fam », etc.). « Cons » correspond à une constante nécessaire au calcul de la référence, « e0 » est un terme d'erreur (variance aléatoire au niveau des items, supposée être distribuée normalement (N), de moyenne 0 et de variance  $\Omega_e$ ). « IGLS Deviance » est un score d'adéquation du modèle aux données, accompagné entre parenthèses du nombre d'items dans l'échantillon.

Tableau 123 : Résultat des tests de significativité

	$\chi^2(1, N = 1225)$	$p <$
<b>AoA</b>	1.484	.23
<b>Cert</b>	18.326	.0001
<b>Fam</b>	0.346	.56
<b>AoA.Cert</b>	2.241	.14
<b>AoA.Fam</b>	0.774	.38
<b>Cert.Fam</b>	2.556	.11
<b>AoA.Cert.Fam</b>	2.042	.16

Note. AoA pour âge d'acquisition estimé ; Cert pour degré de certitude pour l'estimation de l'âge d'acquisition ; Fam pour degré de familiarité.

L'analyse révèle que la quantité de réponses par item varie en fonction du degré de certitude sur l'estimation donnée pour l'âge d'acquisition. Plus le degré de certitude augmente et plus la quantité de réponses par item augmente (augmentation moyenne de 1.1 réponses par degré de certitude supplémentaire ; effet significatif à un seuil  $p < .0001$ ). Aucun autre facteur n'a d'influence significative, aucune interaction n'est significative.

Il existe donc un biais quantitatif dans les données, puisque le nombre de réponses fournies par les sujets pour chaque item dépendrait de la difficulté qu'ils ont eue à évaluer l'âge auquel ils ont appris le mot. Ce phénomène pourrait s'expliquer par le choix des participants de se déconnecter préférentiellement au moment où la difficulté pour répondre devient plus grande, ce facteur déclenchant s'ajoutant peut-être à un

degré de lassitude élevé face à une tâche répétitive. Toutefois, il faut mettre en avant le fait qu'aucune relation n'apparaît entre la quantité de réponses recueillie et les 2 facteurs d'intérêt principal, l'âge d'acquisition estimé et la familiarité. Les données recueillies pour ces 2 facteurs sont donc valables.

#### Vérification de la conformité des données avec les normes existantes

La deuxième procédure de validation est une comparaison des données de notre base avec chacune des 4 bases prises comme référence (bases de Alario et col. (1999), Ferrand et col. (2003), Bonin et col. (2003), et de Chalard et col. (2003)), pour les mots que ces bases ont en commun. Les résultats de cette comparaison, obtenus par une analyse de régression linéaire, sont récapitulés dans le Tableau 2. Ils montrent généralement une forte corrélation ( $R^2$  ajusté de .58 à .7 pour l'âge d'acquisition estimé, de .54 pour une comparaison avec l'âge d'acquisition réel, et de .07 à .43 pour la familiarité).

Tableau 124 : Comparaison statistique des données fournies par notre base avec celles fournies par 4 bases de données de référence

		Alario et col.	Ferrand et col.	Bonin et col.	Chalard et col.
Base Genève	N mots	49	86	59	36
	N rép	19.82	20.88	20.20	20.14
	%	4	7.02	4.82	2.94
Âge acquis.	$\beta$	0.84	0.84	0.76	0.74
	$R^2$ ajusté	0.7	0.7	0.58	0.54
Fam	$\beta$	0.66	-	0.3	-
	$R^2$ ajusté	0.43	-	0.07	-

Note. N mots pour le nombre de mots de notre base en commun avec chaque base comparée, N rép pour la taille moyenne de l'échantillon par mot, % le pourcentage de chevauchement entre notre base et chacune des 4 bases.  $\beta$  est le coefficient de régression, et  $R^2$  le coefficient de détermination.

Parallèlement, les 4 bases sont comparées entre elles par analyses de régression linéaire simple (Tableau 3 –  $R^2$  ajusté de .91 à .94 pour l'âge d'acquisition estimé, et de .11 à .29 pour une comparaison entre âge d'acquisition estimé et réel. Pour la comparaison des scores de familiarité, seules les bases d'Alario et col. et de Bonin et col. pouvaient être utilisées, mais d'une part le nombre d'items communs entre les deux bases était trop faible ( $N = 6$ ), et d'autre part seulement 2 items sur 6 (« léopard » et « tonneau ») avaient une information de familiarité dans la base d'Alario et col. (Alario et col. vs Bonin et col. : respectivement 2.45 vs. 1.37, et 2.4 vs. 1.27)).

Tableau 125 : Comparaison des scores d'âge d'acquisition fournis par les 4 bases de données de référence

$R^2$ ajusté \ N	Alario et col.	Ferrand et col.	Bonin et col.	Chalard et col.
Alario et col.		<u>28</u>	<u>6</u>	<u>226</u>
Ferrand et col.	0.94		<u>21</u>	<u>21</u>
Bonin et col.	-	0.91		<u>2</u>
Chalard et col.	0.29	0.11	-	

Note. N est le nombre de mots que les bases ont en commun 2 à 2, et R2 est le coefficient de détermination.

Enfin, une comparaison entre les résultats des analyses de régression impliquant notre base et ceux des analyses de régression n'impliquant pas notre base est effectuée. Elle montre que les données d'âge d'acquisition estimé contenues dans notre base ont des caractéristiques intermédiaires entre les données d'âge d'acquisition estimé par Alario et col., Ferrand et col., ou Bonin et col. (avec lesquelles elles corréleront moins bien que les données de ces bases de référence entre elles), et les données d'âge réel mesurées par Chalard et col. (avec lesquelles elles corréleront mieux que les données des bases de référence précédentes ne le font). Ce résultat indique que l'estimation de l'âge d'acquisition fournie par notre base est plus proche de l'âge réel d'acquisition que les estimations fournies par les bases d'Alario et col. et Ferrand et col. (pas de statistiques pour Bonin et col.). Concernant les écarts entre notre base et les bases de référence pour la familiarité, l'explication qui peut être avancée concerne les différences de procédure de recueil entre études (échelles, tri des données brutes, conditions de passation, précision des acceptions sémantiques, etc.). Hélas, le nombre de mots en commun entre les bases d'Alario et col. et Bonin et col. est trop faible pour faire une analyse statistique. On remarque toutefois un écart notable entre les 2 items en commun bien que la procédure de recueil soit similaire dans ces 2 études.

#### Ciblage de la moyenne

L'examen du ciblage de la moyenne fournit une indication de la congruence des réponses des participants. Concernant l'estimation de l'âge d'acquisition, l'écart-type moyen est de 2.7 (Min = 0.9, Max = 10.9, E.T. = 1.2). Il augmente avec l'âge d'acquisition estimé ( $\beta = .78$ ). Concernant le degré de familiarité, il est de 1.1 (Min = 0, Max = 2.9, E.T. = 0.6). Il diminue lorsque le degré de certitude augmente ( $\beta = -.85$ ).

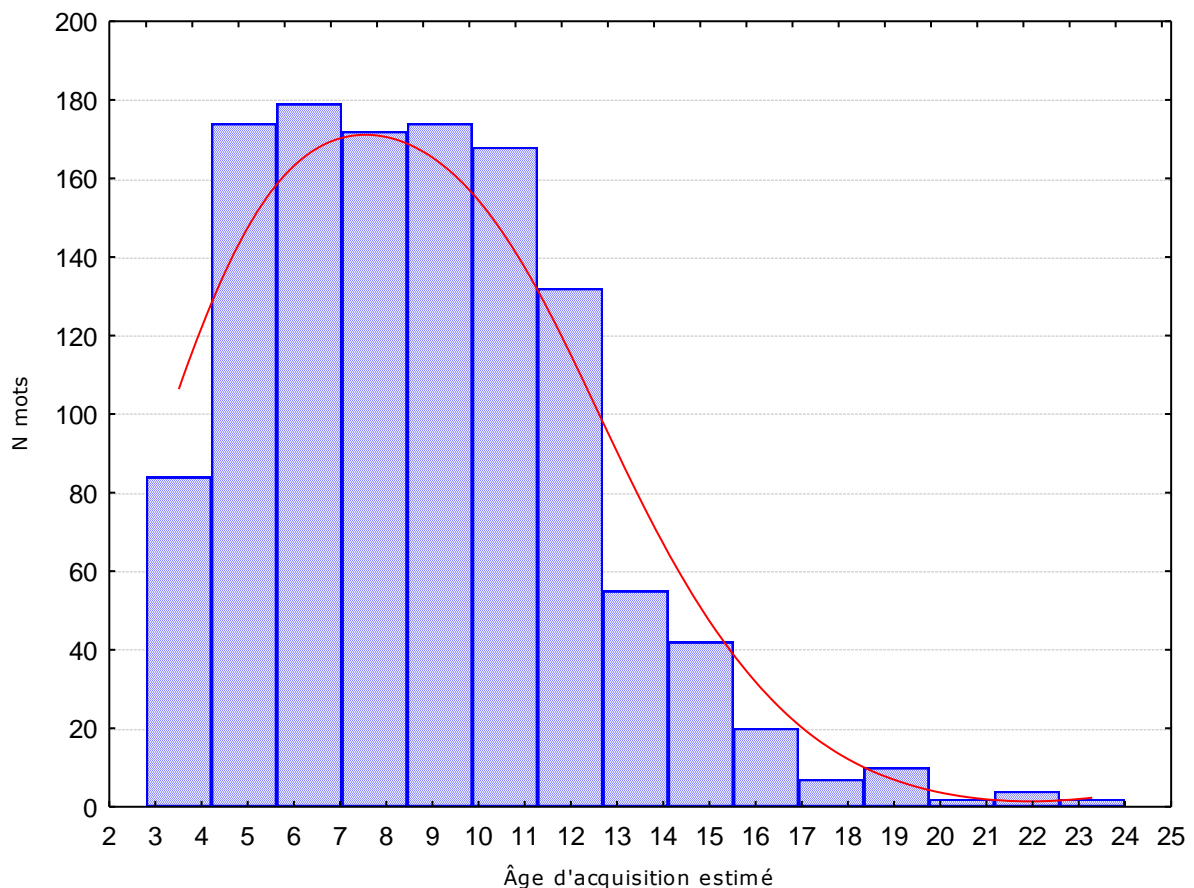
Cette information permet de confirmer une congruence plutôt bonne des réponses des participants, tant pour l'estimation de l'âge d'acquisition que pour le degré de certitude.

Conclusions sur la validité des données de la base

L'information proposée par notre base apparaît être plutôt fiable. Premièrement, aucun biais impliquant la procédure de recueil en lien avec l'âge d'acquisition ou la familiarité n'a pu être mis en évidence. Deuxièmement, l'âge d'acquisition estimé dans notre base s'approche plus fortement de l'âge réel d'acquisition que l'estimation trouvée dans les autres bases de référence. Concernant la familiarité, il est plus difficile de se prononcer étant donné qu'il manque une valeur de référence (corrélation entre les bases d'Alario et col. et de Bonin et col.) à laquelle comparer le résultat des corrélations entre notre base et ces 2 bases. Cependant, considérant les résultats sur la validité du recueil et sur la validité de l'âge d'acquisition, d'une part, et d'autre part ceux relatifs au ciblage des réponses sur la valeur moyenne, il n'y a pas de raisons de douter de la fiabilité du facteur familiarité.

Présentation des données contenues dans la baseExamen quantitatif : quelques distributionsÂge d'acquisition estimé

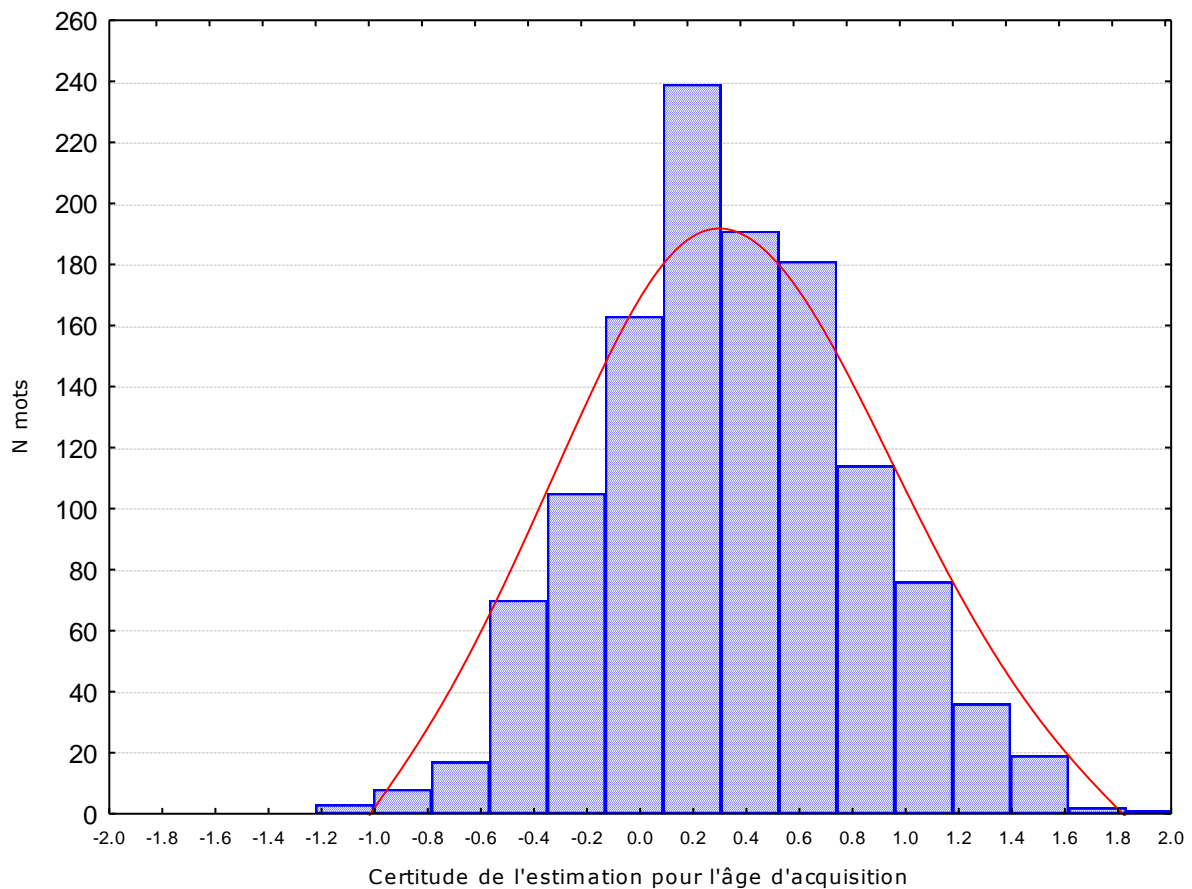
Figure 77. Distribution du nombre de mots en fonction de l'âge d'acquisition estimé (courbe : distance pondérée des moindres carrés)



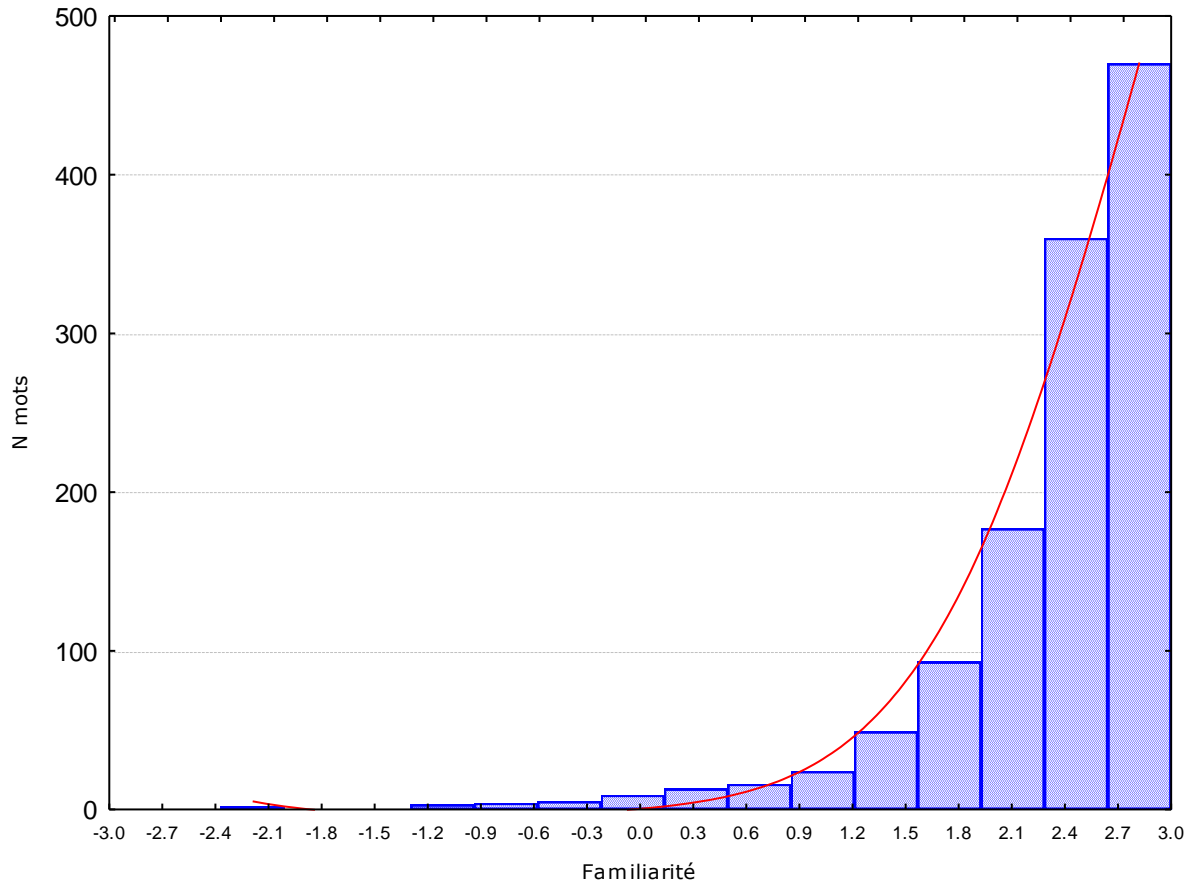
Le vocabulaire testé est en majorité estimé appris entre 4 et 12 ans (moyenne : 8.74 ans ; écart-type : 3.47 ans). Il semble, d'après cette distribution, que vers 16-17 ans, le lexique mental d'un individu de langue maternelle française soit à peu près en place (chute brutale de la quantité estimée de nouveaux mots appris vers 12-13 ans, et atteinte d'un pallier vers 16-17 ans).

#### Degré de certitude de l'estimation de l'âge d'acquisition

Figure 78. Distribution du nombre de mots en fonction du degré de certitude sur l'âge d'acquisition estimé (courbe : distance pondérée des moindres carrés)



Les participants se sont sentis plutôt certains qu'incertains de leur évaluation des âges d'acquisition (moyenne positive : 0.34 ; écart-type : 0.49). La dispersion autour de la moyenne, symétrique, indique cependant qu'il existe aussi bien des réponses estimées avec une forte certitude que des réponses estimées avec une incertitude assez marquée.

Familiarité des participants avec les mots testésFigure 79. Distribution du nombre de mots en fonction de la familiarité (courbe : distance pondérée des moindres carrés)

Les participants se sont sentis familiers à très familiers avec la majorité des mots testés (moyenne : 2.34 ; écart-type : 0.67).

Examen qualitatif des réponsesÂge d'acquisition estimé

Les valeurs obtenues pour l'âge d'acquisition estimé, la certitude sur l'âge d'acquisition estimé, et la familiarité, présentent une grande plausibilité psychologique, ainsi qu'une finesse certaine. Concernant l'âge d'acquisition, ces qualités se distinguent plus aisément dans le cas d'homophones, pour lesquels des différences pertinentes se remarquent en fonction du sens. Par exemple : [« air » (*mélodie*) : 5.63 ans / « air » (*gaz*) : 6.07 / « aire » (*surface*) : 9.12 / « ère » (*période*) : 9.91 / « erre » (*errer*) : 10.39] ; [« balle » (*ballon*) : 3.3 ans / « balle » (*munition*) : 6.18 / « balle » (*ballot*) : 9.67 / « balle » (*monnaie*) : 10.71] ; [« ton » (*tien*) : 3.94 ans / « thon » (*poisson*) : 6.62 / « ton » (*tonalité*) : 9.62 / « ton » (*coloris*) : 9.76] ; [« pain » (*aliment*) : 4.13

ans / « pin » (végétal) : 7.25 / « pain » (coup) : 9.3 / « pain » (masse) : 11.72] ; [« loup » (mammifère) : 4.5 ans / « loup » (masque) : 10.09 / « loup » (poisson) : 12.32] ; [« porc » (animal) : 5.77 ans / « port » (maritime) : 6.21 / « port » (porter) : 10.56 / « porc » (débauché) : 11.11 / « port » (silhouette) : 12.78 / « pore » (orifice) : 13]. Les informations se révèlent être très intéressantes d'un point de vue développemental, tant sur le plan de l'acquisition du lexique, que sur ce qu'elles indiquent de la maturation de la personnalité et de l'évolution du vécu de l'enfant au cours de son histoire (accès aux concepts, centres d'intérêt, etc.). Un examen de la distribution des mots pour l'âge moyen d'acquisition estimé révèle que les premiers mots acquis, de 2.8 à 4 ans, sont majoritairement monosyllabiques, et véhiculent des concepts simples ou élémentaires. Soit ils sont directement liés à l'enfant à travers son anatomie (« dent », « œil », « doigt », « tête », « pied », etc.), la notion d'individualité (« moi », « toi », « tu », « il », « ma », « ton », « est », « suis », etc.), les caractéristiques du corps (« grand », « long », « gros », « âge », etc.), le comportement (« sage », « cri », « rire », etc.), le ressenti (« aïe » / « ouille » (douleur), « froid », « aime », etc.), soit ils sont en rapport étroit avec l'environnement dans lequel l'enfant évolue, comme les objets qu'il manipule (« eau », « lait », « balle », « jouet », « lit », etc.), les phénomènes qui surviennent dans cet environnement (« nuit », « noir », « jour », « neige », « pluie », « vent », « bruit », etc.), les êtres vivants qu'il y rencontre (« homme », « femme », « père », « chat », « fleur », « arbre », etc.). À l'opposé, les mots estimés acquis le plus tardivement, de 17 à 24 ans, sont monosyllabiques ou bisyllabiques (cette étude ne teste pas de mots de plus de 2 syllabes), ne sont pas forcément plus complexes au niveau formel, mais surtout, font référence à des éléments caractérisés par leur abstraction. Celle-ci peut être liée à l'éloignement des objets/êtres/phénomènes désignés, hors du quotidien de l'individu (« roussette », « phénol », « calice », « anche », « attique », « palette », « serveur », « Brig » (ville), « crack » (drogue), etc.) comme au degré d'élaboration des concepts que véhiculent ces mots (« lemme », « tore », « orbe », « tripe » (cigare), « ose » (glucide), etc.).

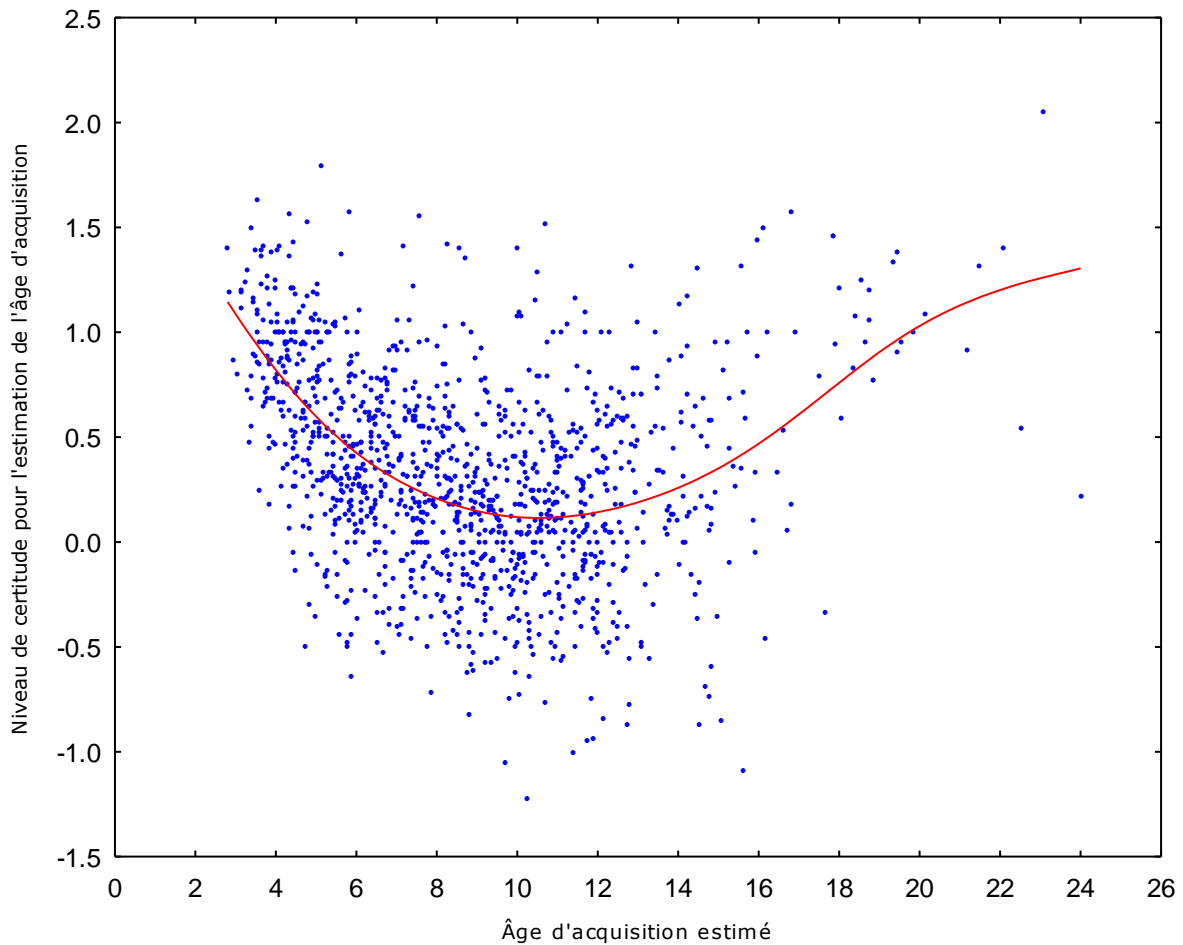
#### Degré de certitude de l'estimation de l'âge d'acquisition

Concernant le score de certitude sur l'estimation de l'âge d'acquisition, l'examen de la distribution des réponses moyennes par item pour les 1225 mots, en fonction de l'âge d'acquisition estimé et du degré de certitude fourni pour l'estimation de l'âge (Figure 6), montre que la période de la vie où les participants pensent avoir appris la plus grande quantité des mots présentés (environ 4 à 12 ans) est aussi celle pour laquelle le degré de certitude est le plus bas. Ce phénomène peut être lié à l'augmentation importante de la quantité de mots nouvellement mémorisés durant cette période, ce qui les rend



plus difficiles à caractériser individuellement et à localiser dans le temps. Il peut aussi être lié à un phénomène développemental, un lexique mental en phase d'élaboration subissant peut-être, au gré de l'intégration de nouveaux éléments, un remaniement interne aboutissant à une mémoire plus confuse.

Figure 80. Distribution des réponses moyennes par item en fonction de l'âge d'acquisition estimé et du degré de certitude de l'estimation (courbe : distance pondérée des moindres carrés)



Malgré ce fait, le degré de certitude de l'estimation reste théoriquement utilisable pour évaluer la netteté du souvenir, ou du souvenir imaginé, logique, plausible, reconstruit, etc., ainsi que la cohérence des répondants. Par exemple, d'après cet indice, l'âge d'acquisition estimé pour le mot « *pain* » (*masse*) serait moins certain que celui estimé pour le mot « *jeu* » (*divertissement*), puisque le premier est marqué d'un score de certitude moyen plutôt faible (-0.94) et dispersé (de -3 à 3 ; E.T. : 1.7), le second d'un score de certitude moyen plutôt élevé (1.63) et plus homogène (de -1 à 3, E.T. : 1.3). Parallèlement, l'âge d'acquisition moyen est peu ciblé pour le mot « *pain* » (*masse*)

(minimum : 4 ans ; maximum : 26 ans ; E.T. : 6.4), alors qu'il apparaît plutôt ciblé pour le mot « jeu » (*divertissement*) (de 1 à 6 ans, E.T. : 1.3).

#### Familiarité des participants avec les mots testés

Concernant la familiarité, il semble exister un lien très fort entre ce facteur et l'âge d'acquisition estimé, visible sans autre analyse : les mots non familiers tendent tous à avoir des valeurs élevées d'âge d'acquisition estimé. L'observation qualitative des données suggère donc qu'un mot appris très tôt dans l'enfance, du fait de sa référence à des éléments conceptuellement simples et de sa grande antériorité (on est plus familier avec des objets que l'on connaît depuis longtemps), devrait être considéré comme très familier. Ceci se vérifie mathématiquement à partir des scores de la base de données (pour les 57 mots estimés appris avant 4 ans, le degré de familiarité moyen est de 2.92).

#### Conclusions sur la méthode de recueil et sur la qualité de l'information obtenue

L'examen quantitatif des données confirme les avantages de la méthode de recueil adoptée dans cette recherche. Face à une information difficile à obtenir, la méthode s'avère efficace (échantillonnage de la population sur un vaste territoire de la francophonie) et économique (recueil rapide d'une quantité importante d'information avec un coût nul en terme d'expérimentation directe).

L'examen qualitatif des données montre que l'avantage quantitatif n'a pas lieu au détriment de la qualité. La méthode est précise (l'information recueillie est très fine, comme le montre la comparaison des scores pour les homophones) et paraît fiable (l'information recueillie est cohérente, et consistante avec celle proposée par d'autres études).

De l'insuffisance d'un contrôle expérimental basé sur la fréquence d'occurrence à la nécessité de disposer de bases de données lexicales pour des facteurs psychologiques

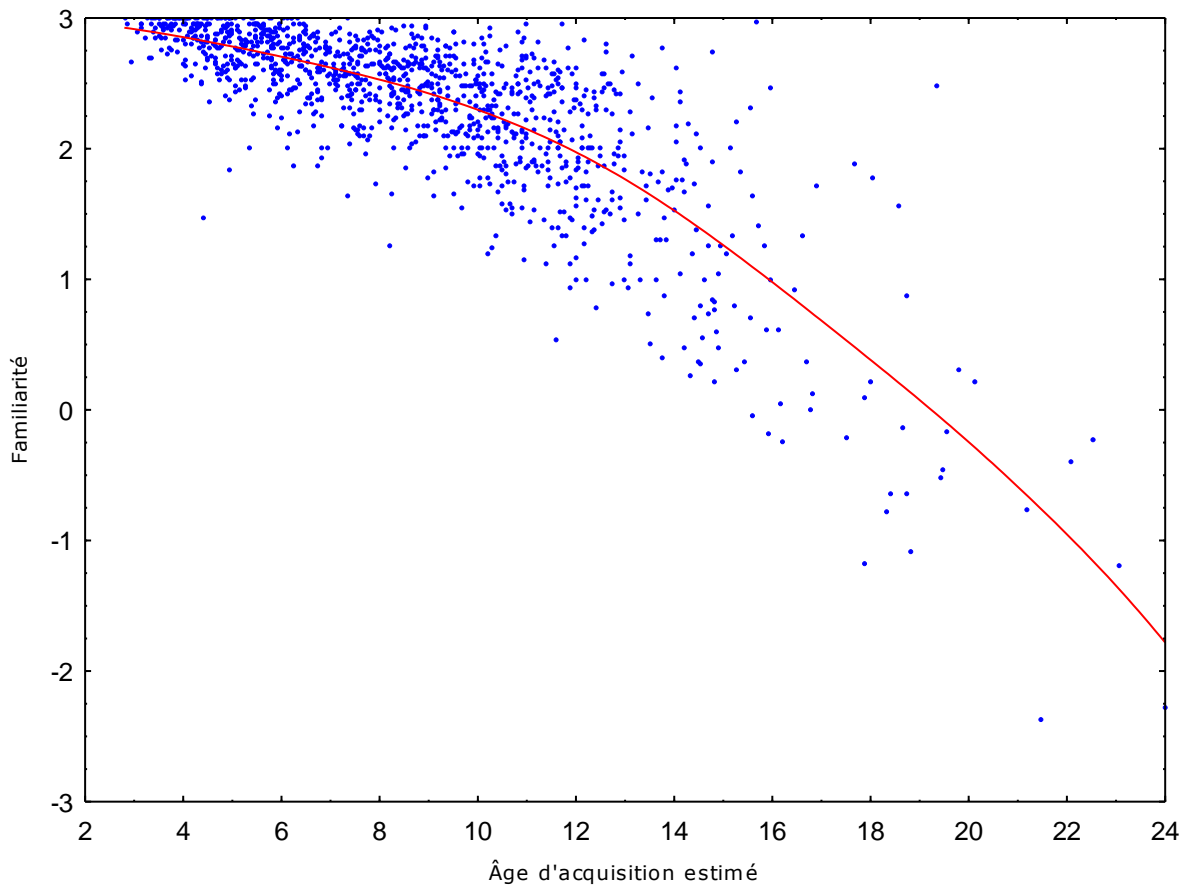
L'article part d'une problématique relative au contrôle du matériel expérimental lexical à l'aide de la fréquence d'occurrence, dimension habituellement utilisée à cette fin et qui se trouve être logiquement corrélée avec l'âge d'acquisition et la familiarité. Si l'âge d'acquisition et la familiarité sont des dimensions réellement influentes sur le traitement lexical, et que la fréquence permette de contrôler la situation parce qu'elle offre un contrôle de ces deux dimensions, alors il est nécessaire de connaître le degré de chevauchement entre fréquence, âge d'acquisition et familiarité, pour mieux évaluer le niveau de contrôle général de ces facteurs par la fréquence. La question est donc de savoir quelle proportion résiduelle de l'âge d'acquisition et de la familiarité reste non

contrôlée suite au contrôle de la fréquence, pour apporter une justification supplémentaire à l'importance d'avoir à disposition des bases de données renseignant l'âge d'acquisition et la familiarité.

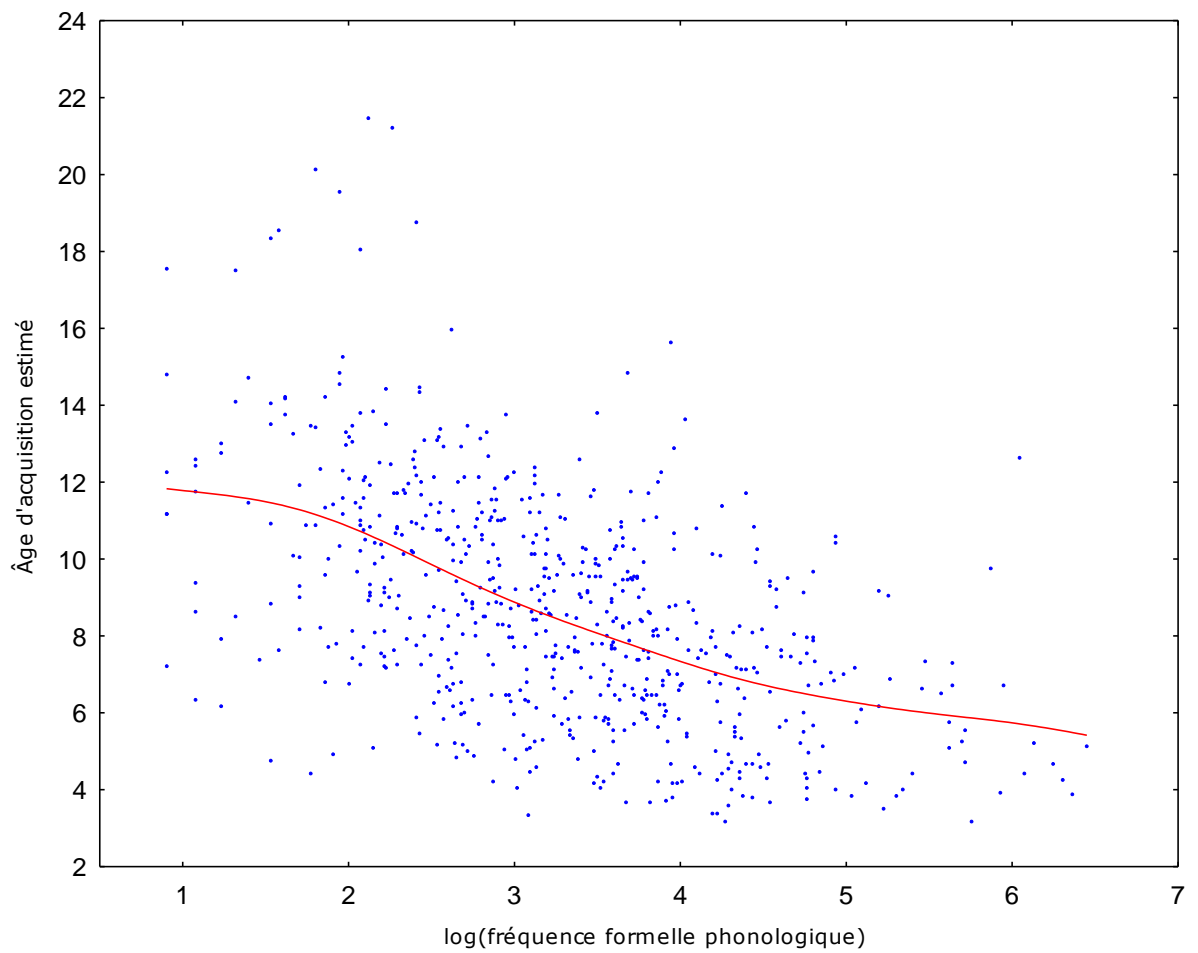
Pour répondre à cette question, la relation entre l'âge d'acquisition estimé, la familiarité et la fréquence formelle phonologique des mots (calculée à partir de BRULEX – se reporter à la section « Examen statistique des relations entre facteurs » pour le détail) est explorée dans cette section. Le facteur « certitude » n'est pas étudié ici, étant par ailleurs un indicateur secondaire relatif au facteur âge d'acquisition (pour rappel, il sert à donner un indice sur la validité du score d'âge d'acquisition, avec lequel il est faiblement corrélé ( $R^2$  ajusté = .040)). De même, l'indice de certitude est très faiblement corrélé avec la fréquence formelle phonologique des mots ( $R^2$  ajusté = .016) et avec la familiarité ( $R^2$  ajusté = .003)).

#### Présentation des distributions des réponses en fonction des facteurs 2 à 2

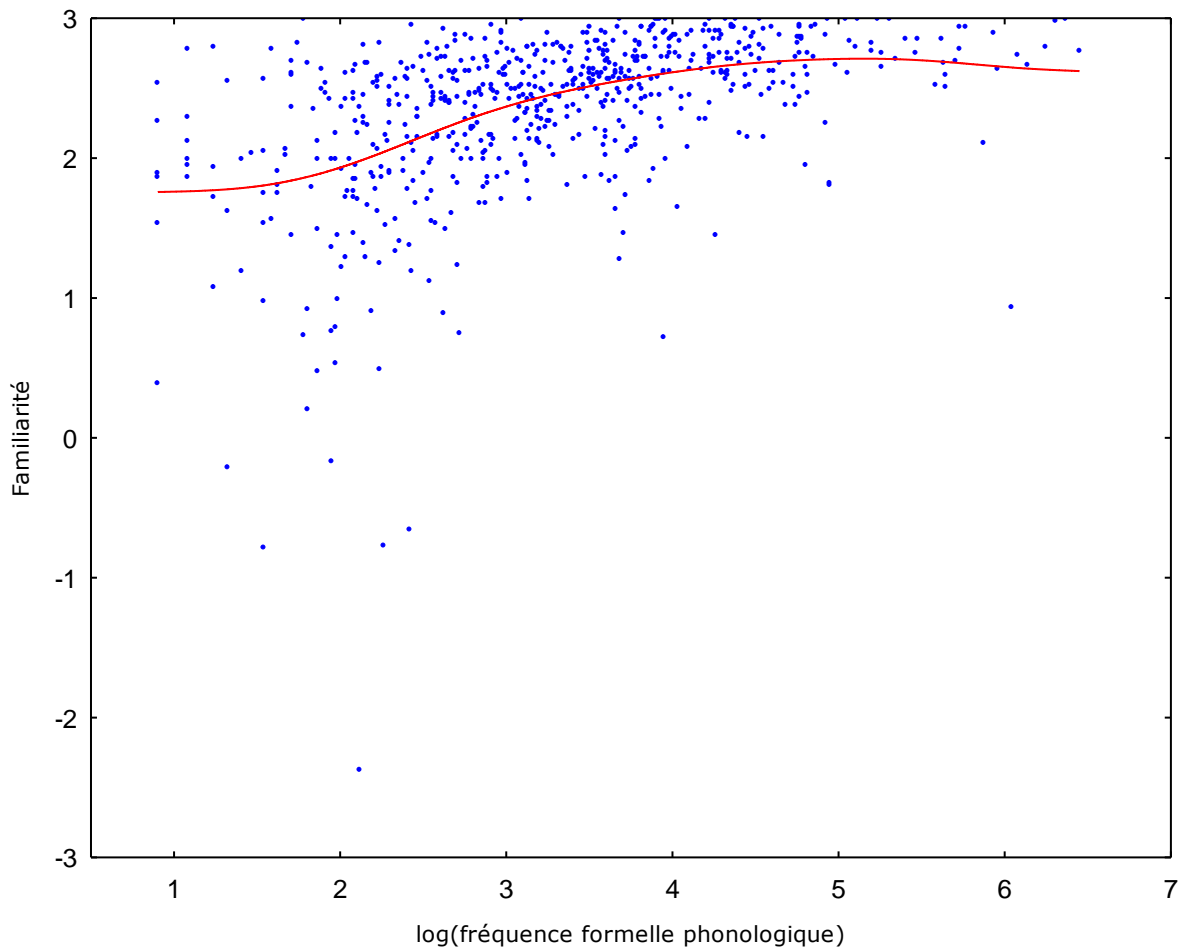
Remarque : les valeurs présentées dans les graphiques qui font intervenir le facteur fréquence sont des moyennes pour 653 formes phonologiques du corpus (se reporter plus loin à la section « Analyses statistiques » pour les explications). La Figure 7, qui ne graphe pas le facteur fréquence, présente la distribution des réponses moyennes par item pour les 1225 mots du corpus.

Relation Âge d'acquisition / FamiliaritéFigure 81. Distribution des réponses moyennes par item en fonction de l'âge d'acquisition estimé et de la familiarité (courbe : distance pondérée des moindres carrés)

Il semble que plus un mot a été appris tôt durant l'ontogenèse, plus l'individu se sente familier avec.

Relation Âge d'acquisition / Fréquence d'occurrenceFigure 82. Distribution des réponses moyennes par item en fonction de la fréquence d'occurrence et de l'âge d'acquisition (courbe : distance pondérée des moindres carrés)

Il semble que plus un mot est fréquent, plus il a été appris tôt.

Relation Familiarité / Fréquence d'occurrenceFigure 83. Distribution des réponses moyennes par item en fonction de la fréquence d'occurrence et de la familiarité (courbe : distance pondérée des moindres carrés)

Il semble que plus un individu soit familier avec un mot, plus ce dernier soit fréquent.

Il n'est cependant pas possible de conclure visuellement, à partir de ces distributions, sur la relation réelle entre facteurs, étant donné leur colinéarité supposée (plus un mot est fréquent, plus il y a de chances pour qu'il ait été appris tôt, et plus il y a de chances pour que l'on soit familier avec). Il est donc nécessaire de procéder à une analyse statistique à l'aide d'un modèle prenant simultanément en considération les trois facteurs et leur covariance, ainsi que leurs interactions. Le but est de formaliser la relation entre facteurs (effet principal de chacun sur les 2 autres).

Examen statistique des relations entre facteurs

Il était nécessaire de tester 2 modèles afin d'estimer comment chacune des 3 dimensions est expliquée par les 2 autres, tout en contrôlant la covariance des

prédicteurs et l'incidence des interactions simples et multiples entre les 3 facteurs sur la variable dépendante. Cependant, un troisième modèle a été estimé pour connaître l'effet sur chaque dimension des interactions des prédicteurs. L'équation de chaque modèle testé est présentée et explicitée, accompagnée d'un tableau récapitulant les résultats du test statistique ( $\text{Chi}^2_{(\text{degré de liberté, N observations})}$ , seuil de significativité). Une représentation graphique de la distribution des données (centrées sur la moyenne pour chacun des deux facteurs représentés), sur laquelle figurent la droite de régression (estimation de la relation entre les deux facteurs par le modèle<sup>139</sup>) et l'intervalle de confiance de l'estimation (seuil de confiance à 95%), est insérée dans le texte de description qui suit la présentation des résultats pour chaque modèle.

Plusieurs précisions doivent être apportées relativement aux données utilisées pour ces analyses statistiques.

Concernant la quantité d'items utilisés, premièrement, une limite a été imposée parce que la base de données BRULEX ne fournit pas la fréquence pour tous les mots de notre base de données (par exemple, les verbes conjugués comme hante ou aima, ou certains noms rares comme gnou, sari, erg). Deuxièmement, une fréquence des formes orales des mots doit être calculée à partir de BRULEX (somme des fréquences formelles écrites des homophones), ce travail s'insérant dans une recherche plus générale qui portait sur la reconnaissance des mots parlés. À cause des 2 premiers points, le nombre total de mots analysables dans notre base de données est de 653 formes phonologiques. Troisièmement, les données comportementales disponibles au moment de l'analyse dans l'étude générale de laquelle sont issues les analyses présentées ci-après, portaient sur 151 formes phonologiques. C'est donc cette quantité d'items qui a été utilisée dans les analyses présentées. Elle est amplement suffisante d'un point de vue statistique pour faire ressortir des phénomènes robustes.

Concernant la nature des données analysées, premièrement, à cause du passage d'une information lexicale à une information formelle phonologique, les valeurs d'âge d'acquisition estimé et de familiarité correspondent maintenant à une moyenne par forme phonologique. Deuxièmement, la variable fréquence n'est pas qu'une somme des fréquences de comptage des différents homophones calculée depuis BRULEX, mais un logarithme en base 10 de cette somme de fréquences. Une telle transformation est nécessaire pour « linéariser » le facteur, autrement distribué selon une loi logarithmique

---

<sup>139</sup> La droite de régression peut donc apparaître mal orientée par rapport au nuage de points, dans les graphiques. Elle correspond en effet à une partie de l'information dans l'équation, alors que le nuage de points correspond à la totalité de l'information modélisée dans l'équation.

(sans cette précaution, l'analyse de régression linéaire sur laquelle se base la méthode utilisée ici est techniquement invalide).

Une analyse de type corrélatif entre le logarithme en base 10 de la fréquence formelle phonologique d'occurrence calculée à partir de BRULEX, la valeur moyenne d'âge d'acquisition par forme phonologique, et la valeur moyenne de familiarité par forme phonologique est donc effectuée pour 151 formes phonologiques. Dans chaque modèle, les valeurs de chaque variable indépendante ont été centrées sur la moyenne du facteur pour permettre l'interprétation directe des effets dans la dimension de la variable dépendante.

#### Effet des prédicteurs de l'âge d'acquisition estimé sur cette dimension

Équation 17 : Équation du modèle étudiant l'influence des prédicteurs de l'âge d'acquisition

$$\hat{a}g_{n^{\circ} \text{ item}} \sim N(\mathbf{XB}, \Omega)$$

$$\hat{a}g_{n^{\circ} \text{ item}} = \beta_{0n^{\circ} \text{ item}} \text{Cons} + \mathbf{-0.001(0.259)} \text{LgFc}_{n^{\circ} \text{ item}} + \mathbf{-2.921(0.605)} \text{famc}_{n^{\circ} \text{ item}} + \mathbf{-0.207(0.094)} \text{LgFc} \cdot \hat{a}gc_{n^{\circ} \text{ item}} + \mathbf{-1.299(0.735)} \text{LgFc} \cdot \text{famc}_{n^{\circ} \text{ item}} + \mathbf{0.693(0.199)} \hat{a}gc \cdot \text{famc}_{n^{\circ} \text{ item}} + \mathbf{1.069(0.170)} \text{LgFc} \cdot \hat{a}gc \cdot \text{famc}_{n^{\circ} \text{ item}}$$

$$\beta_{0n^{\circ} \text{ item}} = \mathbf{8.632(0.185)} + e_{0n^{\circ} \text{ item}}$$

$$[e_{0n^{\circ} \text{ item}}] \sim N(0, \Omega_e) : \Omega_e = [\mathbf{3.335(0.384)}]$$

**-2\*loglikelihood(IGLS Deviance) = 610.414(151 of 151 cases in use)**

Note.  $\hat{a}g$  et  $\hat{a}gc$  pour l'âge d'acquisition estimé (c pour centré sur la moyenne) ; LgFc pour le logarithme en base 10 de la fréquence d'occurrence ; famc pour la Familiarité. Se reporter à la note accompagnant l'Équation 1 pour le détail des autres éléments du modèle.

Tableau 126 : Résultat des tests de significativité

	Chi <sup>2</sup> <sub>(1, N = 151)</sub>	p <
<b>famc</b>	23.3	<b>.0001</b>
<b>LgFc</b>	0	1
<b>famc * LgFc</b>	3.1	.08
<b>famc * <math>\hat{a}gc</math></b>	12.1	<b>.0005</b>
<b>LgFc * <math>\hat{a}gc</math></b>	4.9	<b>.028</b>
<b>LgFc * <math>\hat{a}gc</math> * famc</b>	39.7	<b>.0001</b>

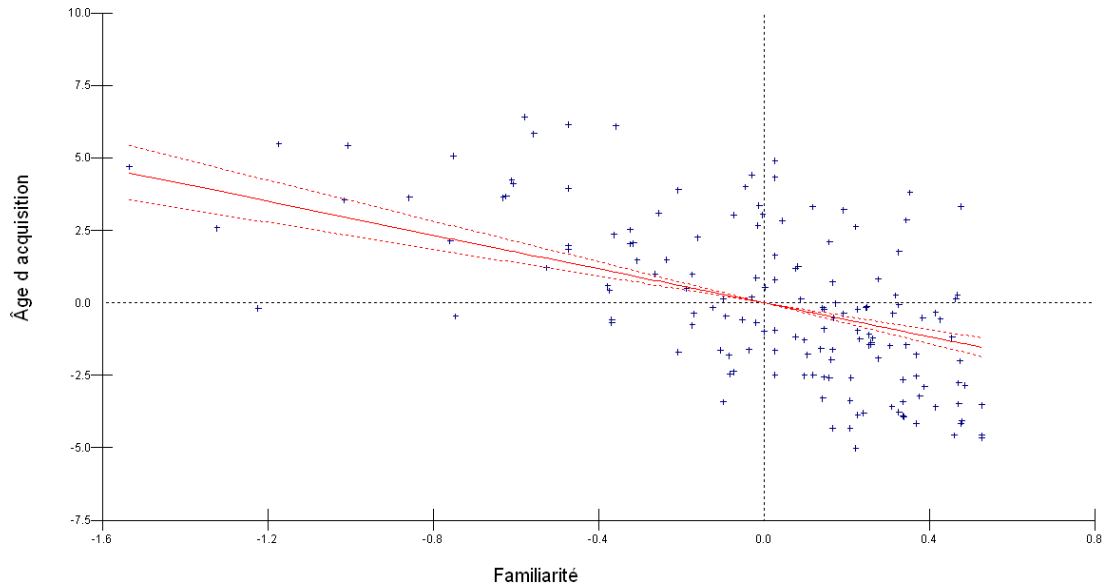
Note. famc pour degré de familiarité (scores centrés sur la moyenne de la dimension) ; LgFc pour logarithme en base 10 de la fréquence d'occurrence du mot (valeurs centrées sur la moyenne de la dimension) ;  $\hat{a}gc$  pour âge d'acquisition estimé (scores centrés sur la moyenne de la dimension).

Statistiquement, l'âge d'acquisition estimé est fortement relié à la familiarité, avec laquelle il est négativement corrélé (Figure 10), et ce indépendamment de la fréquence



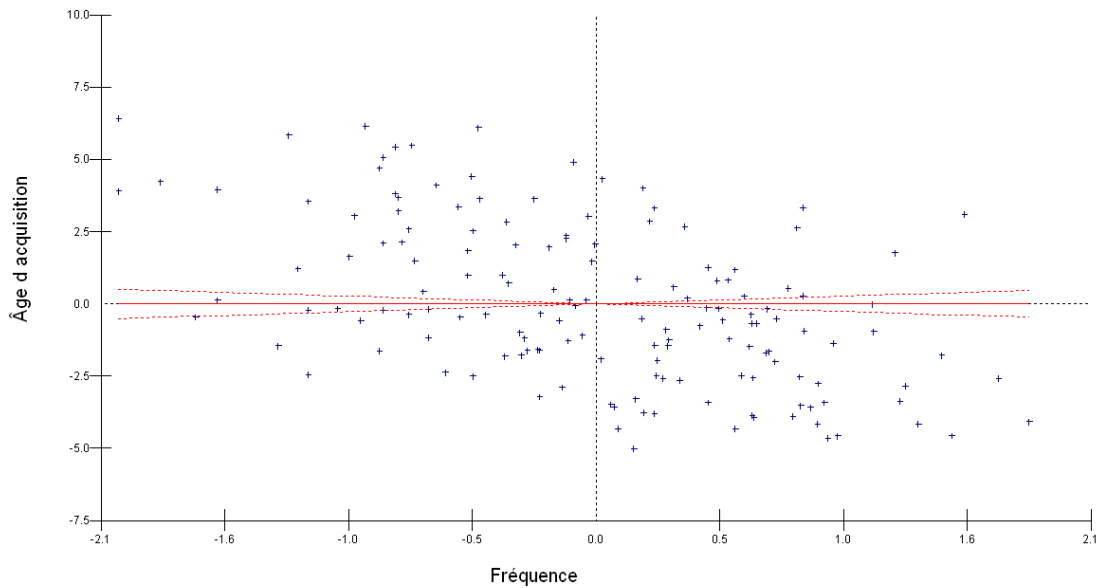
d'occurrence et des interactions que les trois facteurs peuvent entretenir entre eux (Tableau 4).

Figure 84. Relation entre âge d'acquisition estimé et familiarité



Plus un individu est familier avec un mot, et plus le mot a été appris tôt (en moyenne, l'âge d'acquisition d'un mot diminue de 2.92 années par degré de familiarité supplémentaire -  $\text{Chi}^2_{(1, N = 151)} = 23.3$  ;  $p < .0001$ ). L'âge d'acquisition estimé dépend également des interactions qu'il fait avec les autres facteurs (Tableau 2). Par contre, la fréquence d'occurrence n'influence pas directement l'âge d'acquisition (Figure 11 -  $\text{Chi}^2_{(1, N = 151)} = 0$  ;  $p = 1$ ). Si elle le fait, c'est par le biais d'interactions.

Figure 85. Relation entre âge d'acquisition estimé et fréquence d'occurrence



Effet des prédicteurs de la fréquence d'occurrence sur cette dimension

Équation 18 : Équation du modèle étudiant l'influence des prédicteurs de la fréquence

$$\lg Fq_{n^{\circ} \text{ item}} \sim N(\mathbf{XB}, \Omega)$$

$$\lg Fq_{n^{\circ} \text{ item}} = \beta_{0n^{\circ} \text{ item}} \text{Cons} + 0.887(0.191) \text{famc}_{n^{\circ} \text{ item}} + \\ 0.000(0.026) \hat{\text{âgc}}_{n^{\circ} \text{ item}} + 0.074(0.029) \lg Fc \cdot \hat{\text{âgc}}_{n^{\circ} \text{ item}} + \\ 0.616(0.228) \lg Fc \cdot \text{famc}_{n^{\circ} \text{ item}} + -0.189(0.063) \hat{\text{âgc}} \cdot \text{famc}_{n^{\circ} \text{ item}} + \\ -0.279(0.056) \lg Fc \cdot \hat{\text{âgc}} \cdot \text{famc}_{n^{\circ} \text{ item}}$$

$$\beta_{0n^{\circ} \text{ item}} = 2.826(0.058) + e_{0n^{\circ} \text{ item}}$$

$$[e_{0n^{\circ} \text{ item}}] \sim N(0, \Omega_e) : \Omega_e = [0.330(0.038)]$$

$$-2 * \log \text{likelihood (IGLS Deviance)} = 261.048(151 \text{ of } 151 \text{ cases in use})$$

Note.  $\lg Fq$  et  $\lg Fc$  pour le logarithme en base 10 de la Fréquence d'occurrence (c pour centré sur la moyenne) ;  $\text{famc}$  pour la Familiarité ;  $\hat{\text{âgc}}$  pour l'âge d'acquisition estimé. Se reporter à la note accompagnant l'Équation 1 pour le détail des autres éléments du modèle.

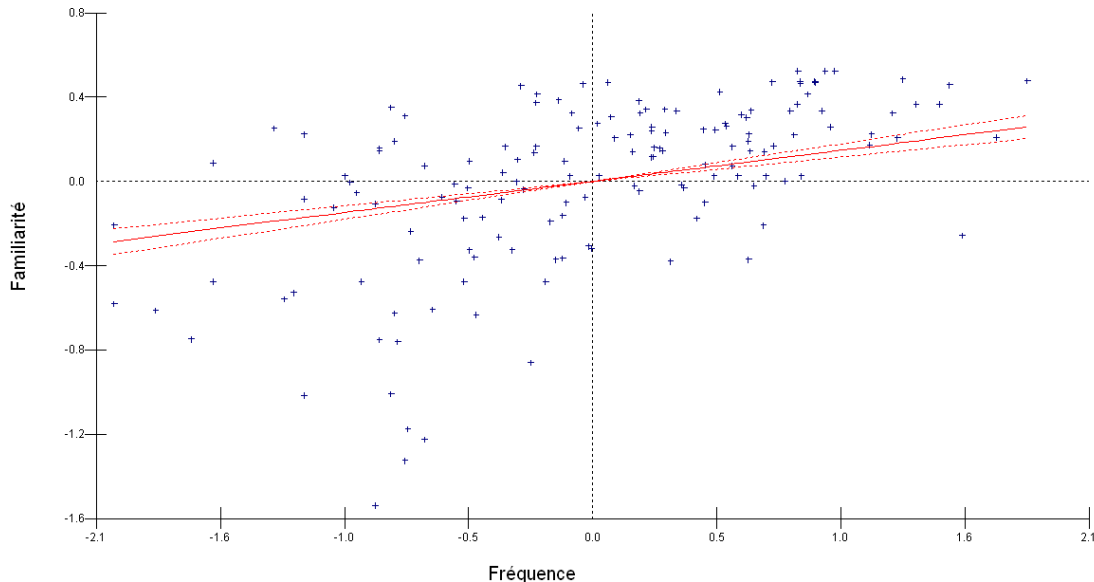
Tableau 127 : Résultat des tests de significativité

	Chi <sup>2</sup> <sub>(1, N = 151)</sub>	p <
<b>famc</b>	21.5	<b>.0001</b>
<b>âgc</b>	0	1
<b>âgc * LgFc</b>	6.3	<b>.013</b>
<b>famc * âgc</b>	8.9	<b>.003</b>
<b>LgFc * famc</b>	7.3	<b>.007</b>
<b>LgFc * âgc * famc</b>	25.2	<b>.0001</b>

Note.  $\text{famc}$  pour degré de familiarité (scores centrés sur la moyenne de la dimension) ;  $\text{LgFc}$  pour logarithme en base 10 de la fréquence d'occurrence du mot (valeurs centrées sur la moyenne de la dimension) ;  $\hat{\text{âgc}}$  pour âge d'acquisition estimé (scores centrés sur la moyenne de la dimension).

La fréquence d'occurrence dépend fortement de la familiarité (Figure 12), avec laquelle elle est positivement corrélée, et ce indépendamment de l'âge d'acquisition et des interactions que les trois facteurs entretiennent entre eux.

Figure 86. Relation entre la familiarité et la fréquence d'occurrence



Plus un mot est fréquent, plus le sentiment de familiarité avec ce mot est élevé (en moyenne, le logarithme de la fréquence augmente de 0.887 unité par unité de familiarité supplémentaire -  $\text{Chi}^2_{(1, N = 151)} = 21.5$  ;  $p < .0001$ ). La fréquence dépend également des 4 interactions modélisées (Tableau 5). Par contre, on retrouve l'absence de relation entre fréquence et âge d'acquisition observée dans le modèle précédent (Équation 2 et Figure 11 -  $\text{Chi}^2_{(1, N = 151)} = 0$  ;  $p = 1$ ).

#### Effet des prédicteurs de la familiarité sur cette dimension

#### Équation 19 : Équation du modèle étudiant l'influence des prédicteurs de la familiarité

$$\text{fam}_{n^\circ \text{ item}} \sim \mathbf{N}(\mathbf{XB}, \Omega)$$

$$\text{fam}_{n^\circ \text{ item}} = \beta_{0n^\circ \text{ item}} \text{Cons} + \mathbf{0.141(0.030)} \lg \text{Fc}_{n^\circ \text{ item}} + \\ \mathbf{-0.046(0.009)} \hat{\text{agc}}_{n^\circ \text{ item}} + \mathbf{-0.054(0.011)} \lg \text{Fc} \cdot \hat{\text{agc}}_{n^\circ \text{ item}} + \\ \mathbf{-0.388(0.087)} \lg \text{Fc} \cdot \text{fam}_{n^\circ \text{ item}} + \mathbf{0.100(0.025)} \hat{\text{agc}} \cdot \text{fam}_{n^\circ \text{ item}} + \\ \mathbf{-0.027(0.024)} \lg \text{Fc} \cdot \hat{\text{agc}} \cdot \text{fam}_{n^\circ \text{ item}}$$

$$\beta_{0n^\circ \text{ item}} = \mathbf{2.557(0.022)} + e_{0n^\circ \text{ item}}$$

$$[e_{0n^\circ \text{ item}}] \sim \mathbf{N(0, \Omega_e)} : \Omega_e = [\mathbf{0.052(0.006)}]$$

$$\mathbf{-2 * \log \text{likelihood (IGLS Deviance)} = -17.125(151 \text{ of } 151 \text{ cases in use)}$$

Note. fam et famc pour la Familiarité (c pour centré sur la moyenne) ; lgFc pour le logarithme en base 10 de la Fréquence d'occurrence ; âgc pour l'âge d'acquisition estimé. Se reporter à la note accompagnant l'Équation 1 pour le détail des autres éléments du modèle.

Tableau 128 : Résultat des tests de significativité

	<b>Chi<sup>2</sup><sub>(1, N = 151)</sub></b>	<b>p &lt;</b>
<b>âgc</b>	23.3	<b>.0001</b>
<b>LgFc</b>	21.5	<b>.0001</b>
<b>âgc * LgFc</b>	23.3	<b>.0001</b>
<b>famc * âgc</b>	16.4	<b>.0001</b>
<b>LgFc * famc</b>	19.6	<b>.0001</b>
<b>LgFc * âgc * famc</b>	1.3	.26

Note. famc pour degré de familiarité (scores centrés sur la moyenne de la dimension) ; LgFc pour logarithme en base 10 de la fréquence d'occurrence du mot (valeurs centrées sur la moyenne de la dimension) ; âgc pour âge d'acquisition estimé (scores centrés sur la moyenne de la dimension).

La familiarité dépend aussi bien de l'âge d'acquisition (Figure 10), avec lequel elle est fortement et négativement corrélée (en moyenne, le score de familiarité diminue de 0.046 par année de vie supplémentaire -  $\text{Chi}^2_{(1, N = 151)} = 23.3$  ;  $p < .0001$ ), que de la fréquence (Figure 12), avec laquelle elle est fortement et positivement corrélée (en moyenne, le score de familiarité augmente de 0.141 par unité logarithmique de fréquence supplémentaire -  $\text{Chi}^2_{(1, N = 151)} = 21.5$  ;  $p < .0001$ ), et ce indépendamment de l'âge d'acquisition et des interactions que les trois facteurs peuvent entretenir entre eux. Elle dépend par ailleurs des 3 interactions simples modélisées (Tableau 6).

### Conclusions

Cette série d'analyses statistiques permet de dresser un schéma de la relation entre âge d'acquisition estimé, familiarité et fréquence d'occurrence d'un mot. On est plus familier avec des mots appris tôt dans l'enfance qu'avec des mots appris tardivement (forte corrélation des deux dimensions), et avec des mots utilisés fréquemment qu'avec des mots utilisés rarement (forte corrélation des deux dimensions), mais on n'utilise pas plus ou moins fréquemment les mots en fonction de l'âge auquel ils ont été appris. La fréquence d'occurrence est cependant reliée à l'âge d'acquisition par l'intermédiaire de la familiarité, et par le biais d'interactions (détail dans la présentation des résultats).

Ces résultats suggèrent que lorsqu'un psycholinguiste contrôle la fréquence des mots pour une expérience, il contrôle aussi et sans le savoir une partie de l'influence de la familiarité, et probablement aussi, une partie de l'influence de l'âge d'acquisition. Toutefois, comme le facteur fréquence n'est que partiellement non orthogonal avec les deux autres dimensions, l'influence de l'âge d'acquisition et de la familiarité parasitent encore la situation expérimentale. Comme ces deux dimensions semblent correspondre à

de meilleurs prédicteurs des temps de traitement d'un mot que la fréquence, il y a intérêt à renseigner plus systématiquement ces dimensions psychologiques dans les modèles d'analyse.

### Conclusions de l'étude

Ce travail a permis de proposer une méthode d'enquête pour constituer assez rapidement une base de données lexicale volumineuse pour des facteurs psychologiques difficiles à recueillir. Il a également permis de connaître le motif corrélationnel entre âge d'acquisition estimé, familiarité et fréquence d'occurrence, et ainsi de justifier de l'importance de contrôler ces dimensions dans une étude sur la reconnaissance des mots parlés. Il a enfin permis de disposer des données pour effectuer ce contrôle plus poussé des résultats présentés.

## Annexe 2

### Phrases porteuses des items utilisés dans l'étude préalable (analyse acoustique, Simulations 1 et 2) et les Expériences 1 et 2 (Chapitre 1, Partie IV)

#### A/ Items pour les analyses acoustiques, Simulations 1 et 2, Expérience 1

##### Items G/D et F/D :

1. Nous avons cherché cette rarissime plante grasse à l'endroit où tu m'affirmais l'avoir observée. Hormis **les végétaux succulents** du désert, impossible de trouver un seul spécimen de ce précieux cactus.
2. Carla n'a pas apprécié le repas qu'on lui a servi ce midi. Hormis **le chocolat savoureux** des gâteaux du chef, rien n'était mangeable selon elle.
3. Je ne vois pas ce qu'il y a de novateur dans le type d'architecture que montrent tes dernières photos. Hormis **les bâtiments lumineux** du stadium, le reste ne présente aucune avancée en la matière.
4. Je suis très content du matériel expérimental que nous avons créé pour le prochain test. Hormis **les gabarits différents** des phrases lues, tout y est très homogène.
5. Tu peux me croire : il n'y a aucun bar accueillant qui vaille le détour dans les environs proches. Hormis **le cabaret chaleureux** des "jeunes loups", à 30 km d'ici, tout est triste et ennuyeux.
6. Tous les pays participants au tournoi international d'échecs ont montré beaucoup de professionnalisme dans l'organisation de ces jeux. Hormis **le comité négligent** du club russe, tous sont dignes d'éloges.
7. Je dois reconnaître qu'il n'y a pas grand chose de captivant dans cette petite bourgade. Hormis **le monument colonial** du village, rien ne vaut le détour.
8. Nous avons découvert des écrits rédigés par l'un des enfants royaux, relatant avec détail le déroulement des faits. Hormis **les documents minutieux** du prince, il y a aussi cette lettre de conspiration qui prouve l'assassinat du Roi.
9. Le personnel que nous avons recruté aux cuisines est vraiment trop laxiste. Hormis **le boulanger compétent** du matin, ce sont tous des incapables.

##### Items G/L et L/D :

1. Les intempéries de ces derniers temps nous retardent dans les travaux agricoles. Hormis **le champ labouré** des voisins, aucune parcelle n'est prête pour les semailles.
2. L'éruption volcanique de la nuit dernière a détruit tout le village. Hormis **le feu rougeoyant** de la lave, le spectacle était horripilant.
3. Ton dessert est délicieux, et me rappelle quelque chose. Hormis **le goût savoureux** des fruits, on sent un arôme que je n'arrive pas à définir.
4. Les photos de la NASA ont montré que le plus grand volcan du système solaire était sur Mars. Hormis **le mont gigantesque** de cette planète, deux autres volcans y atteignent 10 km d'altitude.
5. Lorsque l'hiver s'installe, ce jardin devient triste au possible. Hormis **la peau colorée** des poissons du bassin, tout est gris et sinistre.
6. Pour distinguer ce crustacée des espèces voisines, les zoologistes s'appuient sur la disparité anatomique gauche/droite qui existe à certains stades du

développement. Hormis **la queue symétrique** des larves, la forme des antennes est également un critère d'identification.

7. Depuis que la famine a frappé, c'est la désolation chez les paysans du village. Hormis **le pain nourricier** qu'on leur donne, ces pauvres gens n'ont plus rien à manger.
8. J'aime beaucoup ces montagnes et les randonnées que nous y faisons. Hormis **le temps menaçant** qui nous gâche un peu le plaisir, j'apprécie beaucoup mon séjour dans les Alpes.
9. El Niño a entraîné l'apparition de pluies régulières sur cette région autrefois désertique, provoquant une évolution progressive des paysages. Hormis **le lent ravinement** des collines, ces précipitations ont aussi causé des éboulements de terrain.

### Items F/L et L/G :

1. Cette vieille femme a un don miraculeux qui lui permet de soigner les gens. Hormis le **mercredi saint de** ce mois, on peut la voir tous les jours.
2. Le conseil municipal s'attaque à la réfection des façades de bâtiments classés. Hormis le **monument peint de** la ville, tous les édifices du centre vont être nettoyés.
3. Je devrais pouvoir faire concorder les pièces du système d'évacuation des eaux usées. Hormis le **gabarit fin du** tuyau, tout le reste s'ajuste parfaitement.
4. Les spécialistes reconnaissent le rôle du froid sur les adaptations morphologiques des plantes sibériennes. Hormis les **végétaux nains de** la steppe, les plantes alpines constituent un autre groupe d'espèces touchées par ce phénomène.
5. Ce qu'ils servent dans cet établissement n'est plus aussi bon qu'autrefois. Hormis le **chocolat chaud des** galettes, rien de tout ce qu'on a commandé n'était vraiment appétissant.
6. Les aménagements qui ont été réalisés pour la prochaine coupe du monde de football sont vraiment peu esthétiques. Hormis le **bâtiment rond du** stadium, tout ce qui a été construit est du plus mauvais goût.
7. J'ai assisté au concert des "Rockies". Hormis le **hurlement fou des** chanteurs, l'ambiance était assez décevante.
8. Nous sommes restés dans ce pub irlandais toute la soirée à boire bière sur bière. Hormis le **boulangier roux du** village, tout le monde est reparti sur ses deux pieds.
9. Sur cette photo, on était habillés en tenue sud-américaine traditionnelle. Hormis le **pantalons teints des** Incas, ce ne sont pas des vêtements que j'aimerais mettre tous les jours.

### Items F/G :

1. Les sociétés paysannes de cette région ont développé une structure politique originale qui assure la cohésion des groupes éparpillés. Hormis les **comités de** village, il y a aussi un guide spirituel qui joue un rôle éminent.
2. Ces personnes se sont fait la même opinion que toi. Hormis le **monument du** quartier, cette partie de la ville n'a rien d'attrayant.
3. Je ne sais pas quel autre matériel on aurait pu utiliser pour résoudre un tel problème d'assainissement. Hormis ce **gabarit de** tuyaux, rien ne pouvait faire l'affaire.
4. Je crois que j'ai eu une bonne intuition en achetant des habits locaux dès mon arrivée. Hormis les **pantalons du** marché, ces vêtements m'ont bien protégé du froid.
5. On vient de terminer l'aménagement paysager de la propriété. Hormis les **végétaux de** montagne, nous avons quelques plantes de jardin très communes.
6. Je sais parfaitement à quoi attribuer le succès de ce repas campagnard. Hormis le **chocolat des** gâteaux, tout était fait maison.

7. Mon impression sur les projets d'architecture proposés est assez mitigée. Hormis le **bâtiment du** futur musée, les autres conceptions semblent trop ambitieuses.
8. Un véritable scandale a été découvert autour des importations de bœuf britannique. Hormis les **documents du** gouvernement, d'autres dossiers secrets ont été trouvés au siège de la commission européenne.
9. Ma femme a sympathisé avec tous les commerçants. Hormis le **boulangier du** village, tous sont vraiment très gentils.

### Items G/G:

1. Ce sont les éleveurs de porc qui, les premiers, se sont organisés pour se faire entendre du gouvernement. Hormis le comité **de ces** agriculteurs, c'est l'ensemble des paysans qui s'est ensuite mobilisé.
2. La sonde qui devait atteindre Mars a une nouvelle fois échappé au contrôle de la NASA. Hormis l'échec **de la** mission, c'est tout le financement du futur programme spatial américain qui va être remis en cause.
3. Paul a été malade toute la nuit à cause des gâteaux au chocolat qu'il a mangés hier. Hormis le chocolat **de ces** gâteaux, je ne vois pas ce qui aurait pu lui donner une telle crise de foie.
4. Une fois encore, les pays européens ont été incapables de parler d'une même voix. Hormis le trio **de la** communauté, France, Allemagne, Grande Bretagne, de nombreux participants s'opposent toujours au projet.
5. Je n'ai pas vu une seule architecture qui ait un peu de caractère. Hormis le monument **de ce** quartier, les bâtiments de la ville sont inintéressants au possible.
6. On apprend qu'un cambriolage a eu lieu la nuit dernière à Londres. Hormis les bijoux **de la** couronne, des pièces d'une valeur inestimable ont été dérobées.
7. On nous envoie faire un reportage sur les pratiques magiques dans diverses régions d'Afrique. Hormis les rites **de ces** pays, nous devons aborder l'impact de la mondialisation sur leur culture.
8. Le chaman a dit que cette préparation médicinale était essentiellement composée de plantes himalayennes. Hormis les végétaux **de la** montagne, des ingrédients secrets sont aussi employés.
9. Il y a eu une réunion des élus locaux pour définir la marche à suivre auprès du gouvernement. Hormis le maire **de la** commune, tous sont tombés d'accord sur l'idée d'une pétition adressée au Premier ministre.

### Items L/L:

1. Avant, il y avait beaucoup de chats sauvages dans le quartier. Hormis le **chat roux** des voisins, tous se sont fait enlever par la SPA.
2. Cette fille est totalement tombée sous le charme. Hormis les **mots doux** qu'elle écrit, elle lui offre des cadeaux de grande valeur.
3. On est en train de préparer un repas léger. Hormis le **riz long** et les poissons, il n'y aura rien d'autre à manger.
4. J'adore les petits déjeuners à la campagne. Hormis le **pain chaud** du boulanger, le miel et la confiture sont un vrai régal.
5. L'an passé, la région a été inondée par un raz de marée dévastateur. Hormis les **champs nus** qu'il faudrait replanter, ce qui restait de forêt continue à dépérir.
6. La tenue qu'elle a choisie de porter pour ce safari se veut avant tout fonctionnelle. Hormis les **gants fins** de flanelle, son ensemble paraît effectivement adéquat.
7. Ton adversaire n'a aucun code d'honneur. Hormis les **coups bas** qu'il te porte, ses propos dépassent la mesure.
8. Cette attitude des félins est facile à comprendre, mais s'accompagne en général d'autres signes très parlants. Hormis le **dos rond** et le hérissément des poils, un grognement rauque peut par exemple avertir l'intrus qu'il doit quitter le territoire.



9. On vient d'aménager un jardin potager sur notre nouvelle propriété. Hormis les **choux nains** et les tomates de printemps, nous cultivons beaucoup de légumes rares et anciens.

## B/ Items pour le test perceptif de détection/localisation de frontières

### Items G<sub>1</sub>D<sub>2</sub>, mot trisyllabique, items F<sub>1</sub>D<sub>2</sub> et F<sub>2</sub>D<sub>3</sub> :

1. [**lesvégé**], [**végétaux**], [**gétauxsu**], [**tauxsuccu**], [**succulents**] : Nous avons cherché cette rarissime plante grasse à l'endroit où tu m'affirmais l'avoir observée. Hormis **les végétaux succulents** du désert, impossible de trouver un seul spécimen de ce précieux cactus.
2. [**lechoco**], [**chocolat**], [**colatsa**], [**latsavou**] : Carla n'a pas apprécié le repas qu'on lui a servi ce midi. Hormis **le chocolat savou**reux des gâteaux du chef, rien n'était mangeable selon elle.
3. [**lesbâti**], [**bâtiments**], [**timentslu**], [**mentslumi**], [**lumineux**] : Je ne vois pas ce qu'il y a de novateur dans le type d'architecture que montrent tes dernières photos. Hormis **les bâtiments lumineux** du stadium, le reste ne présente aucune avancée en la matière.
4. [**lesgaba**], [**gabarits**], [**baritsdi**], [**ritsdiffé**], [**différents**] : Je suis très content du matériel expérimental que nous avons créé pour le prochain test. Hormis **les gabarits différents** des phrases lues, tout y est très homogène.
5. [**lecaba**], [**cabaret**], [**barecha**], [**retchaleu**] : Tu peux me croire : il n'y a aucun bar accueillant qui vaille le détour dans les environs proches. Hormis **le cabaret chaleu**reux des "jeunes loups", à 30 km d'ici, tout est triste et ennuyeux.
6. [**lecomi**], [**comité**], [**miténé**], [**ténégli**], [**négligent**] : Tous les pays participants au tournoi international d'échecs ont montré beaucoup de professionnalisme dans l'organisation de ces jeux. Hormis **le comité négligent** du club russe, tous sont dignes d'éloges.
7. [**lemonu**], [**monument**], [**numentco**], [**mentcolo**], [**colonial**] : Je dois reconnaître qu'il n'y a pas grand chose de captivant dans cette petite bourgade. Hormis **le monument colonial** du village, rien ne vaut le détour.
8. [**lesdocu**], [**documents**], [**cumentsmi**], [**mentsminu**], [**minutieux**] : Nous avons découvert des écrits rédigés par l'un des enfants royaux, relatant avec détail le déroulement des faits. Hormis **les documents minutieux** du prince, il y a aussi cette lettre de conspiration qui prouve l'assassinat du Roi.
9. [**leboulan**], [**boulangier**], [**langercom**], [**gercompé**] : Le personnel que nous avons recruté aux cuisines est vraiment trop laxiste. Hormis **le boulangier compé**tent du matin, ce sont tous des incapables.

### Items L<sub>1</sub>D<sub>2</sub> et mot trisyllabique :

1. [**champlabou**], [**labouré**] : Les intempéries de ces derniers temps nous retardent dans les travaux agricoles. Hormis le **champ labouré** des voisins, aucune parcelle n'est prête pour les semailles.
2. [**feurougeoy**], [**rougeoyant**] : L'éruption volcanique de la nuit dernière a détruit tout le village. Hormis le **feu rougeoyant** de la lave, le spectacle était horripilant.
3. [**goûtsavou**], [**savoureux**] : Ton dessert est délicieux, et me rappelle quelque chose. Hormis le **goût savoureux** des fruits, on sent un arôme que je n'arrive pas à définir.
4. [**montgigan**], [**gigantesque**] : Les photos de la NASA ont montré que le plus grand volcan du système solaire était sur Mars. Hormis le **mont gigantesque** de cette planète, deux autres volcans y atteignent 10 km d'altitude.

5. **[peucolo], [colorée]** : Lorsque l'hiver s'installe, ce jardin devient triste au possible. Hormis la **peau colorée** des poissons du bassin, tout est gris et sinistre.
6. **[queuesymé], [symétrique]** : Pour distinguer ce crustacée des espèces voisines, les zoologistes s'appuient sur la disparité anatomique gauche/droite qui existe à certains stades du développement. Hormis la **queue symétrique** des larves, la forme des antennes est également un critère d'identification.
7. **[painnourri], [nourricier]** : Depuis que la famine a frappé, c'est la désolation chez les paysans du village. Hormis le **pain nourricier** qu'on leur donne, ces pauvres gens n'ont plus rien à manger.
8. **[tempsmena], [menaçant]** : J'aime beaucoup ces montagnes et les randonnées que nous y faisons. Hormis le **temps menaçant** qui nous gâche un peu le plaisir, j'apprécie beaucoup mon séjour dans les Alpes.
9. **[lentravine], [ravinement]** : El Niño a entraîné l'apparition de pluies régulières sur cette région autrefois désertique, provoquant une évolution progressive des paysages. Hormis le **lent ravinement** des collines, ces précipitations ont aussi causé des éboulements de terrain.

### Mot trisyllabique et item F<sub>2</sub>L<sub>3</sub> :

1. **[mercredi], [credisaint]** : Cette vieille femme a un don miraculeux qui lui permet de soigner les gens. Hormis le **mercredi saint** de ce mois, on peut la voir tous les jours.
2. **[numentpeint]** : Le conseil municipal s'attaque à la réfection des façades de bâtiments classés. Hormis le **monument peint** de la ville, tous les édifices du centre vont être nettoyés.
3. **[baritfin]** : Je devrais pouvoir faire concorder les pièces du système d'évacuation des eaux usées. Hormis le **gabarit fin** du tuyau, tout le reste s'ajuste parfaitement.
4. **[gétauxnains]** : Les spécialistes reconnaissent le rôle du froid sur les adaptations morphologiques des plantes sibériennes. Hormis les **végétaux nains** de la steppe, les plantes alpines constituent un autre groupe d'espèces touchées par ce phénomène.
5. **[colatchaud]** : Ce qu'ils servent dans cet établissement n'est plus aussi bon qu'autrefois. Hormis le **chocolat chaud** des galettes, rien de tout ce qu'on a commandé n'était vraiment appétissant.
6. **[timentrond]** : Les aménagements qui ont été réalisés pour la prochaine coupe du monde de football sont vraiment peu esthétiques. Hormis le **bâtiment rond** du stadium, tout ce qui a été construit est du plus mauvais goût.
7. **[hurlement], [lementfou]** : J'ai assisté au concert des "Rockies". Hormis le **hurlement fou** des chanteurs, l'ambiance était assez décevante.
8. **[langerroux]** : Nous sommes restés dans ce pub irlandais toute la soirée à boire bière sur bière. Hormis le **bou langer roux** du village, tout le monde est reparti sur ses deux pieds.
9. **[pantalon], [talonteint]** : Sur cette photo, on était habillés en tenue sud-américaine traditionnelle. Hormis le **pantalon teint** des Incas, ce ne sont pas des vêtements que j'aimerais mettre tous les jours.

### Items F<sub>2</sub>G<sub>3</sub> :

1. **[mitésde]** : Les sociétés paysannes de cette région ont développé une structure politique originale qui assure la cohésion des groupes éparpillés. Hormis les **comités de** village, il y a aussi un guide spirituel qui joue un rôle éminent.
2. **[numentdu]** : Ces personnes se sont fait la même opinion que toi. Hormis le **monument du** quartier, cette partie de la ville n'a rien d'attrayant.
3. **[baritde]** : Je ne sais pas quel autre matériel on aurait pu utiliser pour résoudre un tel problème d'assainissement. Hormis ce **gabarit de** tuyaux, rien ne pouvait faire l'affaire.

4. **[talonsdu]** : Je crois que j'ai eu une bonne intuition en achetant des habits locaux dès mon arrivée. Hormis les pantalons du marché, ces vêtements m'ont bien protégé du froid.
5. **[gétauxde]** : On vient de terminer l'aménagement paysager de la propriété. Hormis les végétaux de montagne, nous avons quelques plantes de jardin très communes.
6. **[colatdes]** : Je sais parfaitement à quoi attribuer le succès de ce repas campagnard. Hormis le chocolat des gâteaux, tout était fait maison.
7. **[timentdu]** : Mon impression sur les projets d'architecture proposés est assez mitigée. Hormis le bâtiment du futur musée, les autres conceptions semblent trop ambitieuses.
8. **[cumentsdu]** : Un véritable scandale a été découvert autour des importations de bœuf britannique. Hormis les documents du gouvernement, d'autres dossiers secrets ont été trouvés au siège de la commission européenne.
9. **[langerdu]** : Ma femme a sympathisé avec tous les commerçants. Hormis le boulanger du village, tous sont vraiment très gentils.

## Annexe 3

### Étude préalable : Analyse acoustique statique des syllabes

Équation 20 : Analyse de la durée

$$\text{Durée}_{\text{Syllabe, Item}} \sim N(XB, \Omega)$$

$$\begin{aligned} \text{Durée}_{\text{Syllabe, Item}} = & \beta_{0\text{Syllabe, Item}} \text{Cons} + -60.756(9.052)D_{\text{Syllabe, Item}} + \\ & -64.465(9.052)F_{\text{Syllabe, Item}} + -113.990(7.467)G_{\text{Syllabe, Item}} + \\ & 0.000(0.000)L_{\text{Syllabe, Item}} \end{aligned}$$

$$\beta_{0\text{Syllabe, Item}} = 196.413(5.487) + \mu_{0\text{Item}} + e_{0\text{Syllabe, Item}}$$

$$\left[ \mu_{0\text{Item}} \right] \sim N(0, \Omega_u) : \Omega_u = \left[ 303.245(176.997) \right]$$

$$\left[ e_{0\text{Syllabe, Item}} \right] \sim N(0, \Omega_e) : \Omega_e = \left[ 1260.766(198.165) \right]$$

$$-2 * \text{loglikelihood(IGLS Deviance)} = 1648.143(162 \text{ of } 162 \text{ cases in use})$$

Équation 21 : Analyse de la f0 moyenne

$$f0_{\text{Syllabe, Item}} \sim N(XB, \Omega)$$

$$\begin{aligned} f0_{\text{Syllabe, Item}} = & \beta_{0\text{Syllabe, Item}} \text{Cons} + 3.997(7.120)D_{\text{Syllabe, Item}} + \\ & -32.032(7.120)F_{\text{Syllabe, Item}} + -22.748(5.879)G_{\text{Syllabe, Item}} + \\ & 0.000(0.000)L_{\text{Syllabe, Item}} \end{aligned}$$

$$\beta_{0\text{Syllabe, Item}} = 270.735(4.343) + \mu_{0\text{Item}} + e_{0\text{Syllabe, Item}}$$

$$\left[ \mu_{0\text{Item}} \right] \sim N(0, \Omega_u) : \Omega_u = \left[ 210.286(111.245) \right]$$

$$\left[ e_{0\text{Syllabe, Item}} \right] \sim N(0, \Omega_e) : \Omega_e = \left[ 768.177(120.596) \right]$$

$$-2 * \text{loglikelihood(IGLS Deviance)} = 1571.435(162 \text{ of } 162 \text{ cases in use})$$

Avec :  $N(XB, \Omega)$  = présupposé sur la distribution : le vecteur réponse se distribue normalement (moyenne  $XB$  (partie fixe), variables aléatoires organisées selon matrice de covariance  $\Omega$  (partie aléatoire)) ;  $\text{Cons}$  = constante ;  $\beta_0$  = référence ou moyenne générale ;  $\mu_0$  = variance résiduelle au niveau des items (partie aléatoire) ;  $e_0$  = variance résiduelle au niveau des syllabes (partie aléatoire) ;  $-2 * \text{loglikelihood}$  = degré de correspondance du modèle aux données ;  $D, F, L, G$  les catégories des syllabes

## Annexe 4

Étude préalable : Analyse acoustique dynamique  
des syllabes par réseaux neuromimétiques  
(Simulation 1)Équation 22 : Analyse des données de classification correcte<sup>140</sup>

$$\text{rep}_{\text{syllabe, item}} \sim N(\chi B, \Omega)$$

$$\begin{aligned} \text{rep}_{\text{syllabe, item}} = & \beta_{0\text{syllabe, item}} \text{Cons} + 0.537(0.117)\text{FD.F1}_{\text{syllabe, item}} + 0.300(0.117)\text{FD.D2}_{\text{syllabe, item}} + \\ & 0.365(0.117)\text{FG.F1}_{\text{syllabe, item}} + 0.723(0.117)\text{FG.G2}_{\text{syllabe, item}} + 0.457(0.117)\text{FL.F1}_{\text{syllabe, item}} + \\ & 0.712(0.117)\text{FL.L2}_{\text{syllabe, item}} + 0.905(0.117)\text{GD.G1}_{\text{syllabe, item}} + 0.276(0.117)\text{GD.D2}_{\text{syllabe, item}} + \\ & 0.731(0.117)\text{GG.G1}_{\text{syllabe, item}} + 0.315(0.117)\text{GG.G2}_{\text{syllabe, item}} + 0.306(0.117)\text{GL.G1}_{\text{syllabe, item}} + \\ & 0.119(0.117)\text{GL.L2}_{\text{syllabe, item}} + 0.725(0.117)\text{LD.L1}_{\text{syllabe, item}} + 0.106(0.117)\text{LD.D2}_{\text{syllabe, item}} + \\ & 0.787(0.117)\text{LG.L1}_{\text{syllabe, item}} + 0.297(0.117)\text{LG.G2}_{\text{syllabe, item}} + 0.462(0.117)\text{LL.L1}_{\text{syllabe, item}} + \\ & 0.549(0.117)\text{LL.L2}_{\text{syllabe, item}} + -0.011(0.038)\text{FD}_{\text{item}} + -0.029(0.038)\text{FG}_{\text{item}} + \\ & -0.045(0.036)\text{FL}_{\text{item}} + -0.040(0.038)\text{GD}_{\text{item}} + -0.014(0.038)\text{GG}_{\text{item}} + 0.017(0.036)\text{GL}_{\text{item}} + \\ & -0.025(0.036)\text{LD}_{\text{item}} + -0.027(0.036)\text{LG}_{\text{item}} + 0.000(0.000)\text{LL}_{\text{item}} + 0.251(0.048)\text{F1}_{\text{syllabe, item}} + \\ & 0.104(0.048)\text{L1}_{\text{syllabe, item}} + 0.113(0.048)\text{G1}_{\text{syllabe, item}} + 0.245(0.048)\text{D2}_{\text{syllabe, item}} + \\ & 0.129(0.048)\text{L2}_{\text{syllabe, item}} + 0.169(0.048)\text{G2}_{\text{syllabe, item}} \end{aligned}$$

$$\beta_{0\text{syllabe, item}} = 0.019(0.032) + u_{0\text{item}} + e_{0\text{syllabe, item}}$$

$$[u_{0\text{item}}] \sim N(0, \Omega_u) : \Omega_u = [-0.009(0.001)]$$

$$[e_{0\text{syllabe, item}}] \sim N(0, \Omega_e) : \Omega_e = [0.090(0.005)]$$

-2\*loglikelihood(IGLS Deviance) = 126.258(648 of 648 cases in use)

<sup>140</sup> L'erreur standard ( $\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{N}}$ ) est homogène. Ce phénomène est lié à la production des données par réseau neuromimétique.

Équation 23 : Analyse des données de classification incorrecte<sup>140</sup> page 464

$$\text{rep}_{\text{syllabe, item}} \sim N(X\beta, \Omega)$$

$$\begin{aligned} \text{rep}_{\text{syllabe, item}} = & \beta_{0\text{syllabe, item}} \text{Cons} + 0.113(0.062)\text{FD}_{\text{item}} + 0.104(0.062)\text{FG}_{\text{item}} + 0.092(0.057)\text{FL}_{\text{item}} + \\ & 0.105(0.062)\text{GD}_{\text{item}} + 0.021(0.062)\text{GG}_{\text{item}} + -0.143(0.057)\text{GL}_{\text{item}} + 0.030(0.057)\text{LD}_{\text{item}} + \\ & 0.029(0.057)\text{LG}_{\text{item}} + 0.000(0.000)\text{LL}_{\text{item}} + 0.630(0.066)\text{F1}_{\text{syllabe, item}} + 0.784(0.066)\text{L1}_{\text{syllabe, item}} + \\ & 0.813(0.066)\text{G1}_{\text{syllabe, item}} + 0.423(0.066)\text{D2}_{\text{syllabe, item}} + 0.655(0.066)\text{L2}_{\text{syllabe, item}} + \\ & 0.598(0.066)\text{G2}_{\text{syllabe, item}} + -0.837(0.128)\text{FD.L1}_{\text{syllabe, item}} + -0.775(0.128)\text{FD.G1}_{\text{syllabe, item}} + \\ & -0.312(0.128)\text{FD.L2}_{\text{syllabe, item}} + -0.654(0.128)\text{FD.G2}_{\text{syllabe, item}} + -0.778(0.128)\text{FG.L1}_{\text{syllabe, item}} + \\ & -0.636(0.128)\text{FG.G1}_{\text{syllabe, item}} + -0.276(0.128)\text{FG.D2}_{\text{syllabe, item}} + -0.710(0.128)\text{FG.L2}_{\text{syllabe, item}} + \\ & -0.609(0.128)\text{FL.L1}_{\text{syllabe, item}} + -0.783(0.128)\text{FL.G1}_{\text{syllabe, item}} + -0.400(0.128)\text{FL.D2}_{\text{syllabe, item}} + \\ & -0.557(0.128)\text{FL.G2}_{\text{syllabe, item}} + -0.646(0.128)\text{GD.F1}_{\text{syllabe, item}} + -0.849(0.128)\text{GD.L1}_{\text{syllabe, item}} + \\ & -0.719(0.128)\text{GD.L2}_{\text{syllabe, item}} + -0.186(0.128)\text{GD.G2}_{\text{syllabe, item}} + -0.412(0.128)\text{GG.F1}_{\text{syllabe, item}} + \\ & -0.579(0.128)\text{GG.L1}_{\text{syllabe, item}} + -0.175(0.128)\text{GG.D2}_{\text{syllabe, item}} + -0.496(0.128)\text{GG.L2}_{\text{syllabe, item}} + \\ & -0.190(0.128)\text{GL.F1}_{\text{syllabe, item}} + -0.488(0.128)\text{GL.L1}_{\text{syllabe, item}} + 0.210(0.128)\text{GL.D2}_{\text{syllabe, item}} + \\ & -0.251(0.128)\text{GL.G2}_{\text{syllabe, item}} + -0.393(0.128)\text{LD.F1}_{\text{syllabe, item}} + -0.753(0.128)\text{LD.G1}_{\text{syllabe, item}} + \\ & -0.502(0.128)\text{LD.L2}_{\text{syllabe, item}} + -0.382(0.128)\text{LD.G2}_{\text{syllabe, item}} + -0.470(0.128)\text{LG.F1}_{\text{syllabe, item}} + \\ & -0.695(0.128)\text{LG.G1}_{\text{syllabe, item}} + -0.010(0.128)\text{LG.D2}_{\text{syllabe, item}} + -0.616(0.128)\text{LG.L2}_{\text{syllabe, item}} + \\ & 0.050(0.128)\text{LL.F1}_{\text{syllabe, item}} + -0.707(0.128)\text{LL.G1}_{\text{syllabe, item}} + -0.267(0.128)\text{LL.D2}_{\text{syllabe, item}} + \\ & -0.490(0.128)\text{LL.G2}_{\text{syllabe, item}} \end{aligned}$$

$$\beta_{0\text{syllabe, item}} = -0.039(0.045) + u_{0\text{item}} + e_{0\text{syllabe, item}}$$

$$[u_{0\text{item}}] \sim N(0, \Omega_u) : \Omega_u = [-0.009(0.001)]$$

$$[e_{0\text{syllabe, item}}] \sim N(0, \Omega_e) : \Omega_e = [0.085(0.005)]$$

-2\*loglikelihood(IGLS Deviance) = 97.185(648 of 648 cases in use)

Équation 24 : Analyse de la comparaison spécifique des taux d'activation correspondant aux confusions entre catégories attendues et reconnues (cas de confusions entre catégories)<sup>140 page 464</sup>

$$\text{rep}_{\text{syllabe, item}} \sim N(\mathbf{XB}, \Omega)$$

$$\begin{aligned} \text{rep}_{\text{syllabe, item}} = & \beta_{0\text{syllabe, item}} \text{Cons} + -0.049(0.044)\text{FD}_{\text{item}} + 0.053(0.034)\text{FG}_{\text{item}} + 0.047(0.034)\text{FL}_{\text{item}} + \\ & -0.026(0.045)\text{GD}_{\text{item}} + 0.062(0.034)\text{GG}_{\text{item}} + -0.041(0.044)\text{GL}_{\text{item}} + 0.006(0.045)\text{LD}_{\text{item}} + \\ & -0.006(0.045)\text{LG}_{\text{item}} + 0.000(0.000)\text{LL}_{\text{item}} + 0.371(0.053)\text{F1}_{\text{syllabe, item}} + 0.290(0.053)\text{L1}_{\text{syllabe, item}} + \\ & 0.329(0.051)\text{G1}_{\text{syllabe, item}} + 0.123(0.071)\text{D2}_{\text{syllabe, item}} + 0.248(0.056)\text{L2}_{\text{syllabe, item}} + \\ & 0.279(0.060)\text{G2}_{\text{syllabe, item}} + 0.223(0.143)\text{FD.L2}_{\text{syllabe, item}} + 0.229(0.145)\text{GD.G2}_{\text{syllabe, item}} + \\ & 0.374(0.150)\text{GL.D2}_{\text{syllabe, item}} + -0.074(0.145)\text{LD.G2}_{\text{syllabe, item}} + 0.292(0.150)\text{LG.D2}_{\text{syllabe, item}} + \\ & 0.275(0.142)\text{LL.F1}_{\text{syllabe, item}} + 0.485(0.150)\text{FD.D2}_{\text{syllabe, item}} + 0.409(0.150)\text{GD.D2}_{\text{syllabe, item}} + \\ & 0.082(0.143)\text{GL.L2}_{\text{syllabe, item}} + 0.221(0.150)\text{LD.D2}_{\text{syllabe, item}} + 0.190(0.145)\text{LG.G2}_{\text{syllabe, item}} + \\ & 0.300(0.142)\text{LL.L1}_{\text{syllabe, item}} \end{aligned}$$

$$\beta_{0\text{syllabe, item}} = -0.005(0.039) + u_{0\text{item}} + e_{0\text{syllabe, item}}$$

$$\begin{bmatrix} u_{0\text{item}} \end{bmatrix} \sim N(0, \Omega_u) : \Omega_u = \begin{bmatrix} -0.015(0.001) \end{bmatrix}$$

$$\begin{bmatrix} e_{0\text{syllabe, item}} \end{bmatrix} \sim N(0, \Omega_e) : \Omega_e = \begin{bmatrix} 0.138(0.008) \end{bmatrix}$$

$$-2 * \text{loglikelihood(IGLS Deviance)} = 371.872(648 \text{ of } 648 \text{ cases in use})$$

## Annexe 5

### Analyse complémentaire relative à l'excitabilité des unités de sortie dans la Simulation 1

Le terme « excitabilité » fait référence à la propension du réseau à fournir un type de réponse préférablement à un autre. Il s'agit donc d'évaluer un biais dans sa structure.

Analyse de l'excitabilité des unités de sortie liée au type de syllabe et à sa position dans l'item

Tableau 129 : Taux moyens d'activation des unités de sortie du réseau (excitabilité)

Syllabe	Taux d'activation pour les réponses correctes	Taux d'activation pour les réponses incorrectes
<b>F<sub>1</sub></b>	0.25	0.63
<b>L<sub>1</sub></b>	0.1	0.78
<b>G<sub>1</sub></b>	0.11	0.81
<b>D<sub>2</sub></b>	0.25	0.42
<b>L<sub>2</sub></b>	0.13	0.66
<b>G<sub>2</sub></b>	0.17	0.6

Tableau 130 : Test des contrastes (comparaison de l'excitabilité des unités de sortie selon le type de syllabe et sa position)

	Réponses correctes		Réponses incorrectes	
	p<	Chi <sup>2</sup> <sub>(1, N=648)</sub>	p<	Chi <sup>2</sup> <sub>(1, N=648)</sub>
<b>F<sub>1</sub>/L<sub>1</sub></b>	<b>.02</b>	5.9	<b>.09</b>	2.8
<b>F<sub>1</sub>/G<sub>1</sub></b>	<b>.02</b>	5.2	<b>.05</b>	4.0
<b>F<sub>1</sub>/D<sub>2</sub></b>	.92	0.0	<b>.02</b>	5.6
<b>F<sub>1</sub>/L<sub>2</sub></b>	<b>.04</b>	4.2	.78	0.1
<b>F<sub>1</sub>/G<sub>2</sub></b>	.17	1.9	.71	0.1
<b>L<sub>1</sub>/G<sub>1</sub></b>	.88	0.0	.75	0.1
<b>L<sub>1</sub>/D<sub>2</sub></b>	<b>.02</b>	5.6	<b>0</b>	16.9
<b>L<sub>1</sub>/L<sub>2</sub></b>	.68	0.2	.14	2.2
<b>L<sub>1</sub>/G<sub>2</sub></b>	.27	1.2	<b>.03</b>	4.5
<b>G<sub>1</sub>/D<sub>2</sub></b>	<b>.03</b>	4.9	<b>0</b>	19.7
<b>G<sub>1</sub>/L<sub>2</sub></b>	.80	0.1	<b>.07</b>	3.2
<b>G<sub>1</sub>/G<sub>2</sub></b>	.35	0.9	<b>.01</b>	6.0
<b>D<sub>2</sub>/L<sub>2</sub></b>	<b>.05</b>	3.7	<b>.01</b>	6.4
<b>D<sub>2</sub>/G<sub>2</sub></b>	.21	1.6	<b>.06</b>	3.7
<b>L<sub>2</sub>/G<sub>2</sub></b>	.50	0.5	.53	0.4

Les unités F<sub>1</sub> et D<sub>2</sub> présentent une réactivité différente, respectivement par rapport aux syllabes [G<sub>1</sub> ou L<sub>1</sub>] et aux syllabes [G<sub>2</sub> ou L<sub>2</sub>].



Les autres unités de sortie ( $L_1$ ,  $G_1$ ,  $L_2$ ,  $G_2$ ) présentent la même excitabilité moyenne (pour les réponses correctes et pour les réponses incorrectes).

Ce phénomène provient très probablement du nombre d'exemples, 2 fois moindre pour les syllabes  $F_1$  et  $D_2$ .

Pas de différence liée à la position pour les réponses correctes ( $F_1/D_2$  ns.,  $G_1/G_2$  ns.,  $L_1/L_2$  ns.).

Par contre, une asymétrie  $S_1/S_2$  (différence liée à la position) existe pour les confusions, pour  $F_1/D_2$  et  $G_1/G_2$ .

#### Analyse de l'excitabilité des unités de sortie liée au type de frontière

Tableau 131 : Taux moyens d'activation des unités de sortie (excitabilité) en fonction de la frontière (valeurs négatives : activation diminuée de x lorsque ce type d'item est présenté)

Frontière	Taux d'activation pour les réponses correctes	Taux d'activation pour les réponses incorrectes
<b>FD</b>	-0.011	0.11
<b>FG</b>	-0.029	0.1
<b>FL</b>	-0.045	0.09
<b>GD</b>	-0.04	0.11
<b>GG</b>	-0.014	0.02
<b>GL</b>	0.017	-0.14
<b>LD</b>	-0.025	0.03
<b>LG</b>	-0.027	0.03
<b>LL</b>	réf	réf

Tableau 132 : Test des contrastes (comparaison de l'excitabilité des unités de sortie selon le type de frontière)

Contraste	Réponses correctes		Réponses incorrectes	
	p<	Chi <sup>2</sup> <sub>(1, N=648)</sub>	p<	Chi <sup>2</sup> <sub>(1, N=648)</sub>
<b>FD/FG</b>	.62	0.2	.88	0.0
<b>FD/FL</b>	.35	0.9	.71	0.1
<b>FD/GD</b>	.42	0.7	.88	0.0
<b>FD/GG</b>	.93	0.0	.14	2.2
<b>FD/GL</b>	.46	0.6	<b>.0001</b>	17.2
<b>FD/LD</b>	.70	0.2	.15	2.1
<b>FD/LG</b>	.67	0.2	.17	1.9
<b>FD/LL</b>	.78	0.1	<b>.07</b>	3.4
<b>FG/FL</b>	.66	0.2	.83	0.1
<b>FG/GD</b>	.76	0.1	1	0.0
<b>FG/GG</b>	.69	0.2	.15	2.1
<b>FG/GL</b>	.22	1.5	<b>.0001</b>	16.0
<b>FG/LD</b>	.92	0.0	.23	1.5
<b>FG/LG</b>	.96	0.0	.19	1.7
<b>FG/LL</b>	.45	0.6	<b>.09</b>	2.9
<b>FL/GD</b>	.90	0.0	.84	0.0
<b>FL/GG</b>	.42	0.7	.25	1.3

Contraste	Réponses correctes		Réponses incorrectes	
	p<	Chi <sup>2</sup> <sub>(1, N=648)</sub>	p<	Chi <sup>2</sup> <sub>(1, N=648)</sub>
<b>FL/GL</b>	.09	2.9	<b>.0001</b>	16.8
<b>FL/LD</b>	.60	0.7	.31	1.0
<b>FL/LG</b>	.64	0.2	.31	1.0
<b>FL/LL</b>	.22	1.5	.11	2.6
<b>GD/GG</b>	.48	0.5	.15	2.1
<b>GD/GL</b>	.11	2.5	<b>.0001</b>	18.6
<b>GD/LD</b>	.68	0.2	.19	1.7
<b>GD/LG</b>	.73	0.1	.22	1.5
<b>GD/LL</b>	.29	1.1	.09	2.9
<b>GG/GL</b>	.39	0.7	<b>.005</b>	8.2
<b>GG/LD</b>	.78	0.1	.89	0.0
<b>GG/LG</b>	.72	0.1	.89	0.0
<b>GG/LL</b>	.71	0.1	.73	0.1
<b>GL/LD</b>	.26	1.3	<b>.006</b>	7.8
<b>GL/LG</b>	.24	1.4	<b>.006</b>	7.7
<b>GL/LL</b>	.63	0.2	<b>.013</b>	6.2
<b>LD/LG</b>	.95	0.0	<b>.0001</b>	1.0
<b>LD/LL</b>	.50	0.5	.61	0.3
<b>LG/LL</b>	.46	0.6	.61	0.3

L'excitabilité des unités de sortie n'est pas modulée en fonction du type de frontière, sauf pour la frontière GL : pour les réponses correctes, une activité différente a lieu dans le réseau (valeur positive), par rapport aux autres catégories de frontières (valeurs inférieures ou égales à 0). Pour une réponse incorrecte, GL diminue l'activation de 0.14 lorsque les autres types de frontières l'augmentent.

Le réseau serait donc légèrement biaisé. Néanmoins, il s'avère à même d'effectuer la classification des syllabes.

## Annexe 6

### Étude préalable : Analyse dynamique de la différence entre items bisyllabiques par des réseaux neuromimétiques (Simulation 2)

Équation 25 : Effet de la quantité de syllabes similaires à une même position dans les 2 items comparés, sur la performance de discrimination des items par les réseaux

$$Y_{rés_{j-Mesur, i-It}} \sim N(XB, \Omega)$$

$$Y_{rés_{j-Mesur, i-It}} = \beta_{0j-Mesur, i-It} \text{Cons} + 0.032(0.010)1SYL_{i-It} + 0.566(0.012)2SYLS_{i-It}$$

$$\beta_{0j-Mesur, i-It} = 0.030(0.007) + \mu_{0i-It} + e_{0j-Mesur, i-It}$$

$$\left[ \mu_{0i-It} \right] \sim N(0, \Omega_u) : \Omega_u = \left[ 0.032(0.001) \right]$$

$$\left[ e_{0j-Mesur, i-It} \right] \sim N(0, \Omega_e) : \Omega_e = \left[ 0.019(0.000) \right]$$

$$-2 * \loglikelihood(IGLS Deviance) = -33481.060(34992 \text{ of } 34992 \text{ cases in use})$$

Avec :  $Y_{rés}$  = le degré d'activation de l'unité de sortie ;  $0syl$  = aucune syllabe en commun ;  $1syl$  = 1 syllabe en commun sur la première ou sur la deuxième syllabe ;  $2syl$  = 2 syllabes en commun

Équation 26 : Effet du type de syllabe en commun à une même position dans les 2 items comparés, sur la performance de discrimination des items par les réseaux<sup>141</sup>

$$Y_{rés_{j-Mesur, i-It}} \sim N(XB, \Omega)$$

$$Y_{rés_{j-Mesur, i-It}} = \beta_{0j-Mesur, i-It} \text{Cons} + 0.011(0.017)F1_{i-It} + 0.089(0.017)G1_{i-It} + 0.033(0.017)L1_{i-It} + \\ -0.014(0.017)D2_{i-It} + -0.007(0.017)G2_{i-It} + 0.078(0.017)L2_{i-It} + \\ 0.590(0.036)F1.D2_{i-It} + 0.841(0.036)F1.G2_{i-It} + 0.264(0.036)F1.L2_{i-It} + \\ 0.699(0.036)G1.D2_{i-It} + 0.319(0.036)G1.G2_{i-It} + 0.217(0.036)G1.L2_{i-It} + \\ 0.532(0.036)L1.D2_{i-It} + 0.579(0.036)L1.G2_{i-It} + 0.485(0.036)L1.L2_{i-It}$$

$$\beta_{0j-Mesur, i-It} = 0.030(0.006) + \mu_{0i-It} + e_{0j-Mesur, i-It}$$

$$\left[ \mu_{0i-It} \right] \sim N(0, \Omega_u) : \Omega_u = \left[ 0.027(0.001) \right]$$

$$\left[ e_{0j-Mesur, i-It} \right] \sim N(0, \Omega_e) : \Omega_e = \left[ 0.019(0.000) \right]$$

$$-2 * \loglikelihood(IGLS Deviance) = -33783.140(34992 \text{ of } 34992 \text{ cases in use})$$

<sup>141</sup> Erreur standard homogène : lié à la production des données par réseau neuromimétique.

Avec : [F, D, L, G] = le type de syllabe en commun ; [1, 2] = la position de la similitude catégorielle dans les 2 items comparés

Équation 27 : Effet du type d’item 1 et 2 comparés, sur la performance de discrimination des items par les réseaux<sup>141 page 464, 142</sup>

$$Y_{rés_{k-Mesur, j-It}} \sim N(XB, \Omega)$$

$$\begin{aligned} Y_{rés_{k-Mesur, j-It}} = & \beta_{0k-Mesur, j-It} \text{Cons} + 0.007(0.887)FD1.FG2_{j-It} + 0.108(0.887)FD1.FL2_{j-It} + 0.005(0.887)FD1.GD2_{j-It} + \\ & 0.023(0.887)FD1.GG2_{j-It} + 0.082(0.887)FD1.GL2_{j-It} + 0.046(0.887)FD1.LD2_{j-It} + 0.011(0.887)FD1.LG2_{j-It} + \\ & 0.014(0.887)FD1.LL2_{j-It} + 0.013(0.887)FG1.FL2_{j-It} + 0.095(0.887)FG1.GD2_{j-It} + 0.049(0.887)FG1.GG2_{j-It} + \\ & 0.029(0.887)FG1.GL2_{j-It} + 0.006(0.887)FG1.LD2_{j-It} + 0.008(0.887)FG1.LG2_{j-It} + 0.005(0.887)FG1.LL2_{j-It} + \\ & 0.005(0.887)FL1.GD2_{j-It} + 0.022(0.887)FL1.GG2_{j-It} + 0.145(0.887)FL1.GL2_{j-It} + 0.078(0.887)FL1.LD2_{j-It} + \\ & 0.037(0.887)FL1.LG2_{j-It} + 0.151(0.887)FL1.LL2_{j-It} + 0.187(0.887)GD1.GG2_{j-It} + 0.029(0.887)GD1.GL2_{j-It} + \\ & 0.004(0.887)GD1.LD2_{j-It} + 0.003(0.887)GD1.LG2_{j-It} + 0.007(0.887)GD1.LL2_{j-It} + 0.148(0.887)GG1.GL2_{j-It} + \\ & 0.023(0.887)GG1.LD2_{j-It} + 0.018(0.887)GG1.LG2_{j-It} + 0.012(0.887)GG1.LL2_{j-It} + 0.089(0.887)GL1.LD2_{j-It} + \\ & 0.038(0.887)GL1.LG2_{j-It} + 0.033(0.887)GL1.LL2_{j-It} + 0.079(0.887)LD1.LG2_{j-It} + 0.030(0.887)LD1.LL2_{j-It} + \\ & 0.086(0.887)LG1.LL2_{j-It} + 0.619(0.887)FD1.FD2_{j-It} + 0.877(0.887)FG1.FG2_{j-It} + 0.384(0.887)FL1.FL2_{j-It} + \\ & 0.807(0.887)GD1.GD2_{j-It} + 0.433(0.887)GG1.GG2_{j-It} + 0.416(0.887)GL1.GL2_{j-It} + 0.583(0.887)LD1.LD2_{j-It} + \\ & 0.637(0.887)LG1.LG2_{j-It} + 0.627(0.887)LL1.LL2_{j-It} \end{aligned}$$

$$\beta_{0k-Mesur, j-It} = -0.002(0.887) + u_{0j-It} + e_{0k-Mesur, j-It}$$

$$\begin{bmatrix} u_{0j-It} \end{bmatrix} \sim N(0, \Omega_u) : \Omega_u = \begin{bmatrix} 0.025(0.001) \end{bmatrix}$$

$$\begin{bmatrix} e_{0k-Mesur, j-It} \end{bmatrix} \sim N(0, \Omega_e) : \Omega_e = \begin{bmatrix} 0.019(0.000) \end{bmatrix}$$

$$-2 * \text{loglikelihood(IGLS Deviance)} = -33849.890(34992 \text{ of } 34992 \text{ cases in use})$$

<sup>142</sup> Remarque : l’erreur standard présentée dans l’équation pour chaque coefficient est visiblement aberrante (en plus d’être homogène pour tous les coefficients estimés, elle est supérieure à la valeur d’activation moyenne). Ce problème est probablement lié à la manière dont les données analysées sont générées (mathématiquement par un réseau). Par conséquent, cette information n’a pas été exploitée dans les représentations graphiques correspondantes.

## Annexe 7

### Expérience 1 : Analyse des données comportementales

Équation 28 : Effet de la quantité de syllabes similaires à une même position dans les 2 items comparés, sur la performance de discrimination des items par les auditeurs (taux de réponses « similaire »)

$$\left. \begin{aligned} \text{réponse}_{N^{\circ}\text{Compar}, n^{\circ}\text{ sujet}} &\sim \text{Binomial}(\text{denom}_{N^{\circ}\text{Compar}, n^{\circ}\text{ sujet}}, \pi_{N^{\circ}\text{Compar}, n^{\circ}\text{ sujet}}) \\ \text{réponse}_{N^{\circ}\text{Compar}, n^{\circ}\text{ sujet}} &= \pi_{N^{\circ}\text{Compar}, n^{\circ}\text{ sujet}} + e_{1N^{\circ}\text{Compar}, n^{\circ}\text{ sujet}} \text{bcons}^* \end{aligned} \right\}$$

$$\text{logit}(\pi_{N^{\circ}\text{Compar}, n^{\circ}\text{ sujet}}) = \beta_{0n^{\circ}\text{ sujet}} \text{cons} + 0.486(0.043)1 \text{ syl}_{N^{\circ}\text{Compar}, n^{\circ}\text{ sujet}} + 1.257(0.054)2 \text{ syls}_{N^{\circ}\text{Compar}, n^{\circ}\text{ sujet}}$$

$$\beta_{0n^{\circ}\text{ sujet}} = -0.558(0.085) + u_{0n^{\circ}\text{ sujet}}$$

$$\left[ u_{0n^{\circ}\text{ sujet}} \right] \sim N(0, \Omega_u) : \Omega_u = \left[ 0.265(0.061) \right]$$

$$\text{bcons}^* = \text{bcons} \left[ \pi_{N^{\circ}\text{Compar}, n^{\circ}\text{ sujet}} (1 - \pi_{N^{\circ}\text{Compar}, n^{\circ}\text{ sujet}}) / \text{denom}_{N^{\circ}\text{Compar}, n^{\circ}\text{ sujet}} \right]^{0.5}$$

$$\left[ e_{1N^{\circ}\text{Compar}, n^{\circ}\text{ sujet}} \right] \sim (0, \Omega_e) : \Omega_e = \left[ 1.000(0.000) \right]$$

Avec : 1 syl pour 1 syllabe de même catégorie théorique dans les deux items, 2 syls pour les deux syllabes identiques.

Calcul du pourcentage de réponses Oui à partir des valeurs fournies par le modèle :

$1/(\text{invln}(-(\text{référence}+x))+1)$ , avec référence =  $\beta_0$  et x = valeur fournie pour le facteur.

Équation 29 : Effet du type de similitude à une position donnée dans les 2 items comparés, sur la performance de discrimination des items par les auditeurs (taux de réponses « similaire »)

$$\begin{aligned} \text{réponse}_{ij} &\sim \text{Binomial}(\text{denom}_{ij}, \pi_{ij}) \\ \text{logit}(\pi_{ij}) &= \beta_{0j} \text{cons} + 0.211(0.077)F1_{ij} + 0.936(0.078)G1_{ij} + 0.242(0.077)L1_{ij} + \\ &\quad 0.404(0.077)G2_{ij} + 0.479(0.076)L2_{ij} + 0.641(0.077)D2_{ij} + \\ &\quad 1.672(0.199)F1.G2_{ij} + -0.099(0.161)F1.L2_{ij} + 0.764(0.174)F1.D2_{ij} + \\ &\quad -0.035(0.167)G1.G2_{ij} + -0.504(0.162)G1.L2_{ij} + 0.741(0.201)G1.D2_{ij} + \\ &\quad -0.073(0.161)L1.G2_{ij} + 0.390(0.164)L1.L2_{ij} + 0.315(0.165)L1.D2_{ij} \\ \beta_{0j} &= -0.558(0.086) + u_{0j} \\ [u_{0j}] &\sim N(0, \Omega_u) : \Omega_u = [0.277(0.064)] \end{aligned}$$

$$\text{var}(\text{réponse}_{ij} | \pi_{ij}) = 0.935(0.012)\pi_{ij}(1 - \pi_{ij})/\text{denom}_{ij}$$

Avec F, G, L et D les catégories de syllabes, le chiffre 1 ou 2 pour la position dans l’item, le premier [catégorie][chiffre] pour le bisyllabe /mama/1, le deuxième pour le bisyllabe /mama/2.

Équation 30 : Effet du type de frontière sur la performance de discrimination des items par les auditeurs (taux de réponses « similaire »)

$$\begin{aligned} \left. \begin{aligned} \text{réponse}_{N^{\text{e}} \text{Compar}, n^{\text{e}} \text{Sujet}} &\sim \text{Binomial}(\text{denom}_{N^{\text{e}} \text{Compar}, n^{\text{e}} \text{Sujet}}, \pi_{N^{\text{e}} \text{Compar}, n^{\text{e}} \text{Sujet}}) \\ \text{réponse}_{N^{\text{e}} \text{Compar}, n^{\text{e}} \text{Sujet}} &= \pi_{N^{\text{e}} \text{Compar}, n^{\text{e}} \text{Sujet}} + e_{1N^{\text{e}} \text{Compar}, n^{\text{e}} \text{Sujet}} \text{bcons}^* \end{aligned} \right\} \\ \text{logit}(\pi_{N^{\text{e}} \text{Compar}, n^{\text{e}} \text{Sujet}}) &= \beta_{0n^{\text{e}} \text{Sujet}} \text{cons} + 2.494(0.214)FG1.FG2_{N^{\text{e}} \text{Compar}, n^{\text{e}} \text{Sujet}} + 0.212(0.182)FG1.LD2_{N^{\text{e}} \text{Compar}, n^{\text{e}} \text{Sujet}} + \\ &\quad 0.363(0.180)FG1.LG2_{N^{\text{e}} \text{Compar}, n^{\text{e}} \text{Sujet}} + -0.271(0.190)FG1.LL2_{N^{\text{e}} \text{Compar}, n^{\text{e}} \text{Sujet}} + 0.797(0.178)FL1.FL2_{N^{\text{e}} \text{Compar}, n^{\text{e}} \text{Sujet}} + \\ &\quad 0.638(0.179)FL1.LD2_{N^{\text{e}} \text{Compar}, n^{\text{e}} \text{Sujet}} + -0.631(0.201)FL1.LG2_{N^{\text{e}} \text{Compar}, n^{\text{e}} \text{Sujet}} + 1.102(0.180)FL1.LL2_{N^{\text{e}} \text{Compar}, n^{\text{e}} \text{Sujet}} + \\ &\quad 2.525(0.216)GD1.GD2_{N^{\text{e}} \text{Compar}, n^{\text{e}} \text{Sujet}} + 1.761(0.189)GD1.GG2_{N^{\text{e}} \text{Compar}, n^{\text{e}} \text{Sujet}} + 1.297(0.182)GD1.GL2_{N^{\text{e}} \text{Compar}, n^{\text{e}} \text{Sujet}} + \\ &\quad 0.437(0.180)GD1.LD2_{N^{\text{e}} \text{Compar}, n^{\text{e}} \text{Sujet}} + -0.151(0.188)GD1.LG2_{N^{\text{e}} \text{Compar}, n^{\text{e}} \text{Sujet}} + -0.046(0.186)GD1.LL2_{N^{\text{e}} \text{Compar}, n^{\text{e}} \text{Sujet}} + \\ &\quad 1.513(0.184)GG1.GG2_{N^{\text{e}} \text{Compar}, n^{\text{e}} \text{Sujet}} + 0.148(0.183)GG1.LD2_{N^{\text{e}} \text{Compar}, n^{\text{e}} \text{Sujet}} + -0.131(0.187)GG1.LG2_{N^{\text{e}} \text{Compar}, n^{\text{e}} \text{Sujet}} + \\ &\quad -0.487(0.196)GG1.LL2_{N^{\text{e}} \text{Compar}, n^{\text{e}} \text{Sujet}} + 1.118(0.180)GL1.GL2_{N^{\text{e}} \text{Compar}, n^{\text{e}} \text{Sujet}} + 0.813(0.179)GL1.LD2_{N^{\text{e}} \text{Compar}, n^{\text{e}} \text{Sujet}} + \\ &\quad 0.542(0.179)GL1.LL2_{N^{\text{e}} \text{Compar}, n^{\text{e}} \text{Sujet}} + 1.406(0.183)LD1.LD2_{N^{\text{e}} \text{Compar}, n^{\text{e}} \text{Sujet}} + 0.518(0.179)LD1.LG2_{N^{\text{e}} \text{Compar}, n^{\text{e}} \text{Sujet}} + \\ &\quad 0.805(0.179)LD1.LL2_{N^{\text{e}} \text{Compar}, n^{\text{e}} \text{Sujet}} + 0.781(0.178)LG1.LG2_{N^{\text{e}} \text{Compar}, n^{\text{e}} \text{Sujet}} + 1.319(0.182)LL1.LL2_{N^{\text{e}} \text{Compar}, n^{\text{e}} \text{Sujet}} + \\ &\quad 1.822(0.191)FD1.FD2_{N^{\text{e}} \text{Compar}, n^{\text{e}} \text{Sujet}} + 0.558(0.179)FD1.FG2_{N^{\text{e}} \text{Compar}, n^{\text{e}} \text{Sujet}} + 0.925(0.179)FD1.FL2_{N^{\text{e}} \text{Compar}, n^{\text{e}} \text{Sujet}} + \\ &\quad 1.423(0.183)FD1.LD2_{N^{\text{e}} \text{Compar}, n^{\text{e}} \text{Sujet}} + -0.056(0.186)FD1.LG2_{N^{\text{e}} \text{Compar}, n^{\text{e}} \text{Sujet}} + 0.845(0.179)FD1.LL2_{N^{\text{e}} \text{Compar}, n^{\text{e}} \text{Sujet}} + \\ &\quad -0.312(0.191)FL1.FG2_{N^{\text{e}} \text{Compar}, n^{\text{e}} \text{Sujet}} + 0.702(0.179)GD1.FD2_{N^{\text{e}} \text{Compar}, n^{\text{e}} \text{Sujet}} + 1.477(0.184)GD1.FG2_{N^{\text{e}} \text{Compar}, n^{\text{e}} \text{Sujet}} + \\ &\quad -0.350(0.192)GD1.FL2_{N^{\text{e}} \text{Compar}, n^{\text{e}} \text{Sujet}} + -0.093(0.186)GG1.FD2_{N^{\text{e}} \text{Compar}, n^{\text{e}} \text{Sujet}} + 1.586(0.186)GG1.FG2_{N^{\text{e}} \text{Compar}, n^{\text{e}} \text{Sujet}} + \\ &\quad -0.582(0.199)GG1.FL2_{N^{\text{e}} \text{Compar}, n^{\text{e}} \text{Sujet}} + 0.861(0.179)GL1.FD2_{N^{\text{e}} \text{Compar}, n^{\text{e}} \text{Sujet}} + 0.734(0.178)GL1.FG2_{N^{\text{e}} \text{Compar}, n^{\text{e}} \text{Sujet}} + \\ &\quad 0.413(0.180)GL1.FL2_{N^{\text{e}} \text{Compar}, n^{\text{e}} \text{Sujet}} + 0.445(0.180)GL1.GG2_{N^{\text{e}} \text{Compar}, n^{\text{e}} \text{Sujet}} + 0.000(0.185)LL1.LG2_{N^{\text{e}} \text{Compar}, n^{\text{e}} \text{Sujet}} + \\ &\quad 0.000(0.000)GL1.LG2_{N^{\text{e}} \text{Compar}, n^{\text{e}} \text{Sujet}} \\ \beta_{0n^{\text{e}} \text{Sujet}} &= -0.765(0.156) + u_{0n^{\text{e}} \text{Sujet}} \\ [u_{0n^{\text{e}} \text{Sujet}}] &\sim N(0, \Omega_u) : \Omega_u = [0.309(0.071)] \end{aligned}$$

$$\text{bcons}^* = \text{bcons}[\pi_{N^{\text{e}} \text{Compar}, n^{\text{e}} \text{Sujet}}(1 - \pi_{N^{\text{e}} \text{Compar}, n^{\text{e}} \text{Sujet}})/\text{denom}_{N^{\text{e}} \text{Compar}, n^{\text{e}} \text{Sujet}}]^{0.5}$$

$$[e_{1N^{\text{e}} \text{Compar}, n^{\text{e}} \text{Sujet}}] \sim (0, \Omega_e) : \Omega_e = [0.932(0.012)]$$

## Annexe 8

### Expérience 2 : Analyse des données comportementales

Équation 31 : Analyse générale de la répartition des catégories de réponse obtenues relativement à la situation présentée (1 modèle par catégorie de réponse : P<sub>0</sub>, P<sub>1</sub>, P<sub>2</sub>)

$$\left. \begin{aligned} \text{RépP0}_{item, sujet} &\sim \text{Binomial}(\text{denom}_{item, sujet}, \pi_{item, sujet}) \\ \text{RépP0}_{item, sujet} &= \pi_{item, sujet} + e_{1item, sujet} \text{bcons1}^* \end{aligned} \right\}$$

$$\text{logit}(\pi_{item, sujet}) = \beta_{0sujet} \text{cons1} + 0.512(0.079)P0 \text{ ds } it_{item, sujet} +$$

$$0.265(0.080)P1 \text{ ds } it_{item, sujet} +$$

$$0.000(0.000)P2 \text{ ds } it_{item, sujet}$$

$$\beta_{0sujet} = -0.634(0.076) + u_{0sujet}$$

$$[u_{0sujet}] \sim N(0, \Omega_u) : \Omega_u = [0.123(0.035)]$$

$$\text{bcons1}^* = \text{bcons1} [\pi_{item, sujet} (1 - \pi_{item, sujet}) / \text{denom}_{item, sujet}]^{0.5}$$

$$[e_{1item, sujet}] \sim (0, \Omega_e) : \Omega_e = [1.000(0.000)]$$

$$\left. \begin{aligned} \text{RépP1}_{item, sujet} &\sim \text{Binomial}(\text{denom}_{item, sujet}, \pi_{item, sujet}) \\ \text{RépP1}_{item, sujet} &= \pi_{item, sujet} + e_{1item, sujet} \text{bcons1}^* \end{aligned} \right\}$$

$$\text{logit}(\pi_{item, sujet}) = \beta_{0sujet} \text{cons1} + 0.586(0.090)P0 \text{ ds } it_{item, sujet} +$$

$$1.229(0.087)P1 \text{ ds } it_{item, sujet} +$$

$$0.000(0.000)P2 \text{ ds } it_{item, sujet}$$

$$\beta_{0sujet} = -1.377(0.073) + u_{0sujet}$$

$$[u_{0sujet}] \sim N(0, \Omega_u) : \Omega_u = [0.033(0.018)]$$

$$\text{bcons1}^* = \text{bcons1} [\pi_{item, sujet} (1 - \pi_{item, sujet}) / \text{denom}_{item, sujet}]^{0.5}$$

$$[e_{1item, sujet}] \sim (0, \Omega_e) : \Omega_e = [1.000(0.000)]$$

$$\left. \begin{aligned} \text{RépP2}_{item, sujet} &\sim \text{Binomial}(\text{denom}_{item, sujet}, \pi_{item, sujet}) \\ \text{RépP2}_{item, sujet} &= \pi_{item, sujet} + e_{1item, sujet} \text{bcons1}^* \end{aligned} \right\}$$

$$\text{logit}(\pi_{item, sujet}) = \beta_{0sujet} \text{cons1} + -1.081(0.086)P0 \text{ ds } it_{item, sujet} +$$

$$-1.724(0.098)P1 \text{ ds } it_{item, sujet} +$$

$$0.000(0.000)P2 \text{ ds } it_{item, sujet}$$

$$\beta_{0sujet} = -0.193(0.070) + u_{0sujet}$$

$$[u_{0sujet}] \sim N(0, \Omega_u) : \Omega_u = [0.093(0.033)]$$

$$\text{bcons1}^* = \text{bcons1} [\pi_{item, sujet} (1 - \pi_{item, sujet}) / \text{denom}_{item, sujet}]^{0.5}$$

$$[e_{1item, sujet}] \sim (0, \Omega_e) : \Omega_e = [1.000(0.000)]$$



Tableau 133 : Grilles de calcul des moyennes à partir des résultats du modèle

Contrôle	Frontière					
	RP <sub>1</sub>			RP <sub>2</sub>		
RP <sub>0</sub>	DF	L	G	DF	L	G
mot=Σ :	F <sub>1</sub> D <sub>2</sub> =Σ :	L <sub>1</sub> D <sub>2</sub> =Σ :	G <sub>1</sub> D <sub>2</sub> =Σ :	F <sub>2</sub> D <sub>3</sub> =Σ :	F <sub>2</sub> L <sub>3</sub> =Σ :	F <sub>2</sub> G <sub>3</sub> =Σ :
Moyenne	Moyenne	Moyenne	Moyenne	Moyenne	Moyenne	Moyenne
-	Frontier	Frontier	Frontier	Frontier	Frontier	Frontier
-	-	L	G	Ps <sub>2</sub> S <sub>3</sub>	Ps <sub>2</sub> S <sub>3</sub>	Ps <sub>2</sub> S <sub>3</sub>
-	-	-	-	-	L	G
-	-	-	-	-	Ps <sub>2</sub> S <sub>3</sub> .L	Ps <sub>2</sub> S <sub>3</sub> .G

Équation 32 : Calcul du pourcentage de réponses à partir des valeurs fournies par le modèle

$$1/(\text{invln}(-(\beta_x + \Sigma(\text{coefficients des facteurs}))) + 1)$$

avec  $\text{invln}$  l'inverse du logarithme népérien,  $\beta_x$  la référence.

Équation 33 : Analyse détaillée de la répartition des réponses en fonction du type de frontière testée

$$\left. \begin{aligned} MV\_RESP_{1jk} &\sim \text{Binomial}(\text{denom}_{1jk}, \pi_{1jk}) \\ MV\_RESP_{1jk} &= \pi_{1jk} + u_{ijk}RP0.bcons1_{ijk} \\ MV\_RESP_{2jk} &\sim \text{Binomial}(\text{denom}_{2jk}, \pi_{2jk}) \\ MV\_RESP_{2jk} &= \pi_{2jk} + u_{ijk}RP1.bcons1_{ijk} \\ MV\_RESP_{3jk} &\sim \text{Binomial}(\text{denom}_{3jk}, \pi_{3jk}) \\ MV\_RESP_{3jk} &= \pi_{3jk} + u_{ijk}RP2.bcons1_{ijk} \end{aligned} \right\}$$

$$\begin{aligned} \text{logit}(\pi_{1jk}) &= \beta_{3k}RP0.cons1_{ijk} - 1.043(0.123)RP0.frontier_{ijk} + 0.176(0.153)RP0.Ps2s3_{ijk} - 0.025(0.157)RP0.L_{ijk} + 2.293(0.156)RP0.G_{ijk} + \\ &\quad - 1.043(0.241)RP0.Ps2s3.L_{ijk} + -0.661(0.213)RP0.Ps2s3.G_{ijk} \\ \beta_{3k} &= -0.122(0.080) + v_{3k} \\ \text{logit}(\pi_{2jk}) &= \beta_{4k}RP1.cons1_{ijk} + 1.455(0.116)RP1.frontier_{ijk} - 1.995(0.153)RP1.Ps2s3_{ijk} - 0.202(0.139)RP1.L_{ijk} - 2.699(0.178)RP1.G_{ijk} + \\ &\quad 0.105(0.217)RP1.Ps2s3.L_{ijk} + 2.658(0.243)RP1.Ps2s3.G_{ijk} \\ \beta_{4k} &= -0.791(0.067) + v_{4k} \\ \text{logit}(\pi_{3jk}) &= \beta_{5k}RP2.cons1_{ijk} + 0.000(0.000)RP2.frontier_{ijk} + 0.000(0.000)RP2.Ps2s3_{ijk} + 0.000(0.000)RP2.L_{ijk} + 0.000(0.000)RP2.G_{ijk} + \\ &\quad 0.000(0.000)RP2.Ps2s3.L_{ijk} + 0.000(0.000)RP2.Ps2s3.G_{ijk} \\ \beta_{5k} &= 0.000(0.000) + v_{5k} \end{aligned}$$

$$\begin{bmatrix} v_{3k} \\ v_{4k} \\ v_{5k} \end{bmatrix} \sim N(0, \Omega_v) : \Omega_v = \begin{bmatrix} 0.173(0.047) & & \\ -0.075(0.028) & 0.049(0.023) & \\ 0.000(0.000) & 0.000(0.000) & 0.000(0.000) \end{bmatrix}$$

$$\begin{aligned} RP0.bcons1_{ijk} &= RP0.bcons1_{ijk}[\pi_{4jk}(1 - \pi_{4jk})/\text{denom}_{4jk}]^{0.5} \\ RP1.bcons1_{ijk} &= RP1.bcons1_{ijk}[\pi_{4jk}(1 - \pi_{4jk})/\text{denom}_{4jk}]^{0.5} \\ RP2.bcons1_{ijk} &= RP2.bcons1_{ijk}[\pi_{4jk}(1 - \pi_{4jk})/\text{denom}_{4jk}]^{0.5} \end{aligned}$$

$$\begin{bmatrix} u_{ijk} \\ u_{ijk} \\ u_{ijk} \end{bmatrix} \sim (0, \Omega_u) : \Omega_u = \begin{bmatrix} 1.000(0.000) & & \\ -0.581(0.009) & 1.000(0.000) & \\ 0.000(0.000) & 0.000(0.000) & 0.000(0.000) \end{bmatrix}$$

Avec : RP<sub>[012]</sub> : type de réponse ; [DFLG] : type de syllabe ; Frontier : présence d'une frontière dans l'item (réf = P<sub>0</sub>) ; Ps<sub>2</sub>S<sub>3</sub> : position de la frontière entre les 2 dernières syllabes de l'item (réf = Ps<sub>1</sub>S<sub>2</sub>)

$$\begin{aligned}
 & \left. \begin{aligned}
 MV\_RESP_{1jk} &\sim \text{Binomial}(\text{denom}_{1jk}, \pi_{1jk}) \\
 MV\_RESP_{1jk} &= \pi_{1jk} + u_{0jk} RP0.bcons1_{1jk}^*
 \end{aligned} \right\} \\
 & \left. \begin{aligned}
 MV\_RESP_{2jk} &\sim \text{Binomial}(\text{denom}_{2jk}, \pi_{2jk}) \\
 MV\_RESP_{2jk} &= \pi_{2jk} + u_{1jk} RP2.bcons1_{1jk}^*
 \end{aligned} \right\} \\
 & \left. \begin{aligned}
 MV\_RESP_{3jk} &\sim \text{Binomial}(\text{denom}_{3jk}, \pi_{3jk}) \\
 MV\_RESP_{3jk} &= \pi_{3jk} + u_{2jk} RP1.bcons1_{1jk}^*
 \end{aligned} \right\} \\
 \text{logit}(\pi_{1jk}) &= \beta_{3k} RP0.consl_{1jk} - 1.043(0.123) RP0.frontier_{1jk} + 0.176(0.153) RP0.Ps2s3_{1jk} - 0.025(0.157) RP0.L_{1jk} + 2.293(0.156) RP0.G_{1jk} + \\
 & \quad -1.043(0.241) RP0.Ps2s3.L_{1jk} + -0.661(0.213) RP0.Ps2s3.G_{1jk} \\
 \beta_{3k} &= -0.122(0.080) + v_{3k} \\
 \text{logit}(\pi_{2jk}) &= \beta_{4k} RP2.consl_{1jk} - 0.898(0.169) RP2.frontier_{1jk} + 2.253(0.182) RP2.Ps2s3_{1jk} + 0.464(0.203) RP2.L_{1jk} + 0.262(0.210) RP2.G_{1jk} + \\
 & \quad 0.272(0.246) RP2.Ps2s3.L_{1jk} + -2.139(0.267) RP2.Ps2s3.G_{1jk} \\
 \beta_{4k} &= -1.274(0.081) + v_{4k} \\
 \text{logit}(\pi_{3jk}) &= \beta_{5k} RP1.consl_{1jk} + 0.000(0.000) RP1.frontier_{1jk} + 0.000(0.000) RP1.Ps2s3_{1jk} + 0.000(0.000) RP1.L_{1jk} + 0.000(0.000) RP1.G_{1jk} + \\
 & \quad 0.000(0.000) RP1.Ps2s3.L_{1jk} + 0.000(0.000) RP1.Ps2s3.G_{1jk} \\
 \beta_{5k} &= 0.000(0.000) + v_{5k}
 \end{aligned}$$

$$\begin{bmatrix} v_{3k} \\ v_{4k} \\ v_{5k} \end{bmatrix} \sim N(0, \Omega_v) : \Omega_v = \begin{bmatrix} 0.170(0.046) & & \\ -0.132(0.038) & 0.113(0.038) & \\ 0.000(0.000) & 0.000(0.000) & 0.000(0.000) \end{bmatrix}$$

$$\begin{aligned}
 RP0.bcons1_{1jk}^* &= RP0.bcons1_{1jk} [\pi_{4jk}(1 - \pi_{4jk}) / \text{denom}_{4jk}]^{0.5} \\
 RP2.bcons1_{1jk}^* &= RP2.bcons1_{1jk} [\pi_{4jk}(1 - \pi_{4jk}) / \text{denom}_{4jk}]^{0.5} \\
 RP1.bcons1_{1jk}^* &= RP1.bcons1_{1jk} [\pi_{4jk}(1 - \pi_{4jk}) / \text{denom}_{4jk}]^{0.5}
 \end{aligned}$$

$$\begin{bmatrix} u_{0jk} \\ u_{1jk} \\ u_{2jk} \end{bmatrix} \sim (0, \Omega_u) : \Omega_u = \begin{bmatrix} 1.000(0.000) & & \\ -0.476(0.011) & 1.000(0.000) & \\ 0.000(0.000) & 0.000(0.000) & 0.000(0.000) \end{bmatrix}$$

$$\left. \begin{aligned}
 MV\_RESP_{1jk} &\sim \text{Binomial}(\text{denom}_{1jk}, \pi_{1jk}) \\
 MV\_RESP_{1jk} &= \pi_{1jk} + u_{1jk} RP1.bcons1_{1jk} \\
 MV\_RESP_{2jk} &\sim \text{Binomial}(\text{denom}_{2jk}, \pi_{2jk}) \\
 MV\_RESP_{2jk} &= \pi_{2jk} + u_{2jk} RP2.bcons1_{2jk} \\
 MV\_RESP_{3jk} &\sim \text{Binomial}(\text{denom}_{3jk}, \pi_{3jk}) \\
 MV\_RESP_{3jk} &= \pi_{3jk} + u_{3jk} RP0.bcons1_{3jk}
 \end{aligned} \right\}$$

$$\text{logit}(\pi_{1jk}) = \beta_{3k} RP1.consl_{1jk} + 1.455(0.116) RP1.frontier_{1jk} + -1.995(0.153) RP1.Ps2s3_{1jk} + -0.202(0.139) RP1.L_{1jk} + -2.699(0.178) RP1.G_{1jk} + 0.105(0.217) RP1.Ps2s3.L_{1jk} + 2.658(0.243) RP1.Ps2s3.G_{1jk}$$

$$\beta_{3k} = -0.791(0.068) + v_{3k}$$

$$\text{logit}(\pi_{2jk}) = \beta_{4k} RP2.consl_{1jk} + -0.898(0.169) RP2.frontier_{1jk} + 2.253(0.182) RP2.Ps2s3_{1jk} + 0.464(0.203) RP2.L_{1jk} + 0.262(0.210) RP2.G_{1jk} + 0.272(0.246) RP2.Ps2s3.L_{1jk} + -2.139(0.267) RP2.Ps2s3.G_{1jk}$$

$$\beta_{4k} = -1.274(0.082) + v_{4k}$$

$$\text{logit}(\pi_{3jk}) = \beta_{5k} RP0.consl_{1jk} + 0.000(0.000) RP0.frontier_{1jk} + 0.000(0.000) RP0.Ps2s3_{1jk} + 0.000(0.000) RP0.L_{1jk} + 0.000(0.000) RP0.G_{1jk} + 0.000(0.000) RP0.Ps2s3.L_{1jk} + 0.000(0.000) RP0.Ps2s3.G_{1jk}$$

$$\beta_{5k} = 0.000(0.000) + v_{5k}$$

$$\begin{bmatrix} v_{3k} \\ v_{4k} \\ v_{5k} \end{bmatrix} \sim N(0, \Omega_v) : \Omega_v = \begin{bmatrix} 0.056(0.024) & & \\ 0.050(0.022) & 0.122(0.040) & \\ 0.000(0.000) & 0.000(0.000) & 0.000(0.000) \end{bmatrix}$$

$$RP1.bcons1_{1jk} = RP1.bcons1_{1jk} [\pi_{4jk}(1 - \pi_{4jk}) / \text{denom}_{4jk}]^{0.5}$$

$$RP2.bcons1_{1jk} = RP2.bcons1_{1jk} [\pi_{4jk}(1 - \pi_{4jk}) / \text{denom}_{4jk}]^{0.5}$$

$$RP0.bcons1_{1jk} = RP0.bcons1_{1jk} [\pi_{4jk}(1 - \pi_{4jk}) / \text{denom}_{4jk}]^{0.5}$$

$$\begin{bmatrix} u_{1jk} \\ u_{2jk} \\ u_{3jk} \end{bmatrix} \sim (0, \Omega_u) : \Omega_u = \begin{bmatrix} 1.000(0.000) & & \\ -0.430(0.012) & 1.000(0.000) & \\ 0.000(0.000) & 0.000(0.000) & 0.000(0.000) \end{bmatrix}$$

## Annexe 9

### Expérience 3 (4 & 9) : Matériel linguistique

Tableau 134 : Mots expérimentaux

nI		I	
nE	E	nE	E
dʁē	ʁātʃ	kε	plas
sœj	fεv	cle	pʁεs
ʒqē	gεn	gʁē	ʃaʁ
gʁwē	gʁāʒ	fʁō	kas
fwē	ɲjes	tʁō	bes
plqi	ʒāʁ	kʁwa	fʁ
djø	ʒut	dʁa	gaʁ
pʁwa	suʃ	lu	pas
ljē	ʁiʁ	pō	bak
swē	ʁol	gu	fas
vjø	soʒ	fly	pys
dœj	ʃez	lē	buk
ʒwε	dob	ʒy	pis
bwe	sɥit	tʁu	laʁ
bqε	bjεʁ	flo	bat

Avec : I = mot enchâssable en position initiale (superposition lexicale virtuelle - nI pour non enchâssable) ; E = mot renfermant un mot enchâssé initial (superposition lexicale réelle - nE = sans mot enchâssé)

Tableau 135 : Mots et Pseudo-mots distracteurs (monosyllabiques et bisyllabiques)

Pseudo-mots monosyllabiques		
I	nI	illégal
gʁō	zyʁ	zya
kʁa	vʁal	dʁεf
fʁo	ksa	mjy
sty	glø	dʒi
spa	flō	tʁwē
pʁy	vʁø	kʁwi
kli	tja	bjy
bla	zjo	dʒø
fla	zaʁ	dzy
gli	zuʁ	vʁy
pʁa	ʒεk	bʒa
gʁø	dqi	ʒli
tʁy	gyʁ	tso
tʁa	ʁjo	fty
kla	ʁja	mti
bʁā	tʁø	jyt

<b>Pseudo-mots monosyllabiques</b>		
<b>I</b>	<b>nI</b>	<b>illégal</b>
glo	fək	fɛyl
glu	vɛo	gjo
fle	zis	ʒjy
fɛa	kly	lju

Avec : Illégal = syllabe illégale en Français selon BRULEX.

<b>Mots bisyllabiques</b>			
baɛda	gɛɔna	ɛɔfyɜ	skoɔjoz
bebe	gɛɔnad	ɛɔkɛ	soda
bedwɛ	gudɛɔ	ɛɔlik	sodjɔm
blasfɛm	kalsjɔm	ɛɔljɛf	ʃɔnal
blazɔ	kanin	pɛɛdɛi	ʃɔvij
bɛasɛɛ	klavje	pɛalin	spɛɛsjat
bɛɔve	kɛavat	poɔjɔm	stɛykyɛ
flotɛɛ	kutyɛ	pɛolɔg	stydjo
fɔɛsɛps	mazut	pɛɛtɛɛ	suɛsje
fukyɛ	mufɔ	potɛɜ	tibja
fymje	mwatje	pɛavɛɔ	tijɛl
gamel	nave	ʃɛtje	timys
godas	nevɛoz	ʃɛɛnje	tɛɛfik
gɛɛfɛɜ	noze	sɛɛvjɛt	vɛtuz
gonad	ɛɔɜɛɛ	skɔlet	zymel
<b>Pseudo-mots bisyllabiques</b>			
baɛvjɛt	gove	ɛɔfɛm	skonal
bebja	goze	ɛɔjɛl	sodɛɔ
bedjo	gunin	ɛɔvat	ʃɔdwɛ
bladas	kaltje	ɛɔzjɛt	ʃɔlet
blasbe	katɛɛ	pɛɛtɛɛ	sozut
bɛatuz	klanad	pɛɛfyɜ	spɛɛzɔ
bɛɛfik	kɛalik	pɛɛmel	stɛyknɛ
flodɛi	kusjɔm	ponjɛɛ	styvje
fɛɛvɛɔ	maflɔ	pɛokɛ	suɛda
futje	mutɛɜ	posje	tinad
fynɛoz	mwaɜɛɛ	pɛamɛl	tive
gamje	naljɛf	ʃɛɛjɔm	tivij
gɛɛsɛɛ	nɛkyɛ	ʃɛɛda	tɛɛljoz
gɛɛlin	noɔg	sɛɛɛjɔm	vɛsɛps
gɛɛtyɛ	ɛɔtyɛ	skɔmys	zyfɛɜ

## Annexe 10

### Expérience 3 : Équation du modèle et tests de significativité de l'effet des covariables sur les temps de réponse

Équation 34 : Équation du modèle étudié

$$TR_{It, \text{Suj}} \sim N(XB, \Omega)$$

$$TR_{It, \text{Suj}} = \beta_{0It, \text{Suj}} \text{cons} + 28.969(10.478)M_{It, \text{Suj}} + 25.433(10.677)R_{It, \text{Suj}} + 13.799(9.209)H_{It, \text{Suj}} + -57.590(13.441)M.R_{It, \text{Suj}} + -2.480(13.108)M.H_{It, \text{Suj}} + -32.574(13.259)R.H_{It, \text{Suj}} + 15.581(18.691)M.R.H_{It, \text{Suj}} + 0.490(0.031)PRD_{It, \text{Suj}}$$

$$\beta_{0It, \text{Suj}} = 807.361(12.336) + u_{0\text{Suj}} + e_{0It, \text{Suj}}$$

$$[u_{0\text{Suj}}] \sim N(0, \Omega_u) : \Omega_u = [4042.230(892.555)]$$

$$[e_{0It, \text{Suj}}] \sim N(0, \Omega_e) : \Omega_e = [17579.160(440.343)]$$

$$-2 * \log\text{likelihood (IGLS Deviance)} = 40918.260(3234 \text{ of } 4224 \text{ cases in use})$$

Avec : TR = temps de réponse (V.D., en ms.) ; M = lexicalité de l'item (réf = pseudo-mot) ; R = débit d'élocution rapide (réf = L) ; H = hauteur de voix haute (réf = B) ; PRD = position acoustique du point de reconnaissance pour les mots et de divergence pour les pseudo-mots (en ms.)

Tableau 136 : Tests de significativité des covariables et de leurs interactions avec les facteurs expérimentaux

	Coefficient	Chi <sup>2</sup> <sub>(1, N=3234)</sub>	p<
M	28.969	7.644	<b>.006</b>
M.R	-57.59	18.36	<b>2<sup>E-05</sup></b>
M.H	-2.48	0.036	.85
M.R.H	15.581	0.695	.41
PRD	0.49	244.649	<b>0</b>

## Annexe 11

# Analyses complémentaires évaluant la relation entre performance des participants et caractéristiques acoustiques des stimuli (Expérience 1)

### Analyses statistiques complémentaires par régression linéaire

Le but de ces analyses est :

- ▷ de connaître la/les dimension(s) physique(s) qui a/ont déterminé les réponses des auditeurs dans le test perceptif, en estimant l'importance de chaque régresseur testé à travers 3 analyses :
  - ∩ analyse sur l'ensemble des réponses « similaire »
  - ∩ analyse des réponses correctes pour les comparaisons intra catégorie
  - ∩ analyse des réponses correctes pour les comparaisons inter catégories
- ▷ d'estimer la concordance des réponses avec les caractéristiques physiques des items
- ▷ d'évaluer les catégories de frontière qui présentent une non concordance afin de discuter l'origine de ces distorsions perceptives.

Plus généralement, il s'agit donc d'estimer le degré de justification que trouve la performance des participants (V.D.) dans le signal acoustique (V.I.), et de discuter les écarts observés entre réalité physique et performance obtenue : cela peut-il avoir un sens psychologique ?

Nous procédons à l'aide d'analyses de régression linéaire simples et de modèles de régression linéaire multiples.

### Aspects techniques

V.D. = réponse correcte (« similaire » pour les comparaisons intra catégorie ; « différent » pour les comparaisons inter catégories - moyenne sur 7 mesures par combinaison d'items).

V.I. = cinq régresseurs répartis en deux groupes (mélodie et temps), décrivant des caractéristiques physiques des items. Les régresseurs sont calculés pour chaque combinaison d'items. Ils représentent différentes façons possibles, et probables d'un point de vue logique, d'exploiter les dimensions mélodique et temporelle dans un contraste.

## Présentation des régresseurs

Groupe mélodie : 3 V.I.

▷  $\Delta f_0$

modulation moyenne inter syllabique inter items du fondamental de voix. Cet indice permet d'évaluer grossièrement l'importance des contrastes mélodiques inter syllabiques moyens entre 2 items. Il se calcule ainsi :

$\Delta f_0 = \text{ABS}[(f_{0\text{moy}}\text{Syl}_1\text{I}_1 - f_{0\text{moy}}\text{Syl}_2\text{I}_1) - (f_{0\text{moy}}\text{Syl}_1\text{I}_2 - f_{0\text{moy}}\text{Syl}_2\text{I}_2)]$ , avec  $f_{0\text{moy}}$  la fréquence fondamentale moyenne par syllabe (Hz).

▷  $\text{dtw}^{143}$

écart entre la forme de 2 courbes mélodiques. Calcul de l'indice : P. Courriou (CNRS et Université de Provence, France).

▷ « composite » :

force des contrastes mélodiques, exprimée par l'écart inter items des amplitudes mélodiques par unité de temps. Cet indice évalue l'intensité de la modulation mélodique en normalisant celle-ci par unité de temps. Il renseigne sur la perceptibilité d'une modulation de  $f_0$  (une variation forte de  $f_0$ , mais étalée sur une durée longue peut être plus difficile à remarquer qu'une variation identique sur une durée brève). Cet indice se calcule ainsi :

« composite » =  $\text{ABS}\left\{\frac{(f_{0\text{moy}}\text{Syl}_1\text{I}_1 - f_{0\text{moy}}\text{Syl}_2\text{I}_1)}{(\text{DuréeSyl}_1\text{I}_1 + \text{DuréeSyl}_2\text{I}_1)} - \frac{(f_{0\text{moy}}\text{Syl}_1\text{I}_2 - f_{0\text{moy}}\text{Syl}_2\text{I}_2)}{(\text{DuréeSyl}_1\text{I}_2 + \text{DuréeSyl}_2\text{I}_2)}\right\}$ .

Groupe temps : 2 V.I.

▷  $\Delta\text{durée}$

Il s'agit de l'écart de durée entre 2 items d'une combinaison. Cet indice évalue grossièrement un contraste de durée entre items. Il est incapable de renseigner sur les aspects rythmiques (des items ayant des syllabes de durées différentes peuvent avoir une durée identique). Il se calcule ainsi :

$\Delta\text{durée} = \text{ABS}[\text{DuréeI}_1 - \text{DuréeI}_2]$  ; ABS valeur absolue, durée en ms.

▷ « Rythme »

Cet indice, calculé sur la base d'une addition des contrastes syllabiques inter items est utilisé pour combler la lacune de l'indice  $\Delta\text{durée}$ . Il évalue la différence rythmique entre items, et se calcule ainsi :

$\text{Rythme} = [\text{ABS}[\text{DuréeSyl}_1\text{I}_1 - \text{DuréeSyl}_1\text{I}_2] + \text{ABS}[\text{DuréeSyl}_2\text{I}_1 - \text{DuréeSyl}_2\text{I}_2]]$ , avec I item, Syl syllabe, Durée en ms).

## Méthode d'analyse

Les régresseurs choisis présentent une forte colinéarité entre eux au sein d'un même groupe (mélodie, temps). Ils ne peuvent donc pas être testés ensembles dans un même

---

<sup>143</sup> « Dynamic time warping ». Méthode de comparaison, dite de « programmation dynamique », utilisée en reconnaissance automatique de la parole pour évaluer l'écart entre un mot et son prototype. Correspond mathématiquement à la différence entre deux vecteurs uni ou multidimensionnels. L'algorithme optimise la correspondance des vecteurs en évacuant la dimension temporelle. Il permet d'obtenir un indice de distance (sans dimension) mesurant les écarts de forme, ici entre courbes de  $f_0$  des items comparés (0 pour 2 courbes identiques, valeur proche 0 pour 2 courbes originellement identiques dont une est étirée temporellement, de même que pour 2 courbes homothétiques. Valeurs différentes de 0, d'autant plus élevées que l'écart entre courbes est important). La description mathématique de l'algorithme peut être trouvée dans (Calliope, 1989)



modèle de régression. Pour cette raison, des analyses de régression linéaire simple sont effectuées. Dans certains cas, où deux V.I. non colinéaires ont une influence conjointe remarquable, une analyse de régression linéaire multiple est effectuée.

Pour évaluer les seuils de significativité, la technique du « bootstrap » a été employée ((Davidson & MacKinnon, 1993), pour pallier à la trop faible quantité de données (N = 36 pour les comparaisons par catégorie). Évaluation de 10000 t (minimum théorique +1).

### Résultats obtenus

#### Analyse générale sur l'ensemble des réponses « similaire »

Tableau 137 : Analyse sur l'ensemble des réponses « similaire » (comparaisons intra et inter catégories)<sup>144</sup>

	$\Delta f0$	Dtw	Composite	$\Delta$ durée	Rythme	$\Delta f0 + \text{rythme}$
<b>R</b>	0.25	0.02	0.21	0.5	0.53	0.57
<b>R<sup>2</sup> ajusté</b>	0.06	0.0002	0.04	0.25	0.28	0.33
<b>F<sub>(1,1618)</sub></b>	104	0.4	74.7	540.9	643.4	F <sub>(2,1617)</sub> 399.3
<b>p&lt;</b>	<b>.0001</b>	.54	<b>.0001</b>	<b>.0001</b>	<b>.0001</b>	<b>.0001</b>

Ce sont essentiellement les aspects rythmiques qui ont influencé la perception des auditeurs (28% de la variation des données expliquée), puis la différence de f0 moyenne (6% de variance expliquée).

Les régresseurs testés ne permettent pas, en moyenne, d'expliquer plus de 33% de la variance (modèle  $\Delta f0 + \text{rythme}$ ).

#### Analyse des relations régresseurs/réponses correctes pour les comparaisons intra catégorie

Pour les effets significatifs, les indices renseignés dans le tableau ci-dessous sont dans cet ordre : p, R, R<sup>2</sup> ajusté, F<sub>(1, 34)</sub> (pour  $\Delta f0$ , Dtw, Composite,  $\Delta$ durée, Rythme) et F<sub>(2,33)</sub> (pour Multiple). Pour les effets non significatifs, seul le p est indiqué. Les résultats les plus explicatifs des données pour une comparaison de frontière (évaluation basée à partir du coefficient de détermination) sont à chaque fois surlignés (orange : régression simple ; bleu : modèle de régression multiple).

<sup>144</sup> Coefficient de détermination : R<sup>2</sup>ajusté = 1-[(SC résidus/dl)/(SC totale/dl)].

Tableau 138 : Résultats de l'analyse des réponses correctes pour les comparaisons intra catégorie

	$\Delta f0$	Dtw	Composite	$\Delta$ durée	Rythme	Multiple
<b>FD/FD (0.74)</b>	<.37	<.15	<.33	<b>&lt;.032</b> , 0.4, 0.11, 5.4	<b>&lt;.0002</b> , 0.51, 0.24, 12.2	Dtw+rythme <b>&lt;.0005</b> , 0.61, 0.34, 9.9
<b>FG/FG (0.85)</b>	<b>&lt;.003</b> , 0.49, 0.22, 11	<b>0</b> , 0.67, 0.44, 28.3	<b>&lt;.012</b> , 0.43, 0.16, 7.6	<.98	<b>&lt;.048</b> , 0.35, 0.09, 4.6	
<b>FL/FL (0.51)</b>	<.1	<b>0</b> , 0.69, 0.46, 30.6	<b>&lt;.04</b> , 0.35, 0.1, 4.8	<b>0</b> , 0.81, 0.64, 62.8	<b>0</b> , 0.77, 0.58, 49.2	Dtw+ $\Delta$ durée <b>0</b> , 0.83, 0.68, 37.7
<b>GD/GD (0.85)</b>	<.75	<.84	<.85	<.4	<b>&lt;.041</b> , 0.35, 0.1, 4.8	
<b>GG/GG (0.68)</b>	<.97	<b>&lt;.005</b> , 0.47, 0.2, 9.6	<.46	<b>0</b> , 0.85, 0.71, 86.6	<b>0</b> , 0.85, 0.72, 90.9	Dtw+rythme <b>0</b> , 0.89, 0.78, 61.3
<b>GL/GL (0.59)</b>	<b>&lt;.025</b> , 0.38, 0.12, 5.8	<b>&lt;.004</b> , 0.48, 0.21, 10.3	<b>&lt;.05</b> , 0.33, 0.09, 4.3	<.054	<b>&lt;.02</b> , 0.4, 0.14, 6.7	Dtw+rythme <b>&lt;.00005</b> , 0.68, 0.43, 14.1
<b>LD/LD (0.65)</b>	<b>&lt;.03</b> , 0.37, 0.11, 5.3	<b>&lt;.006</b> , 0.46, 0.19, 9.1	<b>&lt;.04</b> , 0.37, 0.11, 5.4	<.34	<.53	
<b>LG/LG (0.5)</b>	<.9	<.061	<.81	<b>&lt;.0001</b> , 0.67, 0.44, 28.4	<b>0</b> , 0.75, 0.54, 42.5	
<b>LL/LL (0.63)</b>	<.2	<.69	<.21	<b>&lt;.013</b> , 0.43, 0.16, 7.7	<b>&lt;.002</b> , 0.51, 0.24, 11.8	$\Delta f0$ +Rythme <b>&lt;.00006</b> , 0.67, 0.41, 13.2
<b>Moy. R</b>	<b>0.22</b>	<b>0.38</b>	<b>0.23</b>	<b>0.42</b>	<b>0.51</b>	
<b>Moy. R<sup>2</sup>aj</b>	<b>0.06</b>	<b>0.18</b>	<b>0.05</b>	<b>0.24</b>	<b>0.29</b>	
<b>Moy. R significatifs</b>	<b>0.41</b>	<b>0.55</b>	<b>0.37</b>	<b>0.63</b>	<b>0.56</b>	
<b>Moy. R<sup>2</sup>aj significatifs</b>	<b>0.15</b>	<b>0.30</b>	<b>0.12</b>	<b>0.41</b>	<b>0.33</b>	

En général, les réponses des auditeurs sont fondées aussi bien sur les dimensions mélodiques que temporelles. La spécificité dépend de chaque frontière, qui semblent perçues différemment (tableau ci-dessous).

Tableau 139 : Dimension acoustique fondant la réponse des participants en fonction du type de frontière

Frontière avec un	Plutôt rythmique	Plutôt mélodique	Les deux
<b>Début mot lex. pluri.</b>	GD	LD	FD
<b>Monosyl. grammatic.</b>	LG	FG	GG
<b>Monosyllabe lexical</b>			FL, GL, LL

Tableau 140 : Répartition des réponses les mieux et les moins expliquées, pour chaque catégorie de frontière

Comparaison		FD	FG	FL	GD	GG	GL	LD	LG	LL
<b>Taux de réponses expliquées</b>	%	25	37	35	9	53	25	12	27	26
	Cat <sub>1</sub>	-	+	+	--	++	-	--	-	-
<b>Taux de réponses inexpliquées</b>	%	49	48	16	76	15	34	53	23	37
	Cat <sub>2</sub>	+	+	--	++	--	-	+	--	-
<b>Total</b>	%	74	85	51	85	68	59	65	50	63
	Cat <sub>(1-2)</sub>	--	0	+++	----	++++	0	---	+	0

Avec : % le taux de réponses expliqué par les données acoustiques (taux de réponses expliquées : [taux de réponses exactes] \* [coefficient de détermination du régresseur/modèle le plus explicatif]. Taux de réponses inexpliquées : [taux de réponses exactes] - [taux de réponses expliquées])

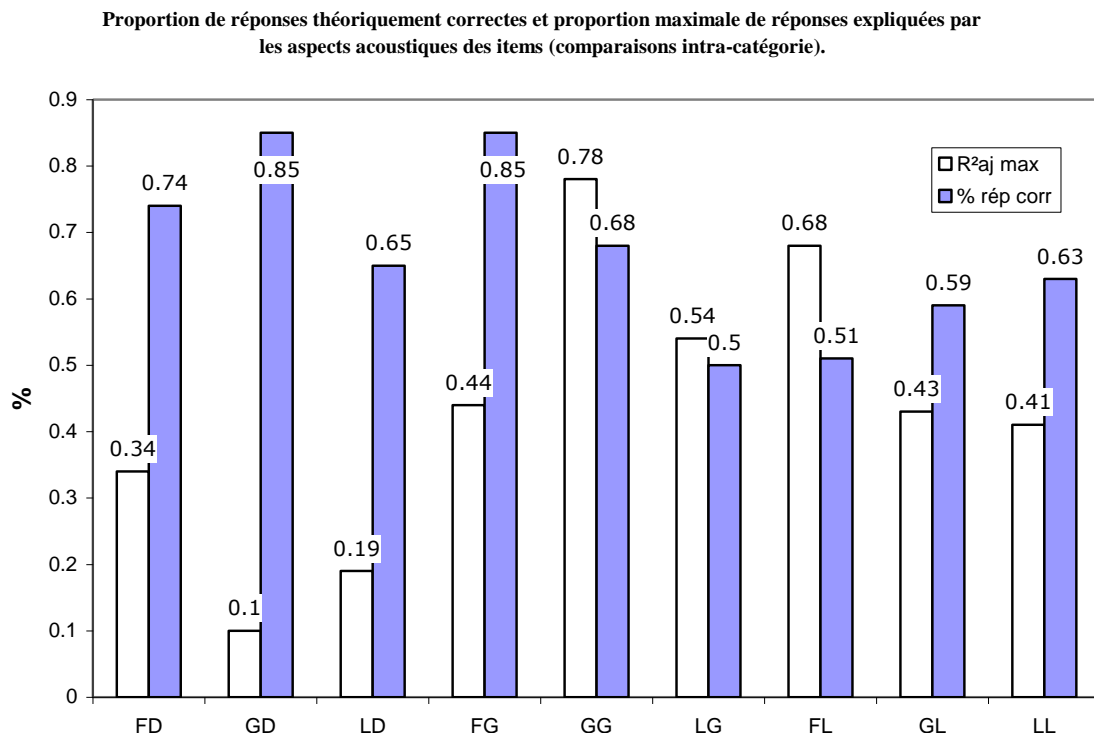
Cat<sub>1</sub> et Cat<sub>2</sub> la catégorisation du % précédent (de -- à ++) sur la distribution des % des 9 comparaisons intra catégorie. Cat<sub>(1-2)</sub> = [Cat<sub>1</sub>] - [Cat<sub>2</sub>], score de classement des frontières tenant compte de la performance des participants et de la quantité de réponses expliquées par les régresseurs (de ---- à ++++).

Sur la base de ce système d'évaluation (Tableau 140), les frontières peuvent être réparties en 3 groupes :

- ▷ comportement bien expliqué par les indices acoustiques (FL, GG, LG)
- ▷ comportement mal expliqué par les indices acoustiques (FD, GD, LD)
- ▷ comportement insuffisamment expliqué par les indices acoustiques (FG, GL, LL) Le cas de FG est cependant ambigu : il pourrait être classé avec [FD, GD, LD] compte tenu que de bonnes performances ont lieu pour cette frontière.

Ceci peut être figuré plus aisément à l'aide du graphique ci-dessous.

Figure 87 : Proportion de réponses expliquées par les caractéristiques acoustiques des items, pour les comparaisons intra catégorie



On peut remarquer que toutes les frontières impliquant un début de mot lexical plurisyllabique (GD, LD, FD) sont dans un même groupe (comportement mal expliqué par les données acoustiques).

Le système perceptif détecterait les débuts de mots lexicaux plurisyllabiques par la prosodie d'une manière qui semble robuste (performances élevées), mais qui s'explique mal par une approche analytique telle qu'effectuée ici. Les données de la Simulation 2 montrent en effet que le signal acoustique renferme assez d'information pour qu'un système artificiel présente des performances comparables aux auditeurs humains. C'est donc la preuve que les régresseurs choisis sont inadéquats (mais quels autres indices considérer ?), et que la méthode d'analyse est trop caricaturale.

#### Conclusions de l'analyse des relations régresseurs/réponses correctes pour les comparaisons intra catégorie

▷ Dimension temporelle seule

Le meilleur taux d'explication des données par un régresseur est observé avec la dimension temporelle ( $R^2$ ajusté maximum = 0.72 (GG/GG, rythme)).

▷ Dimension mélodique seule

∩ dtw est une variable importante pour les comparaisons intra catégorie

∩  $\Delta f_0$  et « Composite » ont une importance moindre. La perception n'utilise donc pas un contraste moyen de  $f_0$  (par syllabe, par item), mais l'évolution continue de la  $f_0$  (forme mélodique).

La dimension physique la plus importante pour regrouper les items d'une même catégorie apparaît globalement être la dimension temporelle.

▷ Combinaison des dimensions temporelles et mélodiques

∩ Les modèles de régression multiple montrent une utilisation conjointe de la dimension mélodique et temporelle dans 5 cas sur 9 (jusqu'à 78% de variance expliquée pour GG/GG (dtw + rythme)).

∩ Corrélation entre indices acoustiques et comportement est en moyenne plus élevée avec Rythme,  $\Delta$ durée et dtw (coefficients de détermination moyens).

Les résultats de ces analyses sont cependant incomplets ou peu fiables, au vu de la mise en rapport des données du Tableau 140, page 486 et des données de la Simulation 2. Ceci pose donc un problème d'interprétation pour objectiver le phénomène par une démarche analytique.

Analyse des relations régresseurs/réponses correctes pour les comparaisons inter  
catégoriesTableau 141 : Résultats de l'analyse des réponses correctes pour les comparaisons inter  
catégories

Comparaison (taux)	$\Delta f0$	Dtw	Composite	$\Delta$ durée	Rythme	Multiple
FD/FG (0,55)	<.94	<b>&lt;.045</b> 0.35, 0.09, 4.7	<.87	<b>&lt;.0001</b> 0.57, 0.31, 16.6	<b>&lt;.0009</b> 0.54, 0.27, 13.9	
FD/FL (0,46)	.85	<b>&lt;.0025</b> 0.51, 0.24, 11.7	<.41	<b>0</b> 0.79, 0.62, 57.6	<b>0</b> 0.82, 0.67, 70.8	
FD/GD (0,51)	<.33	<.95	<.44	<.74	<.76	
FD/GG (0,70)	<.22	<.25	<.24	<.41	<.38	
FD/GL (0,47)	<.97	<.46	<.82	<b>&lt;.032</b> 0.37, 0.11, 5.2	<.086	
FD/LD (0,34)	<.12	<b>&lt;.015</b> 0.4, 0.14, 6.6	<.085	<.21	<b>&lt;.02</b> 0.39, 0.13, 6.1	Dtw+rythme <b>&lt;.002</b> 0.56, 0.28, 7.7
FD/LG (0,69)	<.73	.44	<.27	<b>&lt;.0045</b> 0.47, 0.2, 9.6	<b>&lt;.0004</b> 0.54, 0.27, 13.9	
FD/LL (0,48)	<.52	<.9	<.53	<b>0</b> 0.68, 0.44, 28.9	<b>0</b> 0.67, 0.43, 27.4	
FG/FL (0,74)	<.23	<.34	<.18	<.5	<.56	
FG/GD (0,32)	<.16	<b>&lt;.0001</b> 0.64, 0.4, 23.8	<.23	<.58	<b>&lt;.02</b> 0.39, 0.13, 6.2	
FG/GG (0,30)	<b>&lt;.011</b> 0.42, 0.16, 7.5	<.62	<.86	<.72	<.73	
FG/GL (0,50)	<.19	<.63	<.2	<.47	<.92	
FG/LD (0,63)	<b>&lt;.004</b> 0.48, 0.21, 10.4	<.11	<b>&lt;.0007</b> 0.55, 0.28, 14.7	<b>&lt;.03</b> 0.37, 0.11, 5.5	<b>&lt;.022</b> 0.39, 0.13, 6.2	Comp+ryth <b>&lt;.00003</b> 0.69, 0.44, 14.6
FG/LG (0,59)	<b>&lt;.025</b> 0.38, 0.12, 5.9	<.95	<.11	<b>0</b> 0.65, 0.41, 25	<b>&lt;.0001</b> 0.67, 0.43, 27.4	Rythme+ $\Delta f0$ <b>0</b> 0.74, 0.52, 19.8
FG/LL (0,73)	<.36	<.22	<.74	<b>&lt;.0004</b> 0.58, 0.32, 17.1	<b>&lt;.0003</b> 0.57, 0.3, 16.3	
FL/GD (0,75)	<.14	<.26	<b>&lt;.048</b> 0.33, 0.08, 4.2	<.34	<.33	
FL/GG (0,79)	<.63	<.51	<.44	<.65	<.61	
FL/GL (0,58)	<.46	<.96	<.67	<b>&lt;.031</b> 0.37, 0.11, 5.4	<b>&lt;.043</b> 0.35, 0.1, 4.7	
FL/LD (0,53)	<1	<b>&lt;.0002</b> 0.6, 0.34, 18.9	<.91	<b>&lt;.001</b> 0.53, 0.26, 13.1	<b>&lt;.0002</b> 0.59, 0.33, 18.3	Dtw+rythme <b>0</b> 0.74, 0.52, 19.8

Comparaison (taux)	$\Delta f0$	Dtw	Composite	$\Delta$ durée	Rythme	Multiple
FL/LG (0,80)	<.31	<.71	<.24	<.004 0.49, 0.22, 10.8	<.02 0.41, 0.14, 6.9	
FL/LL (0,41)	<.011 0.44, 0.17, 8	<.83	<.01 0.44, 0.17, 7.9	0 0.71, 0.49, 34.6	<.0001 0.65, 0.4, 24.5	$\Delta$ durée+ $\Delta f0$ 0 0.82, 0.66, 34.6
GD/GG (0,26)	<.045 0.34, 0.09, 4.5	<.79	<.46	0 0.62, 0.37, 21.6	<.0001 0.58, 0.32, 17.5	
GD/GL (0,37)	<.015 0.42, 0.15, 7.2	<.68	<.082	<.0001 0.63, 0.38, 22.1	<.0003 0.61, 0.35, 20	
GD/LD (0,58)	<.99	<.48	<.74	<.38	<.081	
GD/LG (0,71)	<.013 0.41, 0.15, 7	<.015 0.41, 0.15, 7	<.33	<.0005 0.57, 0.3, 16.1	0 0.64, 0.4, 23.7	Rythme+dtw <.00002 0.7, 0.46, 15.8
GD/LL (0,69)	<.083	<.35	<.21	<.17	<.14	
GG/GL (0,57)	<.32	<.36	<.47	<.54	<.29	
GG/LD (0,64)	<.38	<.12	<.21	<.0004 0.56, 0.29, 15.4	<.065	Dtw+ $\Delta$ durée <.00005 0.67, 0.42, 13.4
GG/LG (0,71)	<.2	<.009 0.44, 0.17, 8.1	<.73	<.004 0.48, 0.21, 10.4	<.011 0.44, 0.17, 8	
GG/LL (0,77)	<.37	<.21	<.17	<.001 0.53, 0.26, 13.6	<.002 0.52, 0.25, 12.5	
GL/LD (0,48)	<.66	<.28	<.85	<.031 0.37, 0.11, 5.4	<.13	Dtw+ $\Delta$ durée <.011 0.49, 0.19, 5.2
GL/LG (0,68)	<.72	<.0013 0.53, 0.26, 13.1	<.84	<.039 0.35, 0.1, 4.8	<.0011 0.52, 0.25, 12.6	Dtw+ $\Delta$ durée <.0005 0.61, 0.34, 9.9
GL/LL (0,55)	<.08	<.98	<.074	<.0008 0.52, 0.25, 12.6	<.0015 0.52, 0.23, 11.7	
LD/LG (0,56)	<.52	<.93	<.87	<.13	<.005 0.47, 0.2, 9.7	
LD/LL (0,49)	<.093	.75	<.24	0 0.71, 0.49, 34	0 0.69, 0.46, 30.3	$\Delta$ durée+ $\Delta f0$ 0 0.8, 0.61, 28.8
LG/LL (0,68)	<.38	<.78	<.27	<.27	<.02 0.4, 0.13, 6.3	
Moy. R significatif	0.41	0.49	0.50	0.54	0.54	
Moy. R <sup>2</sup> aj significatif	0.15	0.22	0.23	0.29	0.28	

Conclusions de l'analyse des relations régresseurs/réponses correctes pour les comparaisons inter catégories

La dimension temporelle est en priorité responsable des confusions des catégories d'items :

- ▷  $\Delta$ durée ou rythme sont impliqués dans 3 fois plus de cas corrélation que dtw,  $\Delta f_0$  ou composite (45 contre 18)
- ▷ Les taux maximaux de corrélation sont plus élevés pour la dimension temporelle que pour la dimension mélodique ( $R^2$ ajusté jusqu'à 0.67 vs. 0.40)
- ▷ Seulement deux comparaisons apparaissent comme uniquement traitées par la dimension mélodique dans ces analyses

Comme dans le cas des comparaisons intra catégorie, la dimension physique qui conduit le plus souvent à une confusion des catégories apparaît globalement être la dimension temporelle. Les aspects rythmiques semblent donc plus prégnants à la perception, plus faciles à appréhender consciemment, que les aspects mélodiques. Il faut toutefois réserver la conclusion étant donné que les régresseurs utilisés sont peut-être mal adaptés pour la résolution du problème posé.

Tableau 142 : Pourcentage de confusions expliquées pour chaque comparaison de catégories de frontière

	Taux de confusions	$R^2$ ajusté	Cat <sub>1</sub>	%	Cat <sub>2</sub>	Cat <sub>(1+2)</sub>
<b>FD/FG</b>	44.8	31	-	13.9	-	--
<b>FD/FL</b>	54	67	++	36.2	++	++++
<b>FD/GD</b>	48.4	0	---	0.0	---	-----
<b>FD/GG</b>	29.8	0	---	0.0	---	-----
<b>FD/GL</b>	52.4	11	--	5.8	--	----
<b>FD/LD</b>	65.9	28	-	18.5	-	--
<b>FD/LG</b>	30.6	27	-	8.3	--	---
<b>FD/LL</b>	52	44	+	22.9	+	++
<b>FG/FL</b>	25.4	0	---	0.0	---	-----
<b>FG/GD</b>	67.1	40	+	26.8	+	++
<b>FG/GG</b>	69.4	16	--	11.1	-	---
<b>FG/GL</b>	49.2	0	---	0.0	---	-----
<b>FG/LD</b>	36.5	44	+	16.1	-	0
<b>FG/LG</b>	40.1	52	++	20.9	+	+++
<b>FG/LL</b>	26.2	32	-	8.4	--	---
<b>FL/GD</b>	24.7	8	--	2.0	--	----
<b>FL/GG</b>	20.6	0	---	0.0	---	-----
<b>FL/GL</b>	41.3	11	--	4.5	--	----
<b>FL/LD</b>	46.8	52	++	24.3	+	+++
<b>FL/LG</b>	19.8	22	-	4.4	--	---
<b>FL/LL</b>	58.3	66	++	38.5	++	++++
<b>GD/GG</b>	73	37	+	27.0	+	++
<b>GD/GL</b>	63	38	+	23.9	+	++
<b>GD/LD</b>	41.9	0	---	0.0	---	-----

	Taux de confusions	R <sup>2</sup> ajusté	Cat <sub>1</sub>	%	Cat <sub>2</sub>	Cat <sub>(1+2)</sub>
<b>GD/LG</b>	28.6	46	+	13.2	-	0
<b>GD/LL</b>	30.8	0	---	0.0	---	-----
<b>GG/GL</b>	42.1	0	---	0.0	---	-----
<b>GG/LD</b>	35	42	+	14.7	-	0
<b>GG/LG</b>	29	21	-	6.1	--	---
<b>GG/LL</b>	22.2	26	-	5.8	--	---
<b>GL/LD</b>	51.2	19	--	9.7	--	----
<b>GL/LG</b>	44.4	34	-	15.1	-	--
<b>GL/LL</b>	31.8	25	-	8.0	--	---
<b>LD/LG</b>	43.9	20	--	8.8	--	----
<b>LD/LL</b>	51	61	++	31.1	++	++++
<b>LG/LL</b>	31.8	13	--	4.1	--	----

Avec : Cat<sub>1</sub> le classement de la comparaison de frontières en fonction du R<sup>2</sup>ajusté (de --- à +++)

% le taux de réponses expliqué par les données acoustiques ([taux de réponses exactes] \* [coefficient de détermination du régresseur/modèle le plus explicatif]). Cat<sub>2</sub> le classement de la comparaison de frontières en fonction de la quantité de réponses expliquées (de --- à +++). Cat<sub>(1+2)</sub> = [Cat<sub>1</sub>]+[Cat<sub>2</sub>], le score de classement de la comparaison de frontières en fonction du degré de confusion des deux frontières comparées expliqué par les régresseurs (de ----- à ++++++).

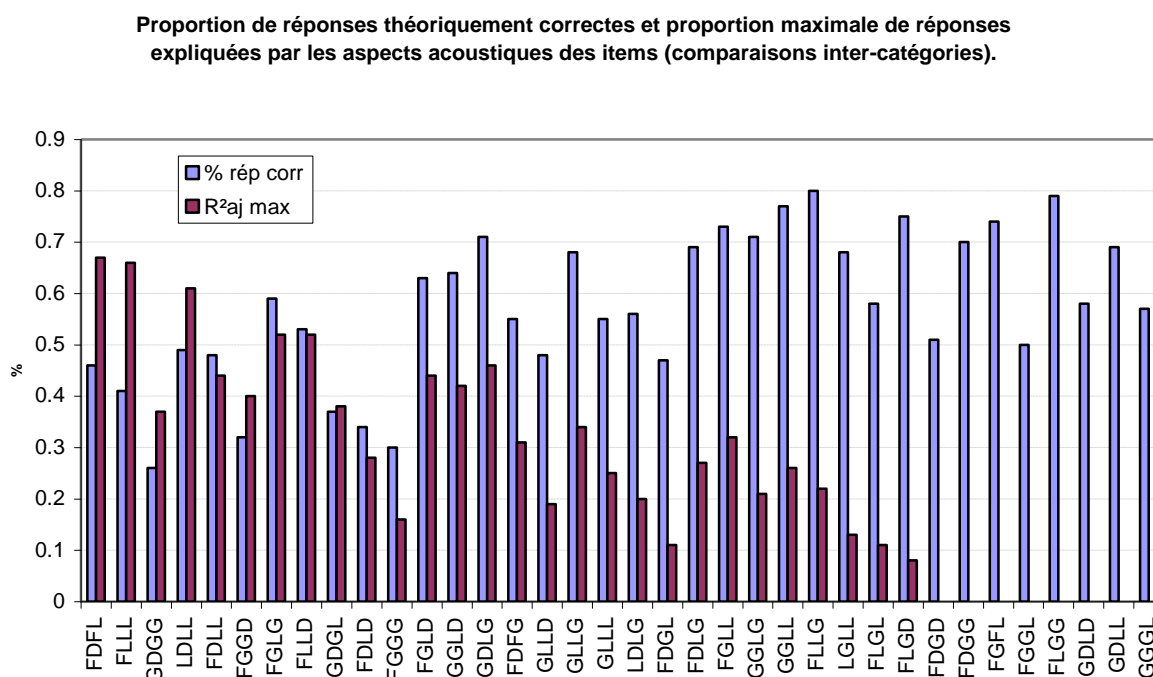
À partir du Tableau 142 ci-dessus, il est possible de répartir les frontières en 3 groupes :

- ▷ Frontières pour lesquelles le comportement est bien expliqué par les indices acoustiques (1/4 des comparaisons : FD / [FL, LL], FG / [GD, LG], FL / [LD, LL], GD / [GG, GL], LD / LL)
- ▷ Frontières pour lesquelles le comportement est mal expliqué par les indices acoustiques (2/3 des comparaisons : FD / [FG, GD, GG, GL, LD, LG], FG / [FL, GG, GL, LL], FL / [GD, GG, GL, LG], GD / [LD, LL], GG/[GL, LG, LL], GL / [LD, LG, LL], LD / LG, LG / LL)
- ▷ Frontières pour lesquelles le comportement est insuffisamment expliqué par les indices acoustiques (1/12 des comparaisons : FG / LD, GD / LG, GG / LD)

Ceci peut être figuré plus aisément à l'aide du graphique ci-dessous.



**Figure 88 : Proportion de réponses expliquées par les caractéristiques acoustiques des items, pour les comparaisons inter catégories**



La majorité des erreurs est donc faiblement justifiée par les régresseurs utilisés dans ces analyses. Percevoir la disparité de deux items peut être le fruit d'une évaluation sur la base d'éléments ou de configurations d'éléments qui diffèrent pour chaque cas, alors que les analyses ne peuvent réussir à mettre en évidence qu'une relation univoque, linéaire, entre les données et les réponses.

Le fait qu'un classificateur neuromimétique parvienne à effectuer une classification avec des performances voisines à meilleures que les auditeurs humains, démontre que les données renferment l'information, et que son utilisation peut se faire via des procédures non linéaires, qui ne peuvent être appréhendées avec les techniques d'analyse statistique utilisées.

### Conclusions générales des analyses complémentaires

- ▷ Les auditeurs fondent essentiellement leurs réponses sur l'utilisation de contrastes temporels de nature rythmique pour se prononcer sur le degré de ressemblance des items.
- ▷ Ils se basent plus rarement sur mélodie, plus difficile à analyser consciemment, moins manifeste et trop complexe.
- ▷ Une proportion importante du comportement reste difficilement explicable par la démarche analytique engagée pour expliquer les observations comportementales. Ce défaut peut être lié à :

- ∩ l'inadéquation des régresseurs choisis, qui auraient pourtant logiquement dû suffire pour décrire au mieux les caractéristiques pertinentes des items utilisées par les participants dans l'exécution de la tâche.
- ∩ l'utilisation d'une méthode d'analyse inapte à explorer un autre type de relation entre V.I. et V.D. qu'une relation linéaire. La prosodie est probablement exploitée de manière plus complexe et dynamique.

La conclusion est donc que la relation entre paramètres acoustiques et réaction comportementale est beaucoup trop complexe pour être étudiée à partir de mesures descriptives simples (moyennes, contrastes mélodiques ou rythmiques, etc.) et de techniques d'analyse statistique de relations linéaires entre V.I. et V.D.

Les phénomènes contrastifs jouent cependant un rôle important dans la comparaison des items, et dans leur caractérisation.

## Annexe 12

### Expérience 4 : Analyse acoustique des stimuli et estimation de l'influence des paramètres prosodiques sur les temps de réponse

Équation 35 : Évaluation de la relation entre facteurs expérimentaux (V.I.) et paramètres prosodiques (V.D.)

$$f_{0It} \sim N(\mathbf{XB}, \Omega)$$

$$f_{0It} = \beta_{0It} \text{Cons} + 8.933(0.517)I_{It} + -3.257(0.517)E_{It} + -1.676(0.731)I.E_{It}$$

$$\beta_{0It} = 139.239(0.365) + e_{0It}$$

$$[e_{0It}] \sim N(0, \Omega_e) : \Omega_e = [80.093(2.312)]$$

$$-2 * \loglikelihood(\text{IGLS Deviance}) = 17330.550(2400 \text{ of } 2400 \text{ cases in use})$$

$$\text{Intens}_{It} \sim N(\mathbf{XB}, \Omega)$$

$$\text{Intens}_{It} = \beta_{0It} \text{Cons} + -0.318(0.202)I_{It} + -4.879(0.202)E_{It} + -1.906(0.285)I.E_{It}$$

$$\beta_{0It} = 75.011(0.143) + e_{0It}$$

$$[e_{0It}] \sim N(0, \Omega_e) : \Omega_e = [12.184(0.352)]$$

$$-2 * \loglikelihood(\text{IGLS Deviance}) = 12811.290(2400 \text{ of } 2400 \text{ cases in use})$$

$$\text{Durée}_{It} \sim N(\mathbf{XB}, \Omega)$$

$$\text{Durée}_{It} = \beta_{0It} \text{Cons} + -79.167(3.905)I_{It} + 100.171(3.905)E_{It} + 118.350(5.522)I.E_{It}$$

$$\beta_{0It} = 387.221(2.761) + e_{0It}$$

$$[e_{0It}] \sim N(0, \Omega_e) : \Omega_e = [4574.234(132.047)]$$

$$-2 * \loglikelihood(\text{IGLS Deviance}) = 27038.570(2400 \text{ of } 2400 \text{ cases in use})$$

Tableau 143 : Tests de significativité du rapport entre facteurs lexicaux et prosodiques

	Effet des V.I. sur f0		Effet des V.I. sur dB		Effet des V.I. sur Durée	
	Chi <sup>2</sup> <sub>(1, N=2400)</sub>	p<	Chi <sup>2</sup> <sub>(1, N=2400)</sub>	p<	Chi <sup>2</sup> <sub>(1, N=2400)</sub>	p<
I	298.909	0	2.494	.12	411.044	0
E	39.736	3 <sup>E-10</sup>	586.122	0	658.09	0
I.E	5.263	.022	44.732	3 <sup>E-11</sup>	459.315	0

Figure 89 : Relations entre facteurs lexicaux et fréquence fondamentale

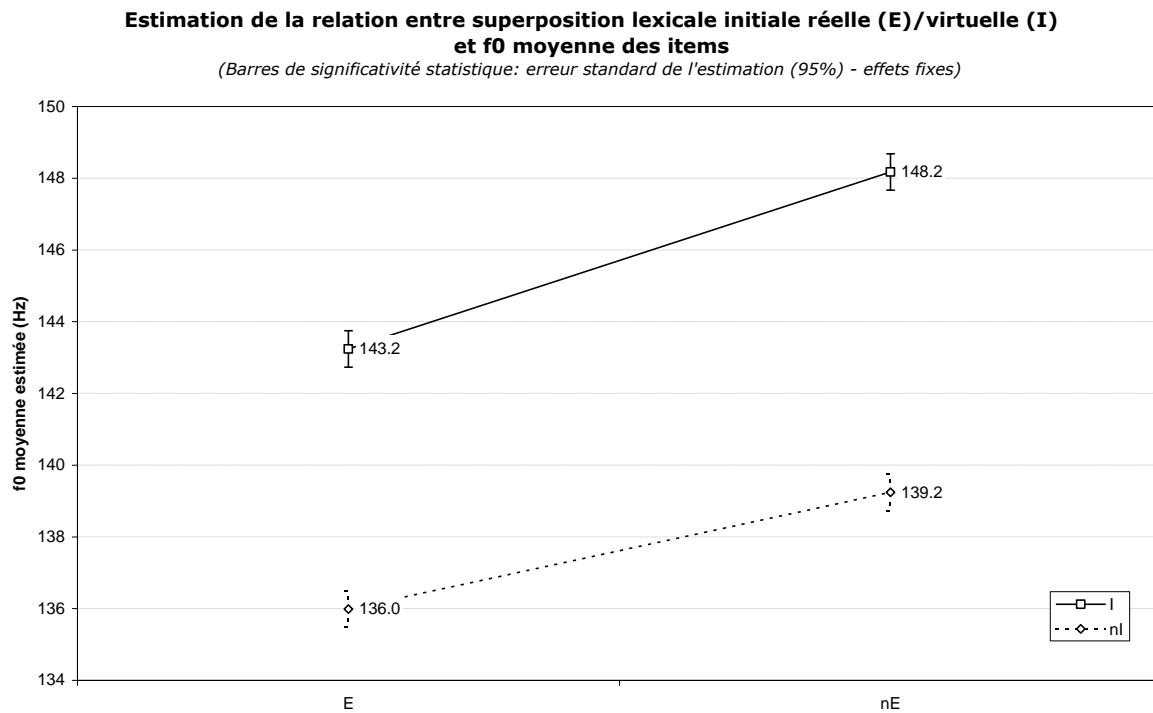


Figure 90 : Relations entre facteurs lexicaux et intensité

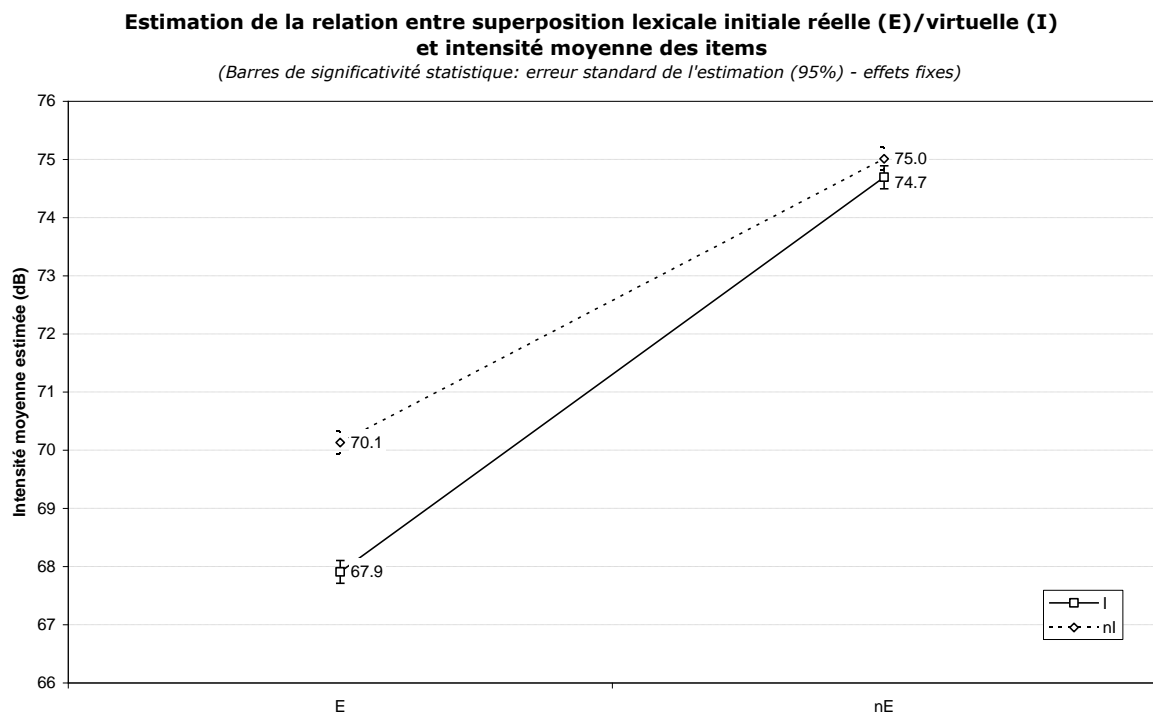
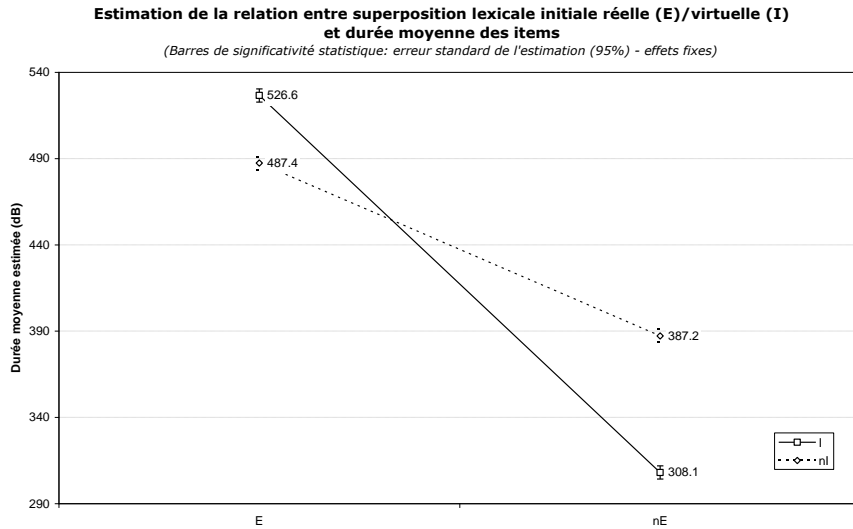


Figure 91 : Relations entre facteurs expérimentaux et durée



## Analyse de l'influence des facteurs prosodiques sur les temps de réponse

Équation 36 : Équation du modèle

$$\begin{aligned}
 &TR_{It, Suj} \sim N(\mathbf{XB}, \Omega) \\
 &TR_{It, Suj} = \beta_{0It, Suj} \mathbf{Cons} + 7.258(1.779)f0_{It, Suj} + 0.169(1.753)dB_{It, Suj} + -3.364(0.589)f0_{It, Suj} \cdot dB_{It, Suj} + \\
 &\quad -9.515(1.899)I.f0_{It, Suj} + 2.196(1.996)I.dB_{It, Suj} + -9.835(1.795)E.f0_{It, Suj} + \\
 &\quad -1.184(2.341)E.dB_{It, Suj} + 2.769(0.615)I.f0_{It, Suj} \cdot dB_{It, Suj} + 3.186(0.607)E.f0_{It, Suj} \cdot dB_{It, Suj} + \\
 &\quad 10.692(1.935)I.E.f0_{It, Suj} + 2.091(2.386)I.E.dB_{It, Suj} + -2.684(0.652)I.E.f0_{It, Suj} \cdot dB_{It, Suj} \\
 &\beta_{0It, Suj} = 570.693(8.903) + u_{0Suj} + e_{0It, Suj} \\
 &[u_{0Suj}] \sim N(\mathbf{0}, \Omega_u) : \Omega_u = [2474.734(597.853)] \\
 &[e_{0It, Suj}] \sim N(\mathbf{0}, \Omega_e) : \Omega_e = [10407.300(324.280)] \\
 &-2 * \log \text{likelihood (IGLS Deviance)} = 25489.060(2100 \text{ of } 2400 \text{ cases in use})
 \end{aligned}$$

Tableau 144 : Tests de significativité des covariables et de leurs interactions

	Coefficient	Chi <sup>2</sup> <sub>(1, N=2100)</sub>	p <
f0	7.258	16.644	5 <sup>E-05</sup>
dB	0.169	0.009	.93
f0.dB	-3.364	32.667	2 <sup>E-08</sup>
I.f0	-9.515	25.102	6 <sup>E-07</sup>
I.dB	2.196	1.211	.28
E.f0	-9.835	30.01	5 <sup>E-08</sup>
E.dB	-1.184	0.256	.62
I.f0.dB	2.769	20.278	7 <sup>E-06</sup>
E.f0.dB	3.186	27.511	2 <sup>E-07</sup>
I.E.f0	10.692	30.518	4 <sup>E-08</sup>
I.E.dB	2.091	0.768	.39
I.E.f0.dB	-2.684	16.923	4 <sup>E-05</sup>

Ces analyses montrent que les facteurs prosodiques ont un effet indépendant des facteurs lexicaux manipulés, sur les temps de réponse des participants.

## Annexe 13

### Expérience 4 : Équation du modèle et tests de significativité de l'effet des covariables sur les temps de réponse

Équation 37 : Équation du modèle étudié

$$TR_{It, Suj} \sim N(XB, \Omega)$$

$$TR_{It, Suj} = \beta_{0It, Suj} \text{Cons} + -117.881(26.208)I_{It, Suj} + -45.893(8.928)E_{It, Suj} + -125.096(14.675)I.E_{It, Suj} + -7.726(12.109)NPhIt_{It, Suj} + 5.838(11.085)PU_{It, Suj} + 48.305(10.156)NPhIt.PU_{It, Suj} + -0.359(0.097)PR_{It, Suj} + 0.742(0.122)DPU_{It, Suj} + 0.153(0.087)Dur_{It, Suj} + 0.003(0.001)PR.DPU_{It, Suj} + -0.006(0.001)PR.Dur_{It, Suj} + -0.001(0.000)DPU.Dur_{It, Suj} + 0.460(0.144)NPhIt.PR_{It, Suj} + -0.344(0.115)NPhI.DPU_{It, Suj} + -0.078(0.096)NPh.Dur_{It, Suj} + -0.003(0.002)NPhPRDPU_{It, Suj} + 0.000(0.001)NPhPRDur_{It, Suj} + 0.000(0.001)NPhDPUDu_{It, Suj} + 0.000(0.000)PRDPUDur_{It, Suj} + 0.000(0.000)NPPRDPUdu_{It, Suj} + -36.974(4.143)Fq_{It, Suj} + 4.766(1.506)AoA_{It, Suj} + 22.688(12.581)Fam_{It, Suj} + 6.638(1.965)Fq.AoA_{It, Suj} + -16.142(16.260)Fq.Fam_{It, Suj} + 3.935(2.645)AoA.Fam_{It, Suj} + -4.468(3.191)FqAoAFam_{It, Suj}$$

$$\beta_{0It, Suj} = 710.561(18.599) + u_{0Suj} + e_{0It, Suj}$$

$$[u_{0Suj}] \sim N(0, \Omega_u) : \Omega_u = [2586.914(614.354)]$$

$$[e_{0It, Suj}] \sim N(0, \Omega_e) : \Omega_e = [8406.880(261.950)]$$

**-2\*loglikelihood(IGLS Deviance) = 25050.430(2100 of 2400 cases in use)**

Avec : *TR* = temps de réponse (V.D.) ; *Cons* = constante (moyenne générale) ; *I* = mot expérimental enchâssable en position initiale (V.I. – réf = *nI*) ; *E* = mot expérimental avec enchâssement initial (V.I. – réf = *nE*) ; *PR* = position du point de reconnaissance (mesurée en ms. dans le signal acoustique à partir du début du mot) ; *DPU* = position du début du point d'unicité (mesurée en ms. dans le signal acoustique à partir du début du mot) ; *Dur* = durée du mots ou du pseudo-mot ; *NPhIt* = nombre de phonèmes de l'item ; *PU* = position du point d'unicité en nombre de phonèmes (calculé à partir des valeurs fournies par BRULEX) ; *FqFMot* = fréquence formelle phonologique - logarithme en base 10 de la somme des fréquences lexicales ayant une forme phonologique identique à l'item (calculé à partir de valeurs de BRULEX) ; *AoA* = âge d'acquisition estimé moyen par forme phonologique ; *Fam* = familiarité conceptuelle moyenne par forme phonologique

Tableau 145 : Tests de significativité des covariables et de leurs interactions

	Coefficient	Chi <sup>2</sup> (1, N=2100)	p<
NPhIt	-7.7263	0.407	.53
PU	5.838	0.227	.64
NPhIt.PU	48.305	22.622	<b>2<sup>E-06</sup></b>
PR	-0.35862	13.544	<b>.00024</b>
DPU	0.74194	37.16	<b>2<sup>E-09</sup></b>
Dur	0.15251	3.074	.08
PR.DPU	0.0033196	15.075	<b>.0002</b>
PR.Dur	-0.0060102	60.164	<b>9<sup>E-15</sup></b>
DPU.Dur	-0.0012311	9.032	<b>.0027</b>
NPhIt.PR	0.45959	10.208	<b>.0014</b>
NPhI.DPU	-0.34409	8.948	<b>.0028</b>
NPh.Dur	-0.077838	0.66	.42
NPhPRDPU	-0.0028181	1.932	.17
NPhPRDur	0.00041477	0.126	.73
NPhDPUDu	0.00044475	0.439	.51
PRDPUDur	1.29 <sup>E-05</sup>	10.126	<b>.0015</b>
NPPRDPUdu	2.73 <sup>E-05</sup>	12.494	<b>.0005</b>
Fq	-36.974	79.628	<b>5<sup>E-19</sup></b>
AoA	4.7664	10.012	<b>.0016</b>

	Coefficient	Chi <sup>2</sup> <sub>(1, N=2100)</sub>	p<
Fam	22.688	3.252	.072
Fq.AoA	6.6376	11.414	<b>.0008</b>
Fq.Fam	-16.142	0.986	.33
AoA.Fam	3.9355	2.214	.14
FqAoAFam	-4.4675	1.96	.17

## Annexe 14

### Expérience 5 : Matériel linguistique

Tableau 146 : Mots expérimentaux et pseudo-mots distracteurs

Matériel expérimental		Matériel distracteur	
E	nE	E	nE
dʁag	dʁɔg	bʁap	gʁøb
lɛb	lɔb	bɑv	kɛl
ʒøn	ʒon	vʏk	nɑl
gal	gœl	tum	bep
kʁik	kʁak	dʁyn	bʁøp
kal	kəl	suz	daz
tāk	tik	lõʁ	gyd
vwal	vʝəl	fʁwak	fʝøv
fʁød	fʁod	gʁoʒ	stam
ʒut	ʒat	fɛz	ʁœd
kʁɑp	kʁup	fʁɑp	fʁum
tʁus	tʁɑs	bʁɛʒ	flef
pʁiz	pʁoʒ	tʁɛs	fʁan
gam	gɔm	ʒɑz	nag
kõt	kɔt	dõʁ	gyf
kaz	koz	mõp	zet
kud	kɔd	nɛʁ	løf
kʁɛt	kʁut	bløʒ	stib
ʒʏʒ	ʒoʒ	pɑʒ	zyb
twal	tʝil	pʝil	pʝɔb
pʁɛs	pʁɛs	plɛg	fʁyl
bʁiz	bʁõz	pʁeg	kliʒ
tʁip	tʁap	gʁaʒ	pʁyv
tʁɛt	tʁɑt	vʁɛk	bluʁ

Avec : E = présence d'un mot enchâssé initial dans l'item (nE pour mot sans enchâssement)



## Annexe 15

### Expérience 5 : Équation du modèle et tests de significativité de l'effet des covariables sur les temps de réponse

Équation 38 : Équation du modèle étudié

$$TR_{It, Suj} \sim N(\mathbf{XB}, \Omega)$$

$$TR_{It, Suj} = \beta_{0It, Suj} \text{Cons} + \mathbf{-18.674(8.606)Ench}_{It, Suj} + \mathbf{-0.782(0.280)Ordr}_{It, Suj} + \mathbf{86.733(45.677)Lat}_{Suj} + \mathbf{15.150(16.615)Sex}_{Suj} + \mathbf{401.529(86.584)Dur}_{It, Suj} + \mathbf{-82.809(32.415)Npho}_{It, Suj} + \mathbf{548.330(77.670)PR}_{It, Suj} + \mathbf{61.922(28.655)PUPho}_{It, Suj} + \mathbf{-30.251(10.678)FrForMo}_{It, Suj} + \mathbf{-10.844(3.170)AoA}_{It, Suj} + \mathbf{-99.174(14.201)Fam}_{It, Suj} + \mathbf{-18.716(183.226)Dur} \cdot \mathbf{Npho}_{It, Suj} + \mathbf{933.042(1408.821)Dur} \cdot \mathbf{PR}_{It, Suj} + \mathbf{-100.779(166.995)Npho} \cdot \mathbf{PR}_{It, Suj} + \mathbf{-8276.084(3023.551)Dur} \cdot \mathbf{Npho} \cdot \mathbf{PR}_{It, Suj} + \mathbf{11.565(4.327)FrForMo} \cdot \mathbf{AoA}_{It, Suj} + \mathbf{127.167(22.689)FrForMo} \cdot \mathbf{Fam}_{It, Suj} + \mathbf{19.575(4.244)AoA} \cdot \mathbf{Fam}_{It, Suj} + \mathbf{-5.250(9.494)FrForMo} \cdot \mathbf{AoA} \cdot \mathbf{Fam}_{It, Suj}$$

$$\beta_{0It, Suj} = \mathbf{739.486(47.611)} + u_{0Suj} + e_{0It, Suj}$$

$$[u_{0Suj}] \sim N(\mathbf{0}, \Omega_u) : \Omega_u = [\mathbf{1485.861(485.075)}]$$

$$[e_{0It, Suj}] \sim N(\mathbf{0}, \Omega_e) : \Omega_e = [\mathbf{9981.563(542.927)}]$$

**-2\*loglikelihood(IGLS Deviance) = 8575.424(708 of 768 cases in use)**

Avec : TR = temps de réponse (V.D., en ms.) ; Ench = enchâssement (V.I. - condition E ; réf = nE) ; Ordr = ordre de présentation des items ; Lat = latéralité manuelle des participants (réf. = G) ; Sex = sexe des participants (réf. = F) ; Dur = durée des mots (ms.) ; NPho = nombre de phonèmes du mot stimulus ; PR = position acoustique du point de reconnaissance du mot stimulus (ms.) ; PUPho = position phonétique du point d'unicité ; FrForMo = logarithme de la fréquence formelle phonologique ; AoA = âge d'acquisition estimé ; Fam = familiarité conceptuelle.

Tableau 147 : Tests de significativité des covariables et de leurs interactions

	Coefficient	Chi <sup>2</sup> <sub>(1, N=708)</sub>	p<
Ordr	-0.782	7.828	<b>.0052</b>
Lat	86.733	3.606	<b>.058</b>
Sex	15.15	0.831	<b>.37</b>
Dur	401.529	21.506	<b>4<sup>E-06</sup></b>
Npho	-82.809	6.526	<b>.011</b>
PR	548.33	49.84	<b>2<sup>E-12</sup></b>
PUPho	61.922	4.67	<b>.031</b>
FrForMo	-30.251	8.025	<b>.0047</b>
AoA	-10.844	11.705	<b>.0007</b>
Fam	-99.174	48.774	<b>3<sup>E-12</sup></b>
Dur.Npho	-18.716	0.01	<b>.93</b>
Dur.PR	933.042	0.439	<b>.51</b>
Npho.PR	-100.779	0.364	<b>.55</b>
Dur.Npho.PR	-8276.084	7.492	<b>.0062</b>
FrForMo.AoA	11.565	7.143	<b>.0076</b>
FrForMo.Fam	127.167	31.412	<b>3<sup>E-08</sup></b>
AoA.Fam	19.575	21.272	<b>4<sup>E-06</sup></b>
FrForMo.AoA.Fam	-5.25	0.306	<b>.59</b>

## Annexe 16

### Expérience 6 : Matériel linguistique

Tableau 148 : Mots expérimentaux et pseudo-mots distracteurs

Matériel expérimental				Matériel distracteur			
M1		nM1		M1		nM1	
M2	nM2	M2	nM2	M2	nM2	M2	nM2
tēbal	kezin	glōbyl	ʃənal	klāʃək	bakmid	ʒaskəʃ	dəkəʃð
tādās	vādāʒ	skādāl	ʃəvlyk	gəkɛfkyb	ʃjēzid	fɛybət	kēʃad
fāfaɕ	sātās	stādaɕ	bluzð	plɛfēt	figzaj	kɛugag	fɛəlāʃ
balās	pādyl	køfly	fjasko	ʒyʒgā	dimkjal	stiniʃ	flōduɕ
bālʃø	køkel	vylgɛɕ	fłøɛt	maɕspð	bɛizlyl	fjəsɔwaf	dɛotav
fukmi	kutym	pɕəskɛi	fɕaze	səmlət	fiskud	pɕəmsðʒ	ftilit
ʃamwa	tuknā	ʒalu	fɕusəɕ	kedē	myzgev	byktāk	kɕiskum
gəkaʒ	fatig	flameʃ	gɕəna	gðbyʃ	tɕðkəɛls	slinɛɕ	gɛlmyn
pikət	pudkøʒ	klima	ʒɛstjð	vipat	sykmez	kəʃtop	zenad
paɕwa	tɛɕmos	fɕomaʒ	koləɔ	bləklād	tufdap	tunlo	səpnid
lisās	gəɕguy	fleməɕ	kəavat	nɛfpjð	polfaʒ	djasklu	kwāzy
suksi	kapsyl	stigmat	myskad	supfyg	lynpaʃ	kjatəɔ	boktij
gitəɕ	kəɕnej	danwa	pəstje	pðpsit	tɕipʃiv	zilēʒ	bɕātēs
kəɕbəɔ	kədjəɔ	kɛspe	pɕēsip	fujbyʃ	bilkiv	nəkpwēt	ʃlédəʃ
pɕefas	pɕədkɛi	despət	statɕet	gɕyvət	blevkal	psalgam	dəvlag
vɛɕgla	pɔwəɕəɕ	tikɛ	stɕyktyɕ	buʃte	dwasiv	skəɔvɛij	djyɕʃø
bɕasle	gɕanit	klavsē	stydjə	gɛdut	nəskym	vaktɕu	bikgɕe
fəɕmyl	tɕibyn	spikɛit	tɕafik	dujkyɕ	feplin	psəʒəb	ʃēdaʒ
ʃəɕpāt	fɛɕmje	skləkøʒ	tɕāzat	kəslē	bəɔɔəɔ	pɕəɕful	māskin
mustəʃ	bɕigad	kðsta	tsəvin	ʃəlkud	stilnak	bɕəfɛim	pjətkø
skəɕbyt	paɕtəɕ	pɕatik	vjəlet	bɕənɕi	gusfyʒ	blylāp	kɕipsəɔ
tɕāspəɕ	lwəziɕ	nəvis	zədjək	bɕētɕəp	tybʒəʒ	sfetāp	spaviʃ

Avec : M1/2 = 1ère/2ème syllabe lexicale (nM pour non lexicale)

## Annexe 17

### Expérience 6 : Analyse acoustique des stimuli

Équation 39 : Estimation du rapport entre paramètres prosodiques (f0, intensité et durée) et facteurs expérimentaux pour la première syllabe

$$f0S1nc_{It} \sim N(\mathbf{XB}, \Omega)$$

$$f0S1nc_{It} = \beta_{0It} \text{Cons} + -2.199(0.346)M1_{It} + -2.597(0.346)M2_{It} + 3.422(0.490)M1.M2_{It}$$

$$\beta_{0It} = 130.038(0.245) + e_{0It}$$

$$[e_{0It}] \sim N(0, \Omega_e) : \Omega_e = [40.900(1.107)]$$

$$-2 * \loglikelihood(IGLS \text{ Deviance}) = 17865.720(2728 \text{ of } 2728 \text{ cases in use})$$

$$dBS1nc_{It} \sim N(\mathbf{XB}, \Omega)$$

$$dBS1nc_{It} = \beta_{0It} \text{Cons} + 3.795(0.243)M1_{It} + -0.010(0.243)M2_{It} + -1.200(0.343)M1.M2_{It}$$

$$\beta_{0It} = 70.192(0.172) + e_{0It}$$

$$[e_{0It}] \sim N(0, \Omega_e) : \Omega_e = [20.102(0.544)]$$

$$-2 * \loglikelihood(IGLS \text{ Deviance}) = 15927.910(2728 \text{ of } 2728 \text{ cases in use})$$

$$DurS1nc_{It} \sim N(\mathbf{XB}, \Omega)$$

$$DurS1nc_{It} = \beta_{0It} \text{Cons} + -69.220(3.721)M1_{It} + 0.701(3.721)M2_{It} + 41.949(5.263)M1.M2_{It}$$

$$\beta_{0It} = 316.757(2.631) + e_{0It}$$

$$[e_{0It}] \sim N(0, \Omega_e) : \Omega_e = [4722.165(127.860)]$$

$$-2 * \loglikelihood(IGLS \text{ Deviance}) = 30820.670(2728 \text{ of } 2728 \text{ cases in use})$$

Équation 40 : Estimation du rapport entre paramètres prosodiques ( $f_0$ , intensité et durée) et facteurs expérimentaux pour la deuxième syllabe

$$f_0S_2nc_{It} \sim N(\mathbf{XB}, \Omega)$$

$$f_0S_2nc_{It} = \beta_{0It} \text{Cons} + \mathbf{-0.371(0.182)}M1_{It} + \mathbf{0.792(0.182)}M2_{It} + \mathbf{-1.144(0.257)}M1.M2_{It}$$

$$\beta_{0It} = \mathbf{111.023(0.129)} + e_{0It}$$

$$[e_{0It}] \sim N(0, \Omega_e) : \Omega_e = [\mathbf{11.302(0.306)}]$$

$$\mathbf{-2*loglikelihood(IGLS Deviance) = 14356.960(2728 \text{ of } 2728 \text{ cases in use)}$$

$$dBS_2nc_{It} \sim N(\mathbf{XB}, \Omega)$$

$$dBS_2nc_{It} = \beta_{0It} \text{Cons} + \mathbf{2.740(0.207)}M1_{It} + \mathbf{1.931(0.207)}M2_{It} + \mathbf{-2.047(0.292)}M1.M2_{It}$$

$$\beta_{0It} = \mathbf{66.948(0.146)} + e_{0It}$$

$$[e_{0It}] \sim N(0, \Omega_e) : \Omega_e = [\mathbf{14.575(0.395)}]$$

$$\mathbf{-2*loglikelihood(IGLS Deviance) = 15050.810(2728 \text{ of } 2728 \text{ cases in use)}$$

$$DurS_2nc_{It} \sim N(\mathbf{XB}, \Omega)$$

$$DurS_2nc_{It} = \beta_{0It} \text{Cons} + \mathbf{47.446(7.230)}M1_{It} + \mathbf{-42.560(7.230)}M2_{It} + \mathbf{19.007(10.225)}M1.M2_{It}$$

$$\beta_{0It} = \mathbf{546.605(5.112)} + e_{0It}$$

$$[e_{0It}] \sim N(0, \Omega_e) : \Omega_e = [\mathbf{17824.400(482.623)}]$$

$$\mathbf{-2*loglikelihood(IGLS Deviance) = 34444.280(2728 \text{ of } 2728 \text{ cases in use)}$$

Tableau 149 : Test de significativité de l'effet des facteurs expérimentaux sur les paramètres prosodiques des syllabes constituant les mots expérimentaux

	Effet des V.I. sur $f_0S_1$		Effet des V.I. sur $dBS_1$		Effet des V.I. sur $DurS_1$	
	Chi <sup>2</sup> <sub>(1, N=2728)</sub>	p<	Chi <sup>2</sup> <sub>(1, N=2728)</sub>	p<	Chi <sup>2</sup> <sub>(1, N=2728)</sub>	p<
<b>M<sub>1</sub></b>	40.303	<b>3<sup>E-10</sup></b>	244.343	<b>0</b>	345.998	<b>0</b>
<b>M<sub>2</sub></b>	56.222	<b>7<sup>E-14</sup></b>	0.002	.97	0.035	.86
<b>M<sub>1</sub>.M<sub>2</sub></b>	48.804	<b>3<sup>E-12</sup></b>	12.221	<b>5<sup>E-04</sup></b>	63.537	<b>2<sup>E-15</sup></b>
	Effet des V.I. sur $f_0S_2$		Effet des V.I. sur $dBS_2$		Effet des V.I. sur $DurS_2$	
<b>M<sub>1</sub></b>	4.144	<b>.042</b>	175.64	<b>5<sup>E-40</sup></b>	43.066	<b>6<sup>E-11</sup></b>
<b>M<sub>2</sub></b>	18.939	<b>2<sup>E-05</sup></b>	87.265	<b>1<sup>E-20</sup></b>	34.654	<b>4<sup>E-09</sup></b>
<b>M<sub>1</sub>.M<sub>2</sub></b>	19.754	<b>9<sup>E-06</sup></b>	49.017	<b>3<sup>E-12</sup></b>	3.456	<b>.06</b>

Avec :  $M_1$  = première syllabe ( $S_1$ ) lexicale ;  $M_2$  = deuxième syllabe ( $S_2$ ) lexicale ;  $f_0$  = fréquence fondamentale moyenne de la syllabe (Hz) ;  $dB$  = intensité moyenne de la syllabe (dB) ;  $Dur$  = durée de la syllabe (ms.)

Ces effets sont représentés graphiquement ci-dessous, pour chaque dimension prosodique.

Figure 92 : Effet des facteurs expérimentaux sur la fréquence fondamentale des syllabes

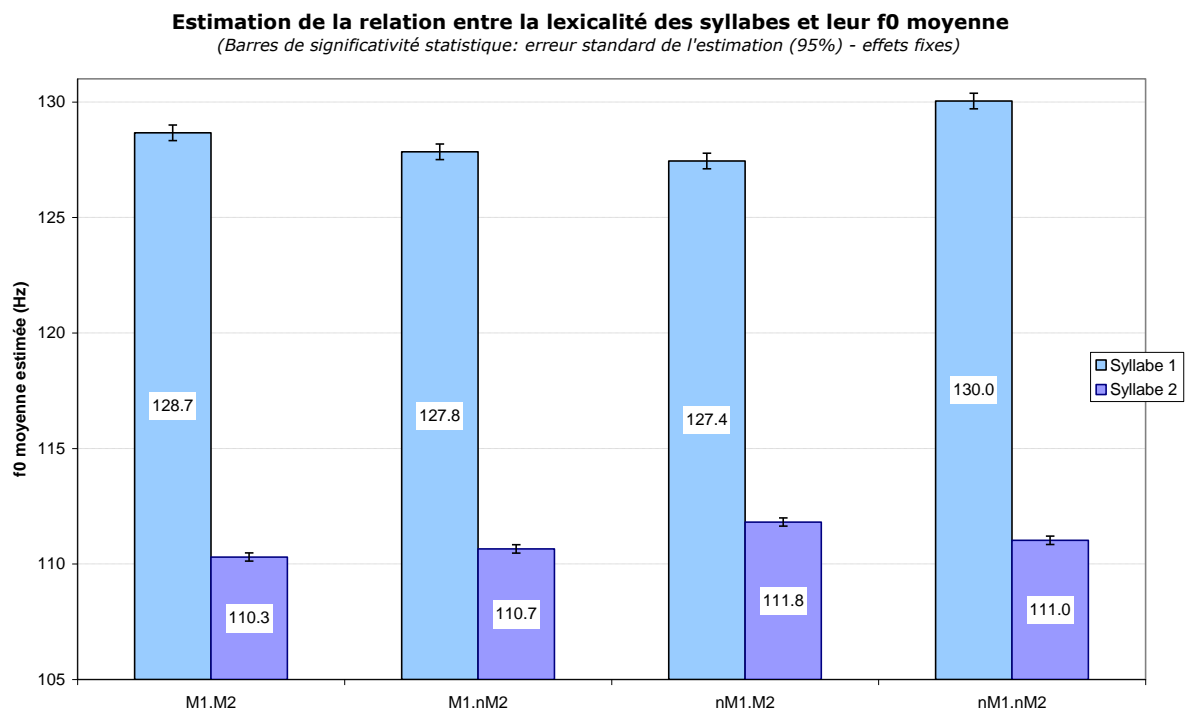


Figure 93 : Effet des facteurs expérimentaux sur l'intensité des syllabes

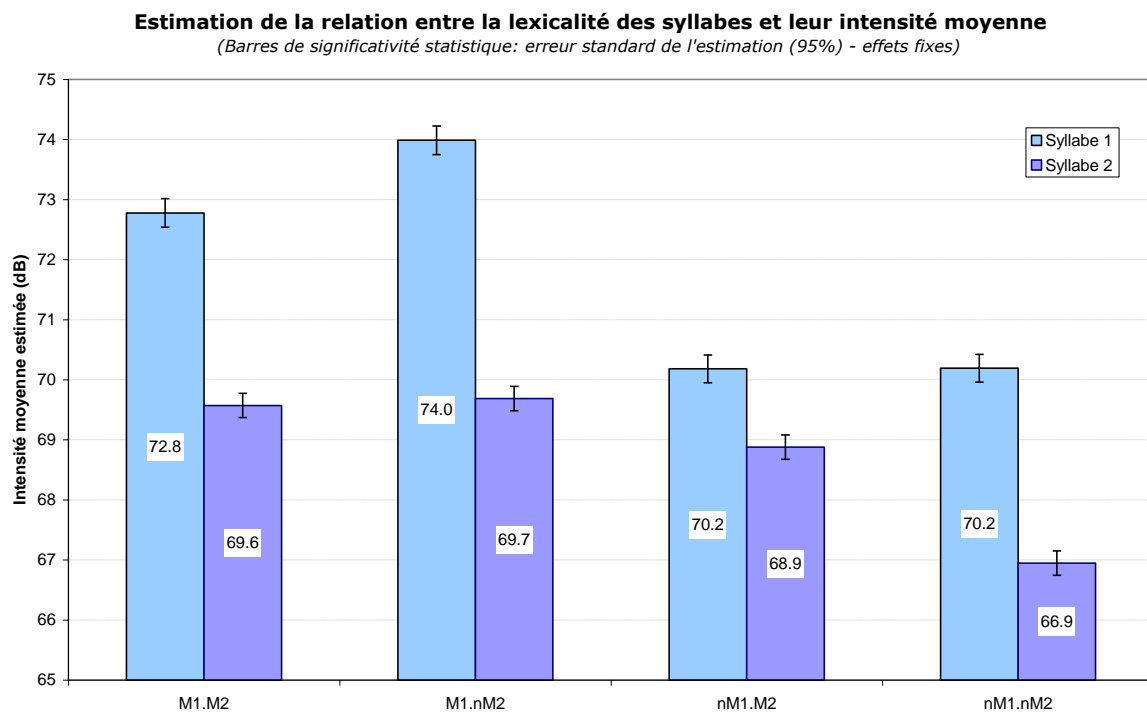
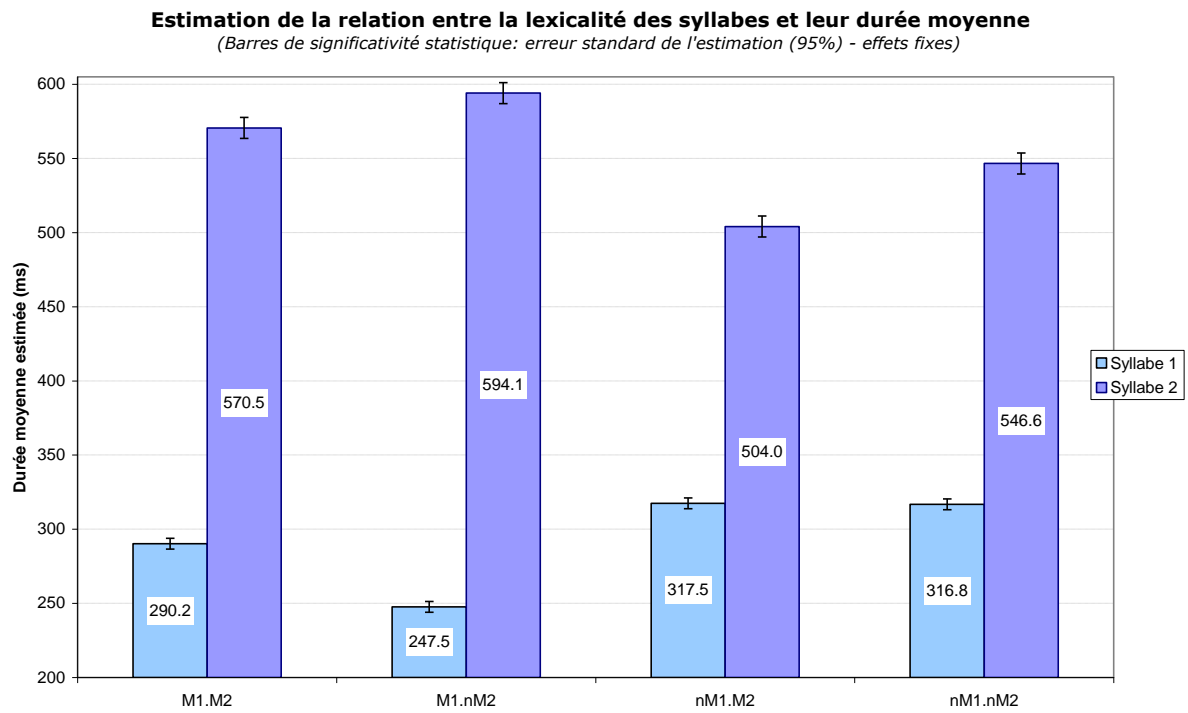


Figure 94 : Effet des facteurs expérimentaux sur la durée des syllabes



## Annexe 18

### Expérience 6 : Équation du modèle et tests de significativité de l'effet des covariables sur les temps de réponse

Équation 41 : Équation du modèle étudié

$TR_{It, Suj} \sim N(\mathbf{XB}, \Omega)$

$$\begin{aligned}
 TR_{It, Suj} = & \beta_{0It, Suj} \mathbf{Cons} + \mathbf{-24.310(11.894)} M1_{It, Suj} + \mathbf{-0.415(11.532)} M2_{It, Suj} + \mathbf{-16.782(19.137)} M1.M2_{It, Suj} + \\
 & \mathbf{-0.042(0.047)} \mathbf{Ordr}_{It, Suj} + \mathbf{7.975(31.000)} \mathbf{Sex}_{Suj} + \mathbf{-34.804(35.151)} \mathbf{Lat}_{Suj} + \mathbf{-0.352(0.678)} \mathbf{fOS1}_{It, Suj} + \\
 & \mathbf{5.842(1.714)} \mathbf{fOS2}_{It, Suj} + \mathbf{2.101(1.168)} \mathbf{dBS1}_{It, Suj} + \mathbf{3.235(1.581)} \mathbf{dBS2}_{It, Suj} + \\
 & \mathbf{0.727(0.138)} \mathbf{DurS1}_{It, Suj} + \mathbf{0.306(0.079)} \mathbf{DurS2}_{It, Suj} + \mathbf{32.150(5.195)} \mathbf{lg(fqS1)}_{It, Suj} + \\
 & \mathbf{-7.141(5.929)} \mathbf{lg(fqS2)}_{It, Suj} + \mathbf{-9.744(7.034)} \mathbf{lg(Fqfrm)}_{It, Suj} + \mathbf{7.592(2.277)} \mathbf{AoA}_{It, Suj} + \\
 & \mathbf{-37.226(13.819)} \mathbf{Fam}_{It, Suj} + \mathbf{-48.604(10.883)} \mathbf{NPhM}_{It, Suj} + \mathbf{-0.075(0.041)} \mathbf{DPU}_{It, Suj} + \\
 & \mathbf{0.233(0.079)} \mathbf{PR}_{It, Suj} + \mathbf{-0.198(0.224)} \mathbf{fOS1} \cdot \mathbf{dBS1}_{It, Suj} + \mathbf{1.530(0.398)} \mathbf{fOS2} \cdot \mathbf{dBS2}_{It, Suj} + \\
 & \mathbf{-6.661(2.962)} \mathbf{lg(Fqfrm)} \cdot \mathbf{AoA}_{It, Suj} + \mathbf{-35.699(21.174)} \mathbf{lg(Fqfrm)} \cdot \mathbf{Fam}_{It, Suj} + \\
 & \mathbf{30.435(7.089)} \mathbf{AoA} \cdot \mathbf{Fam}_{It, Suj} + \mathbf{0.348(5.884)} \mathbf{lg(Fqfrm)} \cdot \mathbf{AoA} \cdot \mathbf{Fam}_{It, Suj} + \\
 & \mathbf{0.000(0.000)} \mathbf{DPU} \cdot \mathbf{PR}_{It, Suj} + \mathbf{0.309(0.073)} \mathbf{NPhM} \cdot \mathbf{DPU}_{It, Suj} + \mathbf{-0.478(0.090)} \mathbf{NPhM} \cdot \mathbf{PR}_{It, Suj} + \\
 & \mathbf{0.000(0.000)} \mathbf{NPhM} \cdot \mathbf{DPU} \cdot \mathbf{PR}_{It, Suj} + \mathbf{0.002(0.001)} \mathbf{DurS1} \cdot \mathbf{DPU}_{It, Suj} + \\
 & \mathbf{-0.001(0.000)} \mathbf{DurS2} \cdot \mathbf{DPU}_{It, Suj} + \mathbf{0.001(0.001)} \mathbf{DurS1} \cdot \mathbf{PR}_{It, Suj} + \mathbf{0.002(0.000)} \mathbf{DurS2} \cdot \mathbf{PR}_{It, Suj} + \\
 & \mathbf{-0.002(0.001)} \mathbf{DurS1} \cdot \mathbf{DurS2}_{It, Suj} + \mathbf{0.000(0.000)} \mathbf{DurS1} \cdot \mathbf{DPU} \cdot \mathbf{PR}_{It, Suj} + \\
 & \mathbf{0.000(0.000)} \mathbf{DurS2} \cdot \mathbf{DPU} \cdot \mathbf{PR}_{It, Suj} + \mathbf{0.000(0.000)} \mathbf{DurS1} \cdot \mathbf{DurS2} \cdot \mathbf{DPU} \cdot \mathbf{PR}_{It, Suj} + \\
 & \mathbf{1.536(1.784)} \mathbf{M1.fOS1}_{It, Suj} + \mathbf{-11.164(3.091)} \mathbf{M2.fOS2}_{It, Suj} + \mathbf{-3.731(1.959)} \mathbf{M1.dBS1}_{It, Suj} + \\
 & \mathbf{-9.585(2.139)} \mathbf{M2.dBS2}_{It, Suj} + \mathbf{0.773(0.175)} \mathbf{M1.DurS1}_{It, Suj} + \mathbf{0.009(0.087)} \mathbf{M2.DurS2}_{It, Suj} + \\
 & \mathbf{1.113(0.456)} \mathbf{M1.fOS1} \cdot \mathbf{dBS1}_{It, Suj} + \mathbf{-5.589(0.880)} \mathbf{M2.fOS2} \cdot \mathbf{dBS2}_{It, Suj} + \\
 & \mathbf{1.128(2.478)} \mathbf{M1.M2.fOS1}_{It, Suj} + \mathbf{22.507(5.143)} \mathbf{M1.M2.fOS2}_{It, Suj} + \mathbf{-1.264(3.815)} \mathbf{M1.M2.dBS1}_{It, Suj} + \\
 & \mathbf{9.381(3.525)} \mathbf{M1.M2.dBS2}_{It, Suj} + \mathbf{-0.059(0.180)} \mathbf{M1.M2.DurS1}_{It, Suj} + \mathbf{0.207(0.093)} \mathbf{M1.M2.DurS2}_{It, Suj} + \\
 & \mathbf{-0.472(0.718)} \mathbf{M1.M2.fOS1} \cdot \mathbf{dBS1}_{It, Suj} + \mathbf{-0.333(1.163)} \mathbf{M1.M2.fOS2} \cdot \mathbf{dBS2}_{It, Suj}
 \end{aligned}$$

$$\beta_{0It, Suj} = \mathbf{893.880(30.708)} + \mathbf{u_{0Suj}} + \mathbf{e_{0It, Suj}}$$

$$[\mathbf{u_{0Suj}}] \sim \mathbf{N(0, \Omega_u)} : \Omega_u = [\mathbf{3121.578(833.625)}]$$

$$[\mathbf{e_{0It, Suj}}] \sim \mathbf{N(0, \Omega_e)} : \Omega_e = [\mathbf{14405.890(404.300)}]$$

$-2 * \loglikelihood(IGLS \text{ Deviance}) = 32030.580(2573 \text{ of } 2728 \text{ cases in use})$

Avec : TR = temps de réponse (V.D.) ;  $M_1$  = première syllabe lexicale (V.I. - réf =  $nM_1$ ) ;  $M_2$  = deuxième syllabe lexicale (V.I. - réf =  $nM_2$ ) ; Ordr = Ordre de présentation des items ; Sex = sexe du participant (réf. = F) ; Lat = Latéralité manuelle (réf. = G) ;  $fOS_{1/2}$  = fréquence fondamentale moyenne de la première/deuxième syllabe (Hz) ;  $dBS_{1/2}$  = intensité moyenne de la première/deuxième syllabe (dB) ;  $DurS_{1/2}$  = durée de la première/deuxième syllabe (ms.) ;  $lg(fqS_{1/2})$  = logarithme en base 10 de la fréquence formelle phonologique de la première/deuxième syllabe (fréquence où la séquence de phonème apparaît en tant que syllabe) ;  $lg(fqfrm)$  = logarithme en base 10 de la fréquence formelle phonologique du mot ; AoA = âge d'acquisition estimé du mot ; Fam = familiarité conceptuelle du mot ; NPhM = nombre de phonèmes composant le mot ; DPU = position acoustique du début du P.U. du mot (ms.) ; PR = position acoustique du point de reconnaissance du mot (ms.).

Tableau 150 : Tests de significativité des covariables et de leurs interactions

	Coefficient	chi <sup>2</sup> (1, N=2573)	p<
Ordr	-0.042	0.794	.38
Sex	7.975	0.066	.80
Lat	-34.804	0.98	.33
fOS <sub>1</sub>	-0.352	0.27	.61
fOS <sub>2</sub>	5.842	11.613	<b>.0007</b>
dBS <sub>1</sub>	2.101	3.239	<b>.072</b>
dBS <sub>2</sub>	3.235	4.188	<b>.041</b>

	Coefficient	chi <sup>2</sup> <sub>(1, N=2573)</sub>	p<
DurS <sub>1</sub>	0.727	27.66	<b>2<sup>E-07</sup></b>
DurS <sub>2</sub>	0.306	14.973	<b>.0002</b>
lg(fqS <sub>1</sub> )	32.15	38.291	<b>7<sup>E-10</sup></b>
lg(fqS <sub>2</sub> )	-7.141	1.451	.23
lg(fqfrm)	-9.744	1.919	.17
AoA	7.592	11.112	<b>.0009</b>
Fam	-37.226	7.256	<b>.0071</b>
NPhM	-48.604	19.947	<b>8<sup>E-06</sup></b>
DPU	-0.075	3.421	<b>.065</b>
PR	0.233	8.675	<b>.0033</b>
f0S <sub>1</sub> .dBS <sub>1</sub>	-0.198	0.783	.38
f0S <sub>2</sub> .dBS <sub>2</sub>	1.53	14.762	<b>.0002</b>
lg(Fqfrm).AoA	-6.661	5.057	<b>.025</b>
lg(Fqfrm).Fam	-35.699	2.843	<b>.092</b>
AoA.Fam	30.435	18.432	<b>2<sup>E-05</sup></b>
lg(Fqfrm).AoA.Fam	0.348	0.004	.95
DPU.PR	0	0.424	.52
NPhM.DPU	0.309	17.793	<b>3<sup>E-05</sup></b>
NPhM.PR	-0.478	28.413	<b>1<sup>E-07</sup></b>
PNPhM.DPU.PR	0	0.135	.72
DurS <sub>1</sub> .DPU	0.002	5.22	<b>.023</b>
DurS <sub>2</sub> .DPU	-0.001	8.771	<b>.0031</b>
DurS <sub>1</sub> .PR	0.001	2.242	.14
DurS <sub>2</sub> .PR	0.002	12.184	<b>.0005</b>
DurS <sub>1</sub> .DurS <sub>2</sub>	-0.002	11.976	<b>.0006</b>
DurS <sub>1</sub> .DPU.PR	0	0	1
DurS <sub>2</sub> .DPU.PR	0	14.929	<b>.0002</b>
DurS <sub>1</sub> .DurS <sub>2</sub> .DPU.PR	0	0.015	.91
M <sub>1</sub> .f0S <sub>1</sub>	1.536	0.741	.39
M <sub>2</sub> .f0S <sub>2</sub>	-11.164	13.046	<b>.0004</b>
M <sub>1</sub> .dBS <sub>1</sub>	-3.731	3.626	<b>.057</b>
M <sub>2</sub> .dBS <sub>2</sub>	-9.585	20.089	<b>8<sup>E-06</sup></b>
M <sub>1</sub> .DurS <sub>1</sub>	0.773	19.582	<b>1<sup>E-05</sup></b>
M <sub>2</sub> .DurS <sub>2</sub>	0.009	0.011	.92
M <sub>1</sub> .f0S <sub>1</sub> .dBS <sub>1</sub>	1.113	5.963	<b>.015</b>
M <sub>2</sub> .f0S <sub>2</sub> .dBS <sub>2</sub>	-5.589	40.381	<b>3<sup>E-10</sup></b>
M <sub>1</sub> .M <sub>2</sub> .f0S <sub>1</sub>	1.128	0.207	.65
M <sub>1</sub> .M <sub>2</sub> .f0S <sub>2</sub>	22.507	19.153	<b>2<sup>E-05</sup></b>
M <sub>1</sub> .M <sub>2</sub> .dBS <sub>1</sub>	-1.264	0.11	.75
M <sub>1</sub> .M <sub>2</sub> .dBS <sub>2</sub>	9.381	7.083	<b>.008</b>
M <sub>1</sub> .M <sub>2</sub> .DurS <sub>1</sub>	-0.059	0.107	.75
M <sub>1</sub> .M <sub>2</sub> .DurS <sub>2</sub>	0.207	4.943	<b>.027</b>
M <sub>1</sub> .M <sub>2</sub> .f0S <sub>1</sub> .dBS <sub>1</sub>	-0.472	0.433	.52
M <sub>1</sub> .M <sub>2</sub> .f0S <sub>2</sub> .dBS <sub>2</sub>	-0.333	0.082	.79



## Annexe 19

## Expérience 7 (&amp; 10) : Matériel linguistique

Tableau 151 : Mots expérimentaux (Expériences 7 &amp; 10)

<b>M<sub>1</sub></b>		<b>nM<sub>1</sub></b>	
<b>M<sub>2</sub></b>	<b>nM<sub>2</sub></b>	<b>M<sub>2</sub></b>	<b>nM<sub>2</sub></b>
tēbal	æzin	glōbyl	ʃənal
fāfak	vādāz	skādal	ʃəvlyk
balās	sātās	stādaκ	bluzō
bāljo	pādyl	κəfly	fjasko
fukmi	kəkəl	vylgək	fbusak
ʃamwa	kutym	pκəskvi	gκəna
gavaz	fyzi	zalu	zestjō
pikat	fatig	klima	kolən
paκwa	pudκəz	fκomaz	kκavat
paκkuk	tκmos	flemaκ	myskad
suksi	gakuj	stigmat	pəstje
gitaκ	kapsyl	κespe	pκēsip
kaκbən	kəknej	despət	statuət
pκefas	pwarak	tike	stkyktyκ
vκgla	gκanit	klavsē	stydjo
bκasle	tκibyn	spikit	tκafik
fκmyl	fκmjje	skleκoz	tκāzat
ʃakpāt	bκigad	kōsta	tsakın
mustaf	pastək	pκatik	vjəlet
tκāspək	lwaziv	fκaka	zodjak

Avec :  $M_{1/2} = 1^{\text{ère}} / 2^{\text{ème}}$  syllabe lexicale (nM pour non lexicale)

Tableau 152 : Pseudo-mots distracteurs

<b>Pseudo-mots DerP (Exp. 7 &amp; 10)</b>				<b>Pseudo-mots DerS (Expérience 7)</b>			
<b>M<sub>1</sub></b>		<b>nM<sub>1</sub></b>		<b>M<sub>1</sub></b>		<b>nM<sub>1</sub></b>	
<b>M<sub>2</sub></b>	<b>nM<sub>2</sub></b>	<b>M<sub>2</sub></b>	<b>nM<sub>2</sub></b>	<b>M<sub>2</sub></b>	<b>nM<sub>2</sub></b>	<b>M<sub>2</sub></b>	<b>nM<sub>2</sub></b>
plāʃu	fκāʃiv	bibκa	blazu	klāʃək	bakmid	zasbət	dκakʃad
səkbe	buʃē	bκəte	klwazə	gκefkyb	ʃjēzaj	fκygag	fκəlāʃ
gκibuʃ	plōbja	bydzu	ljətno	plefēt	figmez	kκuniʃ	fłōduκ
lakdā	tādeb	kκutā	lystκan	zyzɡā	bκizlyl	stitκu	dκotav
dκagu	bestju	danwe	tκistez	maκspō	myzgev	fjosōz	filit
negəm	kōdqe	fκəlē	noza	səmlət	tκəκəls	pκəmtāk	κqiskum
gκigō	klemāp	glute	pκknā	kedē	sykfaz	bykneκ	gelmyn
tāzūκ	kəklyv	mamlo	pjanā	gōkyʃ	tufplin	slitop	sōpnad
plōzē	dwanjē	naκsil	pκevnāl	bləkpat	polkjal	κaflo	kwānid
vəlke	κākys	nəvil	pκynā	nəfpjō	lynvκal	tunklu	boktij

Pseudo-mots DerP (Exp. 7 & 10)				Pseudo-mots DerS (Expérience 7)			
M <sub>1</sub>		nM <sub>1</sub>		M <sub>1</sub>		nM <sub>1</sub>	
M <sub>2</sub>	nM <sub>2</sub>	M <sub>2</sub>	nM <sub>2</sub>	M <sub>2</sub>	nM <sub>2</sub>	M <sub>2</sub>	nM <sub>2</sub>
plakav	kapiv	κøʒto	κøvnno	supfyg	tκipʃiv	djastøn	bκātēs
kølkot	gupid	suvκu	blēdav	pøpsit	bilkiv	kjalēʒ	ʃlēzy
tukmāk	tøksan	skøwē	stenā	fujbyʃ	blepaʃ	nøkpwēt	døvdaʒ
ʃāpaʒ	pētav	gijmō	blizaʃ	gβylād	dwasiv	psalvκij	djuvlag
pørpjø	fetin	κqisō	flaneg	gelē	nøskym	skøswaf	bivgqe
byʃκu	vētjeʒ	flames	kēzjet	dujvøt	fedap	vakzāb	ʃēkin
vivʏʃ	bastiʃ	bløkan	kōvnāκ	κaskud	bøʎnøm	psoful	māsdøʃ
kκwaso	platid	dκesal	skøʎjog	ʃalkyκ	stilnak	bκœfκim	pjosøm
litik	vizav	zasmu	stκatep	bκate	gusfyʒ	blylāp	kκiptκø
mekven	vwazu	skalpeʃ	zefid	bκētκap	tybʒaʒ	sfetāp	spaviʃ

Avec : DerP = mot devenant pseudo-mot par le dernier phonème ; DerS = mot devenant pseudo-mot par la dernière syllabe.

## Annexe 20

### Expérience 7 : Analyse acoustique des stimuli

Équation 42 : Estimation du rapport entre paramètres prosodiques (f0, intensité et durée) et facteurs expérimentaux pour la première syllabe

$$f0S1_{It} \sim N(\mathbf{XB}, \Omega)$$

$$f0S1_{It} = \beta_{0It} \text{Cons} + -2.971(0.346)M1_{It} + -1.757(0.346)M2_{It} + 0.596(0.489)M1.M2_{It}$$

$$\beta_{0It} = 142.035(0.244) + e_{0It}$$

$$[e_{0It}] \sim N(0, \Omega_e) : \Omega_e = [83.620(1.580)]$$

$$-2 * \loglikelihood(IGLS Deviance) = 40679.290(5600 \text{ of } 5600 \text{ cases in use})$$

$$dBS1_{It} \sim N(\mathbf{XB}, \Omega)$$

$$dBS1_{It} = \beta_{0It} \text{Cons} + 3.777(0.210)M1_{It} + -1.717(0.210)M2_{It} + 1.185(0.297)M1.M2_{It}$$

$$\beta_{0It} = 64.871(0.148) + e_{0It}$$

$$[e_{0It}] \sim N(0, \Omega_e) : \Omega_e = [30.798(0.582)]$$

$$-2 * \loglikelihood(IGLS Deviance) = 35085.800(5600 \text{ of } 5600 \text{ cases in use})$$

$$DurS1_{It} \sim N(\mathbf{XB}, \Omega)$$

$$DurS1_{It} = \beta_{0It} \text{Cons} + -45.889(1.835)M1_{It} + 8.608(1.835)M2_{It} +$$

$$7.186(2.596)M1.M2_{It}$$

$$\beta_{0It} = 248.909(1.298) + e_{0It}$$

$$[e_{0It}] \sim N(0, \Omega_e) : \Omega_e = [2358.296(44.568)]$$

$$-2 * \loglikelihood(IGLS Deviance) = 59380.000(5600 \text{ of } 5600 \text{ cases in use})$$

Équation 43 : Estimation du rapport entre paramètres prosodiques (f0, intensité et durée) et facteurs expérimentaux pour la deuxième syllabe

$$f0S2_{It} \sim N(\mathbf{XB}, \Omega)$$

$$f0S2_{It} = \beta_{0It} \text{Cons} + -2.382(0.225)M1_{It} + -0.310(0.225)M2_{It} + 0.606(0.319)M1.M2_{It}$$

$$\beta_{0It} = 123.752(0.159) + e_{0It}$$

$$[e_{0It}] \sim N(0, \Omega_e) : \Omega_e = [35.514(0.671)]$$

$$-2 * \loglikelihood(IGLS Deviance) = 35883.730(5600 \text{ of } 5600 \text{ cases in use})$$

$$dBS2_{It} \sim N(\mathbf{XB}, \Omega)$$

$$dBS2_{It} = \beta_{0It} \text{Cons} + 4.987(0.240)M1_{It} + 2.320(0.240)M2_{It} + -2.226(0.340)M1.M2_{It}$$

$$\beta_{0It} = 59.086(0.170) + e_{0It}$$

$$[e_{0It}] \sim N(\mathbf{0}, \Omega_e) : \Omega_e = [40.389(0.763)]$$

$$-2 * \log\text{likelihood(IGLS Deviance)} = 36604.060(5600 \text{ of } 5600 \text{ cases in use})$$

$$DurS2_{It} \sim N(\mathbf{XB}, \Omega)$$

$$DurS2_{It} = \beta_{0It} \text{Cons} + 40.866(3.878)M1_{It} + -35.236(3.878)M2_{It} + 11.264(5.484)M1.M2_{It}$$

$$\beta_{0It} = 456.241(2.742) + e_{0It}$$

$$[e_{0It}] \sim N(\mathbf{0}, \Omega_e) : \Omega_e = [10527.210(198.946)]$$

$$-2 * \log\text{likelihood(IGLS Deviance)} = 67757.730(5600 \text{ of } 5600 \text{ cases in use})$$

Tableau 153 : Effet des facteurs expérimentaux sur les paramètres prosodiques des syllabes constituant les mots

	Effet des V.I. sur f0S <sub>1</sub>		Effet des V.I. sur dBS <sub>1</sub>		Effet des V.I. sur DurS <sub>1</sub>	
	Chi <sup>2</sup> <sub>(1, N=5600)</sub>	p<	Chi <sup>2</sup> <sub>(1, N=5600)</sub>	p<	Chi <sup>2</sup> <sub>(1, N=5600)</sub>	p<
M <sub>1</sub>	73.913	<b>9<sup>E-18</sup></b>	79.964	<b>4<sup>E-19</sup></b>	625.05	<b>0</b>
M <sub>2</sub>	25.828	<b>4<sup>E-07</sup></b>	324.186	<b>0</b>	21.993	<b>3<sup>E-06</sup></b>
M <sub>1</sub> .M <sub>2</sub>	1.487	.23	66.991	<b>3<sup>E-16</sup></b>	7.664	<b>.006</b>
	Effet des V.I. sur f0S <sub>2</sub>		Effet des V.I. sur dBS <sub>2</sub>		Effet des V.I. sur DurS <sub>2</sub>	
M <sub>1</sub>	111.856	<b>4<sup>E-26</sup></b>	432.443	<b>0</b>	4.217	<b>.041</b>
M <sub>2</sub>	1.894	.17	430.994	<b>0</b>	111.049	<b>6<sup>E-26</sup></b>
M <sub>1</sub> .M <sub>2</sub>	3.634	<b>.06</b>	93.27	<b>5<sup>E-22</sup></b>	82.557	<b>2<sup>E-19</sup></b>

Avec : M<sub>1</sub> = première syllabe (S1) lexicale ; M<sub>2</sub> = deuxième syllabe (S2) lexicale ; f0 = fréquence fondamentale moyenne de la syllabe (Hz) ; dB = intensité moyenne de la syllabe (dB) ; Dur = durée de la syllabe (ms.)

Figure 95 : Effet des facteurs expérimentaux sur la fréquence fondamentale des syllabes

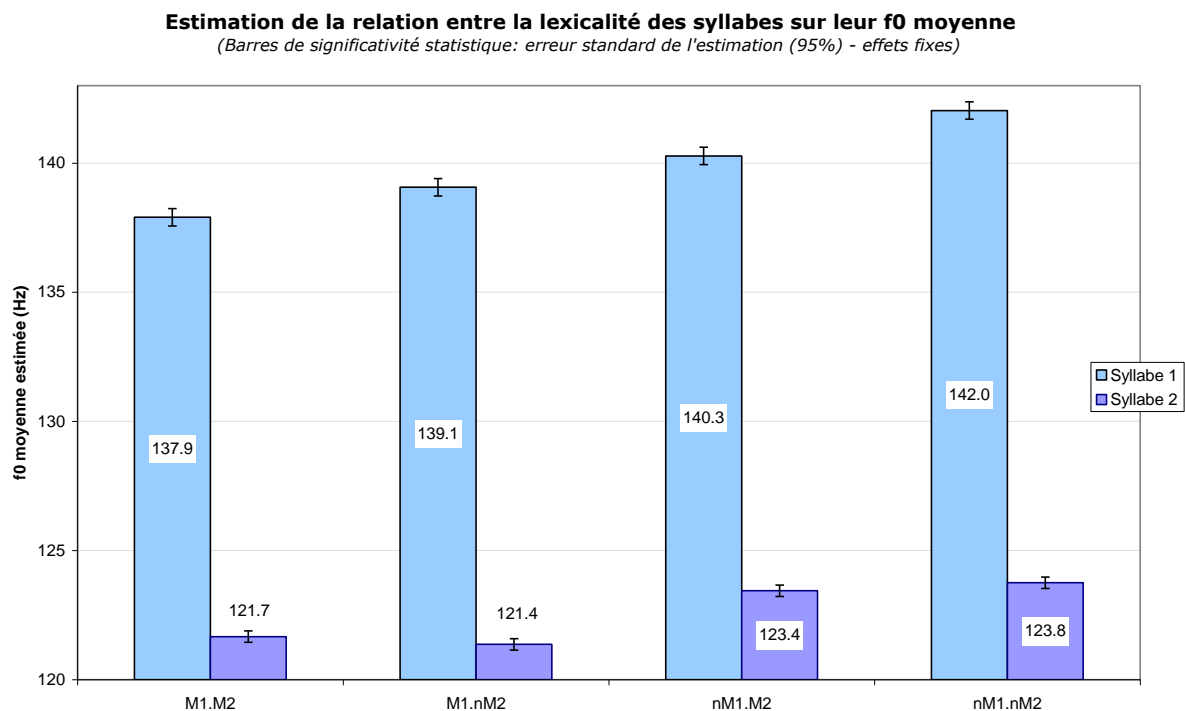


Figure 96 : Effet des facteurs expérimentaux sur l'intensité des syllabes

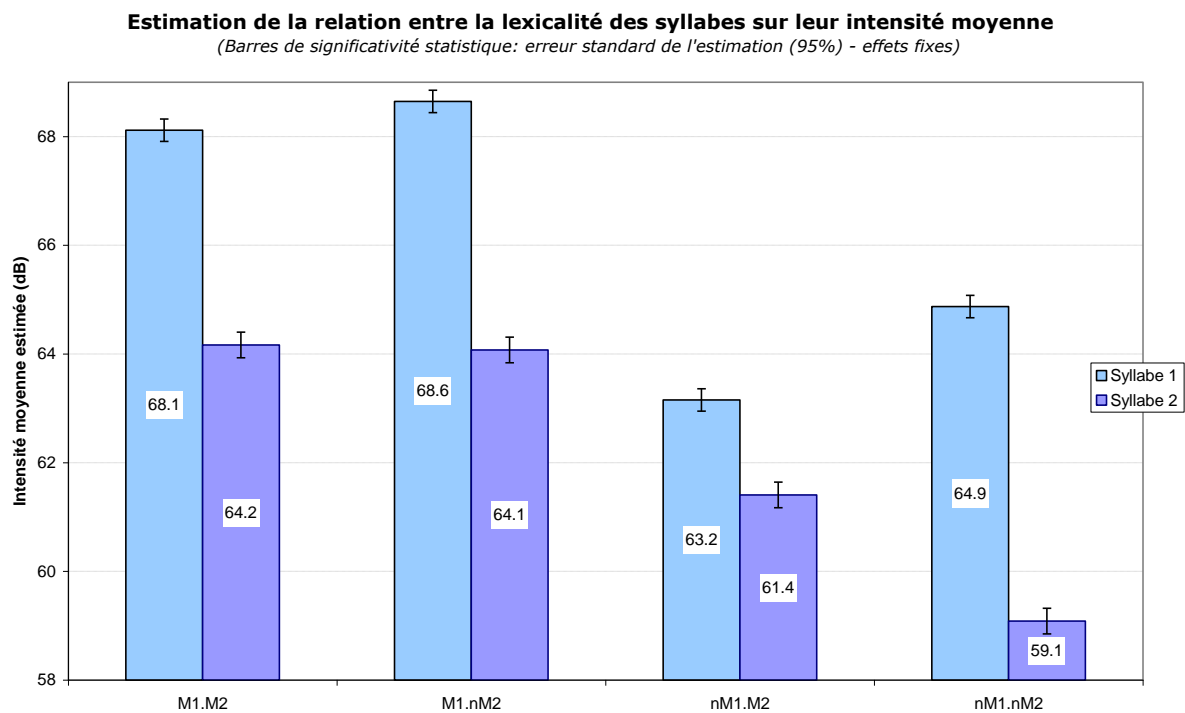
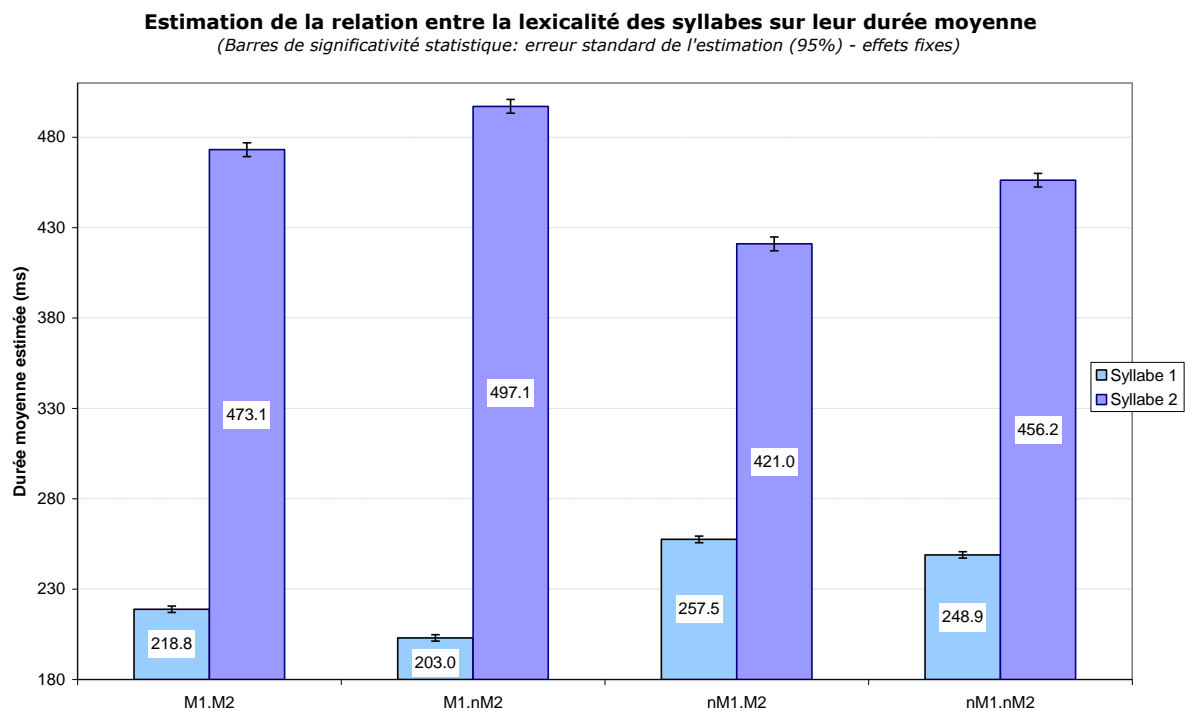


Figure 97 : Effet des facteurs expérimentaux sur la durée des syllabes



## Annexe 21

### Expérience 7 : Équation du modèle et tests de significativité de l'effet des covariables sur les temps de réponse

Équation 44 : Équation du modèle étudié

$$TR_{It, Suj} \sim N(XB, \Omega)$$

$$TR_{It, Suj} = \beta_{0It, Suj} \text{Cons} + -62.032(12.027) M1_{It, Suj} + 16.342(10.432) M2_{It, Suj} + -8.254(14.755) M1.M2_{It, Suj} + -8.401(20.027) TDL_{Suj} + 7.210(19.492) DerP_{Suj} + 1.829(11.253) M1.TDL_{It, Suj} + 10.015(10.886) M1.DerP_{It, Suj} + 12.147(11.371) M2.TDL_{It, Suj} + 5.806(11.015) M2.DerP_{It, Suj} + 47.460(28.040) TDL.DerP_{Suj} + -4.886(15.930) M1.M2.TDL_{It, Suj} + -12.529(15.364) M1.M2.DerP_{It, Suj} + -3.278(15.662) M1.TDL.DerP_{It, Suj} + -8.507(15.815) M2.TDL.DerP_{It, Suj} + 8.400(22.187) M1.M2.TDL.DerP_{It, Suj} + 70.699(23.913) Lat_{Suj} + 28.864(21.961) Sex_{Suj} + 0.155(0.030) Ordre_{It, Suj} + 1.756(0.147) D1c_{It, Suj} + 0.011(0.087) D2c_{It, Suj} + 4.175(0.509) f01c_{It, Suj} + 1.309(0.699) f02c_{It, Suj} + 3.044(0.841) dB1c_{It, Suj} + -0.160(0.787) dB2c_{It, Suj} + -5.875(1.605) AoA_{It, Suj} + -91.961(11.712) Fam_{It, Suj} + 22.590(2.822) FqS1_{It, Suj} + -21.166(4.016) FqS2_{It, Suj} + -2.864(4.307) Fq_{It, Suj} + -17.355(6.444) Nph_{It, Suj} + 0.136(0.057) PR_{It, Suj} + 0.068(0.042) DPU_{It, Suj} + 0.252(0.076) f01c .dB1c_{It, Suj} + 0.272(0.079) f02c .dB2c_{It, Suj} + 15.627(4.058) AoA .Fam_{It, Suj} + 8.617(2.004) AoA .Fq_{It, Suj} + 9.772(13.002) Fam .Fq_{It, Suj} + -9.248(3.518) AoA .Fam .Fq_{It, Suj} + 0.001(0.001) DPU .PR_{It, Suj} + -0.401(0.055) Nph .DPU_{It, Suj} + 0.107(0.060) Nph .PR_{It, Suj} + 0.000(0.000) Nph .DPU .PR_{It, Suj} + 0.000(0.001) D1c .DPU_{It, Suj} + -0.002(0.001) D1c .PR_{It, Suj} + 0.000(0.001) D2c .DPU_{It, Suj} + -0.004(0.001) D2c .PR_{It, Suj} + -0.005(0.001) D1c .D2c_{It, Suj} + 0.000(0.000) D1c .DPU .PR_{It, Suj} + 0.000(0.000) D2c .DPU .PR_{It, Suj} + 0.000(0.000) D1c .D2c .DPU .PR_{It, Suj} + -1.022(0.198) M1.D1c_{It, Suj} + -5.404(0.749) M1.f01c_{It, Suj} + -6.433(1.199) M1.dB1c_{It, Suj} + 0.196(0.067) M2.D2c_{It, Suj} + 5.586(1.562) M2.f02c_{It, Suj} + 5.225(1.321) M2.dB2c_{It, Suj} + -0.530(0.167) M1.f01c .dB1c_{It, Suj} + 0.115(0.144) M2.f02c .dB2c_{It, Suj} + 4.227(0.731) M1.M2.f01c_{It, Suj} + -2.215(1.740) M1.M2.f02c_{It, Suj} + -1.646(1.413) M1.M2.dB1c_{It, Suj} + -3.248(1.375) M1.M2.dB2c_{It, Suj} + -0.168(0.183) M1.M2.D1c_{It, Suj} + -0.431(0.085) M1.M2.D2c_{It, Suj} + -1.074(0.239) M1.M2.f01c .dB1c_{It, Suj} + -0.750(0.179) M1.M2.f02c .dB2c_{It, Suj}$$

$$\beta_{0It, Suj} = 765.581(25.263) + u_{0Suj} + e_{0It, Suj}$$

$$[u_{0Suj}] \sim N(0, \Omega_u) : \Omega_u = [2820.573(499.435)]$$

$$[e_{0It, Suj}] \sim N(0, \Omega_e) : \Omega_e = [10132.940(198.229)]$$

-2\*loglikelihood(IGLS Deviance) = 64093.800(5296 of 5600 cases in use)

Avec : TR = temps de réponse (V.D.) ; M<sub>1</sub> = première syllabe lexicale (V.I. - réf = nM<sub>1</sub>) ; M<sub>2</sub> = deuxième syllabe lexicale (V.I. - réf = nM<sub>2</sub>) ; TDL = tâche de décision lexicale classique (V.I. - réf = gng) ; DerP = pseudo-mots par le dernier phonème de la deuxième syllabe (V.I. - réf = derS) ; Lat = latéralité manuelle (réf. = G) ; Sex = sexe du participant (réf. = F) ; Ordre = ordre de présentation des items ; D<sub>1/2c</sub> = durée de la première/deuxième syllabe (ms.) ; f<sub>0,1/2c</sub> = fréquence fondamentale moyenne de la première/deuxième syllabe (Hz) ; dB<sub>1/2c</sub> = intensité moyenne de la première/deuxième syllabe (dB) ; FqS<sub>1/2</sub> = logarithme en base 10 de la fréquence formelle phonologique de la première/deuxième syllabe (fréquence où la séquence de phonème apparaît en tant que syllabe) ; Fq = logarithme en base 10 de la fréquence formelle phonologique du mot ; NPh = nombre de phonèmes composant le mot ; DPU = position acoustique du début du P.U. du mot (ms.) ; PR = position acoustique du point de reconnaissance du mot (ms.) ; AoA = âge d'acquisition estimé du mot ; Fam = familiarité conceptuelle du mot

Tableau 154 : Tests de significativité des covariables et de leurs interactions

	Coefficient	chi <sup>2</sup> (1, N=5296)	p<
Lat	70.699	8.741	.0032
Sex	28.864	1.728	.19

	Coefficient	chi <sup>2</sup> <sub>(1, N=5296)</sub>	p<
Ordr	0.155	26.415	<b>3<sup>E-07</sup></b>
D <sub>1</sub> c	1.756	143.062	<b>6<sup>E-33</sup></b>
D <sub>2</sub> c	0.011	0.017	.90
f <sub>0</sub> <sub>1</sub> c	4.175	67.231	<b>3<sup>E-16</sup></b>
f <sub>0</sub> <sub>2</sub> c	1.309	3.511	<b>.061</b>
dB <sub>1</sub> c	3.044	13.083	<b>.0003</b>
dB <sub>2</sub> c	-0.16	0.041	.84
AoA	-5.875	13.401	<b>.0003</b>
Fam	-91.961	61.65	<b>5<sup>E-15</sup></b>
FqS <sub>1</sub>	22.59	64.066	<b>2<sup>E-15</sup></b>
FqS <sub>2</sub>	-21.166	27.776	<b>2<sup>E-07</sup></b>
Fq	-2.864	0.442	.51
NPh	-17.355	7.253	<b>.0071</b>
PR	0.136	5.7	<b>.017</b>
DPU	0.068	2.61	.11
f <sub>0</sub> <sub>1</sub> c.dB <sub>1</sub> c	0.252	11.059	<b>.0009</b>
f <sub>0</sub> <sub>2</sub> c.dB <sub>2</sub> c	0.272	11.795	<b>.0006</b>
AoA.Fam	15.627	14.831	<b>.0002</b>
AoA.Fq	8.617	18.485	<b>2<sup>E-05</sup></b>
Fam.Fq	9.772	0.565	.46
AoA.Fam.Fq	-9.248	6.912	<b>.0086</b>
DPU.PR	0.001	4.118	<b>.043</b>
NPh.DPU	-0.401	52.964	<b>4<sup>E-13</sup></b>
NPh.PR	0.107	3.2	<b>.074</b>
NPh.DPU.PR	0	0.346	.56
D <sub>1</sub> c.DPU	0	0.005	.95
D <sub>1</sub> c.PR	-0.002	6.857	<b>.0089</b>
D <sub>2</sub> c.DPU	0	0.077	.79
D <sub>2</sub> c.PR	-0.004	19.832	<b>9<sup>E-06</sup></b>
D <sub>1</sub> c.D <sub>2</sub> c	-0.005	30.016	<b>5<sup>E-08</sup></b>
D <sub>1</sub> c.DPU.PR	0	5.728	<b>.017</b>
D <sub>2</sub> c.DPU.PR	0	1.852	.18
D <sub>1</sub> c.D <sub>2</sub> c.DPU.PR	0	154.042	<b>3<sup>E-35</sup></b>
M <sub>1</sub> .D <sub>1</sub> c	-1.022	26.522	<b>3<sup>E-07</sup></b>
M <sub>1</sub> .f <sub>0</sub> <sub>1</sub> c	-5.404	52.093	<b>6<sup>E-13</sup></b>
M <sub>1</sub> .dB <sub>1</sub> c	-6.433	28.785	<b>9<sup>E-08</sup></b>
M <sub>2</sub> .D <sub>2</sub> c	0.196	8.522	<b>.0036</b>
M <sub>2</sub> .f <sub>0</sub> <sub>2</sub> c	5.586	12.796	<b>.0004</b>
M <sub>2</sub> .dB <sub>2</sub> c	5.225	15.641	<b>8<sup>E-05</sup></b>
M <sub>1</sub> .f <sub>0</sub> <sub>1</sub> c.dB <sub>1</sub> c	-0.53	10.081	<b>.0015</b>
M <sub>2</sub> .f <sub>0</sub> <sub>2</sub> c.dB <sub>2</sub> c	0.115	0.631	.43
M <sub>1</sub> .M <sub>2</sub> .f <sub>0</sub> <sub>1</sub> c	4.227	33.452	<b>8<sup>E-09</sup></b>
M <sub>1</sub> .M <sub>2</sub> .f <sub>0</sub> <sub>2</sub> c	-2.215	1.619	.21
M <sub>1</sub> .M <sub>2</sub> .dB <sub>1</sub> c	-1.646	1.356	.25
M <sub>1</sub> .M <sub>2</sub> .dB <sub>2</sub> c	-3.248	5.577	<b>.019</b>
M <sub>1</sub> .M <sub>2</sub> .D <sub>1</sub> c	-0.168	0.848	.36
M <sub>1</sub> .M <sub>2</sub> .D <sub>2</sub> c	-0.431	25.586	<b>5<sup>E-07</sup></b>
M <sub>1</sub> .M <sub>2</sub> .f <sub>0</sub> <sub>1</sub> c.dB <sub>1</sub> c	-1.074	20.263	<b>7<sup>E-06</sup></b>
M <sub>1</sub> .M <sub>2</sub> .f <sub>0</sub> <sub>2</sub> c.dB <sub>2</sub> c	-0.75	17.496	<b>3<sup>E-05</sup></b>



## Annexe 22

### Expérience 8 : Matériel linguistique

Tableau 155 : Mots expérimentaux et pseudo-mots distracteurs

Matériel expérimental		Matériel distracteur
E	nE	E et nE
<u>j</u> eune	jaune	<u>dru</u> je
<u>c</u> ric	crac	<u>foi</u> ne
<u>l</u> uge	loge	<u>brin</u> re
<u>tra</u> ite	trente	<u>bon</u> pe
<u>to</u> ile	tuile	<u>chou</u> le
<u>t</u> ape	type	<u>rangu</u> e
<u>cas</u> se	cosse	<u>nide</u>
<u>sp</u> ire	spore	<u>pin</u> ve
<u>tron</u> che	tranche	<u>bru</u> pe
<u>tri</u> que	traque	nanle
<u>ble</u> tte	blatte	clige
<u>net</u> te	natte	cheupe
<u>pr</u> ise	prose	vreuk
<u>mo</u> ule	meule	reuze
<u>lou</u> pe	lampe	flinje
<u>son</u> ge	singe	jave
<u>ven</u> te	voûte	kinve
<u>pon</u> te	pente	fléne

Avec : E = mot avec enchâssement initial (nE pour les mots sans enchâssement initial) ; en souligné dans l'item = mot enchâssé (orthographique).

## Annexe 23

### Expérience 8 : Équation du modèle et tests de significativité de l'effet des covariables sur les temps de réponse

#### Équation 45 : Équation du modèle étudié

$$TR_{It, Suj} \sim N(XB, \Omega)$$

$$TR_{It, Suj} = \beta_{0It, Suj} \text{Cons} + -31.572(10.343)E_{It, Suj} + 7.283(16.394)ListIt_{Suj} + -79.328(40.202)Lat_{Suj} + 5.755(19.208)Sexe_{Suj} + -10.346(9.062)FqFrm_{It, Suj} + -6.740(21.927)NGRAPH_{It, Suj} + -78.502(111.219)NPHONS_{It, Suj} + 4.809(17.381)PUGRAPH_{It, Suj} + 58.128(70.923)PUPHONS_{It, Suj} + 4.800(58.923)G.H_{Suj} + 4.909(21.562)NGRAPH.PUGRAPH_{It, Suj} + 311.744(214.156)NPHONS.PUPHONS_{It, Suj} + 209.668(113.445)NGRAPH.NPHONS_{It, Suj} + 112.787(61.444)PUGRAPH.PUPHONS_{It, Suj} + -194.216(109.887)NGRAPH.PUPHONS_{It, Suj} + -157.610(66.419)NPHONS.PUGRAPH_{It, Suj} + -4.479(42.482)NGRAPH.NPHONS.PUGRAPH_{It, Suj} + 0.000(0.000)NGRAPH.NPHONS.PUPHONS_{It, Suj} + 0.000(0.000)NGRAPH.PUGRAPH.PUPHONS_{It, Suj} + 0.000(0.000)NPHONS.PUGRAPH.PUPHONS_{It, Suj} + 0.000(0.000)NGRAPH.NPHONS.PUGRAPH.PUPHONS_{It, Suj}$$

$$\beta_{0It, Suj} = 451.439(65.372) + u_{0Suj} + e_{0It, Suj}$$

$$[u_{0Suj}] \sim N(0, \Omega_u) : \Omega_u = [978.796(402.597)]$$

$$[e_{0It, Suj}] \sim N(0, \Omega_e) : \Omega_e = [7424.727(537.656)]$$

$$-2 * \text{loglikelihood}(IGLS \text{ Deviance}) = 4823.293(408 \text{ of } 468 \text{ cases in use})$$

Avec : TR = temps de réponse (V.D.) ; E = enchâssement (V.I. - réf = nE) ; ListIt = liste d'items ou groupe de participants (inutile - cf. note de bas de page n°105, page 201) ; Lat = latéralité manuelle des participants ; Sexe = sexe des participants ; FqFrm = fréquence formelle graphique (logarithme) ; NGRAPH = nombre de caractères composant le mot ; NPHONS = nombre de phonèmes composant le mot ; PUGRAPH = position du point d'unicité graphique (en nombre de caractères depuis le début du mot) ; PUPHONS = position du point d'unicité phonologique du mot (en nombre de phonèmes depuis le début du mot) ; G.H = latéralité \* sexe

Tableau 156 : Tests de significativité des covariables et de leurs interactions

	Coefficient	Chi <sup>2</sup> <sub>(1, N=408)</sub>	p<
ListIt	7.283	0.197	.66
Lat (1 = gauche)	-79.328	3.894	<b>.049</b>
Sexe	5.755	0.09	.77
FqFrm	-10.346	1.303	.26
NGRAPH	-6.74	0.094	.76
NPHONS	-78.502	0.498	.49
PUGRAPH	4.809	0.077	.79
PUPHONS	58.128	0.672	.42
Lat.Sexe	4.8	0.007	.94
NGRAPH.PUGRAPH	4.909	0.052	.82
NPHONS.PUPHONS	311.744	2.119	.15
NGRAPH.NPHONS	209.668	3.416	<b>.065</b>
PUGRAPH.PUPHONS	112.787	3.369	<b>.067</b>
NGRAPH.PUPHONS	-194.216	3.124	<b>.078</b>

NPHONS.PUGRAPH	-157.61	5.631	<b>.018</b>
NGRAPH.NPHONS.PUGRAPH	-4.479	0.011	.92
NGRAPH.NPHONS.PUPHONS	0	0	1
NGRAPH.PUGRAPH.PUPHONS	0	0	1
NPHONS.PUGRAPH.PUPHONS	0	0	1
NGRAPH.NPHONS.PUGRAPH.PUPHONS	0	0	1

## Annexe 24

### Expérience 9 : Équation du modèle et tests de significativité de l'effet des covariables sur les temps de réponse

Équation 46 : Équation du modèle étudié

$$\begin{aligned}
 TR_{It, Suj} &\sim N(XB, \Omega) \\
 TR_{It, Suj} &= \beta_{0It, Suj} \text{Cons} + 6.727(10.949) \text{Test}_{It, Suj} + -5.470(13.737) P2_{It, Suj} + -2.349(19.881) P3_{It, Suj} + \\
 &\quad -14.997(27.304) P4_{It, Suj} + 8.934(11.314) \text{Ench}_{It, Suj} + -8.588(11.144) \text{Init}_{It, Suj} + \\
 &\quad -13.129(15.479) \text{Test}.\text{Ench}_{It, Suj} + 18.453(15.372) \text{Test}.\text{Init}_{It, Suj} + 3.102(15.670) P2.\text{Ench}_{It, Suj} + \\
 &\quad 1.496(15.363) P2.\text{Init}_{It, Suj} + 1.916(16.027) P3.\text{Ench}_{It, Suj} + -19.299(15.781) P3.\text{Init}_{It, Suj} + \\
 &\quad -14.442(16.706) P4.\text{Ench}_{It, Suj} + -4.481(16.166) P4.\text{Init}_{It, Suj} + -5.054(15.832) \text{Ench}.\text{Init}_{It, Suj} + \\
 &\quad -10.833(21.642) \text{Test}.\text{Ench}.\text{Init}_{It, Suj} + 23.982(21.777) P2.\text{Ench}.\text{Init}_{It, Suj} + \\
 &\quad 18.885(22.167) P3.\text{Ench}.\text{Init}_{It, Suj} + 32.836(22.894) P4.\text{Ench}.\text{Init}_{It, Suj} + -0.140(0.034) \text{Ordr}_{It, Suj} + \\
 &\quad -0.082(0.065) \text{Pos}_{It, Suj} + 0.005(0.037) \text{Dur}_{It, Suj} \\
 \beta_{0It, Suj} &= 398.131(14.247) + u_{0Suj} + e_{0It, Suj}
 \end{aligned}$$

$$[u_{0Suj}] \sim N(0, \Omega_u) : \Omega_u = [6170.361(783.796)]$$

$$[e_{0It, Suj}] \sim N(0, \Omega_e) : \Omega_e = [10415.970(248.944)]$$

-2\*loglikelihood(IGLS Deviance) = 44428.750(3643 of 4200 cases in use)

Avec : TR = temps de réponse mesuré depuis le début de la cible en ms. (V.D.) ; Test = position des cibles insérées dans le signal acoustique correspondant à un mot ou un pseudo-mot ( $P_1$  à  $P_4$  - V.I. - réf =  $P_0$ ) ;  $P_2$  = position de la sonde à 33% de durée du mot, du début du mot (catégoriel - V.I. - réf =  $P_1$ ) ;  $P_3$  = position de la sonde à 66% de durée du mot, du début du mot (catégoriel - V.I. - réf =  $P_1$ ) ;  $P_4$  = position de la sonde à la fin du mot (catégoriel - V.I. - réf =  $P_1$ ) ; Ench = présence d'un enchâssement initial dans le mot stimulus (V.I. - réf =  $nE$ ) ; Init = le mot stimulus est enchâssable à l'initiale d'un autre mot (V.I. - réf =  $nI$ ) ; Ordr = ordre de présentation des items au participant ; Pos = position acoustique de la cible dans les items (en ms.) ; Dur = durée du mot stimulus (en ms.)

Tableau 157 : Tests de significativité des covariables et de leurs interactions

	Coefficient	Chi <sup>2</sup> <sub>(1, N=3643)</sub>	p<
Ordr	-0.14	16.841	5E-05
Pos	-0.082	1.584	.21
Dur	0.005	0.021	.89

## Annexe 25

### Expérience 10 : Équation du modèle et tests de significativité de l'effet des covariables sur les temps de réponse

Équation 47 : Équation du modèle étudié

$$TR_{It, Suj} \sim N(\mathbf{XB}, \Omega)$$

$$TR_{It, Suj} = \beta_{0It, Suj} \text{Cons} + -31.302(7.946) \text{Test}_{It, Suj} + 17.737(7.387) \text{M1}_{It, Suj} + \\ -0.409(7.291) \text{M2}_{It, Suj} + 1.452(7.156) \text{P2}_{It, Suj} + \\ 8.494(12.536) \text{P3}_{It, Suj} + -9.251(11.862) \text{Test.P2}_{It, Suj} + \\ -14.779(18.819) \text{Test.P3}_{It, Suj} + -14.162(10.281) \text{M1.M2}_{It, Suj} + \\ -23.199(10.297) \text{M1.P2}_{It, Suj} + -30.454(10.260) \text{M1.P3}_{It, Suj} + \\ -2.557(10.249) \text{M2.P2}_{It, Suj} + -2.871(10.354) \text{M2.P3}_{It, Suj} + \\ 19.462(14.496) \text{M1.M2.P2}_{It, Suj} + 27.806(14.451) \text{M1.M2.P3}_{It, Suj} + \\ 2.352(3.472) \text{Gr}_{Suj} + -0.994(21.911) \text{Lat}_{Suj} + -16.308(15.335) \text{Sex}_{Suj} + \\ -0.055(0.027) \text{Pos}_{It, Suj} + -0.007(0.024) \text{Dur}_{It, Suj} + \\ 0.000(0.000) \text{Pos} \quad .\text{Dur}_{It, Suj}$$

$$\beta_{0It, Suj} = 342.502(25.146) + u_{0Suj} + e_{0It, Suj}$$

$$[u_{0Suj}] \sim N(0, \Omega_u) : \Omega_u = [2512.439(411.502)]$$

$$[e_{0It, Suj}] \sim N(0, \Omega_e) : \Omega_e = [6267.669(129.361)]$$

**-2\*loglikelihood(IGLS Deviance) = 55570.540(4776 of 4860 cases in use)**

Avec : TR = temps de détection de la cible, mesurés depuis le début de la cible en ms. (V.D.) ; Test = position des cibles insérées dans le signal acoustique correspondant au mot (P<sub>1</sub> à P<sub>3</sub> - V.I. - réf = positions P<sub>1</sub>, P<sub>2</sub>, P<sub>3</sub> non test) ; M<sub>1</sub> = lexicalité de la première syllabe (V.I. - réf = nM<sub>1</sub>) ; M<sub>2</sub> = lexicalité de la deuxième syllabe (V.I. - réf = nM<sub>2</sub>) ; P<sub>2</sub> = deuxième position de la sonde (V.I. - réf = P<sub>1</sub>) ; P<sub>3</sub> = troisième position de la sonde (V.I. - réf = P<sub>1</sub>) ; Gr = liste expérimentale ; Lat = latéralité manuelle des participants (réf = G) ; Sex = sexe des participants (réf = F) ; Pos = position acoustique de la cible dans l'item (en ms.) ; Dur = durée de l'item (en ms.)

Tableau 158 : Tests de significativité des covariables et de leurs interactions

	Coefficient	Chi <sup>2</sup> <sub>(1, N=4776)</sub>	p<
Gr	2.352	0.459	.50
Lat	-0.994	0.002	.97
Sex	-16.308	1.131	.29
Pos	-0.055	4.054	<b>.045</b>
Dur	-0.007	0.08	.78
Pos.dur	0	1.152	.29

## Annexe 26

### Expérience 11 : Matériel linguistique

Tableau 159 : Pseudo-mots expérimentaux

R I				Fq I			
Fq F		R F		Fq F		R F	
item	cible	item	cible	item	cible	item	cible
dzybgle	gle	dzypkja	kja	dzyksty	sty	dzyktɤã	tɤã
dzygzõ	zõ	dzygdɤu	dɤu	dzypkli	kli	dzygbɤã	bɤã
dzygnal	nal	dzygzãp	zãp	dzykfɤa	fɤa	dzyktɤa	tɤa
dzygdaʒ	daʒ	dzygnɔp	nɔp	dzypkɤa	kɤa	dzykfle	fle
dzygzœɤ	zœɤ	dzygdap	dap	dzykspi	spi	dzygʒa	ʒa
dzygnad	nad	dzygzyl	zyl	dzykɤɔ	ɤɔ	dzygda	da
dzygzet	zet	dzygzɤɤ	zɤɤ	dzybglob	glob	dzygna	na
dzyktiv	tiv	dzygzõp	zõp	dzykspe	spe	dzypkɤ	kɤ
dzygzæɤ	zæɤ	dzykɤak	ɤak	dzyksti	sti	dzykko	ko
dzygzif	zif	dzygvɤal	vɤal	dzygʒɔ	ʒɔ	dzygzi	zi
dzygzaʒ	zaʒ	dzygzjas	zjas	dzykɔ	kɔ	dzykti	ti
dzygzãs	zãs	dzygzɤa	zɤa	dzyksta	sta	dzygze	ze
dzygzɤɤ	zɤɤ	dzygzjɔn	zjɔn	dzykstɤa	stɤa	dzygnæɤ	næɤ

Avec : R = rare ; Fq = fréquent ; I = position initiale ; F = position finale ; cible = syllabe à détecter dans l'item.

Tableau 160 : Pseudo-mots distracteurs

item	cible	item	cible
dzyktɤɤ	tɤɤ	dzygzãb	ʒãb
dzykʃɤf	ʃɤf	dzyksjãs	sjãs
dzykkoz	koz	dzykfam	fam
dzygzɤɤ	zɤɤ	dzykfas	fas
dzygnɔt	nɔt	dzykfɤɤ	fɤɤ
dzyksœɤ	sœɤ	dzygnɔɤ	nɔɤ
dzygmõd	mõd	dzykɤjɤ	ɤjɤ
dzykfɤaz	fɤaz	dzykfɤɤi	fɤɤi
dzykful	ful	dzykɤwa	ɤwa
dzykflœv	flœv	dzykfœj	fœj
dzykɤol	ɤol	dzykɤit	ɤit
dzygzãɤ	zãɤ	dzykfãs	fãs
dzykfij	fij	dzygdwa	dwa
dzykɤɤs	ɤɤs	dzybgɤup	gɤup
dzykɤje	ɤje	dzykflam	flam
dzykɤœɤ	ɤœɤ	dzygbwa	bwa
dzygnɤi	nɤi	dzygdut	dut
dzyksɔm	sɔm	dzygbuʃ	buʃ

<b>item</b>	<b>cible</b>	<b>item</b>	<b>cible</b>
dzykfot	fot	dzykfjě	fjě
dzyg3wa	3wa	dzygbɤa	bɤa
dzyktet	tet	dzygbɤɥi	bɤɥi
dzybglwaɤ	glwaɤ	dzyktɤup	tɤup
dzyktuɤ	tuɤ	dzykfis	fis
dzykpa3	pa3	dzyglɥ	ɥ
dzyglwa	lwa	dzygbwat	bwat
dzyrɤɔb	ɤɔb	dzyrɤɛv	ɤɛv

## Annexe 27

### Expérience 11 : Analyse acoustique des stimuli

Équation 48 : Estimation du rapport entre paramètres prosodiques (f0, intensité et durée) et facteurs expérimentaux

$$f0_{It} \sim N(\mathbf{XB}, \Omega)$$

$$f0_{It} = \beta_{0It} \mathbf{Cons} + -1.405(0.165) \text{FrqI}_{It} + 1.787(0.165) \text{FrqF}_{It} + \\ -1.121(0.234) \text{FrqI.FrqF}_{It}$$

$$\beta_{0It} = 118.352(0.117) + e_{0It}$$

$$[e_{0It}] \sim N(0, \Omega_e) : \Omega_e = [2.836(0.139)]$$

$$-2 * \log\text{likelihood(IGLS Deviance)} = 3228.456(832 \text{ of } 832 \text{ cases in use})$$

$$dB_{It} \sim N(\mathbf{XB}, \Omega)$$

$$dB_{It} = \beta_{0It} \mathbf{Cons} + -0.432(0.196) \text{FrqI}_{It} + 0.410(0.196) \text{FrqF}_{It} + \\ 0.170(0.277) \text{FrqI.FrqF}_{It}$$

$$\beta_{0It} = 72.218(0.138) + e_{0It}$$

$$[e_{0It}] \sim N(0, \Omega_e) : \Omega_e = [3.982(0.195)]$$

$$-2 * \log\text{likelihood(IGLS Deviance)} = 3510.847(832 \text{ of } 832 \text{ cases in use})$$

$$\text{Dur}_{It} \sim N(\mathbf{XB}, \Omega)$$

$$\text{Dur}_{It} = \beta_{0It} \mathbf{Cons} + -68.728(3.678) \text{FrqI}_{It} + -22.435(3.678) \text{FrqF}_{It} + \\ 11.019(5.202) \text{FrqI.FrqF}_{It}$$

$$\beta_{0It} = 548.826(2.601) + e_{0It}$$

$$[e_{0It}] \sim N(0, \Omega_e) : \Omega_e = [1407.208(68.994)]$$

$$-2 * \log\text{likelihood(IGLS Deviance)} = 8392.584(832 \text{ of } 832 \text{ cases in use})$$

Tableau 161 : Effet des facteurs expérimentaux sur les facteurs prosodiques de la syllabe

	f0		dB		Durée	
	Chi <sup>2</sup> <sub>(1, N=832)</sub>	p<	Chi <sup>2</sup> <sub>(1, N=832)</sub>	p<	Chi <sup>2</sup> <sub>(1, N=832)</sub>	p<
FrqI	72.42	2 <sup>E-17</sup>	4.86	.027	349.10	0
FrqF	117.09	3 <sup>E-27</sup>	4.39	.036	37.20	2 <sup>E-09</sup>
FrqI.FrqF	23.03	2 <sup>E-06</sup>	0.38	.54	4.49	.034



Figure 98 : Effet des facteurs expérimentaux sur la fréquence fondamentale moyenne des syllabes

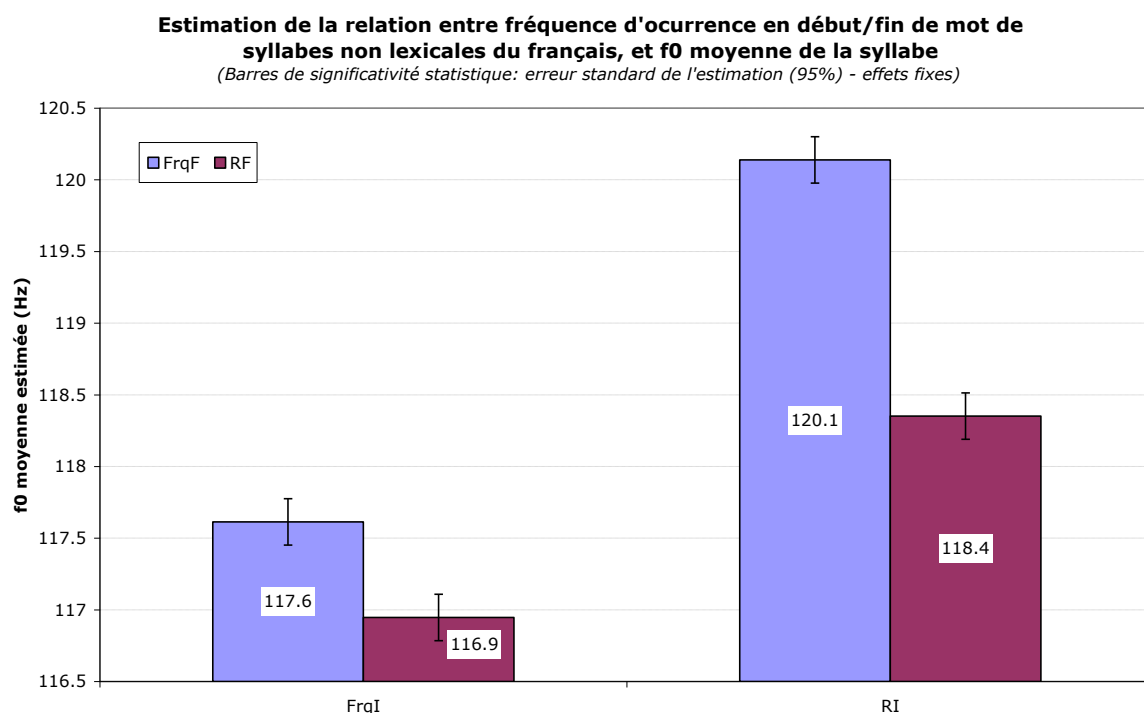


Figure 99 : Effet des facteurs expérimentaux sur l'intensité fondamentale moyenne des syllabes

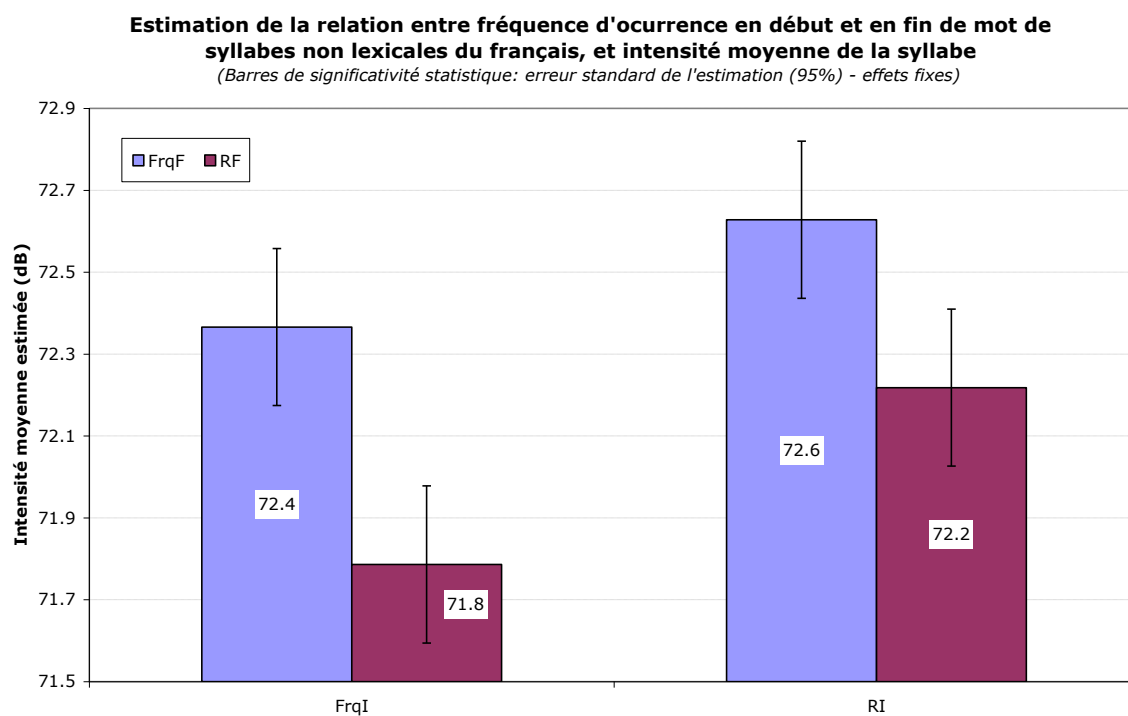
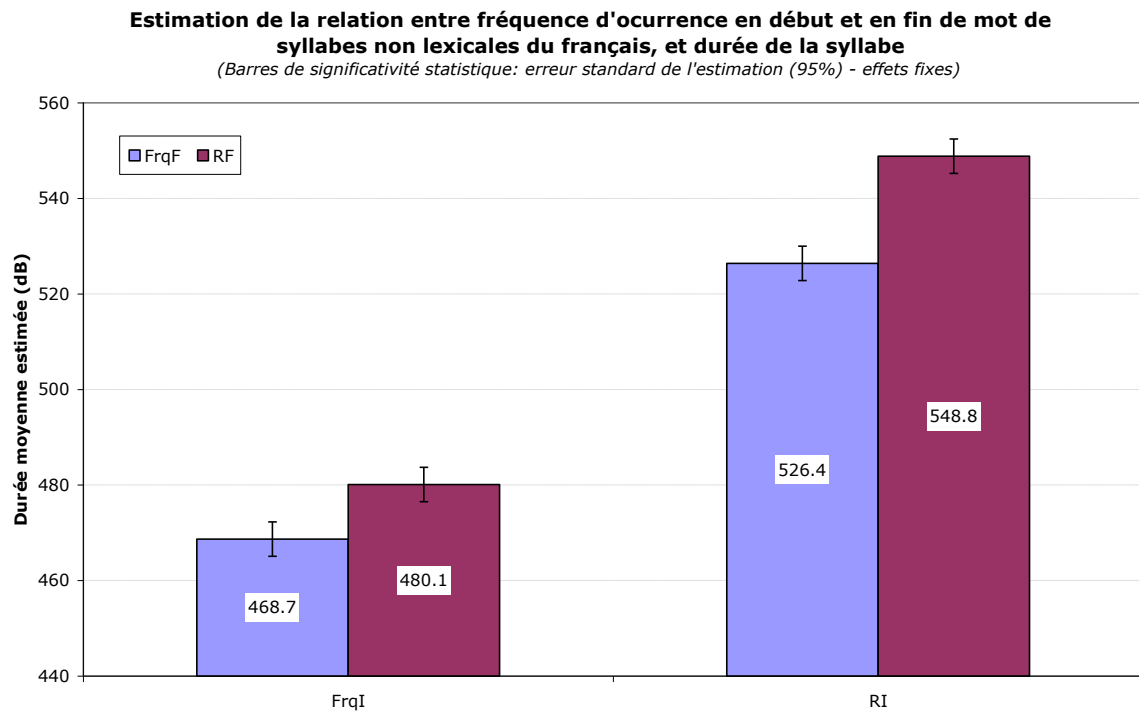


Figure 100 : Effet des facteurs expérimentaux sur la durée des syllabes



## Annexe 28

### Expérience 11 : Équation du modèle et tests de significativité de l'effet des covariables sur les temps de réponse

Équation 49 : Équation du modèle étudié

$$TR_{It, Suj} \sim N(\mathbf{XB}, \Omega)$$

$$TR_{It, Suj} = \beta_{0It, Suj} \mathbf{Cons} + \mathbf{-138.847(54.217)FrqI}_{It, Suj} + \mathbf{-69.555(46.996)FrqF}_{It, Suj} + \mathbf{82.900(71.224)FrqI.FrqF}_{It, Suj} + \mathbf{150.143(113.396)Lat}_{Suj} + \mathbf{48.049(70.426)Sex}_{Suj} + \mathbf{-0.522(0.179)Ordr}_{It, Suj} + \mathbf{45.827(12.498)f0}_{It, Suj} + \mathbf{-27.900(17.430)dB}_{It, Suj} + \mathbf{-2.951(0.864)Dur}_{It, Suj} + \mathbf{42.049(24.489)Npho}_{It, Suj} + \mathbf{4.098(2.332)NmoEv}_{It, Suj} + \mathbf{-10.268(5.968)f0} \cdot \mathbf{dB}_{It, Suj} + \mathbf{1.462(0.632)Dur} \cdot \mathbf{Npho}_{It, Suj} + \mathbf{-95.887(42.439)FrqI.f0}_{It, Suj} + \mathbf{31.105(29.913)FrqI.dB}_{It, Suj} + \mathbf{0.764(2.014)FrqI.Dur}_{It, Suj} + \mathbf{-40.290(21.339)FrqF.f0}_{It, Suj} + \mathbf{20.090(32.802)FrqF.dB}_{It, Suj} + \mathbf{4.038(1.062)FrqF.Dur}_{It, Suj} + \mathbf{-11.233(23.601)FrqI.f0} \cdot \mathbf{dB}_{It, Suj} + \mathbf{23.903(17.131)FrqF.f0} \cdot \mathbf{dB}_{It, Suj} + \mathbf{73.934(48.070)FrqI.FrqF.f0}_{It, Suj} + \mathbf{2.147(47.125)FrqI.FrqF.dB}_{It, Suj} + \mathbf{-1.002(1.981)FrqI.FrqF.Dur}_{It, Suj} + \mathbf{-6.012(28.883)FrqI.FrqF.f0} \cdot \mathbf{dB}_{It, Suj} + \mathbf{-0.426(0.225)f0} \cdot \mathbf{Dur}_{It, Suj} + \mathbf{0.170(0.299)dB} \cdot \mathbf{Dur}_{It, Suj} + \mathbf{0.040(0.172)f0} \cdot \mathbf{dB} \cdot \mathbf{Dur}_{It, Suj} + \mathbf{-2.411(1.897)FrqI.f0} \cdot \mathbf{Dur}_{It, Suj} + \mathbf{0.327(0.709)FrqI.dB} \cdot \mathbf{Dur}_{It, Suj} + \mathbf{-0.052(0.792)FrqI.f0} \cdot \mathbf{dB} \cdot \mathbf{Dur}_{It, Suj} + \mathbf{0.213(0.363)FrqF.f0} \cdot \mathbf{Dur}_{It, Suj} + \mathbf{0.212(0.731)FrqF.dB} \cdot \mathbf{Dur}_{It, Suj} + \mathbf{-0.609(0.357)FrqF.f0} \cdot \mathbf{dB} \cdot \mathbf{Dur}_{It, Suj} + \mathbf{2.197(1.927)FrqI.FrqF.f0} \cdot \mathbf{Dur}_{It, Suj} + \mathbf{-0.646(1.086)FrqI.FrqF.dB} \cdot \mathbf{Dur}_{It, Suj} + \mathbf{0.516(0.897)FrqI.FrqF.f0} \cdot \mathbf{dB} \cdot \mathbf{Dur}_{It, Suj}$$

$$\beta_{0It, Suj} = \mathbf{447.275(121.501)} + u_{0Suj} + e_{0It, Suj}$$

$$[u_{0Suj}] \sim N(0, \Omega_u) : \Omega_u = [\mathbf{8198.764(3041.316)}]$$

$$[e_{0It, Suj}] \sim N(0, \Omega_e) : \Omega_e = [\mathbf{18488.480(981.991)}]$$

**-2\*loglikelihood(IGLS Deviance) = 9229.176(725 of 832 cases in use)**

Avec : *TR* = temps de détection de la cible, mesurés depuis le début de la cible en ms. (V.D.) ; *FrqI* = fréquence d'occurrence de la séquence de phonèmes comme syllabe initiale dans le lexique (V.I. - réf = rare en initiale) ; *FrqF* = fréquence d'occurrence de la séquence de phonèmes comme syllabe finale dans le lexique (V.I. - réf = rare en fin) ; *Lat* = latéralité manuelle des participants (réf = gauche) ; *Sex* = sexe des participants (réf = féminin) ; *Ordr* = ordre de présentation des items ; *f0* = fréquence fondamentale moyenne de la syllabe cible (Hz) ; *dB* = intensité moyenne de la syllabe cible (dB) ; *Dur* = durée de la syllabe cible (ms.) ; *Npho* = nombre de phonèmes composant la syllabe cible ; *NmoEv* = nombre moyen de mots évoqués par la syllabe cible

Tableau 162 : Tests de significativité des covariables et de leurs interactions

	Coefficient	Chi <sup>2</sup> <sub>(1, N=725)</sub>	p<
Lat	150.143	1.753	.19
Sex	48.049	0.465	.50
Ordr	-0.522	8.503	<b>.0036</b>
f0	45.827	13.446	<b>.0003</b>
dB	-27.9	2.562	.11
Dur	-2.951	11.664	<b>.0007</b>
Npho	42.049	2.948	.086

	<b>Coefficient</b>	<b>Chi<sup>2</sup> (1, N=725)</b>	<b>p&lt;</b>
NmoEv	4.098	3.088	.079
f0.dB	-10.268	2.961	.086
Dur.Npho	1.462	5.344	.021
FrqI.f0	-95.887	5.105	.024
FrqI.dB	31.105	1.081	.30
FrqI.Dur	0.764	0.144	.71
FrqF.f0	-40.29	3.565	.06
FrqF.dB	20.09	0.375	.55
FrqF.Dur	4.038	14.457	.0002
FrqI.f0.dB	-11.233	0.227	.64
FrqF.f0.dB	23.903	1.947	.17
FrqI.FrqF.f0	73.934	2.366	.13
FrqI.FrqF.dB	2.147	0.002	.97
FrqI.FrqF.Dur	-1.002	0.256	.62
FrqI.FrqF.f0.dB	-6.012	0.043	.84
f0.Dur	-0.426	3.587	.059
dB.Dur	0.17	0.323	.57
f0.dB.Dur	0.04	0.053	.82
FrqI.f0.Dur	-2.411	1.615	.21
FrqI.dB.Dur	0.327	0.212	.65
FrqI.f0.dB.Dur	-0.052	0.004	.95
FrqF.f0.Dur	0.213	0.344	.56
FrqF.dB.Dur	0.212	0.084	.78
FrqF.f0.dB.Dur	-0.609	2.904	.089
FrqI.FrqF.f0.Dur	2.197	1.3	.26
FrqI.FrqF.dB.Dur	-0.646	0.354	.56
FrqI.FrqF.f0.dB.Dur	0.516	0.334	.57

## Annexe 29

### Expérience 12 : Matériel linguistique

Tableau 163 : Pseudo-mots expérimentaux et mots distracteurs

Matériel expérimental	Matériel distracteur		
	SyFr nInF	Illég	Mots
dap	dve	bɪʃ	gly
dɛu	dzy	buʃ	gɛe
gɥi	gvo	bɛðʃ	gɛi
ʒɛk	ʒda	byʃ	muʃ
kɥa	kvɔ	koz	pœɛ
lɥa	lɛœ	ʃɔk	fɔk
nɛp	mʒa	klã	pli
nɔp	nɛœ	klo	plõ
nɥo	nza	kɔd	plo
tɥo	nze	kɛã	ɛɛv
vɔn	sʃẽ	kɛø	ski
vɛal	tʃẽ	kɛim	sœɛ
zãp	vdœl	fɔt	sup
zjas	vda	fig	staz
zjɔ	vmy	flã	tɛas
zõp	zmy	fɛẽ	tɛi
zɥa	znɔ	gaf	tɛo
zyl	zɛyj	glã	ven

Avec : SyFr nInF = syllabe du Français non initiale et non finale ; Illég = « syllabe » illégale en Français.

## Annexe 30

### Expérience 12 : Analyse acoustique des stimuli

Équation 50 : Estimation du rapport entre paramètres prosodiques (f0, intensité et durée) et facteurs expérimentaux

$$f0 \text{ moy}_{It} \sim N(\mathbf{XB}, \Omega)$$

$$f0 \text{ moy}_{It} = \beta_{0It} \text{Cons} + 1.018(0.172) Fr_{It} + -0.354(0.199) F_{It}$$

$$\beta_{0It} = 107.281(0.099) + e_{0It}$$

$$[e_{0It}] \sim N(0, \Omega_e) : \Omega_e = [2.129(0.145)]$$

$$-2 * \log\text{likelihood(IGLS Deviance)} = 1552.307(432 \text{ of } 432 \text{ cases in use})$$

$$dB \text{ moy}_{It} \sim N(\mathbf{XB}, \Omega)$$

$$dB \text{ moy}_{It} = \beta_{0It} \text{Cons} + 0.378(0.288) Fr_{It} + -0.561(0.333) F_{It}$$

$$\beta_{0It} = 72.513(0.166) + e_{0It}$$

$$[e_{0It}] \sim N(0, \Omega_e) : \Omega_e = [5.986(0.407)]$$

$$-2 * \log\text{likelihood(IGLS Deviance)} = 1998.976(432 \text{ of } 432 \text{ cases in use})$$

$$\text{durée}_{It} \sim N(\mathbf{XB}, \Omega)$$

$$\text{durée}_{It} = \beta_{0It} \text{Cons} + -0.233(6.500) Fr_{It} + 16.736(7.506) F_{It}$$

$$\beta_{0It} = 435.753(3.753) + e_{0It}$$

$$[e_{0It}] \sim N(0, \Omega_e) : \Omega_e = [3042.083(206.988)]$$

$$-2 * \log\text{likelihood(IGLS Deviance)} = 4690.732(432 \text{ of } 432 \text{ cases in use})$$

Tableau 164 : Effet du facteur expérimental sur les facteurs prosodiques des syllabes

	f0		dB		Durée	
	Chi <sup>2</sup> <sub>(1, N=432)</sub>	p<	Chi <sup>2</sup> <sub>(1, N=432)</sub>	p<	Chi <sup>2</sup> <sub>(1, N=432)</sub>	p<
Syllabe fréquente	3.18	.074	2.84	.092	4.97	.026

Figure 101 : Effet du facteur expérimental sur la fréquence fondamentale des syllabes

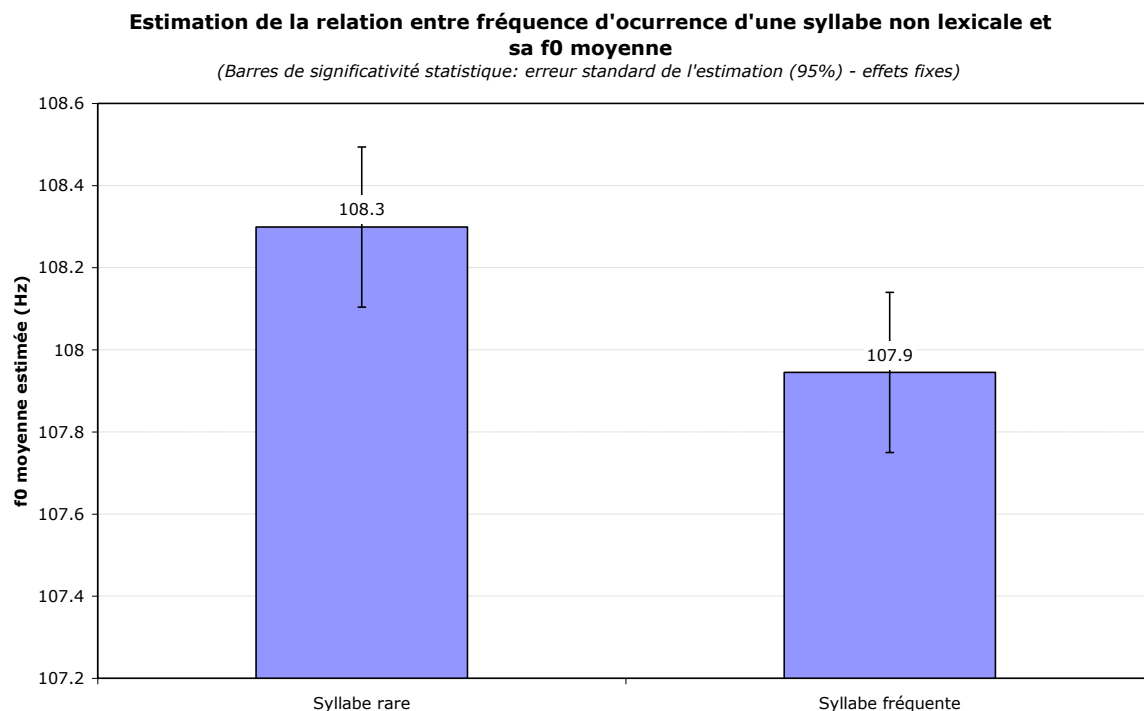


Figure 102 : Effet du facteur expérimental sur l'intensité moyenne des syllabes

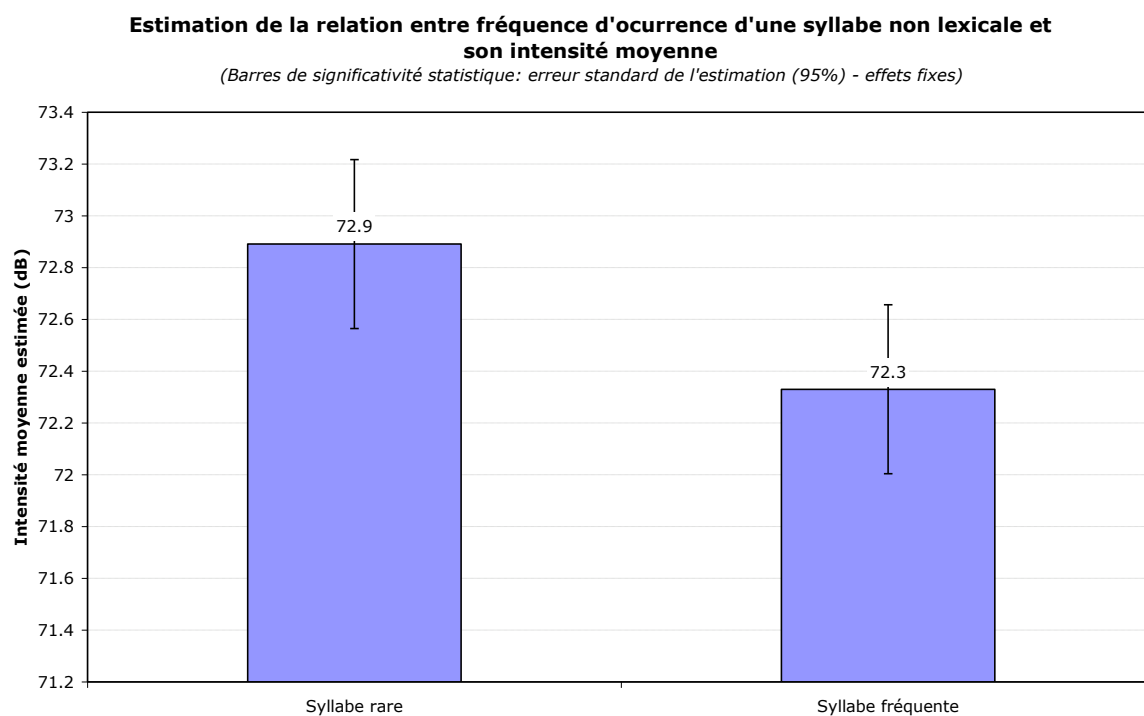
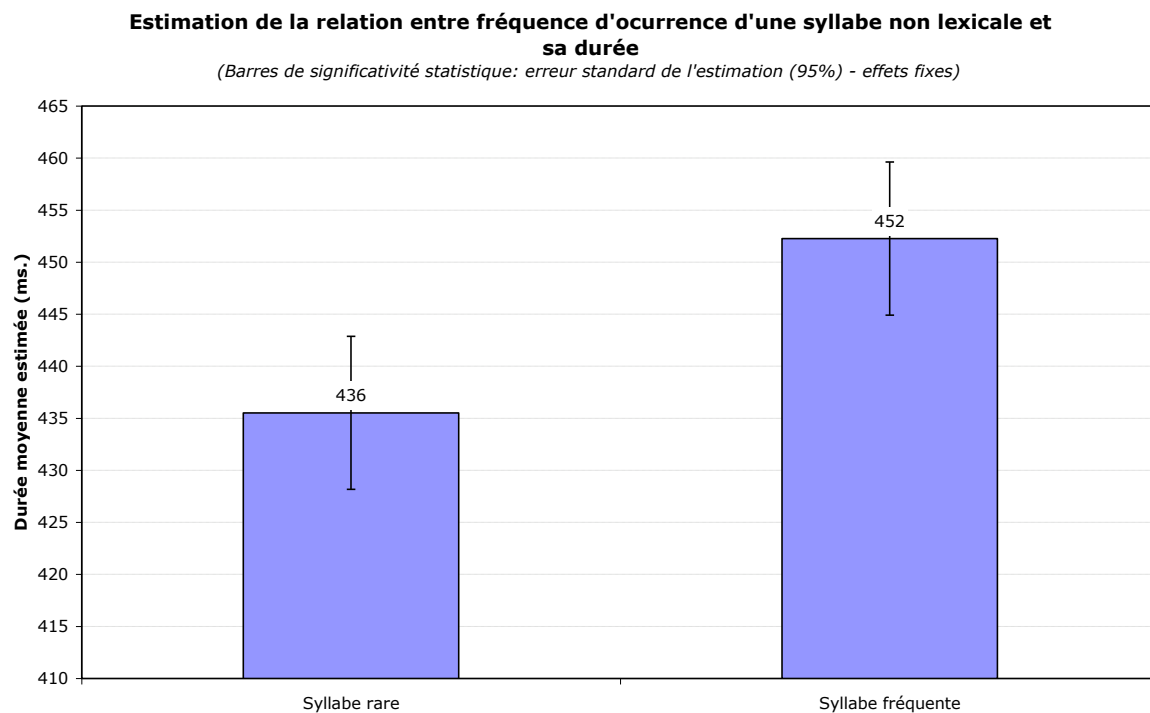


Figure 103 : Effet du facteur expérimental sur la durée des syllabes





## Annexe 31

### Expérience 12 : Équation du modèle et tests de significativité de l'effet des covariables sur les temps de réponse

Équation 51 : Équation du modèle étudié

$$TR_{ij} \sim N(XB, \Omega)$$

$$TR_{ij} = \beta_{0ij} \text{Cons} + -157.802(47.060)F_{ij} + -0.517(3.512)NmoEv_{ij} + \\ -2.545(0.941)dur_{ij} + 199.346(51.507)f0_{ij} + -20.087(12.390)dB_{ij} + \\ -4.731(0.979)dur.f0_{ij} + 0.215(0.244)dur.dB_{ij} + -89.525(18.562)f0.dB_{ij} + \\ 0.319(0.145)dur.f0.dB_{ij} + 2.385(0.988)F.dur_{ij} + -182.962(52.519)F.f0_{ij} + \\ -5.267(15.325)F.dB_{ij} + 5.710(1.181)F.dur.f0_{ij} + -1.548(0.982)F.dur.dB_{ij} + \\ 138.550(29.276)F.f0.dB_{ij} + -0.089(0.409)F.dur.f0.dB_{ij}$$

$$\beta_{0ij} = 824.620(33.161) + u_{0j} + e_{0ij}$$

$$[u_{0j}] \sim N(0, \Omega_u) : \Omega_u = [1333.208(864.601)]$$

$$[e_{0ij}] \sim N(0, \Omega_e) : \Omega_e = [10353.060(1168.376)]$$

$$-2 * \text{loglikelihood(IGLS Deviance)} = 2041.828(168 \text{ of } 198 \text{ cases in use})$$

Avec : TR = temps de détection de la syllabe cible (V.D.) ; F = Fréquence d'occurrence de la syllabe dans le lexique français (V.I. - réf = syllabe rare) ; dur = durée de la syllabe (ms.) ; f0 = fréquence fondamentale moyenne de la syllabe (Hz) ; dB = intensité moyenne de la syllabe (dB) ; NmoEv = nombre de mots évoqués par la syllabe

Tableau 165 : Tests de significativité des covariables et de leurs interactions

	Coefficient	Chi <sup>2</sup> <sub>(1, N=168)</sub>	p<
NmoEv	-0.517	0.022	.89
dur	-2.545	7.31	<b>.0069</b>
f0	199.346	14.979	<b>.0002</b>
dB	-20.087	2.628	.11
dur.f0	-4.731	23.351	<b>2<sup>E-06</sup></b>
dur.dB	0.215	0.775	.38
f0.dB	-89.525	23.263	<b>2<sup>E-06</sup></b>
dur.f0.dB	0.319	4.831	<b>.028</b>
F.dur	2.385	5.829	<b>.016</b>
F.f0	-182.962	12.136	<b>.0005</b>
F.dB	-5.267	0.118	.74
F.dur.f0	5.71	23.379	<b>2<sup>E-06</sup></b>
F.dur.dB	-1.548	2.484	.12
F.f0.dB	138.55	22.397	<b>3<sup>E-06</sup></b>
F.dur.f0.dB	-0.089	0.047	.83

## Annexe 32

### Expérience 12 : Analyses supplémentaires pour départager l'origine de l'effet : lié à la syllabe ou lié à la séquence de phonèmes ?

Équation 52 : Effet de la séquence de phonèmes en tant que syllabe sur les temps de réponse

$$TR_{It, Suj} \sim N(\mathbf{XB}, \Omega)$$

$$TR_{It, Suj} = \beta_{0It, Suj} \text{Cons} + 4.317(3.565) \text{NmoEv}_{It, Suj} + 0.593(0.252) \text{durée}_{It, Suj} + 0.286(11.099) \text{f0 moy}_{It, Suj} + -4.348(5.632) \text{dB moy}_{It, Suj} + 0.725(0.250) \text{durée.f0 moy}_{It, Suj} + 0.217(0.203) \text{durée.dB moy}_{It, Suj} + 2.778(6.184) \text{f0 moy.dB moy}_{It, Suj} + 0.085(0.103) \text{durée.f0 moy.dB moy}_{It, Suj} + -8.936(5.679) \text{syl Nmots}_{It, Suj} + -9.366(28.110) \text{frq syl}_{It, Suj} + 10.103(3.731) \text{syl Nmots.frq syl}_{It, Suj}$$

$$\beta_{0It, Suj} = 675.346(23.895) + u_{0Suj} + e_{0It, Suj}$$

$$[u_{0Suj}] \sim N(0, \Omega_u) : \Omega_u = [1182.055(878.790)]$$

$$[e_{0It, Suj}] \sim N(0, \Omega_e) : \Omega_e = [13113.650(1479.829)]$$

$$-2 * \log\text{likelihood(IGLS Deviance)} = 2079.115(168 \text{ of } 198 \text{ cases in use})$$

Tableau 166 : Tests de significativité des facteurs expérimentaux et des covariables

	Coefficient	Chi <sup>2</sup> <sub>(1, N=168)</sub>	p<
NmoEv	4.317	1.467	.23
durée	0.593	5.514	<b>.019</b>
f0 moy	0.286	0.001	.98
dB moy	-4.348	0.596	.45
durée.f0 moy	0.725	8.425	<b>.0038</b>
durée.dB moy	0.217	1.153	.29
f0 moy.dB moy	2.778	0.202	.66
durée.f0 moy.dB moy	0.085	0.669	.42
syl Nmots	-8.936	2.476	.12
frq syl	-9.366	0.111	.74
syl Nmots.frq syl	10.103	7.332	<b>.0068</b>

Équation 53 : Effet de la séquence de phonèmes en tant que syllabe ou non sur les temps de répons

$$TR_{It, Suj} \sim N(\mathbf{XB}, \Omega)$$

$$\begin{aligned} TR_{It, Suj} = & \beta_{0It, Suj} \text{Cons} + 5.288(4.045) NmoEv_{It, Suj} + 0.405(0.270) \text{durée}_{It, Suj} + \\ & -2.394(10.982) f0 \text{ moy}_{It, Suj} + -3.788(5.725) dB \text{ moy}_{It, Suj} + \\ & 0.407(0.252) \text{durée} \cdot f0 \text{ moy}_{It, Suj} + \\ & -0.053(0.186) \text{durée} \cdot dB \text{ moy}_{It, Suj} + \\ & 4.942(6.308) f0 \text{ moy} \cdot dB \text{ moy}_{It, Suj} + \\ & -0.040(0.096) \text{durée} \cdot f0 \text{ moy} \cdot dB \text{ moy}_{It, Suj} + \\ & -1.821(2.108) \text{seq} N \text{ mots}_{It, Suj} + 10.221(23.791) \text{frq} \text{ seq}_{It, Suj} + \\ & 3.498(2.088) \text{seq} N \text{ mots} \cdot \text{frq} \text{ seq}_{It, Suj} \end{aligned}$$

$$\beta_{0It, Suj} = 688.785(23.363) + u_{0Suj} + e_{0It, Suj}$$

$$[u_{0Suj}] \sim N(0, \Omega_u) : \Omega_u = [1144.928(887.575)]$$

$$[e_{0It, Suj}] \sim N(0, \Omega_e) : \Omega_e = [14050.660(1586.084)]$$

$$-2 * \log \text{likelihood}(\text{IGLS Deviance}) = 2090.084(168 \text{ of } 198 \text{ cases in use})$$

Tableau 167 : Tests de significativité des facteurs expérimentaux et des covariables

	Coefficient	Chi <sup>2</sup> <sub>(1, N=168)</sub>	p<
NmoEv	5.288	1.709	.20
durée	0.405	2.242	.14
f0 moy	-2.394	0.048	.83
dB moy	-3.788	0.438	.51
durée.f0 moy	0.407	2.608	.11
durée.dB moy	-0.053	0.081	.78
f0 moy.dB moy	4.942	0.614	.44
durée.f0 moy.dB moy	-0.04	0.178	.68
seq Nmots	-1.821	0.746	.39
frq seq	10.221	0.185	.67
seq Nmots.frq seq	3.498	2.807	.094

La différence entre le degré d'adéquation aux données de deux modèles non emboîtés est renseignée par l'indice AIC <sup>note 123, page 332</sup>. Le calcul de cet indice est simplifié lorsque les 2 modèles sont de structure identique (différence entre les 2 indices « IGLS Deviance ») : 2090.084 - 2079.115 = 10.969. Cet indice laisse penser que le modèle relatif à la syllabe est plus explicatif que le modèle relatif à la séquence de phonèmes.

## Annexe 33

### Expérience 13 : Matériel linguistique

Tableau 168 : Mots expérimentaux et pseudo-mots distracteurs

Matériel expérimental				Matériel distracteur			
Fq I		R I		Fq I R F	Fq I Fq F	R I Fq F	R I R F
plō	fᵛā	min	pjɛɤ	sty	bᵛā	zō	kja
syk	pas	dyk	lwa	kli	tᵛa	nal	zāp
gᵛo	bᵛa	sak	fᵛō	fᵛa	fle	daʒ	nɔp
mus	fuk	saʒ	ᵛwa	kᵛa	ʒa	zœɤ	dap
bas	fo	kle	fe	spi	da	nad	zyl
suk	fu	ᵛys	ʒwa	glō	na	zet	zuk
bjē	gaɤ	mɔd	fis	spe	kē	tiv	zōp
gᵛā	gᵛi	sɔm	pyɤ	sti	ko	zaɤ	pᵛak
dᵛa	pla	flœɤ	nɥi	ʒo	zi	zif	vᵛal
tᵛu	tᵛās	dwa	tet	kɔ	ti	zaʒ	zjas
kɔs	fy	fᵛwa	mo	sta	ze	zās	zɥa
pᵛes	pᵛe	tᵛē	ʒuk	stᵛa	naɤ	zyɤ	zjɔn

Avec : Fq = fréquent ; R = rare. I = position initiale ; F = position finale.

## Annexe 34

### Expérience 13 : Analyse acoustique des stimuli

Équation 54 : Estimation du rapport entre paramètres prosodiques (f0, intensité et durée) et le facteur expérimental

$$f0_{It} \sim N(\mathbf{XB}, \Omega)$$

$$f0_{It} = \beta_{0It} \mathbf{cons} + 4.735(0.785) \text{FrqI}_{It}$$

$$\beta_{0It} = 146.510(0.555) + e_{0It}$$

$$[e_{0It}] \sim N(0, \Omega_e) : \Omega_e = [110.842(5.842)]$$

$$-2 * \log\text{likelihood(IGLS Deviance)} = 5433.105(720 \text{ of } 720 \text{ cases in use})$$

$$dB_{It} \sim N(\mathbf{XB}, \Omega)$$

$$dB_{It} = \beta_{0It} \mathbf{cons} + 1.210(0.366) \text{FrqI}_{It}$$

$$\beta_{0It} = 69.211(0.259) + e_{0It}$$

$$[e_{0It}] \sim N(0, \Omega_e) : \Omega_e = [24.084(1.269)]$$

$$-2 * \log\text{likelihood(IGLS Deviance)} = 4333.996(720 \text{ of } 720 \text{ cases in use})$$

$$\text{Durée}_{It} \sim N(\mathbf{XB}, \Omega)$$

$$\text{Durée}_{It} = \beta_{0It} \mathbf{cons} + -15.127(8.243) \text{FrqI}_{It}$$

$$\beta_{0It} = 490.873(5.829) + e_{0It}$$

$$[e_{0It}] \sim N(0, \Omega_e) : \Omega_e = [12230.010(644.578)]$$

$$-2 * \log\text{likelihood(IGLS Deviance)} = 8819.658(720 \text{ of } 720 \text{ cases in use})$$

Tableau 169 : Rapport entre la fréquence d'occurrence d'une syllabe lexicale à l'initiale des mots dans le lexique, et les paramètres prosodiques de la syllabe produite

	f0		dB		Durés	
	Chi <sup>2</sup> (1, N=720)	p<	Chi <sup>2</sup> (1, N=720)	p<	Chi <sup>2</sup> (1, N=720)	p<
FrqI	36.41	2 <sup>E-09</sup>	10.94	1 <sup>E-03</sup>	3.37	.067

Figure 104 : Rapport entre le facteur expérimental et la fréquence fondamentale de la syllabe produite

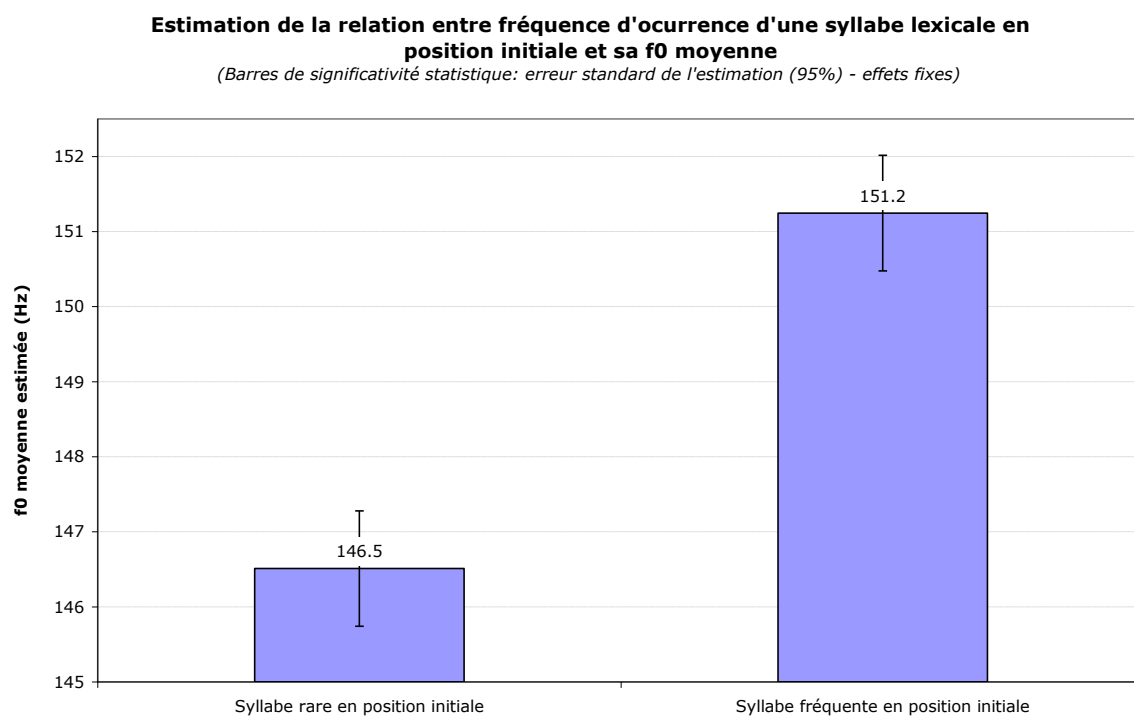


Figure 105 : Rapport entre le facteur expérimental et l'intensité de la syllabe produite

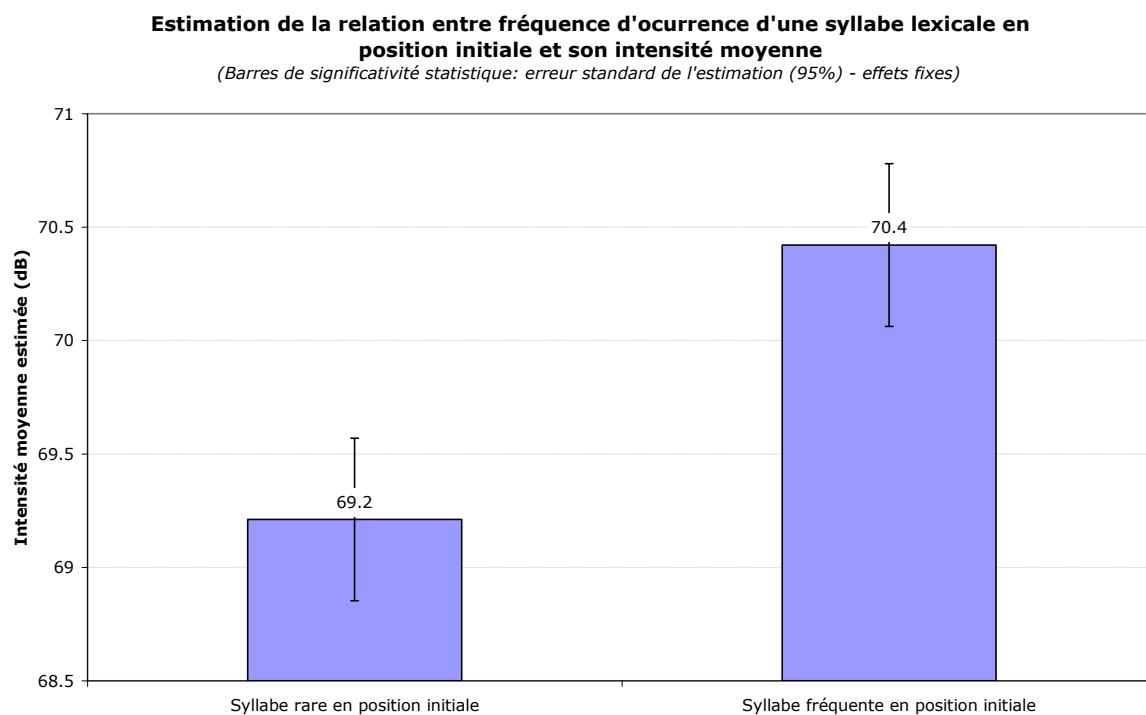
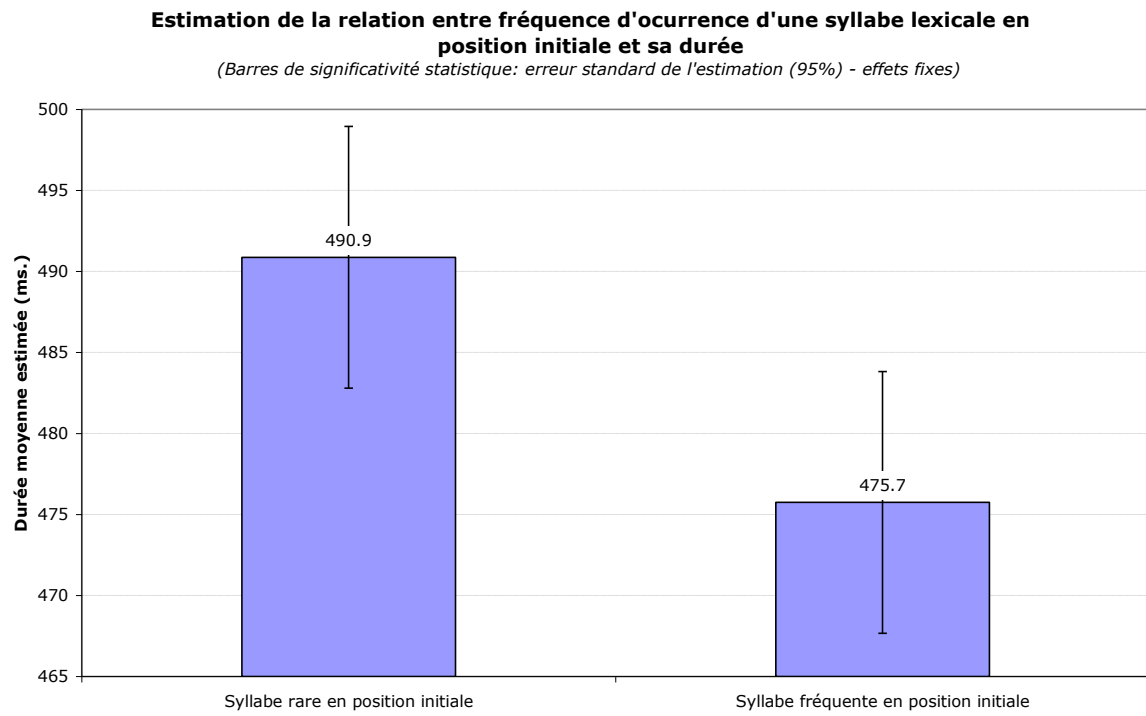


Figure 106 : Rapport entre le facteur expérimental et la durée de la syllabe produite



## Annexe 35

### Expérience 13 : Équation du modèle et tests de significativité de l'effet des covariables sur les temps de réponse

Équation 55 : Équation du modèle étudié

$$TR_{It, Suj} \sim N(XB, \Omega)$$

$$TR_{It, Suj} = \beta_{0It, Suj} \text{cons} + 28.629(23.838)Fi_{It, Suj} + -2.347(1.563)f0_{It, Suj} + 1.874(2.852)dB_{It, Suj} + 0.558(0.146)Dur_{It, Suj} + -0.296(0.216)PR_{It, Suj} + 61.050(44.680)FrLex_{It, Suj} + 14.087(5.193)AoA_{It, Suj} + -29.987(36.757)Fam_{It, Suj} + 22.373(52.149)Lat_{Suj} + -49.184(44.251)sex_{Suj} + 0.094(0.169)Ordr_{It, Suj} + 0.501(0.195)f0_{It, Suj} \cdot dB_{It, Suj} + 0.005(0.014)f0_{It, Suj} \cdot Dur_{It, Suj} + 0.072(0.036)dB_{It, Suj} \cdot Dur_{It, Suj} + -0.001(0.003)f0_{It, Suj} \cdot dB_{It, Suj} \cdot Dur_{It, Suj} + -0.648(1.997)Fi \cdot f0_{It, Suj} + -11.433(4.088)Fi \cdot dB_{It, Suj} + -0.005(0.156)Fi \cdot Dur_{It, Suj} + 0.497(0.567)Fi \cdot f0_{It, Suj} \cdot dB_{It, Suj} + 0.016(0.014)Fi \cdot f0_{It, Suj} \cdot Dur_{It, Suj} + -0.044(0.035)Fi \cdot dB_{It, Suj} \cdot Dur_{It, Suj} + -0.010(0.005)Fi \cdot f0_{It, Suj} \cdot dB_{It, Suj} \cdot Dur_{It, Suj} + 0.001(0.001)Dur_{It, Suj} \cdot PR_{It, Suj} + -5.456(5.986)FrLex_{It, Suj} \cdot AoA_{It, Suj} + -80.138(61.172)FrLex_{It, Suj} \cdot Fam_{It, Suj} + -20.061(11.686)AoA_{It, Suj} \cdot Fam_{It, Suj} + 71.159(15.539)AoA_{It, Suj} \cdot Fam_{It, Suj} \cdot FrLex_{It, Suj} + -27.155(22.092)FrForm Non Init_{It, Suj} + -24.872(46.288)Fi \cdot FrLex_{It, Suj} + 22.019(25.999)Fi \cdot FrForm Non Init_{It, Suj} + 68.919(50.921)FrLex_{It, Suj} \cdot FrForm Non Init_{It, Suj} + -108.336(53.655)Fi \cdot FrLex_{It, Suj} \cdot FrForm Non Init_{It, Suj}$$

$$\beta_{0It, Suj} = 698.072(46.542) + u_{0Suj} + e_{0It, Suj}$$

$$[u_{0Suj}] \sim N(0, \Omega_u) : \Omega_u = [5407.436(2084.980)]$$

$$[e_{0It, Suj}] \sim N(0, \Omega_e) : \Omega_e = [13667.970(751.263)]$$

-2\**loglikelihood*(IGLS Deviance) = 8412.225(677 of 720 cases in use)

Avec : TR = temps de reconnaissance du mot (V.D.) ; FrqI = fréquence d'occurrence de la syllabe en initiale de mot (V.I. - réf = rare) ; f0 = fréquence fondamentale moyenne de la syllabe (Hz) ; dB = intensité moyenne de la syllabe (dB) ; Dur = durée de la syllabe (ms.) ; PR = position acoustique du point de reconnaissance (ms.) ; FrLex = fréquence lexicale de la syllabe (logarithme en base 10) ; AoA = âge moyen estimé d'acquisition du mot ; Fam = familiarité conceptuelle du mot ; Lat = latéralité manuelle des participants ; Sex = sexe des participants ; Ordr = ordre de présentation des items ; FrForm Non Init = fréquence formelle de la syllabe en position non initiale dans le lexique (logarithme en base 10)

Tableau 170 : Tests de significativité des covariables et de leurs interactions

	Coefficient	Chi <sup>2</sup> <sub>(1, N=677)</sub>	p<
f0	-2.347	2.254	.14
dB	1.874	0.432	.52



	<b>Coefficient</b>	<b>Chi<sup>2</sup><sub>(1, N=677)</sub></b>	<b>p&lt;</b>
Dur	0.558	14.553	<b>2<sup>E-04</sup></b>
PR	-0.296	1.872	.18
FrLex	61.05	1.867	.18
AoA	14.087	7.36	<b>.0067</b>
Fam	-29.987	0.666	.42
Lat	22.373	0.184	.67
sex	-49.184	1.235	.27
Ordr	0.094	0.308	.58
f0.dB	0.501	6.61	<b>.011</b>
f0.Dur	0.005	0.126	.73
dB.Dur	0.072	4	<b>.046</b>
f0.dB.Dur	-0.001	0.067	.80
fi.f0	-0.648	0.105	.75
fi.dB	-11.433	7.821	<b>.0052</b>
fi.Dur	-0.005	0.001	.98
fi.f0.dB	0.497	0.768	.39
fi.f0.Dur	0.016	1.31	.26
fi.dB.Dur	-0.044	1.593	.21
fi.f0.dB.Dur	-0.01	4.552	<b>.033</b>
Dur.PR	0.001	0.988	.33
FrLex.AoA	-5.456	0.831	.37
FrLex.Fam	-80.138	1.716	.20
AoA.Fam	-20.061	2.947	<b>.087</b>
FrLex.AoA.Fam	71.159	20.97	<b>5<sup>E-06</sup></b>
FrForm Non Init	-27.155	1.511	.22
Fi.FrLex	-24.872	0.289	.60
Fi.FrForm Non Init	22.019	0.717	.40
FrLex.FrForm Non Init	68.919	1.832	.18
Fi.FrLex.FrForm Non Init	-108.336	4.077	<b>.044</b>

## Annexe 36

### Expérience 14 : Matériel linguistique

Tableau 171 : Mots expérimentaux et pseudo-mots distracteurs

Matériel expérimental				Matériel distracteur			
Fq nI		R nI		RI Fq F	Fq I Fq F	Fq I R F	R I R F
kɛs	pjō	fas	bwa	zō	bɔ̃	sty	kja
ʒwɛ̃	sō	plas	gut	nal	tɕa	kli	zāp
ble	ljō	bɕɥi	ɕiʃ	daʒ	fle	fɕa	nɔp
dɛ̃	plo	fœj	ɕuʒ	zœɕ	ʒa	kɕa	dap
les	ɕɛ̃	vɕɛ	klo	nad	da	spi	zyl
ɕit	lo	ʃwa	klas	zet	na	glɔ	zuɕ
vo	tas	fɕɥi	pje	tiv	kɛ̃	spe	zōp
lās	tɔn	kɕɛ̃	nwaɕ	zaɕ	ko	sti	pɕak
kyɕ	len	sut	fam	zif	zi	ʒɔ	vɕal
tɕak	nɛɕ	ʃās	lād	zaʒ	ti	kɔ	zjas
sil	tjɛɕ	pāt	koz	zās	ze	sta	zɥa
nɛ̃	ni	nɔɕ	ɕoz	zyɕ	naɕ	stɕa	zjɔn

Avec : Fq = fréquent ; R = rare. nI = position non initiale (I pour position initiale) ; F = position finale.

## Annexe 37

### Expérience 14 : Analyse acoustique des stimuli

Équation 56 : Estimation du rapport entre paramètres prosodiques (f0, intensité et durée) et le facteur expérimental

$$f0_{It} \sim N(\mathbf{XB}, \Omega)$$

$$f0_{It} = \beta_{0It} \mathbf{Cons} + -1.344(0.624) \text{FrFIN}_{It}$$

$$\beta_{0It} = 146.885(0.441) + e_{0It}$$

$$[e_{0It}] \sim N(0, \Omega_e) : \Omega_e = [79.391(3.930)]$$

$$-2 * \log\text{likelihood(IGLS Deviance)} = 5885.207(816 \text{ of } 816 \text{ cases in use})$$

$$dB_{It} \sim N(\mathbf{XB}, \Omega)$$

$$dB_{It} = \beta_{0It} \mathbf{Cons} + 1.496(0.333) \text{FrFIN}_{It}$$

$$\beta_{0It} = 69.782(0.235) + e_{0It}$$

$$[e_{0It}] \sim N(0, \Omega_e) : \Omega_e = [22.616(1.120)]$$

$$-2 * \log\text{likelihood(IGLS Deviance)} = 4860.540(816 \text{ of } 816 \text{ cases in use})$$

$$\text{Durée}_{It} \sim N(\mathbf{XB}, \Omega)$$

$$\text{Durée}_{It} = \beta_{0It} \mathbf{Cons} + -42.037(6.458) \text{FrFIN}_{It}$$

$$\beta_{0It} = 525.105(4.566) + e_{0It}$$

$$[e_{0It}] \sim N(0, \Omega_e) : \Omega_e = [8506.918(421.155)]$$

$$-2 * \log\text{likelihood(IGLS Deviance)} = 9699.395(816 \text{ of } 816 \text{ cases in use})$$

Tableau 172 : Rapport ente la fréquence d'occurrence d'une syllabe lexicale en fin de mots dans le lexique, et les paramètres prosodiques de la syllabe

	f0		dB		Durée	
	Chi <sup>2</sup> (1, N=816)	p<	Chi <sup>2</sup> (1, N=816)	p<	Chi <sup>2</sup> (1, N=816)	p<
Frq non initiale	4.64	.032	20.19	8 <sup>E-06</sup>	42.38	8 <sup>E-11</sup>

Figure 107 : Rapport entre le facteur expérimental et la fréquence fondamentale moyenne de la syllabe

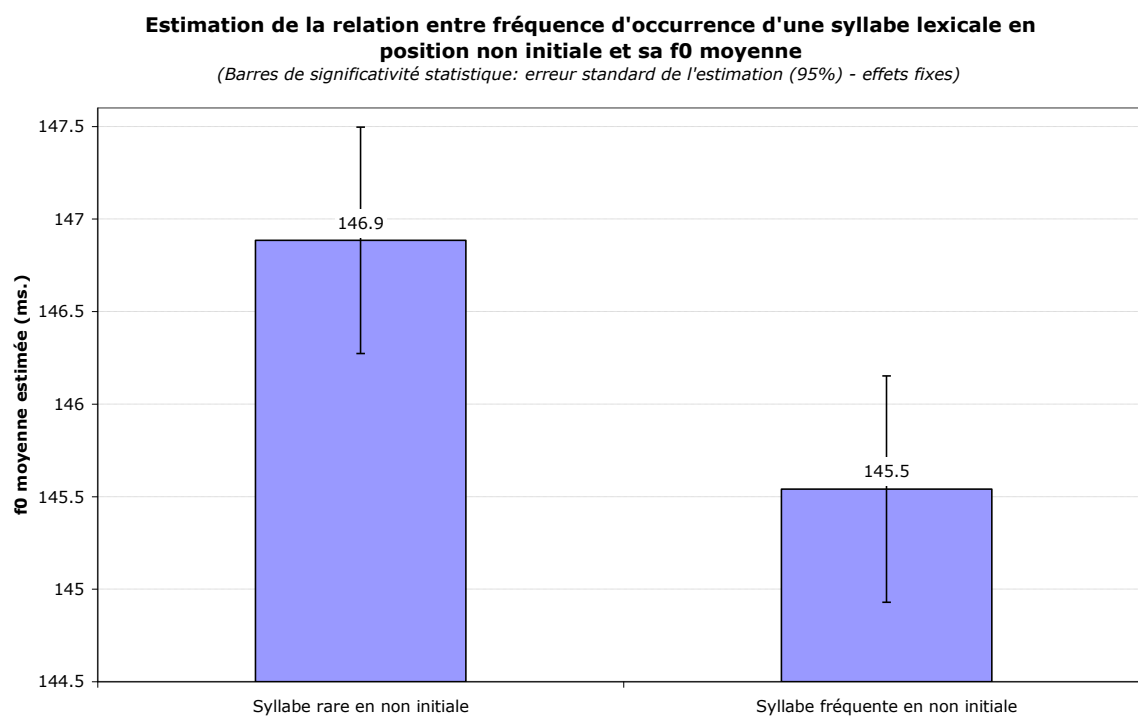


Figure 108 : Rapport entre le facteur expérimental et l'intensité moyenne de la syllabe

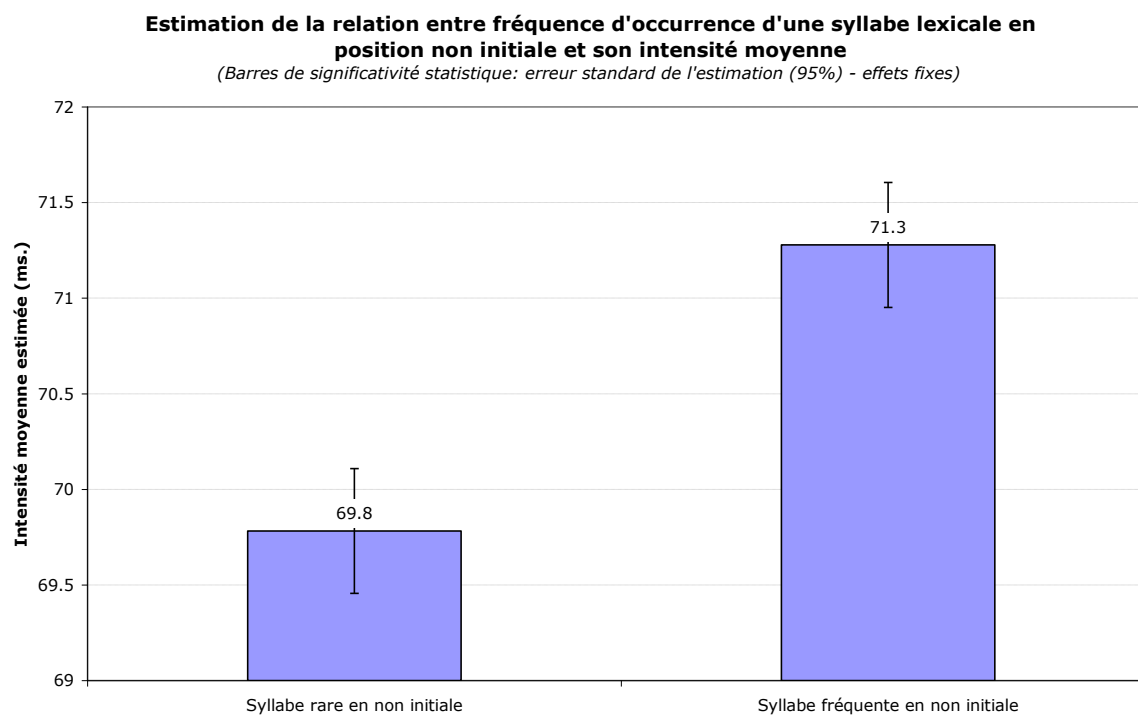
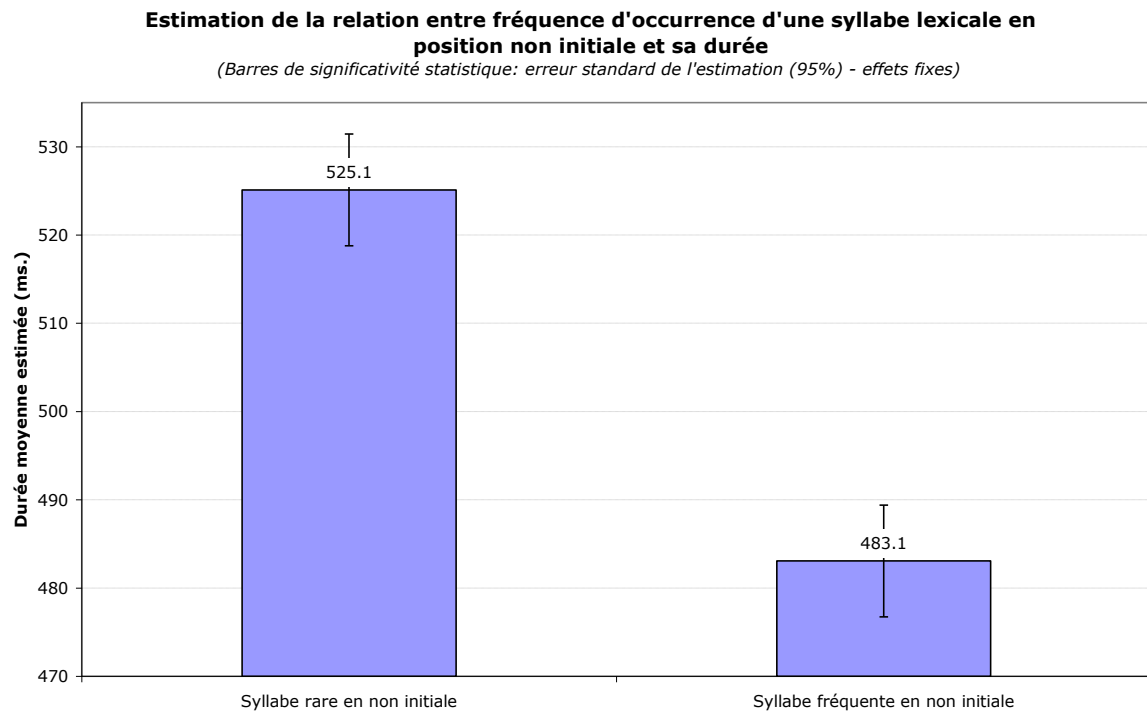


Figure 109 : Rapport entre le facteur expérimental et la durée de la syllabe



## Annexe 38

### Expérience 14 : Équation du modèle et tests de significativité de l'effet des covariables sur les temps de réponse

Équation 57 : Équation du modèle étudié

$$TR_{It, Suj} \sim N(\mathbf{XB}, \Omega)$$

$$TR_{It, Suj} = \beta_{0It, Suj} \text{Cons} + -22.757(20.248) F_{It, Suj} + 0.518(1.126) f0_{It, Suj} + 9.484(2.763) dB_{It, Suj} + -0.092(0.146) Dur_{It, Suj} + 0.672(0.184) PR_{It, Suj} + -24.958(20.929) FL_{It, Suj} + -3.257(2.777) AoA_{It, Suj} + -16.072(28.154) Fam_{It, Suj} + 18.497(12.041) Sex_{Suj} + 0.085(0.127) Ord_{It, Suj} + 0.196(0.335) f0 \cdot dB_{It, Suj} + -0.002(0.012) f0 \cdot Dur_{It, Suj} + -0.022(0.032) dB \cdot Dur_{It, Suj} + 0.001(0.002) f0 \cdot dB \cdot Dur_{It, Suj} + -4.865(1.803) F \cdot f0_{It, Suj} + -4.874(5.181) F \cdot dB_{It, Suj} + 0.011(0.160) F \cdot Dur_{It, Suj} + 0.552(0.601) F \cdot f0 \cdot dB_{It, Suj} + 0.043(0.023) F \cdot f0 \cdot Dur_{It, Suj} + 0.001(0.047) F \cdot dB \cdot Dur_{It, Suj} + -0.008(0.005) F \cdot f0 \cdot dB \cdot Dur_{It, Suj} + 0.001(0.001) Dur \cdot PR_{It, Suj} + -4.240(7.415) FL \cdot AoA_{It, Suj} + -121.514(56.219) FL \cdot Fam_{It, Suj} + -1.431(15.917) AoA \cdot Fam_{It, Suj} + 11.112(10.684) FL \cdot AoA \cdot Fam_{It, Suj} + 31.950(14.157) FI_{It, Suj} + 1.347(24.999) F \cdot FL_{It, Suj} + -10.909(15.760) F \cdot FI_{It, Suj} + -40.618(15.663) FI \cdot FL_{It, Suj} + 90.362(18.124) F \cdot FI \cdot FL_{It, Suj}$$

$$\beta_{0It, Suj} = 661.617(15.989) + u_{0Suj} + e_{0It, Suj}$$

$$[u_{0Suj}] \sim N(0, \Omega_u) : \Omega_u = [412.317(210.637)]$$

$$[e_{0It, Suj}] \sim N(0, \Omega_e) : \Omega_e = [8627.880(456.635)]$$

**-2\*loglikelihood(IGLS Deviance) = 8718.328(731 of 816 cases in use)**

Avec : TR = temps de reconnaissance du mot (V.D.) ; F = fréquence d'occurrence de la syllabe (rare vs. fréquent) en position non initiale dans le lexique (V.I. - réf = rare) ; f0 = fréquence fondamentale moyenne de la syllabe (Hz) ; dB = intensité moyenne de la syllabe (dB) ; Dur = durée de la syllabe (ms.) ; PR = position acoustique du point de reconnaissance du mot (ms.) ; FL = fréquence lexicale (logarithme en base 10) ; AoA = âge moyen d'acquisition estimé du mot ; Fam = familiarité conceptuelle du mot ; Sex = sexe des participants (réf = F) ; Ord = ordre de présentation des items ; FI = fréquence d'occurrence de la syllabe en initiale de mots (logarithme en base 10)

Tableau 173 : Tests de significativité des covariables et de leurs interactions

	Coefficient	Chi <sup>2</sup> (1, N=731)	p<
f0	0.518	0.212	.65
dB	9.484	11.786	<b>.0006</b>
Dur	-0.092	0.398	.53
PR	0.672	13.335	<b>.0003</b>
FL	-24.958	1.422	.24
AoA	-3.257	1.376	.25
Fam	-16.072	0.326	.57
Sex	18.497	2.36	.13

	<b>Coefficient</b>	<b>Chi<sup>2</sup><sub>(1, N=731)</sub></b>	<b>p&lt;</b>
Ordr	0.085	0.443	.51
f0.dB	0.196	0.343	.56
f0.Dur	-0.002	0.021	.89
dB.Dur	-0.022	0.494	.49
f0.dB.Dur	0.001	0.125	.73
F.f0	-4.865	7.279	<b>.007</b>
F.dB	-4.874	0.885	.35
F.Dur	0.011	0.005	.95
F.f0.dB	0.552	0.842	.36
F.f0.Dur	0.043	3.508	<b>.062</b>
F.dB.Dur	0.001	0.001	.98
F.f0.dB.Dur	-0.008	2.432	.12
Dur.PR	0.001	1.803	.18
FL.AoA	-4.24	0.327	.57
FL.Fam	-121.514	4.672	<b>.031</b>
AoA.Fam	-1.431	0.008	.93
FL.AoA.Fam	11.112	1.082	.30
FI	31.95	5.094	<b>.025</b>
F.FL	1.347	0.003	.96
F.FI	-10.909	0.479	.49
FI.FL	-40.618	6.725	<b>.0096</b>
F.FI.FL	90.362	24.858	<b>7E-07</b>

## Annexe 39

### Expériences 13 et 14 : Analyse exploratoire

#### Expérience 13

Équation 58 : Effets de la fréquence des items en tant que syllabe sur les temps de réponse des participants (Expérience 13)

$$TR_{It, Suj} \sim N(\mathbf{XB}, \Omega)$$

$$TR_{It, Suj} = \beta_{0It, Suj} \text{cons} + 1.862(1.324)f0_{It, Suj} + -1.404(3.476)dB_{It, Suj} + \\ 0.948(0.172)Dur_{It, Suj} + -0.559(0.205)PR_{It, Suj} + 20.729(52.185)Lat_{Suj} + \\ -48.203(44.280)sex_{Suj} + 0.018(0.166)Ordr_{It, Suj} + 45.103(22.134)FqI_{It, Suj} + \\ 10.282(18.385)FqnI_{It, Suj} + -0.624(0.350)f0 \quad .dB_{It, Suj} + \\ -0.033(0.012)f0 \quad .Dur_{It, Suj} + 0.069(0.034)dB \quad .Dur_{It, Suj} + \\ -0.004(0.003)f0 \quad .dB \quad .Dur_{It, Suj} + 0.005(0.002)Dur \quad .PR_{It, Suj} + \\ -16.895(27.507)FqI \quad .FqnI_{It, Suj} + -3.735(1.894)f0 \quad .FqI_{It, Suj} + \\ -3.885(1.576)f0 \quad .FqnI_{It, Suj} + -2.047(4.056)dB \quad .FqI_{It, Suj} + \\ -10.385(3.763)dB \quad .FqnI_{It, Suj} + -0.510(0.189)Dur \quad .FqI_{It, Suj} + \\ -0.428(0.196)Dur \quad .FqnI_{It, Suj} + 6.830(3.132)f0 \quad .FqI \quad .FqnI_{It, Suj} + \\ 19.975(6.028)dB \quad .FqI \quad .FqnI_{It, Suj} + 0.082(0.255)Dur \quad .FqI \quad .FqnI_{It, Suj} + \\ 0.728(0.538)f0 \quad .dB \quad .FqI_{It, Suj} + 1.926(0.475)f0 \quad .dB \quad .FqnI_{It, Suj} + \\ 0.037(0.018)f0 \quad .Dur \quad .FqI_{It, Suj} + 0.028(0.022)f0 \quad .Dur \quad .FqnI_{It, Suj} + \\ -0.011(0.044)dB \quad .Dur \quad .FqI_{It, Suj} + -0.057(0.042)dB \quad .Dur \quad .FqnI_{It, Suj} + \\ -0.605(0.826)f0 \quad .dB \quad .FqI \quad .FqnI_{It, Suj} + \\ 0.036(0.033)f0 \quad .Dur \quad .FqI \quad .FqnI_{It, Suj} + \\ 0.115(0.056)dB \quad .Dur \quad .FqI \quad .FqnI_{It, Suj} + \\ -0.001(0.004)f0 \quad .dB \quad .Dur \quad .FqI_{It, Suj} + \\ 0.011(0.006)f0 \quad .dB \quad .Dur \quad .FqnI_{It, Suj} + \\ 0.025(0.009)f0 \quad .dB \quad .Dur \quad .FqI \quad .FqnI_{It, Suj}$$

$$\beta_{0It, Suj} = 674.640(45.119) + u_{0Suj} + e_{0It, Suj}$$

$$[u_{0Suj}] \sim N(0, \Omega_u) : \Omega_u = [5430.605(2087.643)]$$

$$[e_{0It, Suj}] \sim N(0, \Omega_e) : \Omega_e = [12953.940(712.016)]$$

$$-2 * \loglikelihood(IGLS \text{ Deviance}) = 8376.725(677 \text{ of } 720 \text{ cases in use})$$

Tableau 174 : Tests de significativité des variables/covariables (fréquence de la syllabe, Expérience 13)

	Coefficient	Chi <sup>2</sup> <sub>(1, N=677)</sub>	p<
f0	1.862	1.977	.16
dB	-1.404	0.163	.69
Dur	0.948	30.382	<b>4<sup>E-08</sup></b>
PR	-0.559	7.456	<b>.0064</b>
Lat	20.729	0.158	.70
sex	-48.203	1.185	.28
Ordr	0.018	0.011	.92
FqI	45.103	4.152	<b>.042</b>
FqnI	10.282	0.313	.58



	<b>Coefficient</b>	<b>Chi<sup>2</sup><sub>(1, N=677)</sub></b>	<b>p&lt;</b>
f0.dB	-0.624	3.178	.075
f0.Dur	-0.033	6.953	.0084
dB.Dur	0.069	3.999	.046
f0.dB.Dur	-0.004	1.36	.25
Dur.PR	0.005	10.238	.0014
FqI.FqnI	-16.895	0.377	.54
f0.FqI	-3.735	3.891	.049
f0.FqnI	-3.885	6.075	.014
dB.FqI	-2.047	0.255	.62
dB.FqnI	-10.385	7.619	.0058
Dur.FqI	-0.51	7.267	.0071
Dur.FqnI	-0.428	4.755	.03
f0.FqI.FqnI	6.83	4.756	.03
dB.FqI.FqnI	19.975	10.979	.001
Dur.FqI.FqnI	0.082	0.103	.75
f0.dB.FqI	0.728	1.833	.18
f0.dB.FqnI	1.926	16.471	5 <sup>E-05</sup>
f0.Dur.FqI	0.037	4.183	.041
f0.Dur.FqnI	0.028	1.633	.21
dB.Dur.FqI	-0.011	0.06	.81
dB.Dur.FqnI	-0.057	1.826	.18
f0.dB.FqI.FqnI	-0.605	0.536	.47
f0.Dur.FqI.FqnI	0.036	1.236	.27
db.Dur.FqI.FqnI	0.115	4.138	.042
f0.dB.Dur.FqI.FqnI	-0.001	0.057	.82
f0.dB.Dur.FqnI	0.011	2.969	.085
f0.dB.Dur.FqI.FqnI	0.025	8.244	.0041

Équation 59 : Effets du nombre de mots dans lesquels les items sont insérés en tant que syllabe sur les temps de réponse des participants (Expérience 13)

$$TR_{ij} \sim N(XB, \Omega)$$

$$TR_{ij} = \beta_{0ij} \text{cons} + 1.478(2.491)f0_{ij} + 8.224(8.597)dB_{ij} + 0.697(0.268)Dur_{ij} + 0.030(0.210)PR_{ij} + 21.431(52.075)Lat_j + -47.819(44.187)sex_j + 0.059(0.169)Ordr_{ij} + -0.685(2.279)IN_{ij} + -1.431(0.814)nIN_{ij} + 0.728(0.623)f0_{ij}.dB_{ij} + 0.011(0.029)f0_{ij}.Dur_{ij} + -0.046(0.057)dB_{ij}.Dur_{ij} + 0.000(0.017)f0_{ij}.dB_{ij}.Dur_{ij} + -0.001(0.001)Dur_{ij}.PR_{ij} + -0.046(0.089)IN_{ij}.nIN_{ij} + 0.223(0.143)f0_{ij}.IN_{ij} + -0.105(0.103)f0_{ij}.nIN_{ij} + -0.488(0.887)dB_{ij}.IN_{ij} + 0.456(0.290)dB_{ij}.nIN_{ij} + -0.039(0.014)Dur_{ij}.IN_{ij} + 0.019(0.012)Dur_{ij}.nIN_{ij} + -0.004(0.005)f0_{ij}.IN_{ij}.nIN_{ij} + 0.000(0.001)Dur_{ij}.IN_{ij}.nIN_{ij} + -0.067(0.069)f0_{ij}.dB_{ij}.IN_{ij} + 0.114(0.038)f0_{ij}.dB_{ij}.nIN_{ij} + -0.001(0.002)f0_{ij}.Dur_{ij}.IN_{ij} + 0.000(0.001)f0_{ij}.Dur_{ij}.nIN_{ij} + -0.003(0.004)dB_{ij}.Dur_{ij}.IN_{ij} + 0.000(0.003)dB_{ij}.Dur_{ij}.nIN_{ij} + 0.003(0.004)f0_{ij}.dB_{ij}.IN_{ij}.nIN_{ij} + 0.000(0.000)f0_{ij}.Dur_{ij}.IN_{ij}.nIN_{ij} + 0.000(0.000)dB_{ij}.Dur_{ij}.IN_{ij}.nIN_{ij} + 0.000(0.001)f0_{ij}.dB_{ij}.Dur_{ij}.nIN_{ij} + 0.000(0.000)f0_{ij}.dB_{ij}.Dur_{ij}.IN_{ij}.nIN_{ij} + -0.023(0.038)dB_{ij}.IN_{ij}.nIN_{ij}$$

$$\beta_{0ij} = 681.580(48.601) + u_{0ij} + e_{0ij}$$

$$[u_{0ij}] \sim N(0, \Omega_u) : \Omega_u = [5400.779(2079.264)]$$

$$[e_{0ij}] \sim N(0, \Omega_e) : \Omega_e = [13215.280(726.378)]$$

$$-2 * \log\text{likelihood(IGLS Deviance)} = 8389.884(677 \text{ of } 720 \text{ cases in use})$$

Tableau 175 : Tests de significativité des variables/covariables (nombre de mots renfermant la syllabe, Expérience 13)

	Coefficient	Chi <sup>2</sup> <sub>(1, N=677)</sub>	p<
f0	1.478	0.352	.56
dB	8.224	0.915	.34
Dur	0.697	6.754	<b>.01</b>
PR	0.03	0.021	.89
Lat	21.431	0.169	.69
sex	-47.819	1.171	.28
Ordr	0.059	0.124	.73
IN	-0.685	0.09	.77
nIN	-1.431	3.091	<b>.079</b>
f0.dB	0.728	1.364	.25
f0.Dur	0.011	0.148	.71
dB.Dur	-0.046	0.654	.42
f0.dB.Dur	0	0.001	.98
Dur.PR	-0.001	0.78	.38
IN.nIN	-0.046	0.273	.61
f0.IN	0.223	2.427	.12
f0.nIN	-0.105	1.041	.31
dB.IN	-0.488	0.303	.59
dB.nIN	0.456	2.474	.12
Dur.IN	-0.039	7.12	<b>.008</b>
Dur.nIN	0.019	2.388	.13

	Coefficient	Chi <sup>2</sup> <sub>(1, N=677)</sub>	p<
f0.IN.nIN	-0.004	0.709	.40
Dur.IN.nIN	0	0.206	.65
f0.dB.IN	-0.067	0.922	.31
f0.dB.nIN	0.114	9.244	<b>.003</b>
f0.Dur.IN	-0.001	0.136	.72
f0.Dur.nIN	0	0.031	.87
dB.Dur.IN	-0.003	0.375	.55
dB.Dur.nIN	0	0.024	.88
f0.dB.IN.nIN	0.003	0.486	.49
f0.Dur.IN.nIN	0	2.111	.15
db.Dur.IN.nIN	0	0.807	.37
f0.dB.Dur.IN.nIN	0	0.022	.89
f0.dB.Dur.nIN	0	0.133	.72
f0.dB.Dur.IN.nIN	0	0	1
dB.IN.nIN	-0.023	0.375	.55

Équation 60 : Effets de la fréquence des items en tant que séquence de phonèmes sur les temps de réponse des participants (Expérience 13)

$$TR_{It, Suj} \sim N(\mathbf{XB}, \Omega)$$

$$\begin{aligned}
TR_{It, Suj} = & \beta_{0It, Suj} \mathbf{cons} + \mathbf{1.331(1.057)} f0_{It, Suj} + \mathbf{-6.178(3.220)} dB_{It, Suj} + \\
& \mathbf{0.801(0.166)} Dur_{It, Suj} + \mathbf{-0.700(0.198)} PR_{It, Suj} + \mathbf{22.068(52.166)} Lat_{Suj} + \\
& \mathbf{-48.566(44.265)} sex_{Suj} + \mathbf{0.083(0.169)} Ordr_{It, Suj} + \mathbf{-2.367(32.190)} FqsI_{It, Suj} + \\
& \mathbf{4.659(14.076)} FqsnI_{It, Suj} + \mathbf{-0.306(0.440)} f0 \quad .dB_{It, Suj} + \\
& \mathbf{-0.024(0.010)} f0 \quad .Dur_{It, Suj} + \mathbf{0.105(0.045)} dB \quad .Dur_{It, Suj} + \\
& \mathbf{-0.004(0.005)} f0 \quad .dB \quad .Dur_{It, Suj} + \mathbf{0.005(0.001)} Dur \quad .PR_{It, Suj} + \\
& \mathbf{-18.387(64.421)} FqsI \quad .FqsnI_{It, Suj} + \mathbf{-5.818(2.784)} f0 \quad .FqsI_{It, Suj} + \\
& \mathbf{1.337(1.510)} f0 \quad .FqsnI_{It, Suj} + \mathbf{-7.298(11.807)} dB \quad .FqsI_{It, Suj} + \\
& \mathbf{0.382(4.731)} dB \quad .FqsnI_{It, Suj} + \mathbf{-0.352(0.399)} Dur \quad .FqsI_{It, Suj} + \\
& \mathbf{-0.069(0.159)} Dur \quad .FqsnI_{It, Suj} + \mathbf{-2.246(5.480)} f0 \quad .FqsI \quad .FqsnI_{It, Suj} + \\
& \mathbf{-5.989(14.216)} dB \quad .FqsI \quad .FqsnI_{It, Suj} + \\
& \mathbf{-0.594(0.622)} Dur \quad .FqsI \quad .FqsnI_{It, Suj} + \mathbf{0.636(1.121)} f0 \quad .dB \quad .FqsI_{It, Suj} + \\
& \mathbf{1.625(0.382)} f0 \quad .dB \quad .FqsnI_{It, Suj} + \mathbf{0.077(0.038)} f0 \quad .Dur \quad .FqsI_{It, Suj} + \\
& \mathbf{0.022(0.019)} f0 \quad .Dur \quad .FqsnI_{It, Suj} + \mathbf{-0.111(0.121)} dB \quad .Dur \quad .FqsI_{It, Suj} + \\
& \mathbf{0.045(0.068)} dB \quad .Dur \quad .FqsnI_{It, Suj} + \\
& \mathbf{1.890(1.451)} f0 \quad .dB \quad .FqsI \quad .FqsnI_{It, Suj} + \\
& \mathbf{0.163(0.061)} f0 \quad .Dur \quad .FqsI \quad .FqsnI_{It, Suj} + \\
& \mathbf{0.001(0.185)} dB \quad .Dur \quad .FqsI \quad .FqsnI_{It, Suj} + \\
& \mathbf{-0.037(0.017)} f0 \quad .dB \quad .Dur \quad .FqsI_{It, Suj} + \\
& \mathbf{0.035(0.005)} f0 \quad .dB \quad .Dur \quad .FqsnI_{It, Suj} + \\
& \mathbf{-0.005(0.039)} f0 \quad .dB \quad .Dur \quad .FqsI \quad .FqsnI_{It, Suj}
\end{aligned}$$

$$\beta_{0It, Suj} = \mathbf{693.646(45.857)} + u_{0Suj} + e_{0It, Suj}$$

$$[u_{0Suj}] \sim N(\mathbf{0}, \Omega_u) : \Omega_u = [\mathbf{5417.750(2086.186)}]$$

$$[e_{0It, Suj}] \sim N(\mathbf{0}, \Omega_e) : \Omega_e = [\mathbf{13354.390(734.027)}]$$

$$-2 * \loglikelihood(IGLS \text{ Deviance}) = \mathbf{8396.868(677 \text{ of } 720 \text{ cases in use})}$$

Tableau 176 : Tests de significativité des variables/covariables (fréquence de la séquence de phonèmes, Expérience 13)

	Coefficient	Chi <sup>2</sup> <sub>(1, N=677)</sub>	p<
f0	1.331	1.586	.21
dB	-6.178	3.682	<b>.056</b>

	<b>Coefficient</b>	<b>Chi<sup>2</sup><sub>(1, N=677)</sub></b>	<b>p&lt;</b>
Dur	0.801	23.196	<b>2<sup>E-06</sup></b>
PR	-0.7	12.513	<b>.0005</b>
Lat	22.068	0.179	.68
sex	-48.566	1.204	.28
Ordr	0.083	0.241	.63
FqsI	-2.367	0.005	.95
FqsnI	4.659	0.11	.75
f0.dB	-0.306	0.483	.49
f0.Dur	-0.024	6.34	<b>.012</b>
dB.Dur	0.105	5.329	<b>.021</b>
f0.dB.Dur	-0.004	0.603	.44
Dur.PR	0.005	11.204	<b>.0009</b>
FqsI.FqsnI	-18.387	0.081	.78
f0.FqsI	-5.818	4.366	<b>.037</b>
f0.FqsnI	1.337	0.783	.38
dB.FqsI	-7.298	0.382	.54
dB.FqsnI	0.382	0.007	.94
Dur.FqsI	-0.352	0.778	.38
Dur.FqsnI	-0.069	0.189	.67
f0.FqsI.FqsnI	-2.246	0.168	.69
dB.FqsI.FqsnI	-5.989	0.177	.68
Dur.FqsI.FqsnI	-0.594	0.911	.34
f0.dB.FqsI	0.636	0.322	.58
f0.dB.FqsnI	1.625	18.064	<b>3<sup>E-05</sup></b>
f0.Dur.FqsI	0.077	4.124	<b>.043</b>
f0.Dur.FqsnI	0.022	1.379	.25
dB.Dur.FqsI	-0.111	0.841	.36
dB.Dur.FqsnI	0.045	0.45	.51
f0.dB.FqsI.FqsnI	1.89	1.697	.20
f0.Dur.FqsI.FqsnI	0.163	7.02	<b>.0081</b>
db.Dur.FqsI.FqsnI	0.001	0	1
f0.dB.Dur.FqsI	-0.037	4.893	<b>.027</b>
f0.dB.Dur.FqsnI	0.035	50.021	<b>2<sup>E-12</sup></b>
f0.dB.Dur.FqsI.FqsnI	-0.005	0.018	.90

Équation 61 : Effets du nombre de mots dans lesquels les items sont insérés en tant que séquence de phonèmes sur les temps de réponse des participants (Expérience 13)

$$TR_{It, Suj} \sim N(\mathbf{XB}, \Omega)$$

$$TR_{It, Suj} = \beta_{0It, Suj} \mathbf{cons} + 0.412(2.369) f0_{It, Suj} + 30.675(5.929) dB_{It, Suj} + 0.449(0.264) Dur_{It, Suj} + 0.127(0.224) PR_{It, Suj} + 20.846(52.219) Lat_{Suj} + -47.585(44.310) sex_{Suj} + 0.106(0.168) Ord_{It, Suj} + -1.328(0.648) SqIN_{It, Suj} + -0.959(0.803) SqnIN_{It, Suj} + 2.119(0.633) f0_{dB_{It, Suj}} + -0.005(0.029) f0_{Dur_{It, Suj}} + 0.103(0.048) dB_{Dur_{It, Suj}} + 0.002(0.007) f0_{dB_{Dur_{It, Suj}}} + -0.003(0.001) Dur_{PR_{It, Suj}} + -0.073(0.025) SqIN_{SqnIN_{It, Suj}} + 0.197(0.044) f0_{SqnIN_{It, Suj}} + -0.262(0.091) f0_{SqnIN_{It, Suj}} + 0.894(0.370) dB_{SqnIN_{It, Suj}} + 0.859(0.268) dB_{SqnIN_{It, Suj}} + -0.015(0.004) Dur_{SqnIN_{It, Suj}} + 0.002(0.010) Dur_{SqnIN_{It, Suj}} + -0.006(0.003) f0_{SqnIN_{SqnIN_{It, Suj}}} + 0.013(0.017) dB_{SqnIN_{SqnIN_{It, Suj}}} + 0.000(0.000) Dur_{SqnIN_{SqnIN_{It, Suj}}} + 0.017(0.023) f0_{dB_{SqnIN_{It, Suj}}} + 0.139(0.034) f0_{dB_{SqnIN_{It, Suj}}} + -0.003(0.001) f0_{Dur_{SqnIN_{It, Suj}}} + 0.001(0.001) f0_{Dur_{SqnIN_{It, Suj}}} + 0.004(0.002) dB_{Dur_{SqnIN_{It, Suj}}} + 0.006(0.002) dB_{Dur_{SqnIN_{It, Suj}}} + 0.002(0.002) f0_{dB_{SqnIN_{SqnIN_{It, Suj}}}} + 0.000(0.000) f0_{Dur_{SqnIN_{SqnIN_{It, Suj}}}} + 0.000(0.000) dB_{Dur_{SqnIN_{SqnIN_{It, Suj}}}} + 0.000(0.000) f0_{dB_{Dur_{SqnIN_{SqnIN_{It, Suj}}}}} + 0.001(0.000) f0_{dB_{Dur_{SqnIN_{SqnIN_{It, Suj}}}}} + 0.000(0.000) f0_{dB_{Dur_{SqnIN_{SqnIN_{It, Suj}}}}}$$

$$\beta_{0It, Suj} = 705.718(47.608) + u_{0Suj} + e_{0It, Suj}$$

$$[u_{0Suj}] \sim N(0, \Omega_u) : \Omega_u = [5432.036(2090.441)]$$

$$[e_{0It, Suj}] \sim N(0, \Omega_e) : \Omega_e = [13235.440(727.489)]$$

$-2 * \loglikelihood(IGLS \text{ Deviance}) = 8390.977(677 \text{ of } 720 \text{ cases in use})$

Tableau 177 : Tests de significativité des variables/covariables (nombre de mots renfermant la séquence, Expérience 13)

	Coefficient	Chi <sup>2</sup> <sub>(1, N=677)</sub>	p<
f0	0.412	0.03	.87
dB	30.675	26.766	<b>3<sup>E-07</sup></b>
Dur	0.449	2.895	<b>.089</b>
PR	0.127	0.322	.58
Lat	20.846	0.159	.70
sex	-47.585	1.153	.29
Ord	0.106	0.403	.53
SqIN	-1.328	4.205	<b>.041</b>
SqnIN	0.959	1.427	.24
f0.dB	2.119	11.21	<b>.0009</b>
f0.Dur	-0.005	0.027	.87

	<b>Coefficient</b>	<b>Chi<sup>2</sup><sub>(1, N=677)</sub></b>	<b>p&lt;</b>
dB.Dur	0.103	4.573	<b>.033</b>
f0.dB.Dur	0.002	0.084	.78
Dur.PR	-0.003	3.142	<b>.077</b>
SqIN.SqnIN	-0.073	8.621	<b>.0034</b>
f0.SqIN	0.197	20.161	<b>8<sup>E-06</sup></b>
f0.SqnIN	-0.262	8.336	<b>.0039</b>
dB.SqIN	0.894	5.846	<b>.016</b>
dB.SqnIN	0.859	10.255	<b>.0014</b>
Dur.SqIN	-0.015	11.5	<b>.0007</b>
Dur.SqnIN	0.002	0.027	.87
f0.SqIN.SqnIN	-0.006	4.85	<b>.028</b>
dB.SqIN.SqnIN	0.013	0.59	.45
Dur.SqIN.SqnIN	0	1.097	.30
f0.dB.SqIN	0.017	0.591	.45
f0.dB.SqnIN	0.139	17.023	<b>4<sup>E-05</sup></b>
f0.Dur.SqIN	-0.003	20.757	<b>6<sup>E-06</sup></b>
f0.Dur.SqnIN	0.001	0.872	.36
dB.Dur.SqIN	0.004	4.301	<b>.039</b>
dB.Dur.SqnIN	0.006	5.92	<b>.015</b>
f0.dB.SqIN.SqnIN	0.002	2.396	.13
f0.Dur.SqIN.SqnIN	0	3.807	<b>.052</b>
db.Dur.SqIN.SqnIN	0	5.491	<b>.02</b>
f0.dB.Dur.SqIN.SqnIN	0	0.534	.47
f0.dB.Dur.SqIN	0.001	7.458	<b>.0064</b>
f0.dB.Dur.SqnIN	0	0.133	.72

## Expérience 14

Équation 62 : Effets de la fréquence des items en tant que syllabe sur les temps de réponse des participants (Expérience 14)

$$TR_{It, Suj} \sim N(\mathbf{XB}, \Omega)$$

$$TR_{It, Suj} = \beta_{0It, Suj} \text{Cons} + 0.931(1.034) f0_{It, Suj} + 10.536(2.159) dB_{It, Suj} +$$

$$-0.062(0.111) Dur_{It, Suj} + 0.821(0.142) PR_{It, Suj} + 18.526(12.509) Sex_{Suj} +$$

$$0.060(0.124) Ordr_{It, Suj} + 0.928(8.848) FI_{It, Suj} + 41.152(12.599) FnI_{It, Suj} +$$

$$-0.784(0.286) f0 \quad .dB_{It, Suj} + -0.002(0.016) f0 \quad .Dur_{It, Suj} +$$

$$-0.076(0.033) dB \quad .Dur_{It, Suj} + -0.003(0.003) f0 \quad .dB \quad .Dur_{It, Suj} +$$

$$0.001(0.001) Dur \quad .PR_{It, Suj} + 18.833(11.096) FI \quad .FnI_{It, Suj} +$$

$$2.659(1.350) f0 \quad .FI_{It, Suj} + -3.091(1.465) f0 \quad .FnI_{It, Suj} +$$

$$6.343(3.183) dB \quad .FI_{It, Suj} + -2.492(4.102) dB \quad .FnI_{It, Suj} +$$

$$0.007(0.152) Dur \quad .FI_{It, Suj} + -0.035(0.100) Dur \quad .FnI_{It, Suj} +$$

$$1.583(2.479) f0 \quad .FI \quad .FnI_{It, Suj} + -5.072(7.367) dB \quad .FI \quad .FnI_{It, Suj} +$$

$$0.180(0.251) Dur \quad .FI \quad .FnI_{It, Suj} + -1.512(0.688) f0 \quad .dB \quad .FI_{It, Suj} +$$

$$0.701(0.299) f0 \quad .dB \quad .FnI_{It, Suj} + -0.049(0.019) f0 \quad .Dur \quad .FI_{It, Suj} +$$

$$0.027(0.015) f0 \quad .Dur \quad .FnI_{It, Suj} + -0.139(0.051) dB \quad .Dur \quad .FI_{It, Suj} +$$

$$0.032(0.058) dB \quad .Dur \quad .FnI_{It, Suj} +$$

$$2.161(0.539) f0 \quad .dB \quad .FI \quad .FnI_{It, Suj} +$$

$$0.097(0.038) f0 \quad .Dur \quad .FI \quad .FnI_{It, Suj} +$$

$$0.108(0.082) dB \quad .Dur \quad .FI \quad .FnI_{It, Suj} +$$

$$0.016(0.007) f0 \quad .dB \quad .Dur \quad .FI_{It, Suj} +$$

$$0.012(0.006) f0 \quad .dB \quad .Dur \quad .FnI_{It, Suj} +$$

$$-0.035(0.012) f0 \quad .dB \quad .Dur \quad .FI \quad .FnI_{It, Suj}$$

$$\beta_{0It, Suj} = 648.093(11.518) + u_{0Suj} + e_{0It, Suj}$$

$$[u_{0Suj}] \sim N(0, \Omega_u) : \Omega_u = [470.578(227.239)]$$

$$[e_{0It, Suj}] \sim N(0, \Omega_e) : \Omega_e = [8211.803(434.615)]$$

**-2\*loglikelihood(IGLS Deviance) = 8684.334(731 of 816 cases in use)**

Tableau 178 : Tests de significativité des variables/covariables (fréquence de la syllabe, Expérience 14)

	Coefficient	Chi <sup>2</sup> <sub>(1, N=731)</sub>	p<
f0	0.931	0.81	.37
dB	10.536	23.827	<b>2<sup>E-06</sup></b>
Dur	-0.062	0.311	.58
PR	0.821	33.514	<b>8<sup>E-09</sup></b>
sex	18.526	2.194	.14
Ordr	0.06	0.237	.63
FI	0.928	0.011	.92
FnI	41.152	10.669	<b>.0011</b>
f0.dB	-0.784	7.53	<b>.0061</b>
f0.Dur	-0.002	0.025	.88
dB.Dur	-0.076	5.158	<b>.024</b>
f0.dB.Dur	-0.003	0.659	.42
Dur.PR	0.001	0.764	.39
FI.FnI	18.833	2.88	<b>.09</b>

	Coefficient	Chi <sup>2</sup> (1, N=731)	p<
f0.FI	2.659	3.881	.049
f0.FnI	-3.091	4.455	.035
dB.FI	6.343	3.971	.047
dB.FnI	-2.492	0.369	.55
Dur.FI	0.007	0.002	.97
Dur.FnI	-0.035	0.123	.73
f0.FI.FnI	1.583	0.408	.53
dB.FI.FnI	-5.072	0.474	.50
Dur.FI.FnI	0.18	0.515	.48
f0.dB.FI	-1.512	4.832	.028
f0.dB.FnI	0.701	5.504	.019
f0.Dur.FI	-0.049	6.982	9 <sup>E-03</sup>
f0.Dur.FnI	0.027	3.274	.071
dB.Dur.FI	-0.139	7.537	.0061
dB.Dur.FnI	0.032	0.313	.58
f0.dB.FI.FnI	2.161	16.069	7 <sup>E-05</sup>
f0.Dur.FI.FnI	0.097	6.665	.0099
db.Dur.FI.FnI	0.108	1.739	.19
f0.dB.Dur.FI	0.016	5.546	.019
f0.dB.Dur.FnI	0.012	4.354	.037
f0.dB.Dur.FI.FnI	-0.035	7.78	.0053

Équation 63 : Effets du nombre de mots dans lesquels les items sont insérés en tant que syllabe sur les temps de réponse des participants (Expérience 14)

$$TR_{It, Suj} \sim N(XB, \Omega)$$

$$\begin{aligned}
TR_{It, Suj} = & \beta_{0It, Suj} \text{Cons} + 2.904(1.867) f0_{It, Suj} + -7.869(4.502) dB_{It, Suj} + \\
& -0.384(0.296) Dur_{It, Suj} + 0.580(0.147) PR_{It, Suj} + 18.771(12.439) Sex_{Suj} + \\
& 0.006(0.126) Ord_{It, Suj} + 12.723(7.642) IN_{It, Suj} + 2.861(0.749) nIN_{It, Suj} + \\
& -1.115(0.644) f0 \quad .dB_{It, Suj} + -0.019(0.039) f0 \quad .Dur_{It, Suj} + \\
& 0.081(0.052) dB \quad .Dur_{It, Suj} + 0.016(0.006) f0 \quad .dB \quad .Dur_{It, Suj} + \\
& 0.002(0.001) Dur \quad .PR_{It, Suj} + 0.778(0.540) IN \quad .nIN_{It, Suj} + \\
& 1.899(1.070) f0 \quad .IN_{It, Suj} + -0.019(0.102) f0 \quad .nIN_{It, Suj} + \\
& -2.944(3.066) dB \quad .IN_{It, Suj} + -0.510(0.226) dB \quad .nIN_{It, Suj} + \\
& -0.080(0.103) Dur \quad .IN_{It, Suj} + -0.008(0.012) Dur \quad .nIN_{It, Suj} + \\
& -0.056(0.073) f0 \quad .IN \quad .nIN_{It, Suj} + -0.367(0.232) dB \quad .IN \quad .nIN_{It, Suj} + \\
& -0.008(0.008) Dur \quad .IN \quad .nIN_{It, Suj} + -0.376(0.387) f0 \quad .dB \quad .IN_{It, Suj} + \\
& -0.046(0.034) f0 \quad .dB \quad .nIN_{It, Suj} + 0.021(0.013) f0 \quad .Dur \quad .IN_{It, Suj} + \\
& -0.001(0.002) f0 \quad .Dur \quad .nIN_{It, Suj} + -0.010(0.018) dB \quad .Dur \quad .IN_{It, Suj} + \\
& 0.000(0.001) dB \quad .Dur \quad .nIN_{It, Suj} + \\
& 0.007(0.019) f0 \quad .dB \quad .IN \quad .nIN_{It, Suj} + \\
& 0.003(0.001) f0 \quad .Dur \quad .IN \quad .nIN_{It, Suj} + \\
& 0.002(0.001) dB \quad .Dur \quad .IN \quad .nIN_{It, Suj} + \\
& -0.010(0.004) f0 \quad .dB \quad .Dur \quad .IN_{It, Suj} + \\
& 0.001(0.000) f0 \quad .dB \quad .Dur \quad .nIN_{It, Suj} + \\
& -0.001(0.000) f0 \quad .dB \quad .Dur \quad .IN \quad .nIN_{It, Suj}
\end{aligned}$$

$$\beta_{0It, Suj} = 707.340(18.397) + u_{0Suj} + e_{0It, Suj}$$

$$[u_{0Suj}] \sim N(0, \Omega_u) : \Omega_u = [456.668(224.780)]$$

$$[e_{0It, Suj}] \sim N(0, \Omega_e) : \Omega_e = [8493.290(449.511)]$$

$$-2 * \log \text{likelihood(IGLS Deviance)} = 8708.209(731 \text{ of } 816 \text{ cases in use})$$



Le modèle ne converge pas entièrement, y compris en abaissant le seuil de convergence.

Tableau 179 : Tests de significativité des variables/covariables (nombre de mots renfermant la syllabe, Expérience 14)

	Coefficient	Chi <sup>2</sup> <sub>(1, N=731)</sub>	p<
f0	2.904	2.419	.12
dB	-7.869	3.055	.081
Dur	-0.384	1.681	.20
PR	0.58	15.669	<b>8<sup>E-05</sup></b>
sex	18.771	2.277	.14
Ordr	0.006	0.003	.96
IN	12.723	2.771	.096
nIN	2.861	14.597	<b>.0002</b>
f0.dB	-1.115	2.998	.084
f0.Dur	-0.019	0.233	.63
dB.Dur	0.081	2.382	.13
f0.dB.Dur	0.016	7.306	<b>.0069</b>
Dur.PR	0.002	6.247	<b>.013</b>
IN.nIN	0.778	2.074	.15
f0.IN	1.899	3.15	.076
f0.nIN	-0.019	0.036	.85
dB.IN	-2.944	0.922	.34
dB.nIN	-0.51	5.082	<b>.025</b>
Dur.IN	-0.08	0.599	.44
Dur.nIN	-0.008	0.444	.51
f0.IN.nIN	-0.056	0.579	.45
dB.IN.nIN	-0.367	2.499	.12
Dur.IN.nIN	-0.008	1.063	.31
f0.dB.IN	-0.376	0.945	.34
f0.dB.nIN	-0.046	1.784	.19
f0.Dur.IN	0.021	2.336	.13
f0.Dur.nIN	-0.001	0.181	.68
dB.Dur.IN	-0.01	0.313	.58
dB.Dur.nIN	0	0.001	.98
f0.dB.IN.nIN	0.007	0.16	.69
f0.Dur.IN.nIN	0.003	9.948	<b>.0017</b>
db.Dur.IN.nIN	0.002	3.028	.082
f0.dB.Dur.IN	-0.01	5.6	<b>.018</b>
f0.dB.Dur.nIN	0.001	14.653	<b>.0002</b>
f0.dB.Dur.IN.nIN	-0.001	10.78	<b>.001</b>

Équation 64 : Effets de la fréquence des items en tant que séquence de phonèmes sur les temps de réponse des participants (Expérience 14)

$$TR_{It, Suj} \sim N(\mathbf{XB}, \Omega)$$

$$\begin{aligned} TR_{It, Suj} = & \beta_{0It, Suj} \mathbf{Cons} + 2.476(1.399) f0_{It, Suj} + 8.430(2.647) dB_{It, Suj} + \\ & -0.147(0.146) Dur_{It, Suj} + 0.846(0.123) PR_{It, Suj} + 19.109(12.119) Sex_{Suj} + \\ & 0.103(0.126) Ord_{It, Suj} + -15.115(18.706) FqsI_{It, Suj} + \\ & 21.861(10.224) fqsni_{It, Suj} + 0.436(0.379) f0 \quad .dB_{It, Suj} + \\ & 0.018(0.015) f0 \quad .Dur_{It, Suj} + 0.010(0.040) dB \quad .Dur_{It, Suj} + \\ & 0.000(0.004) f0 \quad .dB \quad .Dur_{It, Suj} + 0.001(0.001) Dur \quad .PR_{It, Suj} + \\ & 28.400(27.208) FqsI \quad .fqsni_{It, Suj} + -0.419(2.656) f0 \quad .FqsI_{It, Suj} + \\ & -3.033(1.145) f0 \quad .fqsni_{It, Suj} + 12.647(8.024) dB \quad .FqsI_{It, Suj} + \\ & -0.497(3.137) dB \quad .fqsni_{It, Suj} + 0.195(0.290) Dur \quad .FqsI_{It, Suj} + \\ & -0.024(0.164) Dur \quad .fqsni_{It, Suj} + -1.413(4.585) f0 \quad .FqsI \quad .fqsni_{It, Suj} + \\ & 10.481(11.322) dB \quad .FqsI \quad .fqsni_{It, Suj} + \\ & 0.879(0.536) Dur \quad .FqsI \quad .fqsni_{It, Suj} + \\ & -1.100(0.969) f0 \quad .dB \quad .FqsI_{It, Suj} + \\ & 0.336(0.388) f0 \quad .dB \quad .fqsni_{It, Suj} + \\ & -0.083(0.038) f0 \quad .Dur \quad .FqsI_{It, Suj} + \\ & 0.027(0.016) f0 \quad .Dur \quad .fqsni_{It, Suj} + \\ & 0.101(0.100) dB \quad .Dur \quad .FqsI_{It, Suj} + \\ & 0.001(0.044) dB \quad .Dur \quad .fqsni_{It, Suj} + \\ & 5.203(1.647) f0 \quad .dB \quad .FqsI \quad .fqsni_{It, Suj} + \\ & 0.117(0.055) f0 \quad .Dur \quad .FqsI \quad .fqsni_{It, Suj} + \\ & -0.156(0.155) dB \quad .Dur \quad .FqsI \quad .fqsni_{It, Suj} + \\ & -0.004(0.006) f0 \quad .dB \quad .Dur \quad .FqsI_{It, Suj} + \\ & 0.000(0.007) f0 \quad .dB \quad .Dur \quad .fqsni_{It, Suj} + \\ & -0.019(0.028) f0 \quad .dB \quad .Dur \quad .FqsI \quad .fqsni_{It, Suj} \end{aligned}$$

$$\beta_{0It, Suj} = 664.416(13.771) + u_{0Suj} + e_{0It, Suj}$$

$$[u_{0Suj}] \sim N(0, \Omega_u) : \Omega_u = [423.279(213.344)]$$

$$[e_{0It, Suj}] \sim N(0, \Omega_e) : \Omega_e = [8499.232(449.827)]$$

**-2\*loglikelihood(IGLS Deviance) = 8707.822(731 of 816 cases in use)**

Tableau 180 : Tests de significativité des variables/covariables (fréquence de la séquence de phonèmes, Expérience 14)

	Coefficient	Chi <sup>2</sup> <sub>(1, N=731)</sub>	p<
f0	2.476	3.13	.077
dB	8.43	10.146	.0015
Dur	-0.147	1.015	.32
PR	0.846	47.02	8 <sup>E-12</sup>
sex	19.109	2.486	.12
Ord	0.103	0.677	.42
FqsI	-15.115	0.653	.42
Fqsni	21.861	4.572	.033
f0.dB	0.436	1.32	.26
f0.Dur	0.018	1.331	.25
dB.Dur	0.01	0.066	.80
f0.dB.Dur	0	0.014	.91
Dur.PR	0.001	1.229	.27

	<b>Coefficient</b>	<b>Chi<sup>2</sup><sub>(1, N=731)</sub></b>	<b>p&lt;</b>
FqsI.FqsnI	28.4	1.09	.30
f0.FqsI	-0.419	0.025	.88
f0.FqsnI	-3.033	7.019	<b>.0081</b>
dB.FqsI	12.647	2.484	.12
dB.FqsnI	-0.497	0.025	.88
Dur.FqsI	0.195	0.452	.51
Dur.FqsnI	-0.024	0.022	.89
f0.FqsI.FqsnI	-1.413	0.095	.76
dB.FqsI.FqsnI	10.481	0.857	.36
Dur.FqsI.FqsnI	0.879	2.687	.11
f0.dB.FqsI	-1.1	1.289	.26
f0.dB.FqsnI	0.336	0.751	.39
f0.Dur.FqsI	-0.083	4.917	<b>.027</b>
f0.Dur.FqsnI	0.027	2.919	<b>.088</b>
dB.Dur.FqsI	0.101	1.022	.32
dB.Dur.FqsnI	0.001	0.001	.98
f0.dB.FqsI.FqsnI	5.203	9.974	<b>.0016</b>
f0.Dur.FqsI.FqsnI	0.117	4.586	<b>.033</b>
db.Dur.FqsI.FqsnI	-0.156	1.016	.32
f0.dB.Dur.FqsI	-0.004	0.488	.49
f0.dB.Dur.FqsnI	0	0.001	.98
f0.dB.Dur.FqsI.FqsnI	-0.019	0.486	.49

Équation 65 : Effets du nombre de mots dans lesquels les items sont insérés en tant que séquence de phonèmes sur les temps de réponse des participants (Expérience 14)

$$TR_{It, Suj} \sim N(\mathbf{XB}, \Omega)$$

$$\begin{aligned} TR_{It, Suj} = & \beta_{0It, Suj} \mathbf{Cons} + \mathbf{-9.588(2.586)} f0_{It, Suj} + \mathbf{7.756(3.982)} dB_{It, Suj} + \\ & \mathbf{0.266(0.187)} Dur_{It, Suj} + \mathbf{0.365(0.130)} PR_{It, Suj} + \mathbf{18.301(12.399)} Sex_{Suj} + \\ & \mathbf{0.066(0.124)} Ord_{It, Suj} + \mathbf{9.193(4.119)} sIN_{It, Suj} + \mathbf{1.971(0.407)} snIN_{It, Suj} + \\ & \mathbf{3.341(0.768)} f0 \quad .dB_{It, Suj} + \mathbf{0.128(0.021)} f0 \quad .Dur_{It, Suj} + \\ & \mathbf{0.002(0.046)} dB \quad .Dur_{It, Suj} + \mathbf{-0.026(0.008)} f0 \quad .dB \quad .Dur_{It, Suj} + \\ & \mathbf{0.003(0.001)} Dur \quad .PR_{It, Suj} + \mathbf{0.247(0.101)} sIN \quad .snIN_{It, Suj} + \\ & \mathbf{1.252(0.476)} f0 \quad .sIN_{It, Suj} + \mathbf{-0.298(0.068)} f0 \quad .snIN_{It, Suj} + \\ & \mathbf{-1.747(1.272)} dB \quad .sIN_{It, Suj} + \mathbf{-0.165(0.167)} dB \quad .snIN_{It, Suj} + \\ & \mathbf{-0.056(0.040)} Dur \quad .sIN_{It, Suj} + \mathbf{-0.001(0.005)} Dur \quad .snIN_{It, Suj} + \\ & \mathbf{0.015(0.011)} f0 \quad .sIN \quad .snIN_{It, Suj} + \\ & \mathbf{-0.054(0.040)} dB \quad .sIN \quad .snIN_{It, Suj} + \\ & \mathbf{-0.002(0.001)} Dur \quad .sIN \quad .snIN_{It, Suj} + \\ & \mathbf{-0.046(0.121)} f0 \quad .dB \quad .sIN_{It, Suj} + \\ & \mathbf{0.084(0.016)} f0 \quad .dB \quad .snIN_{It, Suj} + \\ & \mathbf{0.007(0.003)} f0 \quad .Dur \quad .sIN_{It, Suj} + \\ & \mathbf{0.004(0.001)} f0 \quad .Dur \quad .snIN_{It, Suj} + \\ & \mathbf{0.002(0.011)} dB \quad .Dur \quad .sIN_{It, Suj} + \\ & \mathbf{0.006(0.002)} dB \quad .Dur \quad .snIN_{It, Suj} + \\ & \mathbf{0.005(0.004)} f0 \quad .dB \quad .sIN \quad .snIN_{It, Suj} + \\ & \mathbf{0.000(0.000)} f0 \quad .Dur \quad .sIN \quad .snIN_{It, Suj} + \\ & \mathbf{0.001(0.000)} dB \quad .Dur \quad .sIN \quad .snIN_{It, Suj} + \\ & \mathbf{-0.001(0.002)} f0 \quad .dB \quad .Dur \quad .sIN_{It, Suj} + \\ & \mathbf{-0.001(0.000)} f0 \quad .dB \quad .Dur \quad .snIN_{It, Suj} + \\ & \mathbf{0.000(0.000)} f0 \quad .dB \quad .Dur \quad .sIN \quad .snIN_{It, Suj} \end{aligned}$$

$$\beta_{0It, Suj} = \mathbf{692.542(15.531)} + u_{0Suj} + e_{0It, Suj}$$

$$[u_{0Suj}] \sim N(\mathbf{0}, \Omega_u) : \Omega_u = [\mathbf{459.904(223.256)}]$$

$$[e_{0It, Suj}] \sim N(\mathbf{0}, \Omega_e) : \Omega_e = [\mathbf{8172.226(432.521)}]$$

$$-2 * \loglikelihood(IGLS \text{ Deviance}) = \mathbf{8680.584(731 \text{ of } 816 \text{ cases in use})}$$

Tableau 181 : Tests de significativité des variables/covariables (nombre de mots renfermant la séquence de phonèmes, Expérience 14)

	Coefficient	Chi <sup>2</sup> <sub>(1, N=731)</sub>	p<
f0	-9.588	13.75	<b>.0003</b>
dB	7.756	3.794	<b>.052</b>
Dur	0.266	2.014	<b>.16</b>
PR	0.365	7.903	<b>.005</b>
sex	18.301	2.179	<b>.14</b>
Ord	0.066	0.28	<b>.60</b>
sIN	9.193	4.981	<b>.026</b>
snIN	1.971	23.486	<b>2<sup>E-06</sup></b>
f0.dB	3.341	18.911	<b>2<sup>E-05</sup></b>
f0.Dur	0.128	37.225	<b>2<sup>E-09</sup></b>
dB.Dur	0.002	0.001	<b>.98</b>
f0.dB.Dur	-0.026	11.421	<b>.0008</b>
Dur.PR	0.003	18.313	<b>2<sup>E-05</sup></b>

	<b>Coefficient</b>	<b>Chi<sup>2</sup><sub>(1, N=731)</sub></b>	<b>p&lt;</b>
sIN.snIN	0.247	5.971	<b>.015</b>
f0.sIN	1.252	6.913	<b>.009</b>
f0.snIN	-0.298	19.433	<b>2<sup>E-05</sup></b>
dB.sIN	-1.747	1.887	.17
dB.snIN	-0.165	0.977	.33
Dur.sIN	-0.056	1.92	.17
Dur.snIN	-0.001	0.03	.87
f0.sIN.snIN	0.015	1.747	.19
dB.sIN.snIN	-0.054	1.859	.18
Dur.sIN.snIN	-0.002	2.864	<b>.091</b>
f0.dB.sIN	-0.046	0.142	.71
f0.dB.snIN	0.084	25.8	<b>4<sup>E-07</sup></b>
f0.Dur.sIN	0.007	4.581	<b>.033</b>
f0.Dur.snIN	0.004	36.416	<b>2<sup>E-09</sup></b>
dB.Dur.sIN	0.002	0.039	.85
dB.Dur.snIN	0.006	10.179	<b>.0015</b>
f0.dB.sIN.snIN	0.005	2.008	.16
f0.Dur.sIN.snIN	0	12.857	<b>.0004</b>
db.Dur.sIN.snIN	0.001	6.625	<b>.011</b>
f0.dB.Dur.sIN	-0.001	0.364	.55
f0.dB.Dur.snIN	-0.001	21.934	<b>3<sup>E-06</sup></b>
f0.dB.Dur.sIN.snIN	0	2.458	.12

## Annexe 40

### Expérience 15 : Matériel linguistique

Tableau 182 : Mots expérimentaux et pseudo-mots distracteurs

Matériel expérimental (mots)		Matériel distracteur (pseudo-mots)	
M <sub>1</sub>		M <sub>1</sub>	
M <sub>2</sub>	nM <sub>2</sub>	M <sub>2</sub>	nM <sub>2</sub>
bagak	banan	bjegak	kɛ̃nan
butik	bukē	bletik	blekē
kalis	kapyʃ	plɥilis	tɛ̃pyʃ
lisās	lidœk	kɛ̃sās	ʒwadœk
tēbal	tētyk	kɛ̃wabal	dɛ̃tyk
bɛ̃ikəl	bɛ̃igad	pɛ̃wakəl	pɛ̃wagad
mustaʃ	muslin	tɛ̃ɛtaʃ	gɛ̃ylin
tɛ̃mit	tɛ̃mos	bwamit	plɥimos
tukmāt	tuknyk	fɛ̃māt	fɛ̃nyk
gitaɕ	gitun	pnɔtaɕ	gɛ̃ɛtun
vipek	vites	ʒwapek	dɛ̃ates
vēkœk	vēten	plokœk	kɛ̃waten
femyk	fenəl	gɛ̃emyk	bjetys
lɔpē	lotys	bɛ̃ɥipē	gɛ̃wēnəl
ɕuset	ɕutin	gɛ̃yset	ɛ̃ɛtin
sivɛn	sigɔŋ	klāɕɛn	fɛ̃ɔgɔŋ
paɕlwaɕ	paɕtās	ɛ̃ɛlwaɕ	pnɔtās
kalvek	kalsjɔm	fɛ̃ðvek	glāsɔm
tɛ̃āspɔk	tɛ̃āsfyʒ	gupɔk	gufyʒ
kantō	kanlyk	glātō	fɛ̃ɥilyk
ʃalup	ʃamel	tɛ̃ɔlup	tɛ̃ɔmel
kuɛaʒ	kutym	bɛ̃ɛkaʒ	bwatym
delyʒ	dekad	tɛ̃ulyʒ	bɛ̃ɛkad
manɛʒ	mamut	gɛ̃wēneʒ	plomut
paɕjās	palet	dɛ̃ɛsjās	klālet
ʃaɕpāt	ʃaɕzœk	plɛpāt	bɛ̃ɥizœk
baɕbiʃ	baɕda	kɛ̃ɛbiʃ	kɛ̃ɛda
fɔɕmyl	fɔɕseps	fɛ̃ɥimyl	plɔseps
kaɕtuʃ	kaɕlēg	flotuʃ	flolēg
kɔɕniʃ	kɔɕbej	tɛ̃iniʃ	tɛ̃ibej
ɕadaɕ	ɕadjɔm	flydaɕ	kɛ̃idɔm
ɛ̃ɔflœk	ɛ̃ɔdel	stɛ̃iflœk	tɛ̃ɥidel
ɛ̃āgen	ɛ̃ākyn	kɛ̃igen	bɛ̃akyn
poɕjɛk	poɕjet	tɛ̃ɥipjɛk	plɛpjet
sēteʒ	sēdik	nɔteʒ	gnudik

Matériel expérimental (mots)		Matériel distracteur (pseudo-mots)	
M <sub>1</sub>		M <sub>1</sub>	
M <sub>2</sub>	nM <sub>2</sub>	M <sub>2</sub>	nM <sub>2</sub>
lakbē	lakʒes	klubē	kluzes
sɛkfəj	sɛkvœʁ	bʁafəj	stɛivœʁ
fɛkmaʒ	fɛkmje	ɡnumaʒ	nœbvyʁ
makmaj	makbvʁ	dʁamaj	flyʃō
tœpœʁ	tœʃō	plɔpœʁ	tœumje

Avec : M<sub>1</sub> première syllabe lexicale ; M<sub>2</sub> deuxième syllabe lexicale ; nM<sub>2</sub> deuxième syllabe non lexicale.

## Annexe 41

# Expérience 15 : Équation du modèle et tests de significativité de l'effet des covariables sur les temps de réponse

### Équation 66 : Équation du modèle étudié

$$\begin{aligned}
 TR_{It, Suj} &\sim N(XB, \Omega) \\
 TR_{It, Suj} = &\beta_{0It, Suj} \text{Cons} + -115.659(41.133)F_{It, Suj} + 114.691(60.382)M2_{It, Suj} + -129.960(81.732)F.M2_{It, Suj} + -0.953(1.171)f0_{1It, Suj} + \\
 &-1.371(2.750)f0_{2It, Suj} + -18.230(4.774)dB_{1It, Suj} + -6.665(4.655)dB_{2It, Suj} + -4.727(5.255)DurS1_{It, Suj} + \\
 &-4.108(5.235)DurS2_{It, Suj} + -5.687(3.076)F.f0_{1It, Suj} + -9.825(4.048)F.f0_{2It, Suj} + -13.705(8.337)F.dB_{1It, Suj} + \\
 &5.280(10.454)F.dB_{2It, Suj} + 0.858(0.531)F.DurS1_{It, Suj} + -1.904(0.531)F.DurS2_{It, Suj} + -8.177(4.238)M2.f0_{1It, Suj} + \\
 &-1.970(3.886)M2.f0_{2It, Suj} + -16.720(13.015)M2.dB_{1It, Suj} + 8.016(5.732)M2.dB_{2It, Suj} + -1.490(0.720)M2.DurS1_{It, Suj} + \\
 &0.270(0.454)M2.DurS2_{It, Suj} + 7.635(5.643)F.M2.f0_{1It, Suj} + 14.370(5.177)F.M2.f0_{2It, Suj} + 45.305(15.768)F.M2.dB_{1It, Suj} + \\
 &-34.221(13.696)F.M2.dB_{2It, Suj} + 0.787(0.798)F.M2.DurS1_{It, Suj} + -0.089(0.726)F.M2.DurS2_{It, Suj} + \\
 &0.948(0.347)f0_{1, dB1} .DurS1_{It, Suj} + 0.016(0.030)f0_{1, DurS1} .DurS1_{It, Suj} + 0.083(0.051)dB1 .DurS1_{It, Suj} + \\
 &0.001(0.006)f0_{1, dB1} .DurS1_{It, Suj} + -0.312(0.646)f0_{2, dB2} .DurS2_{It, Suj} + -0.059(0.039)f0_{2, DurS2} .DurS2_{It, Suj} + \\
 &-0.209(0.080)dB2 .DurS2_{It, Suj} + -0.019(0.010)f0_{2, dB2} .DurS2_{It, Suj} + -3.753(0.770)F.f01 .dB1_{It, Suj} + \\
 &-0.002(0.040)F.f01 .DurS1_{It, Suj} + -0.322(0.087)F.dB1 .DurS1_{It, Suj} + -0.021(0.009)F.f01 .dB1 .DurS1_{It, Suj} + \\
 &1.401(1.070)F.f02 .dB2_{It, Suj} + 0.321(0.062)F.f02 .DurS2_{It, Suj} + 0.232(0.143)F.dB2 .DurS2_{It, Suj} + \\
 &-0.002(0.016)F.f02 .dB2 .DurS2_{It, Suj} + 1.477(1.030)M2.f01 .dB1_{It, Suj} + 0.050(0.047)M2.f01 .DurS1_{It, Suj} + \\
 &0.417(0.091)M2.dB1 .DurS1_{It, Suj} + -0.020(0.009)M2.f01 .dB1 .DurS1_{It, Suj} + 0.416(0.914)M2.f02 .dB2_{It, Suj} + \\
 &0.107(0.057)M2.f02 .DurS2_{It, Suj} + 0.205(0.097)M2.dB2 .DurS2_{It, Suj} + 0.061(0.014)M2.f02 .dB2 .DurS2_{It, Suj} + \\
 &0.800(1.260)F.M2.f01 .dB1_{It, Suj} + -0.136(0.067)F.M2.f01 .DurS1_{It, Suj} + -0.165(0.111)F.M2.dB1 .DurS1_{It, Suj} + \\
 &0.046(0.013)F.M2.f01 .dB1 .DurS1_{It, Suj} + -0.646(1.469)F.M2.f02 .dB2_{It, Suj} + -0.488(0.110)F.M2.f02 .DurS2_{It, Suj} + \\
 &0.073(0.188)F.M2.dB2 .DurS2_{It, Suj} + -0.039(0.028)F.M2.f02 .dB2 .DurS2_{It, Suj} + 0.259(0.099)f01 .f02_{It, Suj} + \\
 &-0.036(0.417)dB1 .dB2_{It, Suj} + 0.001(0.003)DurS1 .DurS2_{It, Suj} + -0.015(0.015)f01 .f02 .dB1 .dB2_{It, Suj} + \\
 &0.000(0.000)f01 .f02 .DurS1 .DurS2_{It, Suj} + 0.000(0.000)dB1 .dB2 .DurS1 .DurS2_{It, Suj} + \\
 &0.000(0.000)f01 .f02 .dB1 .dB2 .DurS1 .DurS2_{It, Suj} + 0.023(0.018)F.f01 .f02 .dB1 .dB2_{It, Suj} + \\
 &0.000(0.000)F.f01 .f02 .DurS1 .DurS2_{It, Suj} + 0.001(0.000)F.dB1 .dB2 .DurS1 .DurS2_{It, Suj} + \\
 &0.000(0.000)F.f01 .f02 .dB1 .dB2 .DurS1 .DurS2_{It, Suj} + 0.041(0.020)M2.f01 .f02 .dB1 .dB2_{It, Suj} + \\
 &0.000(0.000)M2.f01 .f02 .DurS1 .DurS2_{It, Suj} + 0.000(0.000)M2.dB1 .dB2 .DurS1 .DurS2_{It, Suj} + \\
 &0.000(0.000)M2.f01 .f02 .dB1 .dB2 .DurS1 .DurS2_{It, Suj} + -0.057(0.022)F.M2.f01 .f02 .dB1 .dB2_{It, Suj} + \\
 &0.000(0.000)F.M2.f01 .f02 .DurS1 .DurS2_{It, Suj} + 0.000(0.000)F.M2.dB1 .dB2 .DurS1 .DurS2_{It, Suj} + \\
 &0.000(0.000)F.M2.f01 .f02 .dB1 .dB2 .DurS1 .DurS2_{It, Suj} + 4.466(5.202)DurIt_{It, Suj} + 0.077(0.299)PRIt_{It, Suj} + \\
 &0.187(0.287)PRS2_{It, Suj} + 0.004(0.002)DurIt .PRIt_{It, Suj} + -11.444(25.304)Lat_{suj} + -0.083(0.126)Ord_{It, Suj} + \\
 &0.000(0.000)DurS2 .DurIt .PRIt_{It, Suj} + 0.000(0.000)DurS2 .DurIt .PRS2_{It, Suj} + \\
 &0.000(0.000)PRIt .PRS2 .DurS2_{It, Suj} + 0.000(0.000)PRIt .PRS2 .DurIt_{It, Suj} + \\
 &0.000(0.000)DurS2 .DurIt .PRS2 .PRIt_{It, Suj} + -0.004(0.002)PRS2 .DurS2_{It, Suj}
 \end{aligned}$$

$$\beta_{0It, Suj} = 1595.435(36.196) + u_{0Suj} + e_{0It, Suj}$$

$$[u_{0Suj}] \sim N(0, \Omega_u) : \Omega_u = [3449.692(729.795)]$$

$$[e_{0It, Suj}] \sim N(0, \Omega_e) : \Omega_e = [14300.860(479.625)]$$

$$-2 * \text{loglikelihood(IGLS Deviance)} = 22887.140(1835 \text{ of } 2240 \text{ cases in use})$$

Avec : TR = temps de réponse des participants (V.D.) ; F = frontière lexicale entre les deux syllabes (V.I. - réf = présence d'une frontière) ; M2 = lexicalité de la deuxième syllabe de l'item (V.I. - réf = S2 non lexicale) ; f0<sub>1/2</sub> = fréquence fondamentale moyenne de la première/deuxième syllabe (Hz) ; dB<sub>1/2</sub> = intensité moyenne de la première/deuxième syllabe (dB) ; dur<sub>1/2</sub> = durée de la première/deuxième syllabe (ms.) ; PRi = position acoustique du point de reconnaissance de l'item (ms.) ; PR2 = position acoustique du point de reconnaissance de la deuxième syllabe (ms.) ; Lat = latéralité manuelle des participants (réf = gauche) ; Ord = ordre de présentation des items

Tableau 183 : Tests de significativité des covariables et de leurs interactions, entre elles et avec les variables expérimentales

	Coefficient	Chi <sup>2</sup> <sub>(1, N=1835)</sub>	p<
f0 <sub>1</sub>	-0.953	0.661	.42
f0 <sub>2</sub>	-1.371	0.249	.62
dB <sub>1</sub>	-18.23	14.579	<b>.0002</b>
dB <sub>2</sub>	-6.665	2.05	.16



	Coefficient	Chi <sup>2</sup> (1, N=1835)	p<
DurS <sub>1</sub>	-4.727	0.809	.37
DurS <sub>2</sub>	-4.108	0.616	.44
F.f0 <sub>1</sub>	-5.687	3.418	.065
F.f0 <sub>2</sub>	-9.825	5.891	.016
F.dB <sub>1</sub>	-13.705	2.702	.11
F.dB <sub>2</sub>	5.28	0.255	.62
F.DurS <sub>1</sub>	0.858	2.607	.11
F.DurS <sub>2</sub>	-1.904	12.859	.0004
M <sub>2</sub> .f0 <sub>1</sub>	-8.177	3.723	.054
M <sub>2</sub> .f0 <sub>2</sub>	1.97	0.257	.62
M <sub>2</sub> .dB <sub>1</sub>	-16.72	1.65	.20
M <sub>2</sub> .dB <sub>2</sub>	8.016	1.956	.17
M <sub>2</sub> .DurS <sub>1</sub>	-1.49	4.28	.039
M <sub>2</sub> .DurS <sub>2</sub>	0.27	0.353	.56
F.M <sub>2</sub> .f0 <sub>1</sub>	7.635	1.831	.18
F.M <sub>2</sub> .f0 <sub>2</sub>	14.37	7.704	.006
F.M <sub>2</sub> .dB <sub>1</sub>	45.305	8.255	.005
F.M <sub>2</sub> .dB <sub>2</sub>	-34.221	6.243	.013
F.M <sub>2</sub> .DurS <sub>1</sub>	0.787	0.973	.33
F.M <sub>2</sub> .DurS <sub>2</sub>	-0.089	0.015	.91
f0 <sub>1</sub> .dB <sub>1</sub>	0.948	7.467	.0063
f0 <sub>1</sub> .DurS <sub>1</sub>	0.016	0.28	.60
dB <sub>1</sub> .DurS <sub>1</sub>	0.083	2.647	.11
f0 <sub>1</sub> .dB <sub>1</sub> .DurS <sub>1</sub>	0.001	0.035	.86
f0 <sub>2</sub> .dB <sub>2</sub>	-0.312	0.233	.63
f0 <sub>2</sub> .DurS <sub>2</sub>	-0.059	2.235	.14
dB <sub>2</sub> .DurS <sub>2</sub>	-0.209	6.744	.0095
f0 <sub>2</sub> .dB <sub>2</sub> .DurS <sub>2</sub>	-0.019	3.447	.064
F.f0 <sub>1</sub> .dB <sub>1</sub>	-3.753	23.758	2 <sup>E-06</sup>
F.f0 <sub>1</sub> .DurS <sub>1</sub>	-0.002	0.002	.97
F.dB <sub>1</sub> .DurS <sub>1</sub>	-0.322	13.693	.0003
F.f0 <sub>1</sub> .dB <sub>1</sub> .DurS <sub>1</sub>	-0.021	5.422	.02
F.f0 <sub>2</sub> .dB <sub>2</sub>	1.401	1.714	.20
F.f0 <sub>2</sub> .DurS <sub>2</sub>	0.321	26.598	3 <sup>E-07</sup>
F.dB <sub>2</sub> .DurS <sub>2</sub>	0.232	2.63	.11
F.f0 <sub>2</sub> .dB <sub>2</sub> .DurS <sub>2</sub>	-0.002	0.022	.89
M <sub>2</sub> .f0 <sub>1</sub> .dB <sub>1</sub>	1.477	2.054	.16
M <sub>2</sub> .f0 <sub>1</sub> .DurS <sub>1</sub>	0.05	1.137	.29
M <sub>2</sub> .dB <sub>1</sub> .DurS <sub>1</sub>	0.417	20.836	6 <sup>E-06</sup>
M <sub>2</sub> .f0 <sub>1</sub> .dB <sub>1</sub> .DurS <sub>1</sub>	-0.02	4.561	.033
M <sub>2</sub> .f0 <sub>2</sub> .dB <sub>2</sub>	0.416	0.207	.65
M <sub>2</sub> .f0 <sub>2</sub> .DurS <sub>2</sub>	0.107	3.47	.063
M <sub>2</sub> .dB <sub>2</sub> .DurS <sub>2</sub>	0.205	4.487	.035
M <sub>2</sub> .f0 <sub>2</sub> .dB <sub>2</sub> .DurS <sub>2</sub>	0.061	18.184	3 <sup>E-05</sup>
M <sub>2</sub> .F.f0 <sub>1</sub> .dB <sub>1</sub>	0.8	0.404	.53
M <sub>2</sub> .F.f0 <sub>1</sub> .DurS <sub>1</sub>	-0.136	4.086	.044
M <sub>2</sub> .F.dB <sub>1</sub> .DurS <sub>1</sub>	-0.165	2.197	.1
M <sub>2</sub> .F.f0 <sub>1</sub> .dB <sub>1</sub> .DurS <sub>1</sub>	0.046	13.324	.0003
M <sub>2</sub> .F.f0 <sub>2</sub> .dB <sub>2</sub>	-0.646	0.193	.67
M <sub>2</sub> .F.f0 <sub>2</sub> .DurS <sub>2</sub>	-0.488	19.709	1 <sup>E-05</sup>
M <sub>2</sub> .F.dB <sub>2</sub> .DurS <sub>2</sub>	0.073	0.151	.70
M <sub>2</sub> .F.f0 <sub>2</sub> .dB <sub>2</sub> .DurS <sub>2</sub>	-0.039	2.023	.16
f0 <sub>1</sub> .f0 <sub>2</sub>	0.259	6.808	.0091
dB <sub>1</sub> .dB <sub>2</sub>	-0.036	0.007	.94
DurS <sub>1</sub> .DurS <sub>2</sub>	0.001	0.251	.62
f0 <sub>1</sub> .f0 <sub>2</sub> .dB <sub>1</sub> .dB <sub>2</sub>	-0.015	1.008	.32
f0 <sub>1</sub> .f0 <sub>2</sub> .DurS <sub>1</sub> .DurS <sub>2</sub>	0	2.633	.11
dB <sub>1</sub> .dB <sub>2</sub> .DurS <sub>1</sub> .DurS <sub>2</sub>	0	1.702	.20
f0 <sub>1</sub> .f0 <sub>2</sub> .dB <sub>1</sub> .dB <sub>2</sub> .DurS <sub>1</sub> .DurS <sub>2</sub>	0	0.769	.39

	Coefficient	Chi <sup>2</sup> (1, N=1835)	p<
F.f0 <sub>1</sub> .f0 <sub>2</sub> .dB <sub>1</sub> .dB <sub>2</sub>	0.023	1.639	.21
F.f0 <sub>1</sub> .f0 <sub>2</sub> .DurS <sub>1</sub> .DurS <sub>2</sub>	0	0.117	.74
F.dB <sub>1</sub> .dB <sub>2</sub> .DurS <sub>1</sub> .DurS <sub>2</sub>	0.001	1.782	.19
F.f0 <sub>1</sub> .f0 <sub>2</sub> .dB <sub>1</sub> .dB <sub>2</sub> .DurS <sub>1</sub> .DurS <sub>2</sub>	0	7.445	<b>.0064</b>
M <sub>2</sub> .f0 <sub>1</sub> .f0 <sub>2</sub> .dB <sub>1</sub> .dB <sub>2</sub>	0.041	4.339	<b>.038</b>
M <sub>2</sub> .f0 <sub>1</sub> .f0 <sub>2</sub> .DurS <sub>1</sub> .DurS <sub>2</sub>	0	0.085	.78
M <sub>2</sub> .dB <sub>1</sub> .dB <sub>2</sub> .DurS <sub>1</sub> .DurS <sub>2</sub>	0	0.626	.43
M <sub>2</sub> .f0 <sub>1</sub> .f0 <sub>2</sub> .dB <sub>1</sub> .dB <sub>2</sub> .DurS <sub>1</sub> .DurS <sub>2</sub>	0	1.984	.16
F.M <sub>2</sub> .F.f0 <sub>1</sub> .f0 <sub>2</sub> .dB <sub>1</sub> .dB <sub>2</sub>	-0.057	6.703	<b>.0097</b>
F.M <sub>2</sub> .F.f0 <sub>1</sub> .f0 <sub>2</sub> .DurS <sub>1</sub> .DurS <sub>2</sub>	0	0.687	.41
F.M <sub>2</sub> .F.dB <sub>1</sub> .dB <sub>2</sub> .DurS <sub>1</sub> .DurS <sub>2</sub>	0	0.204	.66
F.M <sub>2</sub> .F.f0 <sub>1</sub> .f0 <sub>2</sub> .dB <sub>1</sub> .dB <sub>2</sub> .DurS <sub>1</sub> .DurS <sub>2</sub>	0	3.299	<b>.07</b>
DurIt	4.466	0.737	.40
PRIt	0.077	0.067	.80
PR <sub>2</sub>	0.187	0.423	.52
DurIt.PRIt	0.004	6.319	<b>.012</b>
Lat	-11.444	0.205	.66
Ordr	-0.083	0.438	.51
DurS <sub>2</sub> .DurIt.PRIt	0	22.075	<b>3<sup>E-06</sup></b>
DurS <sub>2</sub> .DurIt.PR <sub>2</sub>	0	10.719	<b>.0011</b>
PRIt.PR <sub>2</sub> .DurS <sub>2</sub>	0	38.199	<b>7<sup>E-10</sup></b>
PRIt.PR <sub>2</sub> .DurIt	0	18.749	<b>2<sup>E-05</sup></b>
DurS <sub>2</sub> .PRIt.PR <sub>2</sub> .DurIt	0	8.648	<b>.0033</b>
PR <sub>2</sub> .DurS <sub>2</sub>	-0.004	3.083	<b>.08</b>

## Annexe 42

## Expérience 16 : Matériel linguistique

Tableau 184 : Mots expérimentaux et distracteurs, et pseudo-mots distracteurs

Matériel expérimental	Matériel distracteur		
	Mots	Mots	Pseudo-mots
bagak	banan	bjegak	kɛ̃nan
butik	bukē	bletik	blekē
kalis	karyʃ	plqilis	tɛ̃pyʃ
lisās	lidœk	kɛ̃sās	ʒwadœk
tēbal	tētyk	kɛ̃wabal	dɛ̃tyk
bɛ̃kɔl	bɛ̃igad	pɛ̃wakɔl	pɛ̃wagad
mustaʃ	muslin	tɛ̃taʃ	gɛ̃ylin
tɛ̃mit	tɛ̃mos	bwamit	plqimos
tɛ̃māt	tɛ̃nyk	fɛ̃māt	fɛ̃nyk
gitak	gitun	pɛ̃tak	gɛ̃tun
vɛ̃pɛk	vites	ʒwapɛk	dɛ̃ates
vēkœk	vēten	plokœk	kɛ̃waten
femyk	fenɔl	gɛ̃myk	bjetys
lɔpē	lotys	bɛ̃qipē	gɛ̃wēnɔl
ɛ̃uset	ɛ̃utin	gɛ̃yset	ɛ̃tɪn
sɛ̃ɛn	sigɔɲ	klāɛn	fɛ̃ɔgɔɲ
pɛ̃lwak	pɛ̃tās	ɛ̃lɛwak	pɛ̃tās
kalvɛk	kalsjɔm	fɛ̃ɔvɛk	glāsɔm
tɛ̃āspɛk	tɛ̃āsfyʒ	gɛ̃pɛk	gɛ̃fyʒ
kantō	kanlyk	glātō	fɛ̃qilyk
ʃalup	ʃamel	tɛ̃ɔlup	tɛ̃ɔmel
kɛ̃kɔʒ	kutym	bɛ̃kɔʒ	bwatym
delyʒ	dekad	tɛ̃ulyʒ	bɛ̃kɛkad
manɛʒ	mamut	gɛ̃wēnɛʒ	plomut
pɛ̃sjās	palet	dɛ̃sɛjās	klālet
ʃɛ̃pāt	ʃɛ̃ʒœk	plɛ̃pāt	bɛ̃qipɛk
bɛ̃bɪʃ	bɛ̃da	kɛ̃bɪʃ	kɛ̃da
fɛ̃myl	fɛ̃sɛps	fɛ̃qimyl	plɔsɛps
kɛ̃tuʃ	kɛ̃lēg	flotuʃ	flolēg
kɔ̃nɪʃ	kɔ̃bɛʒ	tɛ̃nɪʃ	tɛ̃bɛʒ
ɛ̃dak	ɛ̃dɔm	flydak	kɛ̃idɔm
ɛ̃ɔflœk	ɛ̃ɔdel	stɛ̃ɪflœk	tɛ̃qidɛl
ɛ̃āɛn	ɛ̃ākyn	kɛ̃igen	bɛ̃akyn
pɛ̃pɛk	pɛ̃pɛt	tɛ̃qipɛk	plɛ̃pɛt
sētez	sēdik	nøtez	gnudik
lɛ̃bē	lɛ̃ʒɛs	klubē	kluzɛs

<b>Matériel expérimental</b>	<b>Matériel distracteur</b>		
<b>Mots</b>	<b>Mots</b>	<b>Pseudo-mots</b>	
seʃfœj	seʃvœʁ	bʁafœj	stœivœʁ
fœʁmaz	fœʁmje	gnumaz	nœbʁyʁ
makmaj	makbʁyʁ	dʁamaj	flyʃō
tœʁœʁ	tœʃō	plōʁœʁ	tœumje

## Annexe 43

### Expérience 16 : Équation du modèle et tests de significativité de l'effet des covariables sur les temps de réponse

#### Équation 67 : Équation du modèle étudié

$$TR_{It, suj} \sim N(XB, \Omega)$$

$$TR_{It, suj} = \beta_{0It, suj} \text{Cons} + -376.224(124.517)F_{It, suj} + 6.036(5.638)f0_{1It, suj} + -7.380(4.224)f0_{2It, suj} + 8.048(8.136)dB_{1It, suj} + -39.988(14.810)dB_{2It, suj} + 3.900(1.228)dur_{1It, suj} + -2.758(0.849)dur_{2It, suj} + 19.111(9.341)F.f0_{1It, suj} + 4.677(7.419)F.f0_{2It, suj} + 38.133(24.466)F.dB_{1It, suj} + 48.111(15.533)F.dB_{2It, suj} + 0.932(1.291)F.dur_{1It, suj} + 5.417(1.132)F.dur_{2It, suj} + 0.042(0.760)f0_{1 .dB_{1It, suj}} + 0.132(0.072)f0_{1 .dur_{1It, suj}} + -0.014(0.043)dB_{1 .dur_{1It, suj}} + -0.003(0.007)f0_{1 .dB_{1 .dur_{1It, suj}}} + 0.476(1.231)f0_{2 .dB_{2It, suj}} + 0.190(0.126)f0_{2 .dur_{2It, suj}} + 0.744(0.201)dB_{2 .dur_{2It, suj}} + -0.009(0.034)f0_{2 .dB_{2 .dur_{2It, suj}}} + -4.843(1.980)F.f0_{1 .dB_{1It, suj}} + -0.287(0.101)F.f0_{1 .dur_{1It, suj}} + -0.413(0.145)F.dB_{1 .dur_{1It, suj}} + 0.046(0.016)F.f0_{1 .dB_{1 .dur_{1It, suj}}} + -2.582(1.675)F.f0_{2 .dB_{2It, suj}} + 0.039(0.144)F.f0_{2 .dur_{2It, suj}} + -0.650(0.214)F.dB_{2 .dur_{2It, suj}} + -0.028(0.037)F.f0_{2 .dB_{2 .dur_{2It, suj}}} + 0.733(0.270)f0_{1 .f0_{2It, suj}} + 0.078(0.979)dB_{1 .dB_{2It, suj}} + 0.007(0.005)dur_{1 .dur_{2It, suj}} + 0.053(0.012)f0_{1 .dB_{1 .f0_{2 .dB_{2It, suj}}}} + 0.000(0.000)f0_{1 .dur_{1 .f0_{2 .dur_{2It, suj}}}} + 0.000(0.000)dB_{1 .dur_{1 .dB_{2 .dur_{2It, suj}}}} + 0.000(0.000)f0_{1 .dB_{1 .dur_{1 .f0_{2 .dB_{2 .dur_{2It, suj}}}}} + -0.072(0.024)F.f0_{1 .dB_{1 .f0_{2 .dB_{2It, suj}}} + 0.000(0.000)F.f0_{1 .dur_{1 .f0_{2 .dur_{2It, suj}}} + 0.001(0.000)F.dB_{1 .dur_{1 .dB_{2 .dur_{2It, suj}}} + 0.000(0.000)F.f0_{1 .dB_{1 .dur_{1 .f0_{2 .dB_{2 .dur_{2It, suj}}}}} + -2.474(0.937)PRi_{It, suj} + 2.882(0.928)PR2_{It, suj} + -0.399(0.191)DurI_{It, suj} + 0.001(0.002)DurI .PRi_{It, suj} + 0.026(0.004)PR2 .dur_{2It, suj} + 35.461(61.766)Lat_{suj} + 0.269(0.311)Ordr_{It, suj} + 0.000(0.000)dur2 .DurI .PRi_{It, suj} + 0.000(0.000)dur2 .DurI .PR2_{It, suj} + 0.000(0.000)PR2 .dur_{2It, suj} + 0.000(0.000)PRi .PR2 .DurI_{It, suj} + 0.000(0.000)dur2 .PRi .PR2 .DurI_{It, suj}$$

$$\beta_{0It, suj} = 1451.113(78.181) + u_{0suj} + e_{0It, suj}$$

$$[u_{0suj}] \sim N(0, \Omega_u) : \Omega_u = [17184.510(3715.016)]$$

$$[e_{0It, suj}] \sim N(0, \Omega_e) : \Omega_e = [44666.230(2131.782)]$$

$$-2 * \log\text{likelihood (IGLS Deviance)} = 12792.150(936 \text{ of } 1160 \text{ cases in use})$$

Avec : TR = temps de réponse (V.D., en ms.) ; F = frontière lexicale entre les deux syllabes (V.I. - réf = absence de frontière) ; f0<sub>1/2</sub> = fréquence fondamentale moyenne de la première/deuxième syllabe (Hz) ; dB<sub>1/2</sub> = intensité moyenne de la première/deuxième syllabe (dB) ; dur<sub>1/2</sub> = durée de la première/deuxième syllabe (ms.) ; PRi = position acoustique du point de reconnaissance de l'item (ms.) ; PR<sub>2</sub> = position acoustique du point de reconnaissance de la deuxième syllabe (ms.) ; Lat = latéralité manuelle des participants (réf = gauche) ; Ordr = ordre de présentation des items

Tableau 185 : Tests de significativité des covariables, et de leurs interactions entre elles et avec la variable expérimentale

	Coefficient	Chi <sup>2</sup> <sub>(1, N=936)</sub>	p<
f0 <sub>1</sub>	6.036	1.146	.29
f0 <sub>2</sub>	-7.38	3.053	.081
dB <sub>1</sub>	8.048	0.979	.33
dB <sub>2</sub>	-39.988	7.291	.007

	Coefficient	Chi <sup>2</sup> (1, N=936)	p<
dur <sub>1</sub>	3.9	10.088	<b>.0015</b>
dur <sub>2</sub>	-2.758	10.565	<b>.0012</b>
F.f <sub>01</sub>	19.111	4.186	<b>.041</b>
F.f <sub>02</sub>	4.677	0.397	.53
F.dB <sub>1</sub>	38.133	2.429	.12
F.dB <sub>2</sub>	48.111	9.594	<b>.002</b>
F.dur <sub>1</sub>	0.932	0.521	.48
F.dur <sub>2</sub>	5.417	22.908	<b>2<sup>E-06</sup></b>
f <sub>01</sub> .dB <sub>1</sub>	0.042	0.003	.96
f <sub>01</sub> .dur <sub>1</sub>	0.132	3.41	<b>.065</b>
dB <sub>1</sub> .dur <sub>1</sub>	-0.014	0.102	.75
f <sub>01</sub> .dB <sub>1</sub> .dur <sub>1</sub>	-0.003	0.152	.70
f <sub>02</sub> .dB <sub>2</sub>	0.476	0.15	.70
f <sub>02</sub> .dur <sub>2</sub>	0.19	2.293	.13
dB <sub>2</sub> .dur <sub>2</sub>	0.744	13.759	<b>.0003</b>
f <sub>02</sub> .dB <sub>2</sub> .dur <sub>2</sub>	-0.009	0.075	.79
F.f <sub>01</sub> .dB <sub>1</sub>	-4.843	5.983	<b>.015</b>
F.f <sub>01</sub> .dur <sub>1</sub>	-0.287	8.025	<b>.0047</b>
F.dB <sub>1</sub> .dur <sub>1</sub>	-0.413	8.073	<b>.0045</b>
F.f <sub>01</sub> .dB <sub>1</sub> .dur <sub>1</sub>	0.046	8.09	<b>.0045</b>
F.f <sub>02</sub> .dB <sub>2</sub>	-2.582	2.376	.13
F.f <sub>02</sub> .dur <sub>2</sub>	0.039	0.072	.79
F.dB <sub>2</sub> .dur <sub>2</sub>	-0.65	9.183	<b>.0025</b>
F.f <sub>02</sub> .dB <sub>2</sub> .dur <sub>2</sub>	-0.028	0.556	.46
f <sub>01</sub> .f <sub>02</sub>	0.733	7.367	<b>.0067</b>
dB <sub>1</sub> .dB <sub>2</sub>	0.078	0.006	.94
dur <sub>1</sub> .dur <sub>2</sub>	0.007	1.891	.17
f <sub>01</sub> .f <sub>02</sub> .dB <sub>1</sub> .dB <sub>2</sub>	0.053	20.798	<b>6<sup>E-06</sup></b>
f <sub>01</sub> .f <sub>02</sub> .dur <sub>1</sub> .dur <sub>2</sub>	0	0.001	.98
dB <sub>1</sub> .dB <sub>2</sub> .dur <sub>1</sub> .dur <sub>2</sub>	0	0.394	.54
f <sub>01</sub> .f <sub>02</sub> .dB <sub>1</sub> .dB <sub>2</sub> .dur <sub>1</sub> .dur <sub>2</sub>	0	2.233	.14
F.f <sub>01</sub> .f <sub>02</sub> .dB <sub>1</sub> .dB <sub>2</sub>	-0.072	9.057	<b>.0027</b>
F.f <sub>01</sub> .f <sub>02</sub> .dur <sub>1</sub> .dur <sub>2</sub>	0	0.023	.88
F.dB <sub>1</sub> .dB <sub>2</sub> .dur <sub>1</sub> .dur <sub>2</sub>	0.001	8.259	<b>.0041</b>
F.f <sub>01</sub> .f <sub>02</sub> .dB <sub>1</sub> .dB <sub>2</sub> .dur <sub>1</sub> .dur <sub>2</sub>	0	0.002	.97
PR <sub>i</sub>	-2.474	6.968	<b>.0083</b>
PR <sub>2</sub>	2.882	9.641	<b>.002</b>
Dur <sub>I</sub>	-0.399	4.38	<b>.037</b>
Dur <sub>I</sub> .PR <sub>i</sub>	0.001	0.177	.68
PR <sub>2</sub> .dur <sub>2</sub>	0.026	40.124	<b>3<sup>E-10</sup></b>
Lat	35.461	0.33	.57
Ordr	0.269	0.747	.39
dur <sub>2</sub> .Dur <sub>I</sub> .PR <sub>i</sub>	0	0.366	.55
dur <sub>2</sub> .Dur <sub>I</sub> .PR <sub>2</sub>	0	7.74	<b>.0055</b>
PR <sub>i</sub> .PR <sub>2</sub> .dur <sub>2</sub>	0	8.153	<b>.0043</b>
PR <sub>i</sub> .PR <sub>2</sub> .Dur <sub>I</sub>	0	0.344	.56
dur <sub>2</sub> .PR <sub>i</sub> .PR <sub>2</sub> .Dur <sub>I</sub>	0	0.078	.79

## Index des Équations

Équation 1 : Équation du modèle linéaire de base pour l'ANOVA d'une expérience à mesures répétées (reproduit à partir de (Raaijmakers et al., 1999)).....	120
Équation 2 : Calcul de $F_1$ .....	121
Équation 3 : Calcul de $F_2$ .....	121
Équation 4 : Calcul du degré de corrélation intra classe.....	126
Équation 5 : Équation du modèle vide (exemple issu des données de l'Expérience 4).....	127
Équation 6 : Fonction permettant de calculer l'énergie infra-lexicale $E_i$ .....	391
Équation 7 : Fonction $V_{ij}$ permettant de calculer la valence informationnelle d'un lexème $x_{ij}$ dans la matrice lexicale.....	391
Équation 8 : Fonction de pondération $P(x_{ij})$ d'un lexème $x_{ij}$ dans la matrice lexicale $ij$ .....	392
Équation 9 : Fonction $f_1(x_{ij})$ de prégnance perceptive d'un mot stimulus.....	393
Équation 10 : Fonction $f_2(x_{ij})$ de prégnance perceptive d'un mot stimulus.....	393
Équation 11 : Fonction $f_3(x_{ij})$ de prégnance perceptive d'un mot stimulus.....	394
Équation 12 : Fonction $f_4(x_{ij})$ de prégnance perceptive d'un mot stimulus.....	394
Équation 13 : Fonction $f_{4na}(x_{ij})$ de prégnance perceptive d'un mot non aligné.....	394
Équation 14 : Fonction $f_{5na}(x_{ij})$ de prégnance perceptive d'un mot non aligné.....	395
Équation 15 : Fonction $f_5(x_{ij})$ de prégnance perceptive d'un mot stimulus.....	395
Équation 16 : Équation du modèle étudiant l'effet des facteurs âge d'acquisition, certitude sur l'estimation de l'âge, et familiarité, sur le nombre de réponses produites par les participants.....	437
Équation 17 : Équation du modèle étudiant l'influence des prédicteurs de l'âge d'acquisition.....	451
Équation 18 : Équation du modèle étudiant l'influence des prédicteurs de la fréquence .....	453
Équation 19 : Équation du modèle étudiant l'influence des prédicteurs de la familiarité .....	454
Équation 20 : Analyse de la durée .....	463
Équation 21 : Analyse de la $f_0$ moyenne.....	463
Équation 22 : Analyse des données de classification correcte .....	464
Équation 23 : Analyse des données de classification incorrecte <sup>page 464</sup> .....	465
Équation 24 : Analyse de la comparaison spécifique des taux d'activation correspondant aux confusions entre catégories attendues et reconnues (cas de confusions entre catégories) <sup>page 464</sup>	466
Équation 25 : Effet de la quantité de syllabes similaires à une même position dans les 2 items comparés, sur la performance de discrimination des items par les réseaux .....	470
Équation 26 : Effet du type de syllabe en commun à une même position dans les 2 items comparés, sur la performance de discrimination des items par les réseaux .....	470
Équation 27 : Effet du type d'item 1 et 2 comparés, sur la performance de discrimination des items par les réseaux <sup>page 464</sup> , .....	471
Équation 28 : Effet de la quantité de syllabes similaires à une même position dans les 2 items comparés, sur la performance de discrimination des items par les auditeurs (taux de réponses « similaire ») .....	472
Équation 29 : Effet du type de similitude à une position donnée dans les 2 items comparés, sur la performance de discrimination des items par les auditeurs (taux de réponses « similaire ») .....	473

Équation 30 : Effet du type de frontière sur la performance de discrimination des items par les auditeurs (taux de réponses « similaire ») .....	473
Équation 31 : Analyse générale de la répartition des catégories de réponse obtenues relativement à la situation présentée (1 modèle par catégorie de réponse : $P_0, P_1, P_2$ ) .....	474
Équation 32 : Calcul du pourcentage de réponses à partir des valeurs fournies par le modèle ....	476
Équation 33 : Analyse détaillée de la répartition des réponses en fonction du type de frontière testée .....	476
Équation 34 : Équation du modèle étudié .....	481
Équation 35 : Évaluation de la relation entre facteurs expérimentaux (V.I.) et paramètres prosodiques (V.D.) .....	494
Équation 36 : Équation du modèle .....	496
Équation 37 : Équation du modèle étudié .....	497
Équation 38 : Équation du modèle étudié .....	500
Équation 39 : Estimation du rapport entre paramètres prosodiques ( $f_0$ , intensité et durée) et facteurs expérimentaux pour la première syllabe.....	502
Équation 40 : Estimation du rapport entre paramètres prosodiques ( $f_0$ , intensité et durée) et facteurs expérimentaux pour la deuxième syllabe.....	503
Équation 41 : Équation du modèle étudié .....	506
Équation 42 : Estimation du rapport entre paramètres prosodiques ( $f_0$ , intensité et durée) et facteurs expérimentaux pour la première syllabe.....	510
Équation 43 : Estimation du rapport entre paramètres prosodiques ( $f_0$ , intensité et durée) et facteurs expérimentaux pour la deuxième syllabe.....	510
Équation 44 : Équation du modèle étudié .....	514
Équation 45 : Équation du modèle étudié .....	517
Équation 46 : Équation du modèle étudié .....	519
Équation 47 : Équation du modèle étudié .....	520
Équation 48 : Estimation du rapport entre paramètres prosodiques ( $f_0$ , intensité et durée) et facteurs expérimentaux.....	523
Équation 49 : Équation du modèle étudié .....	526
Équation 50 : Estimation du rapport entre paramètres prosodiques ( $f_0$ , intensité et durée) et facteurs expérimentaux.....	529
Équation 51 : Équation du modèle étudié .....	532
Équation 52 : Effet de la séquence de phonèmes en tant que syllabe sur les temps de réponse..	533
Équation 53 : Effet de la séquence de phonèmes en tant que syllabe ou non sur les temps de réponses .....	534
Équation 54 : Estimation du rapport entre paramètres prosodiques ( $f_0$ , intensité et durée) et le facteur expérimental.....	536
Équation 55 : Équation du modèle étudié .....	539
Équation 56 : Estimation du rapport entre paramètres prosodiques ( $f_0$ , intensité et durée) et le facteur expérimental.....	542
Équation 57 : Équation du modèle étudié .....	545
Équation 58 : Effets de la fréquence des items en tant que syllabe sur les temps de réponse des participants (Expérience 13) .....	547
Équation 59 : Effets du nombre de mots dans lesquels les items sont insérés en tant que syllabe sur les temps de réponse des participants (Expérience 13) .....	549



Équation 60 : Effets de la fréquence des items en tant que séquence de phonèmes sur les temps de réponse des participants (Expérience 13) .....	550
Équation 61 : Effets du nombre de mots dans lesquels les items sont insérés en tant que séquence de phonèmes sur les temps de réponse des participants (Expérience 13) .....	552
Équation 62 : Effets de la fréquence des items en tant que syllabe sur les temps de réponse des participants (Expérience 14) .....	554
Équation 63 : Effets du nombre de mots dans lesquels les items sont insérés en tant que syllabe sur les temps de réponse des participants (Expérience 14) .....	555
Équation 64 : Effets de la fréquence des items en tant que séquence de phonèmes sur les temps de réponse des participants (Expérience 14) .....	557
Équation 65 : Effets du nombre de mots dans lesquels les items sont insérés en tant que séquence de phonèmes sur les temps de réponse des participants (Expérience 14) .....	559
Équation 66 : Équation du modèle étudié .....	563
Équation 67 : Équation du modèle étudié .....	568

## Index des Figures

Figure 1 : Déroulement schématique de la reconnaissance des mots.....	11
Figure 2 : Déroulement détaillé du processus d'accès ascendant au lexique (modèle calculatoire)	20
Figure 3 : Exploration exhaustive du lexique selon un alignement sur chaque segment, lors de la reconnaissance automatique de la séquence « recognize speech » par une machine (reproduit à partir de (Shillcock, 1990)).....	26
Figure 4 : Codage du système de traits dans la couche d'entrée – exemple pour deux des sept traits .....	34
Figure 5 : Organisation de la couche des phonèmes.....	35
Figure 6 : Organisation de la couche des mots .....	36
Figure 7 : Niveau calculatoire dans le modèle TRACE .....	37
Figure 8 : Simulation de la reconnaissance du mot « <i>catalog</i> » par le modèle TRACE. Évolution des courbes d'activation du mot stimulus (« <i>catalog</i> ») et des 2 mots enchâssés qu'il renferme (« <i>cat</i> » & « <i>log</i> »).....	40
Figure 9 : Hiérarchisation de l'information linguistique .....	46
Figure 10 : Architecture du système MICRO.....	56
Figure 11 : Distribution positionnelle des mots monosyllabiques dans des mots anglais de 2 à 6 syllabes.....	63
Figure 12 : Diversité syllabique des mots en fonction de leur taille .....	64
Figure 13 : Distribution des mots du lexique analysé et des syllabes les composant, respectivement en fonction du nombre de syllabes qu'ils renferment / de la position syllabique dans le mot .....	65
Figure 14 : Organisation syllabique du lexique français analysé.....	66
Figure 15 : Distribution positionnelle des syllabes en début et en fin de mots polysyllabiques, en fonction de leur lexicalité.....	67
Figure 16 : Nombre d'entrées lexicales commençant / finissant par une syllabe lexicale / non lexicale .....	68
Figure 17 : Fréquence syllabique formelle en fonction de la lexicalité des syllabes et de leur position dans les mots .....	69
Figure 18 : Facteurs influençant les résultats dans une étude expérimentale sur la reconnaissance des mots parlés.....	80
Figure 19 : Facteurs primaires influençant la production des données .....	85
Figure 20 : Localisation des différentes origines possibles dans un mot .....	95
Figure 21 : Exemple de temps de réponse mesurés depuis le début du mot.....	97
Figure 22 : Exemple de temps de réponse mesurés depuis la fin du mot .....	98
Figure 23 : Exemple de temps de réponse mesurés depuis le début du PUP.....	98
Figure 24 : Oscillogramme et spectrogramme du mot « <i>baisse</i> » .....	102
Figure 25 : Oscillogramme et spectrogramme du mot « <i>bac</i> ».....	102
Figure 26 : Premier exemple d'effets indésirables liés à un filtre mal adapté .....	109
Figure 27 : Deuxième exemple d'effets indésirables liés à un filtre mal adapté .....	110
Figure 28 : Schématisation des notions de centre de gravité et de dispersion autour de la moyenne pour un exemple fictif de distribution .....	113

Figure 29 : État de la distribution après remplacement des valeurs manquantes par la moyenne générale.....	114
Figure 30 : État de la distribution après remplacement des valeurs manquantes par la moyenne par condition .....	115
Figure 31 : Exemple .....	130
Figure 32 : Architecture générale de la séquence acoustique utilisée comme stimulus .....	144
Figure 33 : Durée et f0 moyenne de chaque catégorie de syllabe .....	167
Figure 34 : Courbes de f0 des syllabes D du corpus .....	168
Figure 35 : Courbes de f0 des syllabes F du corpus.....	169
Figure 36 : Courbes de f0 des syllabes L du corpus.....	169
Figure 37 : Courbes de f0 des syllabes G du corpus .....	170
Figure 38 : Profils moyens de f0 / durée pour les 4 catégories de syllabes du corpus .....	170
Figure 39 : Profils moyens de f0 / durée pour les 9 types de frontières étudiés .....	171
Figure 40 : Performances moyennes du classificateur .....	180
Figure 41 : Taux moyens d'activation des unités de sortie en fonction du type de frontière.....	182
Figure 42 : Taux moyen d'activation des unités de sortie estimés par le modèle en fonction du nombre de similitudes catégorielles et positionnelles.....	189
Figure 43 : Taux d'activation de sortie estimé par le modèle, lié à l'effet de la similitude catégorielle d'une syllabe à la même position dans les 2 items comparés .....	190
Figure 44 : Taux d'activation de sortie estimé par le modèle, lié à l'effet de la similitude catégorielle d'une syllabe à la même position dans les 2 items comparés .....	191
Figure 45 : Représentation graphique des données du Tableau 34 .....	192
Figure 46 : Taux d'activation en sortie de réseau, estimé par le modèle pour les comparaisons intra catégories, en fonction du type de frontière considéré.....	193
Figure 47 : Organisation spatiale de la caractérisation prosodique des 9 catégories de frontière .	194
Figure 48 : Nombre moyen de réponses « similaire », en fonction du nombre de syllabes de même catégorie entre les deux items comparés.....	202
Figure 49 : Taux de réponses « similaire » en fonction d'une similitude catégorielle de la première ou la deuxième syllabe.....	203
Figure 50 : Taux de réponses « similaire » en fonction de l'identité catégorielle des syllabes en première et deuxième positions dans les 2 items comparés .....	204
Figure 51 : Taux de réponse « similaire » en fonction du type de frontière, lors de comparaisons intra catégorie .....	206
Figure 52 : Répartition des réponses obtenues en fonction des réponses attendues .....	215
Figure 53 : Temps de réponse moyens estimés en fonction du débit d'élocution et de la hauteur de la voix.....	225
Figure 54 : Temps moyens de reconnaissance des mots en fonction de la présence d'une superposition réelle ou virtuelle dans la situation.....	240
Figure 55 : Temps moyens de reconnaissance des mots, estimés par le modèle, en fonction de la présence d'un enchaînement initial.....	246
Figure 56 : Temps de réponse moyens estimés en fonction de la lexicalité des syllabes du mot stimulus.....	254
Figure 57 : Vue d'ensemble des temps de réponse moyens estimés par condition .....	261
Figure 58 : Temps de réponse moyens estimés en fonction de la lexicalité des syllabes du mot stimulus .....	262

Figure 59 : Temps moyens estimés de reconnaissance des mots dans la condition contextuelle TDL.DerP.....	264
Figure 60 : Temps moyens de reconnaissance des mots, en fonction de la présence d'un mot enchâssé initial .....	270
Figure 61 : Position de la cible dans l'item .....	274
Figure 62 : Présentation des temps moyens estimés de détection des cibles par condition.....	280
Figure 63 : Présentation des écarts à la moyenne générale, pour les temps moyens estimés de détection des cibles par condition .....	281
Figure 64 : Influence de chacun des 2 facteurs expérimentaux sur les temps de détection des cibles, en fonction de la position de celles-ci dans l'item.....	282
Figure 65 : Coefficients estimés relatifs à la lexicalité des syllabes du mot stimulus et à la position de la cible pour chaque facteur expérimental et leur interaction .....	295
Figure 66 : Temps moyens estimés de détection des cibles en fonction de leur position et du type de mot dans lequel elles sont insérées .....	297
Figure 67 : Temps moyens de détection des syllabes cible en fonction de leur fréquence positionnelle .....	309
Figure 68 : Temps de détection moyens estimés des syllabes en fonction de leur fréquence d'occurrence dans le lexique .....	315
Figure 69 : Temps de reconnaissance des mots en fonction de leur fréquence d'occurrence à l'initiale d'autres mots du lexique .....	322
Figure 70 : Temps de reconnaissance des mots en fonction de leur fréquence d'occurrence en position non initiale dans le lexique .....	328
Figure 71 : Temps de reconnaissance des mots en fonction de la présence d'une frontière lexicale .....	348
Figure 72 : Temps de détection des mots en position de 2 <sup>ème</sup> syllabe dans le stimulus, en fonction de la présence d'une frontière lexicale prosodique entre les syllabes de l'item.....	352
Figure 73 : Organisation du traitement du mot « <i>garage</i> » dans SLEXS1 .....	385
Figure 74 : Diagrammes de l'organisation superposante des mots « <i>garage</i> » et « <i>chenal</i> ».....	396
Figure 75. Distribution du nombre de mots traités par participant (courbe : distance pondérée des moindres carrés).....	433
Figure 76. Distribution du nombre de mots en fonction du nombre de réponses par mot .....	435
Figure 77. Distribution du nombre de mots en fonction de l'âge d'acquisition estimé (courbe : distance pondérée des moindres carrés).....	440
Figure 78. Distribution du nombre de mots en fonction du degré de certitude sur l'âge d'acquisition estimé (courbe : distance pondérée des moindres carrés) .....	441
Figure 79. Distribution du nombre de mots en fonction de la familiarité (courbe : distance pondérée des moindres carrés) .....	442
Figure 80. Distribution des réponses moyennes par item en fonction de l'âge d'acquisition estimé et du degré de certitude de l'estimation (courbe : distance pondérée des moindres carrés) .....	444
Figure 81. Distribution des réponses moyennes par item en fonction de l'âge d'acquisition estimé et de la familiarité (courbe : distance pondérée des moindres carrés) .....	447
Figure 82. Distribution des réponses moyennes par item en fonction de la fréquence d'occurrence et de l'âge d'acquisition (courbe : distance pondérée des moindres carrés).....	448
Figure 83. Distribution des réponses moyennes par item en fonction de la fréquence d'occurrence et de la familiarité (courbe : distance pondérée des moindres carrés) .....	449
Figure 84. Relation entre âge d'acquisition estimé et familiarité .....	452
Figure 85. Relation entre âge d'acquisition estimé et fréquence d'occurrence .....	453

Figure 86. Relation entre la familiarité et la fréquence d'occurrence .....	454
Figure 87 : Proportion de réponses expliquées par les caractéristiques acoustiques des items, pour les comparaisons intra catégorie .....	486
Figure 88 : Proportion de réponses expliquées par les caractéristiques acoustiques des items, pour les comparaisons inter catégories .....	492
Figure 89 : Relations entre facteurs lexicaux et fréquence fondamentale .....	495
Figure 90 : Relations entre facteurs lexicaux et intensité .....	495
Figure 91 : Relations entre facteurs expérimentaux et durée .....	496
Figure 92 : Effet des facteurs expérimentaux sur la fréquence fondamentale des syllabes .....	504
Figure 93 : Effet des facteurs expérimentaux sur l'intensité des syllabes .....	504
Figure 94 : Effet des facteurs expérimentaux sur la durée des syllabes .....	505
Figure 95 : Effet des facteurs expérimentaux sur la fréquence fondamentale des syllabes .....	512
Figure 96 : Effet des facteurs expérimentaux sur l'intensité des syllabes .....	512
Figure 97 : Effet des facteurs expérimentaux sur la durée des syllabes .....	513
Figure 98 : Effet des facteurs expérimentaux sur la fréquence fondamentale moyenne des syllabes .....	524
Figure 99 : Effet des facteurs expérimentaux sur l'intensité fondamentale moyenne des syllabes .....	524
Figure 100 : Effet des facteurs expérimentaux sur la durée des syllabes .....	525
Figure 101 : Effet du facteur expérimental sur la fréquence fondamentale des syllabes .....	530
Figure 102 : Effet du facteur expérimental sur l'intensité moyenne des syllabes .....	530
Figure 103 : Effet du facteur expérimental sur la durée des syllabes .....	531
Figure 104 : Rapport entre le facteur expérimental et la fréquence fondamentale de la syllabe produite .....	537
Figure 105 : Rapport entre le facteur expérimental et l'intensité de la syllabe produite .....	537
Figure 106 : Rapport entre le facteur expérimental et la durée de la syllabe produite .....	538
Figure 107 : Rapport entre le facteur expérimental et la fréquence fondamentale moyenne de la syllabe .....	543
Figure 108 : Rapport entre le facteur expérimental et l'intensité moyenne de la syllabe .....	543
Figure 109 : Rapport entre le facteur expérimental et la durée de la syllabe .....	544

## Index des Tableaux

Tableau 1 : Organisation des connexions entre unités .....	36
Tableau 2 : Les différentes composantes de la prosodie du Français .....	43
Tableau 3 : Paramètres prosodiques : relations entre acoustique, production, et perception .....	43
Tableau 4 : Types de syllabes pouvant être théoriquement définis sur le plan acoustique .....	48
Tableau 5 : Performances de classification de l'algorithme PmBc (taux en %) .....	58
Tableau 6 : Performance moyenne des 3 algorithmes dans la catégorisation des syllabes .....	58
Tableau 7 : Présentation des performances de classification des syllabes par un linguiste expert ..	59
Tableau 8 : Proportion de mots enchâssés à l'initiale, en fonction du caractère morphologique et non morphologique des superpositions lexicales .....	70
Tableau 9 : Temps de reconnaissance prédits par quelques modèles psycholinguistiques de la reconnaissance des mots parlés, lors du traitement de superpositions lexicales réelles initiale ou finale .....	75
Tableau 10 : Facteurs manipulés dans les expériences .....	92
Tableau 11 : Récapitulation du sens des effets de chaque facteur, en fonction de l'origine .....	99
Tableau 12 : Effets principaux des 2 facteurs expérimentaux .....	99
Tableau 13 : Synthèse des effets .....	99
Tableau 14 : Avantages et inconvénients de chaque origine .....	105
Tableau 15 : Calcul de l'espérance mathématique de chaque source de variation .....	120
Tableau 16 : Récapitulation des problèmes / solutions liés aux analyses statistiques .....	132
Tableau 17 : Aspects théoriques sous-jacents aux expériences .....	156
Tableau 18 : Aspects techniques mis en œuvre par expérience .....	158
Tableau 19 : Catégories de syllabes utilisées dans le matériel expérimental .....	163
Tableau 20 : Catégories de frontières lexicales déterminées par la combinaison 2 à 2 des 4 types de syllabe .....	164
Tableau 21 : Valeurs moyennes de durée et de f0 par catégorie de syllabe (observé et prédit) ..	166
Tableau 22 : Tests de contrastes .....	167
Tableau 23 : Code utilisé pour définir la catégorie de la syllabe .....	176
Tableau 24 : Répartition des items pour la structuration et le test des réseaux .....	178
Tableau 25 : Performance des 3 réseaux (erreur globale du réseau) .....	178
Tableau 26 : Taux moyens d'activation des unités de sortie (réponses de classification) .....	181
Tableau 27 : Test des contrastes entre taux de classification des syllabes d'une même catégorie à une même position, pour les réponses correctes uniquement .....	182
Tableau 28 : Récapitulatif des confusions observées .....	183
Tableau 29 : Test de significativité des contrastes entre conditions .....	183
Tableau 30 : Présentation des taux de classification, fiabilité et certitude par syllabe et par item, obtenus dans la Simulation 1 .....	184
Tableau 31 : Performance de la classification des syllabes par leur prosodie, selon différentes méthodes .....	186
Tableau 32 : Répartition des combinaisons d'items dans chaque liste .....	187

Tableau 33 : Répartition des listes de comparaisons d'items par réseau .....	188
Tableau 34 : Récapitulation des taux moyens d'activation de l'unité de sortie du réseau estimés par le modèle, par comparaison de catégorie de frontière .....	192
Tableau 35 : Test de significativité des contrastes entre conditions pour les comparaisons intra catégorie.....	192
Tableau 36 : Logique de la situation expérimentale .....	196
Tableau 37 : Répartition des combinaisons d'items dans les 6 listes expérimentales .....	199
Tableau 38 : Récapitulation des moyennes par comparaison de catégorie d'item (matrice de comparaison - taux moyens de réponses « similaire ») .....	205
Tableau 39 : Test de significativité des contrastes entre conditions (contrastés entre comparaisons intra catégorie) .....	205
Tableau 40 : Récapitulation des moyennes par comparaison de catégorie d'item (matrice de comparaison - taux moyens de réponses « similaire ») .....	207
Tableau 41 : Structure de l'item selon la condition expérimentale.....	211
Tableau 42 : Test de significativité des contrastes entre conditions.....	215
Tableau 43 : Répartition des réponses $P_0$ , $P_1$ et $P_2$ par catégorie d'item .....	216
Tableau 44 : Codes adoptés pour la nomenclature des tests de contraste.....	217
Tableau 45 : Test de significativité des contrastes pour les comparaisons entre type d'item par réponse.....	217
Tableau 46 : Test de significativité des contrastes entre conditions pour les comparaisons entre réponses par type d'item.....	217
Tableau 47 : Plan de l'expérience .....	222
Tableau 48 : Plan d'analyse .....	222
Tableau 49 : Tests de significativité des facteurs étudiés .....	224
Tableau 50 : Tests de significativité des contrastes.....	225
Tableau 51 : Facteurs expérimentaux .....	236
Tableau 52 : Nombre moyen de phonèmes par mot pour chaque condition.....	238
Tableau 53 : Fréquence formelle phonologique moyenne par mot pour chaque condition.....	238
Tableau 54 : Tests de significativité pour les facteurs expérimentaux.....	239
Tableau 55 : Tests de significativité des contrastes entre conditions .....	239
Tableau 56 : Valeurs moyennes des paramètres acoustiques par condition .....	245
Tableau 57 : Résultats des ANOVA pour la différence entre les 2 conditions du facteur étudié ....	245
Tableau 58 : Test de significativité du facteur expérimental .....	246
Tableau 59 : Facteurs et conditions .....	249
Tableau 60 : Tests de significativité pour les facteurs expérimentaux.....	253
Tableau 61 : Tests de significativité des contrastes entre conditions .....	253
Tableau 62 : Facteurs lexicaux et types de mots .....	257
Tableau 63 : Facteurs et conditions contextuelles .....	258
Tableau 64 : Tests de significativité pour les 2 facteurs lexicaux .....	260
Tableau 65 : Tests de significativité pour les facteurs contextuels et les interactions entre facteurs (lexicaux et contextuels) .....	261
Tableau 66 : Tests de significativité des contrastes entre conditions pour les facteurs lexicaux...	261

Tableau 67 : Tests de significativité des contrastes entre conditions lexicales dans la condition contextuelle TDL.DerP.....	263
Tableau 68 : Test de significativité du facteur expérimental.....	270
Tableau 69 : Tests de significativité des covariables.....	278
Tableau 70 : Tests de significativité pour les facteurs expérimentaux et leurs interactions.....	279
Tableau 71 : Tests de significativité des contrastes entre conditions.....	280
Tableau 72 : Tests de significativité des contrastes entre position contrôle et positions tests, pour chaque facteur et pour l'interaction.....	282
Tableau 73 : Tests de significativité des contrastes entre facteurs et entre facteurs et leur interaction, pour chaque position.....	282
Tableau 74 : Facteurs lexicaux et types de mots.....	289
Tableau 75 : Tests de significativité pour les facteurs expérimentaux et leurs interactions.....	293
Tableau 76 : Tests de significativité des contrastes entre conditions.....	294
Tableau 77 : Tests de significativité des contrastes entre les modalités des 2 facteurs expérimentaux ou d'un des deux facteurs et de leur interaction, pour chaque position de la cible.....	296
Tableau 78 : Tests de significativité des contrastes entre positions, pour chaque facteur expérimental et pour leur interaction.....	296
Tableau 79 : Facteurs expérimentaux.....	302
Tableau 80 : Fréquence moyenne d'occurrence de chaque type de syllabe expérimentale en initiale, milieu, fin de mot.....	302
Tableau 81 : Répartition des 4 syllabes S <sub>1</sub> à travers les 4 conditions expérimentales.....	304
Tableau 82 : Synthèse du tableau précédent (somme par modalité de chaque facteur, du nombre de syllabes S1 terminant avec l'un des 4 phonèmes).....	304
Tableau 83 : Tests de significativité des facteurs étudiés et de leur interaction.....	308
Tableau 84 : Tests de significativité des contrastes entre conditions.....	308
Tableau 85 : Caractéristiques fonctionnelles de 3 types de syllabes.....	311
Tableau 86 : Conditions expérimentales.....	312
Tableau 87 : Indices caractérisant les syllabes utilisées (moyennes).....	312
Tableau 88 : Test de significativité du facteur expérimental.....	314
Tableau 89 : Description des caractéristiques linguistiques formelles moyennes des syllabes par condition.....	319
Tableau 90 : Test de significativité du facteur expérimental.....	321
Tableau 91 : Description des caractéristiques moyennes des syllabes par condition.....	325
Tableau 92 : Test de significativité du facteur expérimental.....	327
Tableau 93 : Résultats fournis par les 4 modèles d'analyse, pour l'Expérience 13.....	331
Tableau 94 : Résultats fournis par les 4 modèles d'analyse, pour l'Expérience 14.....	332
Tableau 95 : Degrés d'adéquation des modèles aux données pour l'Expérience 13, en fonction du type d'unité considéré et du type d'occurrence dans le lexique.....	332
Tableau 96 : Degrés d'adéquation des modèles aux données pour l'Expérience 14, en fonction du type d'unité considéré et du type d'occurrence dans le lexique.....	332
Tableau 97 : Récapitulatif des effets obtenus dans les Expériences 11 à 14, les Analyses supplémentaires et l'Analyse exploratoire.....	336
Tableau 98 : Récapitulatif des résultats obtenus.....	338
Tableau 99 : Conditions expérimentales.....	343



Tableau 100 : Tests de significativité pour les facteurs expérimentaux .....	347
Tableau 101 : Tests de significativité des contrastes entre conditions.....	347
Tableau 102 : Test de significativité du facteur expérimental .....	352
Tableau 103 : Effet d'une superposition lexicale initiale réelle (E) et/ou virtuelle (I) sur les paramètres prosodiques, dans le cas de mots monosyllabiques .....	357
Tableau 104 : Effet de la lexicalité de la première ( $M_1$ ) et/ou de la deuxième syllabe ( $M_2$ ) d'un mot bisyllabique, sur les paramètres prosodiques caractérisant chacune des 2 syllabes du mot ( $S_1$ et $S_2$ ) .....	359
Tableau 105 : Effet de la fréquence d'occurrence dans les mots du lexique, de syllabes non lexicales en position initiale (FqI) et/ou en position finale (FqF), sur les paramètres prosodiques	362
Tableau 106 : Effet de la fréquence d'occurrence dans les mots du lexique, de syllabes non lexicales en position médiale, sur les paramètres prosodiques.....	364
Tableau 107 : Effet de la fréquence d'occurrence dans les mots du lexique, de syllabes lexicales en position initiale (FqI, Expérience 13) ou en position non initiale (FqnI, Expérience 14), sur les paramètres prosodiques.....	365
Tableau 108 : Sens de variation des temps de réponse des auditeurs en fonction des variations des paramètres prosodiques.....	368
Tableau 109 : Récapitulation des motifs de réponse des participants, obtenus dans les expériences en fonction de la lexicalité des stimuli et des facteurs manipulés .....	369
Tableau 110 : Scores de prégnance perceptive des mots libres « <i>garage</i> », « <i>gars</i> », « <i>rage</i> » et des mots « <i>gars</i> » et « <i>rage</i> » enchâssés dans « <i>garage</i> » .....	388
Tableau 111 : Scores moyens de prégnance perceptive par condition et par modèle, et TR moyens estimés par condition.....	398
Tableau 112 : Résultats des analyses de régression entre facteurs expérimentaux et scores de prégnance perceptive.....	398
Tableau 113 : Résultats de l'analyse de régression multiniveaux Scores/Comportement.....	399
Tableau 114 : Scores moyens de prégnance perceptive par condition et par modèle, et TR moyens estimés par condition.....	399
Tableau 115 : Résultats des analyses de régression entre facteurs expérimentaux et scores de prégnance perceptive.....	400
Tableau 116 : Résultats de l'analyse de régression multiniveaux Scores/Comportement.....	401
Tableau 117 : Scores moyens de prégnance perceptive par condition et par modèle, et TR moyens estimés par condition.....	402
Tableau 118 : Résultats des analyses de régression entre enchâssement et scores de prégnance perceptive .....	402
Tableau 119 : Résultats des analyses de régression multiniveaux Scores/Comportement .....	402
Tableau 120 : Résultats des analyses de régression multiniveaux Scores sans gradient/Comportement.....	403
Tableau 121 : Résultats des analyses de régression multiniveaux Scores/Comportement .....	403
Tableau 122 : Comparaison du pourcentage de variance comportementale expliquée par le modèle SLEXS1 et par les facteurs expérimentaux E ou $M_1$ .....	406
Tableau 123 : Résultat des tests de significativité.....	437
Tableau 124 : Comparaison statistique des données fournies par notre base avec celles fournies par 4 bases de données de référence .....	438
Tableau 125 : Comparaison des scores d'âge d'acquisition fournis par les 4 bases de données de référence .....	439
Tableau 126 : Résultat des tests de significativité.....	451

Tableau 127 : Résultat des tests de significativité.....	453
Tableau 128 : Résultat des tests de significativité.....	455
Tableau 129 : Taux moyens d'activation des unités de sortie du réseau (excitabilité).....	467
Tableau 130 : Test des contrastes (comparaison de l'excitabilité des unités de sortie selon le type de syllabe et sa position).....	467
Tableau 131 : Taux moyens d'activation des unités de sortie (excitabilité) en fonction de la frontière (valeurs négatives : activation diminuée de x lorsque ce type d'item est présenté).....	468
Tableau 132 : Test des contrastes (comparaison de l'excitabilité des unités de sortie selon le type de frontière).....	468
Tableau 133 : Grilles de calcul des moyennes à partir des résultats du modèle .....	476
Tableau 134 : Mots expérimentaux.....	479
Tableau 135 : Mots et Pseudo-mots distracteurs (monosyllabiques et bisyllabiques) .....	479
Tableau 136 : Tests de significativité des covariables et de leurs interactions avec les facteurs expérimentaux.....	481
Tableau 137 : Analyse sur l'ensemble des réponses « similaire » (comparaisons intra et inter catégories).....	484
Tableau 138 : Résultats de l'analyse des réponses correctes pour les comparaisons intra catégorie .....	485
Tableau 139 : Dimension acoustique fondant la réponse des participants en fonction du type de frontière.....	485
Tableau 140 : Répartition des réponses les mieux et les moins expliquées, pour chaque catégorie de frontière .....	486
Tableau 141 : Résultats de l'analyse des réponses correctes pour les comparaisons inter catégories .....	488
Tableau 142 : Pourcentage de confusions expliquées pour chaque comparaison de catégories de frontière.....	490
Tableau 143 : Tests de significativité du rapport entre facteurs lexicaux et prosodiques .....	494
Tableau 144 : Tests de significativité des covariables et de leurs interactions.....	496
Tableau 145 : Tests de significativité des covariables et de leurs interactions.....	497
Tableau 146 : Mots expérimentaux et pseudo-mots distracteurs .....	499
Tableau 147 : Tests de significativité des covariables et de leurs interactions.....	500
Tableau 148 : Mots expérimentaux et pseudo-mots distracteurs .....	501
Tableau 149 : Test de significativité de l'effet des facteurs expérimentaux sur les paramètres prosodiques des syllabes constituant les mots.....	503
Tableau 150 : Tests de significativité des covariables et de leurs interactions.....	506
Tableau 151 : Mots expérimentaux (Expériences 7 & 10) .....	508
Tableau 152 : Pseudo-mots distracteurs.....	508
Tableau 153 : Effet des facteurs expérimentaux sur les paramètres prosodiques des syllabes constituant les mots .....	511
Tableau 154 : Tests de significativité des covariables et de leurs interactions.....	514
Tableau 155 : Mots expérimentaux et pseudo-mots distracteurs .....	516
Tableau 156 : Tests de significativité des covariables et de leurs interactions.....	517
Tableau 157 : Tests de significativité des covariables et de leurs interactions.....	519
Tableau 158 : Tests de significativité des covariables et de leurs interactions.....	520

Tableau 159 : Pseudo-mots expérimentaux .....	521
Tableau 160 : Pseudo-mots distracteurs.....	521
Tableau 161 : Effet des facteurs expérimentaux sur les facteurs prosodiques de la syllabe .....	523
Tableau 162 : Tests de significativité des covariables et de leurs interactions.....	526
Tableau 163 : Pseudo-mots expérimentaux et mots distracteurs .....	528
Tableau 164 : Effet du facteur expérimental sur les facteurs prosodiques des syllabes .....	529
Tableau 165 : Tests de significativité des covariables et de leurs interactions.....	532
Tableau 166 : Tests de significativité des facteurs expérimentaux et des covariables .....	533
Tableau 167 : Tests de significativité des facteurs expérimentaux et des covariables .....	534
Tableau 168 : Mots expérimentaux et pseudo-mots distracteurs .....	535
Tableau 169 : Rapport entre la fréquence d'occurrence d'une syllabe lexicale à l'initiale des mots dans le lexique, et les paramètres prosodiques de la syllabe produite .....	536
Tableau 170 : Tests de significativité des covariables et de leurs interactions.....	539
Tableau 171 : Mots expérimentaux et pseudo-mots distracteurs .....	541
Tableau 172 : Rapport ente la fréquence d'occurrence d'une syllabe lexicale en fin de mots dans le lexique, et les paramètres prosodiques de la syllabe .....	542
Tableau 173 : Tests de significativité des covariables et de leurs interactions.....	545
Tableau 174 : Tests de significativité des variables/covariables (fréquence de la syllabe, Expérience 13).....	547
Tableau 175 : Tests de significativité des variables/covariables (nombre de mots renfermant la syllabe, Expérience 13) .....	549
Tableau 176 : Tests de significativité des variables/covariables (fréquence de la séquence de phonèmes, Expérience 13) .....	550
Tableau 177 : Tests de significativité des variables/covariables (nombre de mots renfermant la séquence, Expérience 13).....	552
Tableau 178 : Tests de significativité des variables/covariables (fréquence de la syllabe, Expérience 14).....	554
Tableau 179 : Tests de significativité des variables/covariables (nombre de mots renfermant la syllabe, Expérience 14) .....	556
Tableau 180 : Tests de significativité des variables/covariables (fréquence de la séquence de phonèmes, Expérience 14) .....	557
Tableau 181 : Tests de significativité des variables/covariables (nombre de mots renfermant la séquence de phonèmes, Expérience 14).....	559
Tableau 182 : Mots expérimentaux et pseudo-mots distracteurs .....	561
Tableau 183 : Tests de significativité des covariables et de leurs interactions, entre elles et avec les variables expérimentales.....	563
Tableau 184 : Mots expérimentaux et distracteurs, et pseudo-mots distracteurs .....	566
Tableau 185 : Tests de significativité des covariables, et de leurs interactions entre elles et avec la variable expérimentale.....	568

## Index des Annexes

Annexe 1.....	423
Annexe 2.....	457
Annexe 3.....	463
Annexe 4.....	464
Annexe 5.....	467
Annexe 6.....	470
Annexe 7.....	472
Annexe 8.....	474
Annexe 9.....	479
Annexe 10.....	481
Annexe 11.....	482
Annexe 12.....	494
Annexe 13.....	497
Annexe 14.....	499
Annexe 15.....	500
Annexe 16.....	501
Annexe 17.....	502
Annexe 18.....	506
Annexe 19.....	508
Annexe 20.....	510
Annexe 21.....	514
Annexe 22.....	516
Annexe 23.....	517
Annexe 24.....	519
Annexe 25.....	520
Annexe 26.....	521
Annexe 27.....	523
Annexe 28.....	526
Annexe 29.....	528
Annexe 30.....	529
Annexe 31.....	532
Annexe 32.....	533
Annexe 33.....	535
Annexe 34.....	536
Annexe 35.....	539
Annexe 36.....	541

Annexe 37..... 542  
Annexe 38..... 545  
Annexe 39..... 547  
Annexe 40..... 561  
Annexe 41..... 563  
Annexe 42..... 566  
Annexe 43..... 568

*Thèse soutenue publiquement le 27 septembre 2005  
à l'Université de Genève, Genève, GE, Suisse.*

*Le manuscrit a reçu la note de 5.5/6.  
La soutenance publique a été notée 5.5/6.*

## Résumé

Reconnaître des mots parlés lors d'une conversation implique de résoudre un problème de localisation de frontières lexicales dans le flux sonore. La démarcation lexicale est envisagée dans ce travail de recherche à travers le guidage ascendant des traitements par la prosodie, et descendant par la structure du lexique.

Les résultats présentés montrent :

- ▷ qu'un code prososyllabique structure le signal de parole et contraint fortement le traitement lexical.
- ▷ que la présence d'une superposition lexicale à l'initiale d'un mot facilite sa reconnaissance.

SLEXS, un modèle mathématique de prégnance perceptive des mots parlés, est proposé pour expliquer la contradiction théorique que représente ce phénomène de facilitation, qui serait lié à un processus récursif de construction du percept.

Finalement, il ressort de ce travail que la compétition lexicale serait un mécanisme discret, voire facultatif, au cours de la reconnaissance des mots parlés.