



## AVERTISSEMENT

Ce document est le fruit d'un long travail approuvé par le jury de soutenance et mis à disposition de l'ensemble de la communauté universitaire élargie.

Il est soumis à la propriété intellectuelle de l'auteur. Ceci implique une obligation de citation et de référencement lors de l'utilisation de ce document.

D'autre part, toute contrefaçon, plagiat, reproduction illicite encourt une poursuite pénale.

Contact : [ddoc-memoires-contact@univ-lorraine.fr](mailto:ddoc-memoires-contact@univ-lorraine.fr)

## LIENS

Code de la Propriété Intellectuelle. articles L 122. 4

Code de la Propriété Intellectuelle. articles L 335.2- L 335.10

[http://www.cfcopies.com/V2/leg/leg\\_droi.php](http://www.cfcopies.com/V2/leg/leg_droi.php)

<http://www.culture.gouv.fr/culture/infos-pratiques/droits/protection.htm>



MEMOIRE présenté pour l'obtention du  
**CERTIFICAT DE CAPACITE D'ORTHOPHONISTE**

Par

**JOBIT Pauline**  
**JOYE Nelly**

**INFLUENCE DU LEXIQUE SUR LA PERCEPTION  
PHONEMIQUE DANS UN CONTEXTE DE SIGNAL  
DEGRADE**

Maîtres de Mémoire

**BOUTON Sophie**  
**LODEHO Anne-Dominique**

Membres du Jury

**COLIN Stéphanie**

**OZIL Marie**

**TRUY Eric**

Date de Soutenance

**27 juin 2013**

---

# ORGANIGRAMMES

---

## 1. Université Claude Bernard Lyon1

Président  
**Pr. GILLY François-Noël**

Vice-président CEVU  
**M. LALLE Philippe**

Vice-président CA  
**M. BEN HADID Hamda**

Vice-président CS  
**M. GILLET Germain**

Directeur Général des Services  
**M. HELLEU Alain**

### 1.1 Secteur Santé :

U.F.R. de Médecine Lyon Est  
Directeur **Pr. ETIENNE Jérôme**

U.F.R d'Odontologie  
Directeur **Pr. BOURGEOIS Denis**

U.F.R de Médecine et de  
Maïeutique - Lyon-Sud Charles  
Mérieux  
Directeur **Pr. BURILLON Carole**

Institut des Sciences Pharmaceutiques  
et Biologiques  
Directeur **Pr. VINCIGUERRA Christine**

Institut des Sciences et Techniques de  
la Réadaptation  
Directeur **Pr. MATILLON Yves**

Comité de Coordination des  
Etudes Médicales (C.C.E.M.)  
**Pr. GILLY François Noël**

Département de Formation et Centre  
de Recherche en Biologie Humaine  
Directeur **Pr. FARGE Pierre**

### 1.2 Secteur Sciences et Technologies :

U.F.R. de Sciences et Technologies  
Directeur **M. DE MARCHI Fabien**

IUFM  
Directeur **M. MOUGNIOTTE Alain**

U.F.R. de Sciences et Techniques  
des Activités Physiques et  
Sportives (S.T.A.P.S.)  
Directeur **M. COLLIGNON Claude**

POLYTECH LYON  
Directeur **M. FOURNIER Pascal**

Institut des Sciences Financières et  
d'Assurance (I.S.F.A.)  
Directeur **M. LEBOISNE Nicolas**

Ecole Supérieure de Chimie Physique  
Electronique de Lyon (ESCPE)  
Directeur **M. PIGNAULT Gérard**

Observatoire Astronomique de  
Lyon **M. GUIDERDONI Bruno**

IUT LYON 1  
Directeur **M. VITON Christophe**

---

---

**2. Institut Sciences et Techniques de Réadaptation FORMATION**

**ORTHOPHONIE**

Directeur ISTR

**Pr. MATILLON Yves**

Directeur de la formation

**Pr. Associé BO Agnès**

Directeur de la recherche

**Dr. WITKO Agnès**

Responsables de la formation clinique

**GENTIL Claire**

**GUILLON Fanny**

Chargée du concours d'entrée

**PEILLON Anne**

Secrétariat de direction et de scolarité

**BADIOU Stéphanie**

**BONNEL Corinne**

**CLERGET Corinne**

---

---

## REMERCIEMENTS

---

Nous remercions Sophie Bouton et Anne-Dominique Lodeho, nos maîtres de mémoire, pour nous avoir proposé ce sujet passionnant. Nous vous remercions sincèrement pour votre suivi actif et bienveillant, ainsi que pour votre écoute. Nous vous sommes très reconnaissantes d'avoir toujours su nous encourager et nous mettre sur la bonne voie. Nous remercions aussi Claire qui nous a permis cette rencontre.

Nous adressons nos remerciements à l'équipe du Laboratoire de Neurosciences Cognitives, DEC, ENS pour nous avoir prêté le matériel nécessaire à la réalisation de notre expérimentation et pour avoir financé nos déplacements.

Nous remercions vivement les candidats qui ont participé à notre étude pour leur aide et leur disponibilité. Les échanges que nous avons eus ont été indispensables à notre réflexion.

Nous adressons nos remerciements à Agnès Witko (directrice de la recherche pour la formation en orthophonie de Lyon 1) pour son encadrement méthodologique.

Nous tenons à remercier Sylvia Topouzkhian (professeur pour la formation en orthophonie de Lyon 1) pour son aide à la naissance du projet.

Nous remercions Patricia Malquarti et Richard Nombalais (professeurs en surdit   pour la formation en orthophonie de Lyon 1) pour leurs conseils prodigu  s au d  but de notre recherche, qui nous ont permis de cibler notre sujet.

Nous adressons nos remerciements    Fabien Seldran (docteur en audioproth  sie    l'h  pital Edouard Herriot) pour ses conseils, ses pr  cieuses explications et son aide bibliographique.

Nous souhaitons remercier aussi Emilie Bernard et Marie Redon (orthophonistes) qui ont partag   avec nous leur regard clinique tout au long du projet.

Merci    nos familles et en particulier    nos parents, qui nous ont soutenues tout au long de ces quatre ann  es et qui nous ont donn   les moyens de r  aliser les   tudes de notre choix dans les meilleures conditions. Merci    Charles pour son aide et ses conseils en informatique.

Merci    nos amies, nous sommes heureuses d'avoir travers   ces quatre ann  es avec elles et d'avoir partag   les bons moments comme les moins bons.

Merci    toi Nelly et    toi Pauline pour ces deux ann  es de partage, pour notre confiance mutuelle. C'est    deux que nous y sommes parvenues.

---

---

# SOMMAIRE

---

<b>ORGANIGRAMMES</b> .....	<b>2</b>
1- UNIVERSITE CLAUDE BERNARD LYON1 .....	<b>ERREUR ! SIGNET NON DEFINI.</b>
1.1. Secteur Santé : .....	<i>Erreur ! Signet non défini.</i>
1.2. Secteur Sciences et Technologies : .....	<i>Erreur ! Signet non défini.</i>
2- INSTITUT SCIENCES ET TECHNIQUES DE READAPTATION FORMATION ORTHOPHONIE	<b>ERREUR ! SIGNET NON DEFINI.</b>
<b>REMERCIEMENTS</b> .....	<b>4</b>
<b>SOMMAIRE</b> .....	<b>5</b>
<b>INTRODUCTION</b> .....	<b>7</b>
<b>PARTIE THEORIQUE</b> .....	<b>9</b>
I. L'IMPLANT COCHLEAIRE : ENJEUX ACTUELS.....	10
1. Le signal sonore.....	10
2. L'audition normale .....	11
3. Le fonctionnement de l'implant cochléaire .....	12
4. L'implant cochléaire : indications et limites .....	13
II. LA PERCEPTION DE LA PAROLE .....	14
1. Perception de la parole, phonèmes et traits articulatoires .....	14
2. Description acoustique et articulatoire des phonèmes du français.....	15
3. Le traitement phonémique de la parole .....	16
4. Le traitement lexical de la parole .....	19
III. LA PERCEPTION DE LA PAROLE EN CONTEXTE DE SIGNAL DEGRADE .....	21
1. La perception de la parole de l'enfant implanté .....	21
2. La perception de la parole de l'adulte implanté .....	22
3. Le vocodeur : principe et application.....	23
<b>PROBLEMATIQUE ET HYPOTHESES</b> .....	<b>25</b>
I. PROBLEMATIQUE .....	26
II. HYPOTHESES .....	27
1. Hypothèse générale .....	27
2. Hypothèses opérationnelles .....	28
<b>PARTIE EXPERIMENTALE</b> .....	<b>29</b>
I. POPULATION .....	30
II. MATERIEL.....	32
III. TACHES .....	36
IV. PASSATION.....	36
<b>PRESENTATION DES RESULTATS</b> .....	<b>38</b>
I. PREAMBULE.....	39
II. ANALYSE POUR LES SCORES DE REPNSES CORRECTES .....	40
1. Résultats des effets principaux.....	40
2. Résultats des interactions .....	43
III. ANALYSE POUR LES SCORES DE TEMPS DE REPONSE.....	47
1. Résultats des effets principaux.....	47
2. Résultats des interactions .....	50
<b>DISCUSSION DES RESULTATS</b> .....	<b>54</b>
I. PREAMBULE.....	55
II. INTERPRETATION DES RESULTATS.....	56
1. L'effet du lexique .....	56
2. L'effet de dégradation du signal.....	61
III. CRITIQUES ET LIMITES DE NOTRE EXPERIMENTATION.....	63
1. Pour notre population.....	63
2. Pour notre passation.....	63

---

---

3. <i>Pour l'analyse des données</i> .....	64
IV. APPORTS DE L'ETUDE POUR LA PRATIQUE ORTHOPHONIQUE.....	65
<b>CONCLUSION.....</b>	<b>69</b>
<b>BIBLIOGRAPHIE.....</b>	<b>70</b>
<b>ANNEXES.....</b>	<b>76</b>
<b>LISTE DES ANNEXES .....</b>	<b>77</b>
ANNEXE I : ANALYSE SPECTRALE DE DEUX PAIRES MINIMALES ISSUES DE NOTRE MATERIEL .....	78
ANNEXE II : REPRESENTATIONS DU MOT « MARRON » EN CONDITIONS NON VOCODEE ET VOCODEE .....	80
ANNEXE III : TABLEAU RECAPITULATIF DU MATERIEL EXPERIMENTAL .....	83
ANNEXE IV : DOCUMENT D'INFORMATION POUR NOTRE RECHERCHE DE POPULATION .....	85
ANNEXE V : TABLEAU DE CONSTRUCTION DES PAIRES .....	86
1. <i>Pour la nasalité</i> .....	86
2. <i>Pour le lieu articulatoire</i> .....	87
3. <i>Pour le mode articulatoire</i> .....	88
4. <i>Pour le voisement</i> .....	89
ANNEXE VI : ORDRE DE PRESENTATION DES TACHES .....	90
ANNEXE VII : CONSIGNES PRESENTEES LORS DE LA PASSATION.....	92
1. <i>Consignes écrites (pour Matlab)</i> .....	92
2. <i>Consignes orales (trame des explications)</i> .....	92
<b>TABLE DES ILLUSTRATIONS.....</b>	<b>94</b>
I. LISTE DES FIGURES.....	94
II. LISTE DES TABLEAUX .....	96
<b>TABLE DES MATIERES .....</b>	<b>97</b>

---

---

## INTRODUCTION

---

*« Lorsqu'il est exposé à un flux de parole, l'auditeur doit établir une relation entre deux univers que sont l'univers physique des sons du langage et l'univers symbolique des énoncés. » (Spinelli et Ferrand, 2005).*

En effet, le processus de perception de la parole demande à l'auditeur non seulement d'analyser les informations acoustiques du signal auditif, mais aussi de mettre en lien ces informations avec des représentations lexicales, morphosyntaxiques et sémantiques.

Dans des conditions d'écoute naturelle, la parole parvient aux auditeurs en étant mélangée aux bruits environnants. De plus, l'enchaînement des phonèmes lors de la production de mots est marqué par des phénomènes de coarticulation mêlant les caractéristiques des phonèmes adjacents. Aussi nous ne pouvons pas négliger les variabilités entre les différents locuteurs, qu'elles touchent l'articulation ou la hauteur fréquentielle, le timbre et le rythme. L'ensemble de ces variations va perturber les informations acoustiques portées par le signal de parole. Nous pouvons alors nous demander comment l'auditeur parvient à identifier les unités discrètes que sont les phonèmes, pour ensuite les assembler afin d'isoler les mots dans le flot continu d'informations sonores qu'est la parole.

Les auteurs ont développé différents modèles cherchant à expliquer les processus permettant la perception de la parole. Dans plusieurs de ces modèles, les connaissances lexicales jouent un rôle fondamental dans la perception et l'analyse du signal de parole. Le modèle TRACE (McClelland et Elman, 1986) est l'un d'entre eux. Il postule que les connaissances lexicales des auditeurs leur permettent de compléter les informations parcellaires qu'ils perçoivent. Ainsi, face aux dégradations et aux perturbations du signal de parole, les individus s'aideraient de leurs connaissances sur la langue pour le décoder.

Dans le but de tester cette théorie, certaines études ont évalué les compétences langagières chez des patients porteurs d'un implant cochléaire. Cette prothèse transmet les informations sonores à l'aide d'impulsions électriques. Si l'implant permet aux personnes présentant une surdité sévère à profonde d'accéder au monde sonore, les informations qu'il transmet sont dégradées par rapport à celles transmises par la cochlée. Ainsi grâce à l'implant, il est possible d'observer les répercussions de la dégradation du signal sonore sur la perception de la parole. Certaines études s'intéressent aux perceptions de participants implantés en les comparant à celles de participants normo-entendants, alors que d'autres utilisent un programme informatique (le vocodeur), qui permet de modéliser les informations sonores transmises par l'implant.

Notre étude fait suite aux recherches de Bouton, Colé et Serniclaes (2012) et Bouton, Serniclaes, Bertoncini et Colé (2012) chez des enfants implantés sourds pré-linguaux et des enfants normo-entendants. Ces études avaient pour but d'étudier les capacités perceptives des enfants implantés et d'analyser l'utilisation qu'ils font de leurs connaissances lexicales. Afin de poursuivre ce questionnement, nous avons réalisé une étude chez des adultes normo-entendants que nous avons mis en situation de signal dégradé au vocodeur. L'objectif de notre étude était de tester la mise en place d'un traitement lexical influençant la perception de la parole en situation de signal dégradé. Le choix d'une population adulte nous permet d'observer l'influence de la composante acoustique dans la perception du signal en éliminant l'influence de la composante développementale du langage.

---

---

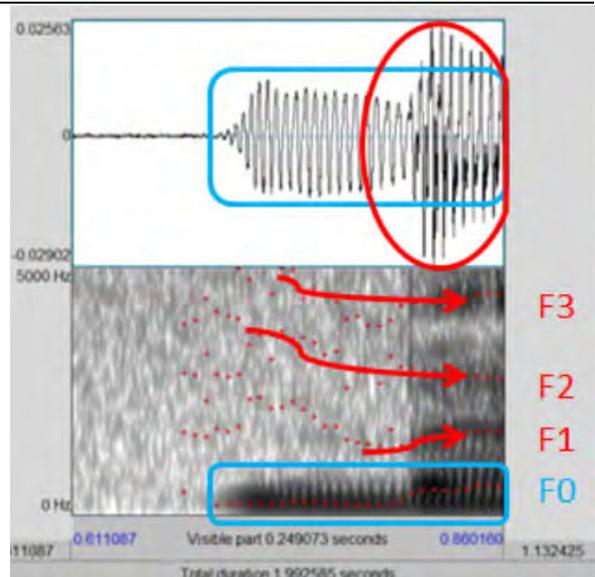
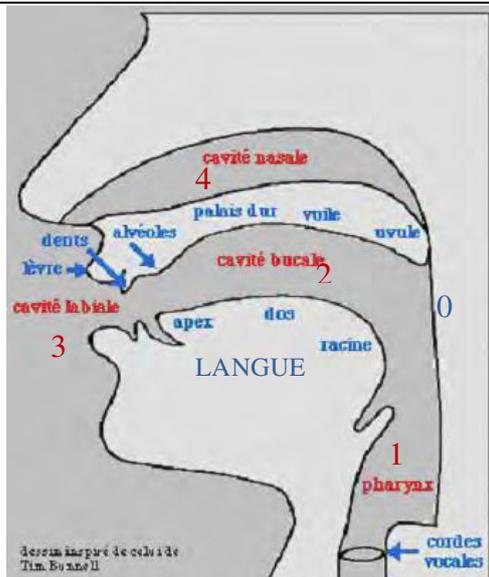
Dans un premier temps, nous présenterons l'ensemble des données théoriques qui nous a permis de définir notre champ de recherche. Nous décrirons les caractéristiques physiques du signal de parole, les éléments permettant sa perception chez les sujets normo-entendants, ainsi que le fonctionnement de l'implant cochléaire et les spécificités du signal acoustique qu'il transmet. Puis, nous exposerons les mécanismes de perception de la parole en explicitant les rôles des différents traitements : acoustique, phonémique et lexical. Les études portant sur la perception de la parole chez les enfants et les adultes implantés seront également évoquées. Enfin nous présenterons le vocodeur.

Dans un second temps, nous développerons les questionnements qui nous ont amenés à la définition de notre hypothèse théorique. Puis, nous détaillerons le choix de la population de notre étude, l'élaboration du matériel et du protocole, et les modalités de passation que nous avons mis en place.

Dans un troisième temps, nous exposerons nos résultats, que nous discuterons, afin de faire le lien avec les données de la littérature et la pratique orthophonique.

---

**Chapitre I**  
**PARTIE THEORIQUE**



a. Représentation schématique des cavités de résonance principales pour la production de la parole

b. Représentation à l'oscillogramme (haut) et au spectrogramme (bas) de la syllabe /ba/ dans le mot «bâche »

Figure 1 : Eléments anatomiques et indices acoustiques mis en jeu dans la production de la parole

- L'onde-source issue de la vibration laryngée (0), est modulée par les quatre cavités représentées ici : (1) la cavité pharyngée, (2) la cavité buccale, (3) la cavité labiale et (4) la cavité nasale. Les résonances générées par ces cavités sont à l'origine des formants visibles sur le spectrogramme et l'oscillogramme (figure b). La figure a. est extraite de <http://www.ph-ludwigsburg.de/html/2b-frnz-s-01/overmann/baf3/phon/3k.htm>
- Les indices de vibration laryngée (F0) sont encadrés en bleu : barre de voisement sur le spectrogramme (bas) ; oscillation périodique sur l'oscillogramme (haut). En rouge apparaissent les formants, avec une envolée de la courbe d'oscillation (cercle rouge en haut) et l'apparition de lignes de fréquences intensifiées sur le spectrogramme (flèches rouges en bas). Les formants sont issus des modulations appliquées à la vibration laryngée par les cavités pharyngée (1), buccale (2) et labiale (3) (figure a). Ils sont nommés par ordre de fréquence : formant le plus grave : F1 = 693 Hz ; formant médian : F2 = 1679 Hz ; formant le plus aigu : F3 = 2716 Hz (données et image issues du logiciel Praat, valeurs des formants à 0,86 s).

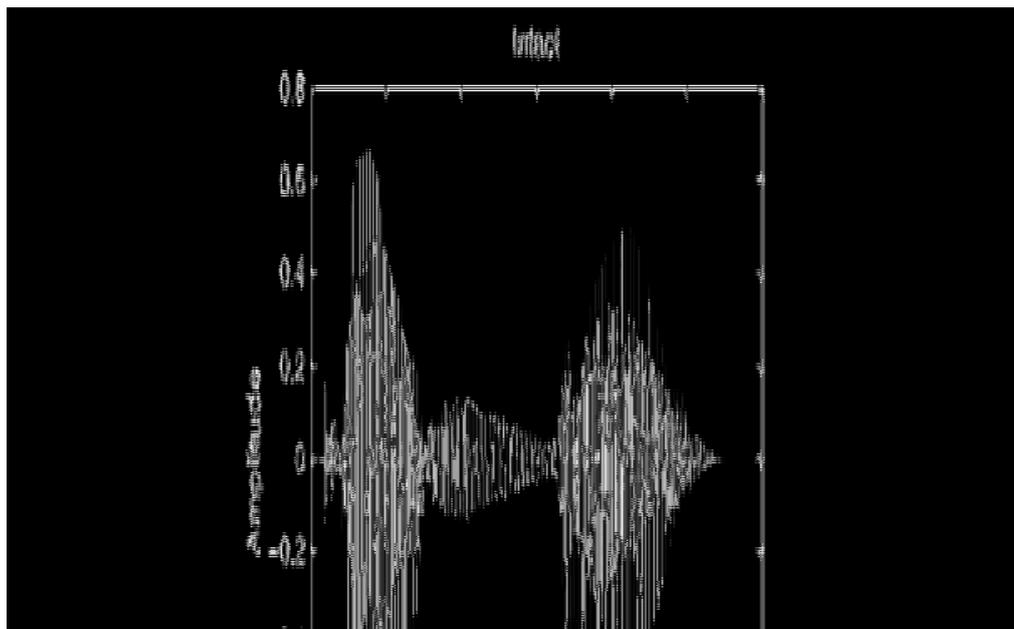


Figure 2 : Oscillogrammes de la production /aba/ (amplitude en (Hz) en fonction du temps en (ms))

Le premier oscillogramme, nommé : « intact », représente la totalité des indices spectraux : l'enveloppe porteuse et la structure temporelle fine. L'oscillogramme en dessous à gauche, nommé : « E-speech », représente les indices d'enveloppe porteuse uniquement. L'oscillogramme en dessous à droite, nommé : « TFS-speech », représente les indices de la structure temporelle fine. Cette figure est extraite de Bertoncini, Serniclaes et Lorenzi, 2009.

---

## I. L'implant cochléaire : enjeux actuels

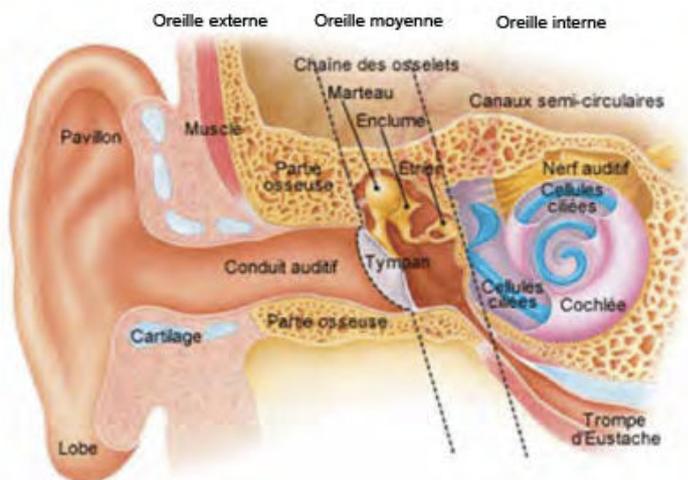
### 1. Le signal sonore

Le signal de parole est une onde, qui varie en fréquence et en amplitude dans le temps. L'amplitude de l'onde va déterminer l'intensité sonore (fort/faible), tandis que la fréquence de l'oscillation va déterminer la sensation de hauteur (grave/aigu). Le déroulement temporel de l'onde permet l'assemblage des unités de sons entre elles. Le signal sonore se représente à l'aide d'un oscillogramme, qui fait apparaître le déroulement de l'onde dans le temps, ou bien à l'aide d'un spectrogramme, qui montre les trois variables : amplitude, fréquence et temps.

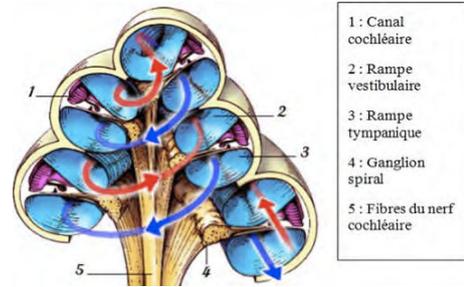
Le signal sonore est le résultat de la modulation d'une onde-source par une série de filtres (Fant, 1960, cité par Spinelli et Ferrand, 2005). L'onde-source correspond à la vibration des cordes vocales, qui produit une onde périodique, avec une fréquence fondamentale. Les filtres interviennent lors du transit de cette onde périodique à travers le conduit oro-naso-pharyngé. Ils correspondent à l'ensemble des mouvements de la langue et de la bouche au cours du passage de l'onde dans le conduit. Ces mouvements génèrent des cavités qui fonctionnent comme des résonateurs (Figure 1a). Ces résonateurs amplifient sélectivement certaines fréquences. Ce sont les formants. Au spectrogramme, les formants apparaissent comme des lignes de fréquence intensifiées. A l'oscillogramme, les formants apparaissent comme une envolée de la courbe d'amplitude et de la fréquence d'oscillation (Figure 1b). Il résulte de ces caractéristiques anatomiques de la production de parole des spécificités acoustiques (Katagiri, 2000). Le signal acoustique peut être décomposé en deux structures (Figure 2) :

- la structure temporelle fine (TFS : temporal fine structure en anglais ou enveloppe harmonique), qui correspond aux vibrations laryngées, de 50 à 800Hz environ.
- l'enveloppe porteuse (aussi appelée enveloppe d'amplitude ou E-speech en anglais), qui correspond aux modulations des filtres et donc aux mouvements des articulateurs, qui produisent des variations de bien moindre fréquence, entre 2 et 50Hz environ.

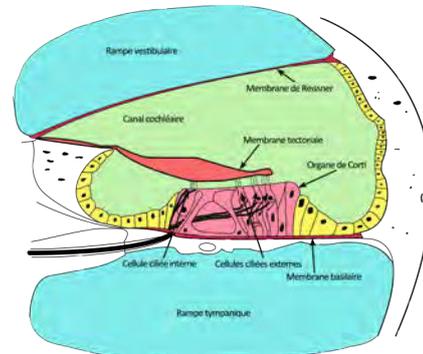
Les informations liées à la fréquence fondamentale, la qualité et le timbre de la voix sont davantage contenues dans la structure fine, donnant des indications sur l'identité du locuteur et la prosodie, tandis que les informations liées au mouvement des articulateurs sont données par l'enveloppe temporelle, qui contient ainsi une grande partie des informations pertinentes pour la reconnaissance de la parole (Katagiri, 2000). Afin d'étudier comment l'implant cochléaire capte et encode le signal acoustique, il est important de comprendre comment l'audition fonctionne chez les sujets normo-entendants. La partie qui suit a été rédigée d'après Loundon et Busquet (2009) et Menner (2005).



**a. Représentation schématique du système auditif dans sa totalité**



**b. Schéma représentant la cochlée**



**c. Schéma de l'organe de Corti sur une coupe de la cochlée**

**Figure 3 : Représentation du système auditif**

Ces trois schémas permettent de visualiser les différents éléments mentionnés dans la partie décrivant l'oreille normale. Ces figures sont extraites des sources : <http://www.audilab.fr/dangers-du-bruit.php>, <http://www.restice.univ-montp2.fr/anlo/co/ANLO-P1-CH4.html> et <http://commons.wikimedia.org/wiki/File:Cochlea-crosssection.png>

- a. Représentation des différentes parties de l'oreille: oreille externe, oreille moyenne et oreille interne.
- b. Le schéma de la cochlée permet de visualiser la situation des deux rampes et du canal cochléaire, mais aussi il permet de voir leur relation.
- c. L'organe de Corti est un tout petit organe composé principalement de cellules sensorielles. Il est situé dans le canal cochléaire.

---

## 2. L'audition normale

Dans le cas d'un système auditif fonctionnel (Figure 3a), les ondes sonores se propagent de l'oreille externe (composée du pavillon et du conduit auditif externe) vers l'oreille interne en passant par l'oreille moyenne. Les ondes arrivent à la membrane tympanique via le conduit auditif externe. Ces ondes acoustiques vont mettre en vibration la membrane du tympan qui à son tour provoque la mise en mouvement de la chaîne des osselets (marteau, enclume et étrier appartenant à l'oreille moyenne). Les stimulations acoustiques se transforment en vibrations mécaniques. L'étrier (le troisième os de la chaîne des osselets) est relié à la membrane de la fenêtre ovale, qui fait partie de la cochlée. Les mouvements de cette membrane via l'activation de l'étrier vont engendrer des variations de pressions qui vont faire bouger les fluides contenus dans la cochlée. La cochlée est un élément de l'oreille interne (Figure 3b). C'est un tube d'environ 35 mm de long enroulé autour d'un axe creux qui contient le nerf auditif. La cochlée est constituée de deux rampes, la rampe tympanique et la rampe vestibulaire, séparées par le canal cochléaire. Elles sont remplies d'un liquide appelé périlymphe. Les mouvements de fluides au sein de la cochlée (induits par les vibrations de la membrane ovale) vont à leur tour provoquer une onde qui va se déplacer le long de la membrane basilaire. Cette membrane appartient à l'organe de Corti, qui est l'organe sensoriel de l'audition (Figure 3c). L'organe de Corti est situé dans la cochlée, au niveau du canal cochléaire. Il est rempli d'endolymphe. A la base de la cochlée (vers la fenêtre ovale) la membrane basilaire est étroite et rigide, alors que vers l'apex, elle est plus souple et large. Cette évolution des propriétés physiologiques de la membrane permet d'avoir une représentation tonotopique de la gamme de fréquences perçue par l'oreille humaine. En effet, à chaque zone de la membrane basilaire correspond une bande de fréquences particulière. Les ondes se propagent sur la membrane, de la base vers l'apex. A la base de la membrane sont situées les hautes fréquences dites aigües et à l'apex, les basses fréquences correspondant à des sons graves. La gamme de fréquences perçue par l'être humain va de 20 Hz à 20 000 Hz. La membrane basilaire est recouverte sur toute sa longueur de cellules ciliées (14 000 en tout, dont 10 500 externes et 3 500 internes). Les cellules ciliées possèdent une partie appelée stéréocils qui opère une sorte de balancement suivant les mouvements de la membrane basilaire consécutifs aux variations de pressions liquidiennes. En fonction des fréquences de la stimulation auditive, des zones spécifiques de la membrane basilaire seront activées et le balancement des cellules ciliées internes provoquera l'émission de substances électrochimiques qui stimuleront les neurones auditifs liés au nerf auditif. Si les cellules ciliées internes ont une fonction de transmission du message, les cellules ciliées externes ont pour rôle d'amplifier les mouvements de la membrane basilaire et d'affiner la résolution fréquentielle.

Ce sont les cellules ciliées qui sont responsables de la transmission des informations vers le système nerveux auditif. Ainsi, lorsque ces organes sont lésés, les informations ne peuvent plus parvenir au nerf auditif et par conséquent au cerveau. Dans ces cas de lésions, le patient sera atteint d'une surdité de perception, en opposition avec une surdité de transmission qui correspond à un trouble touchant l'oreille externe ou moyenne. Par son installation au sein de la cochlée, l'implant cochléaire constitue l'un des systèmes de remédiation des surdités de perception.



**Figure 4 : Représentation de l'installation d'un implant cochléaire**

**Ce schéma permet de visualiser et de situer les différents éléments qui composent l'implant cochléaire (la partie externe et la partie interne). Il permet aussi de voir la relation entre ces différents éléments et les différentes parties anatomiques du système auditif. Cette figure est extraite du site de MED-EL Cochlear : <http://www.speechpathology.com/articles/fitting-techniques-for-pediatric-cochlear-1497-1497>**

---

### 3. Le fonctionnement de l'implant cochléaire

L'implant cochléaire est une prothèse composée d'une partie externe et d'une autre partie implantée, qui stimule le nerf auditif (Figure 4). Cet implant a pour but de transformer et transmettre les informations acoustiques externes en influx électriques. La partie externe comprend un microphone, un processeur vocal et une antenne. Le microphone, qui capte les stimuli sonores, le processeur, qui code ces stimuli acoustiques en stimuli électriques, ainsi qu'une source d'énergie (une pile ou une batterie selon les modèles) sont logés dans un contour d'oreille. Fixée sur l'os pariétal, l'antenne émet des radiofréquences qui permettent le passage transcutané des informations vers une antenne réceptrice située dans la partie interne de l'appareil (Gallégo et Collet, 2008). La partie interne comporte un récepteur-stimulateur. Celui-ci capte les stimuli électriques du processeur vocal grâce à un signal de fréquence radio et réalise un codage spatial des impulsions électriques. Ce codage est destiné à un réseau d'électrodes selon les modèles. Le récepteur-stimulateur est placé sous la peau en face de la partie externe de l'implant, alors que les électrodes sont implantées dans la rampe tympanique appartenant à la cochlée. L'insertion des électrodes se fait chirurgicalement, soit en passant par la fenêtre ronde de la cochlée, soit en réalisant une cochléostomie, c'est-à-dire une incision dans une zone située à la base de la cochlée. La profondeur d'implantation va dépendre de chaque patient. Elle est généralement comprise entre 18 et 26 mm, alors que la cochlée mesure environ 35 mm. Les électrodes auront pour but de stimuler les neurones du nerf auditif. Comme nous l'avons présenté précédemment, la cochlée possède une organisation tonotopique : à chaque partie de la membrane basilaire correspond une fréquence en partant des hautes fréquences à la base de la membrane pour aller vers les basses fréquences à l'apex. Les électrodes vont être réparties afin de reproduire la tonotopie cochléaire. Ainsi à chaque électrode correspond une bande de fréquences du signal sonore.

Pour coder les informations captées par le microphone en stimuli électriques destinés aux électrodes, les ondes sonores subissent de nombreuses modifications. Tout d'abord, la source sonore est traitée par un filtre passe-bande qui va réaliser un séquençage des fréquences situées entre 250 Hz et 8000Hz. Les informations issues de ce filtrage correspondent à l'enveloppe sonore du message. Elles permettent la compréhension des données principales de la parole, mais les éléments prosodiques et le timbre ne seront pas transmis, car ils sont portés par les structures fines qui sont éliminées lors du filtrage. Le message est découpé en bandes de fréquences correspondant au nombre d'électrodes. Le mode de transmission des impulsions électriques va varier selon les modèles (Gallégo et Collet, 2008 ; Micheyl, 2005) :

- Pour la stratégie SPEAK (Spectral Peak Strategy), seules les bandes de fréquence contenant le plus d'énergie seront transmises (jusqu'à dix maximum). La cadence de stimulation est de 250 pulses par seconde et la charge d'activation dépend de l'amplitude de l'enveloppe pour la bande de fréquence donnée.

Or, les électrodes implantées dans la membrane basilaire baignent dans la périlymphe qui est un environnement très conducteur. Un des problèmes majeurs du fonctionnement de l'implant cochléaire est de garder la spécificité fréquentielle de chaque électrode et d'éviter que l'électrode stimulée ne transmette son excitation aux électrodes voisines.

---

Ainsi lorsque deux électrodes voisines doivent être stimulées en même temps il y a un risque important de chevauchement des informations.

- Une autre stratégie a été développée afin d'éviter ce chevauchement. La stratégie CIS (Continuous Interleaved Sampling) correspond à l'activation des électrodes de l'ensemble des bandes de fréquences utilisées dans le signal sonore, mais cette stimulation est alternée : les électrodes reçoivent des pulses électriques les unes après les autres selon un rythme plus élevé (400 pulses par seconde).
- La stratégie ACE (Advanced Combination Encoder) équivaut à une combinaison des deux principes précédents : la sélection des bandes de fréquence et l'alternance de stimulation des électrodes.

L'implant cochléaire est actuellement la prothèse sensorielle la plus performante, cependant son application est limitée par certains facteurs. L'implant cochléaire transmet les éléments auditifs portés par l'enveloppe sonore, alors que les informations communiquées par la structure fine du message sont éliminées par le filtrage. Ainsi la personne implantée se base sur des informations acoustiques partielles pour percevoir la parole.

#### **4. L'implant cochléaire : indications et limites**

L'implantation est destinée aux patients atteints de surdité neurosensorielle profonde à sévère bilatérale (HAS, 2007). Les patients doivent remplir des conditions précises avant de bénéficier de cette technique :

- Un bilan orthophonique doit être proposé. Il permet de vérifier les compétences communicationnelles et langagières du patient, ainsi que son niveau de lecture labiale.
- Au niveau audiométrique, les essais de réhabilitation, avec un appareillage acoustique correctement réglé et récent, ne doivent pas permettre une discrimination de la parole supérieure à 50% pour une intensité de 60 dB (test en audiométrie vocale avec les listes de Fournier ou de Lafon).
- Une réflexion sur l'opération et ses conséquences est nécessaire afin d'obtenir le consentement éclairé du patient et de son entourage, mais aussi pour évaluer leurs capacités adaptatives et motivationnelles.
- Le patient doit bénéficier d'un état physique suffisant pour supporter l'opération et les soins de suite.

Le patient sera entouré d'une équipe pluridisciplinaire regroupant le médecin otorhinolaryngologue, le chirurgien, l'orthophoniste, le psychologue et l'audioprothésiste. Les compétences en perception de la parole des personnes implantées sont variables et dépendent de divers facteurs (Loundon et Busquet, 2009):

- L'âge d'apparition de la surdité

Il existe trois catégories de patients : les patients ayant une surdité pré-linguale, qui sont nés avec une surdité ou qui l'ont développée avant l'émergence du langage oral (avant 2 ans); les patients qui ont développé une surdité au cours de l'acquisition du langage oral

---

(entre 2 et 5 ans) et les patients sourds post-linguaux, qui ont développé une surdité après avoir mis en place le langage oral.

- La durée de la surdité

Les patients qui ont subi une période courte de privation auditive ont des résultats supérieurs en perception de la parole à ceux dont la privation a été plus longue. Une longue période de privation auditive engendre une restructuration cérébrale. Les neurones des aires auditives sont recrutés par d'autres fonctions sensorielles ou cognitives comme la vision (Sharma, Dorman et Spahr, 2002 ; Lee et al., 2001 ; cités par Wilson et Dorman, 2008).

- L'âge d'implantation

La période la plus favorable pour l'implantation s'étendrait jusque vers 3 ans et demi - 4 ans et serait maximale avant 2 ans. Lors de cette période, les voies auditives centrales seraient particulièrement plastiques, et ainsi, beaucoup plus réceptives aux stimuli sonores. Passée cette période et en l'absence d'entrées auditives, la sensibilité serait moins accrue du fait d'une réorganisation cérébrale marquée par un recrutement de certaines aires auditives en faveur d'autres aires non dévolues à l'audition. Par conséquent, une implantation précoce permettrait une maturation corticale plus adéquate et donc favoriserait le développement de la perception de la parole et de l'acquisition du langage oral (Kral et Sharma, 2012).

- La durée d'utilisation de l'implant

D'après Loizou (1998), l'augmentation de la durée de port de l'implant permet de développer les compétences perceptives de la parole aussi bien chez l'adulte que chez l'enfant. L'utilisation quotidienne et prolongée de l'implant est indispensable, car un arrêt de quelques jours engendre une déshabitude. Des patients n'ayant pas porté leur implant pendant quelques jours sont susceptibles de ne plus le supporter et de devoir bénéficier d'un nouveau protocole de réglage afin de se réhabituer.

Pour l'ensemble des patients, un suivi orthophonique est nécessaire afin de préparer et d'accompagner l'implantation (HAS, 2009). En effet la perception de la parole et la compréhension du discours chez le sujet implanté demande de nombreuses compétences. La personne devra comme tout sujet normo-entendant, réaliser à la fois un traitement acoustique, phonémique et lexical du signal sonore. Mais parce que l'implant cochléaire transmet un signal auditif dégradé, ces traitements s'établissent sur des informations acoustiques partielles.

## **II. La perception de la parole**

### **1. Perception de la parole, phonèmes et traits articulatoires**

Le traitement de la parole peut être considéré comme un processus à trois niveaux (McClelland et Elman, 1986) :

---

**Tableau I : Correspondance entre traits articulatoires et indices acoustiques principaux**

Dans ce tableau figurent les différentes caractéristiques acoustiques des consonnes et des voyelles du français. Les indices acoustiques visibles au spectrogramme sont mis en relation avec les traits articulatoires. Une description détaillée de ces indices acoustiques est disponible en Annexe I. Ce tableau est adapté de Serniclaes, Medina, Schepers et Simon (2007).

	Traits	Indices majeurs
<b>Voyelles</b>	Aperture	F1
	Antériorité	F2
	Labialité	F3
	Nasalité	Anti-formants
<b>Consonnes</b>	Mode d'articulation	Durée du bruit de friction Barre d'explosion ou non Durée des transitions de F1, F2 et F3
	Lieu d'articulation	Transition du F2 Transition du F3 Fréquence du spectre d'explosion ou de friction
	Voisement	VOT (Voice Onset Time)
	Nasalité	Anti-formants

---

- 
- Le traitement acoustique : il s'agit d'un niveau de traitement universel et non spécifique à la parole, qui permet de traiter les différences temporelles ou spectro-temporelles entre des sons de n'importe quelle nature.
  - Le traitement phonologique : il s'agit d'un niveau de traitement spécifique à une langue. Il permet de faire la différence entre deux phonèmes d'une langue donnée.
  - Le traitement lexical : les connaissances de l'auditeur dans la langue donnée entrent en jeu dans les processus de reconnaissance. Le niveau de traitement lexical correspond au niveau de l'accès au système sémantique. C'est un niveau de représentation symbolique, qui peut être sonore (langue parlée), gestuel (langue des signes) ou graphique (langage écrit).

Parmi les unités sub-lexicales considérées comme pertinentes dans la perception de la parole, le phonème est une unité récurrente. Le phonème est une unité linguistique abstraite (Dufour et Frauenfelder, 2007), qui permet de faire la différence entre deux unités de sens. Par exemple, les phonèmes /b/ et /m/ sont les unités minimales qui permettent de distinguer les mots /bal/ et /mal/. Chaque langue parlée possède un système fini de phonèmes donnés. La catégorisation des sons en unités phonémiques est elle-même conditionnée par la reconnaissance d'un certain nombre de différences articulatoires dans la production de ces sons. Ces différences articulatoires peuvent être décrites et qualifiées. On parle ainsi de trait articulatoire. Ils permettent notamment de dire si les cordes vocales sont actives lors de la production d'un son (trait de voisement), ou si la fermeture du conduit vocal se fait en avant ou en arrière de la bouche (trait de lieu articulatoire), par exemple.

Dans cette partie nous décrivons les sons du français selon leurs caractéristiques articulatoires et acoustiques, en considérant les phonèmes comme des catégories différenciées par un certain nombre de traits articulatoires. Nous présenterons ensuite le développement de la perception des phonèmes et d'un certain nombre de traits articulatoires chez le sujet normo-entendant. Nous évoquerons enfin l'influence du lexique sur la perception de la parole au regard des expériences menées dans ce domaine.

## **2. Description acoustique et articulatoire des phonèmes du français**

Nous utiliserons dans cette partie la distinction traditionnelle entre phonèmes vocaliques (voyelles) et phonèmes consonantiques (consonnes), pour faciliter la description articulatoire (Tableau I).

Les voyelles, caractérisées par une vibration systématique des cordes vocales, présentent toujours une onde périodique. Cette onde périodique est ensuite amplifiée par au moins trois cavités, qui génèrent trois formants sur le spectrogramme (F1, F2 et F3). Ces cavités correspondent grossièrement à l'espace en arrière de la langue, celui en avant de la langue et celui formé par l'avancement des lèvres (Figure 1). L'ouverture des fosses nasales pour le passage de l'onde vibratoire produit quant à elle une déperdition d'énergie visible sur certaines fréquences du spectrogramme. On appelle cette déperdition un anti-formant, par opposition au formant qui amplifie une fréquence. Il résulte de ces caractéristiques anatomiques de production une catégorisation des voyelles du français selon quatre traits articulatoires : l'aperture (degré d'abaissement de la mandibule), la labialité (l'engagement des lèvres), le lieu articulatoire (endroit le plus étroit de la cavité buccale)

---

et la nasalité (ouverture ou non des fosses nasales par l'abaissement/élévation du velum) (Tableau I sur les « voyelles »).

Les consonnes sont issues de variations de fermeture du conduit vocal. Il y a deux modes de variation: la fermeture complète (occlusion) et la fermeture partielle (constriction). /s/ sera considéré comme une consonne constrictive par opposition à /t/, considérée comme une occlusive. Au moment du resserrement de la langue contre l'arrière des dents, l'air continue de passer lors de la production d'un /s/, tandis que l'émission d'air est momentanément interrompue lors de la production d'un /t/. Il en résulte pour les occlusives une petite explosion au relâchement de l'air. Les occlusives sont donc caractérisées sur le spectrogramme par une interruption suivie d'une forte émission d'énergie sur toutes les fréquences (explosion). Les constrictives ne présentent pas de trace d'explosion. Parmi les constrictives, les fricatives (/s/, /z/, /ʃ/, /ʒ/, /f/, /v/) présentent quant à elles de l'énergie (souffle, son aperiodique) sur une large bande de fréquence au spectrogramme. Une autre variable qui permet d'identifier une consonne est l'endroit dans la bouche où a lieu cette fermeture du conduit. Un /t/ par exemple présente une occlusion juste en arrière des dents tandis qu'un /k/ présente une occlusion bien plus à l'arrière dans la bouche, au niveau de la base de la langue. Ces variations de lieu sont visibles dans le signal acoustique grâce à la transition des F2 et F3, qui donnent une indication sur le déplacement des différents articulateurs lors du passage de la consonne au son suivant. Les consonnes peuvent également présenter ou non une certaine périodicité sur le spectrogramme, selon que les cordes vocales vibrent ou non au début de leur production. On distinguera alors un /p/ (non voisé), d'un /b/ (voisé) car il y aura une vibration des cordes vocales au moment de la fermeture du conduit. Ce phénomène est mesuré grâce à l'indice acoustique de Délai d'Etablissement du Voisement (Serniclaes, Medina, Shepers et Simon, 2007, en Anglais « Voice Onset Time », VOT), qui permet d'indiquer sur le spectrogramme le moment où s'établit le voisement par rapport au début de la consonne. Enfin, comme pour les voyelles, le conduit nasal peut être ouvert durant la production des consonnes, entraînant une déperdition d'énergie. Cette déperdition est visible au spectrogramme comme un anti-formant. C'est ce qui se passe lors de la production d'un /m/ ou d'un /n/. Il résulte de ces quatre variables articulatoires la définition de quatre traits articulatoires des consonnes : le mode d'articulation, le lieu d'articulation, le voisement et la nasalité (Tableau I sur les « consonnes » et pour la description détaillée des indices acoustiques des traits articulatoires des consonnes voir Annexe I avec les Figures 19 et 20).

Nous venons de donner un certain nombre de correspondances entre les caractéristiques articulatoires des sons et leur réalisation acoustique. Cependant, il est important de rappeler qu'il n'y a pas de correspondance stricte entre la forme acoustique d'un son et sa reconnaissance comme élément linguistique. En réalité, les indices se combinent et se chevauchent dans le temps pour permettre l'identification. C'est ce que l'on appelle l'intégration phonétique (Serniclaes et al., 2007).

### **3. Le traitement phonémique de la parole**

Les recherches dans le domaine du développement de la parole ont montré que notre système perceptif est capable de discriminer l'ensemble des traits articulatoires, toutes langues confondues, jusqu'à environ 6 mois (Hoonhorst, Colin, Markessis, Radeau, Deltenre et Serniclaes, 2009a ; Serniclaes et al., 2007). Il assemble et pondère ensuite

---

l'ensemble de ces informations acoustiques pour en retirer de l'invariance (McMurray, Tanenhaus, Aslin et Spivey, 2003) et former des catégories de sons. On parle alors de phénomène de couplage phonétique (Serniclaes et al., 2007). Plusieurs études montrent que les catégories de sons correspondant à la langue d'exposition sont formées autour d'un an (Werker et Tees, 1984 ; Hoonhorst, Colin, Markessis, Radeau, Deltenre et Serniclaes, 2009b). La perception qui résulte de ces couplages est alors dite catégorielle. Elle permet la segmentation de la parole en unités, que l'on appelle phonèmes.

### **3.1. Perception et précision catégorielle : définition**

La première étude de Liberman et collaborateurs (Liberman, Harris, Hoffman et Griffiths, 1957) a permis d'établir que les différences acoustiques entre variants d'une même catégorie phonémique ne sont généralement pas perceptibles (deux variants de /ba/ par exemple), tandis que des différences acoustiques de même amplitude mais entre des sons appartenant à différentes catégories (/ba/ et /da/ par exemple) sont perceptibles. Ce phénomène est connu sous le nom de perception catégorielle (Liberman, Cooper, Shankweiler et Studdert-Kennedy, 1967). La perception catégorielle est définie par le fait que l'auditeur perçoit plus facilement les différences entre deux sons lorsqu'ils appartiennent à deux catégories phonémiques différentes, plutôt que lorsqu'ils font partie d'une même catégorie.

Dès les premières expériences étudiant le phénomène de perception catégorielle, des tâches d'identification et de discrimination ont été utilisées. La définition classique de la perception catégorielle implique que les performances de discrimination puissent être prédites sur la base des résultats d'identification. Traditionnellement, les recherches sur la perception catégorielle utilisent des stimuli variant le long d'un continuum, il est également possible d'utiliser des scores de discrimination et d'identification de paires minimales (Bouton, Serniclaes et al., 2012). Une paire minimale est composée de deux items qui se différencient par au moins un trait phonémique. Par exemple, /ba/ et /da/ ne se différencient que par le lieu d'articulation de la première consonne. Plus précisément, les résultats des tests d'identification et de discrimination de paires minimales permettent d'évaluer deux propriétés : la perception catégorielle et la précision catégorielle. La perception catégorielle correspond à la relation entre la discrimination et l'identification. Elle est d'autant plus forte que la différence entre les deux types de scores est faible. La précision catégorielle correspond aux capacités de discrimination et d'identification. Elle est d'autant plus élevée que la moyenne des deux types de scores est élevée. La capacité de catégorisation d'un individu dépend de ces deux propriétés.

### **3.2. Perception et précision catégorielle : décours développemental**

L'émergence du langage, habileté cognitive complexe, nécessiterait que l'individu soit soumis à une stimulation spécifique pendant une période sensible voire critique. Knudsen (2004) définit les notions de période sensible et de période critique. La période sensible peut être définie comme le laps de temps pendant lequel les effets de l'expérience sur la structuration cérébrale sont majeurs. Des comportements pourraient être altérés en l'absence de stimulation durant cette période. La période critique est une sous-catégorie de la période sensible. Elle correspond à la période pendant laquelle un certain type d'expériences est requis pour permettre à l'individu de se développer normalement.

---

L'absence de stimulation altère irrémédiablement la suite du développement et en l'absence de ces expériences aux périodes critiques, il y a un développement atypique.

Deux études récentes ont permis de clarifier le décours développemental du phénomène de perception catégorielle chez l'enfant normo-entendant (Hoonhorst et al., 2009a; Medina, Hoonhorst, Bogliotti et Serniclaes, 2010). Les catégories phonémiques se forment par exposition à la langue maternelle. Des enfants exposés au français âgés de 4 mois discriminent les paires de stimuli variant sur le voisement en respectant les frontières universelles (situées à -30 et +30ms de VOT), alors que les enfants français âgés de 8 mois discriminent les stimuli en respectant la frontière caractéristique du français (situé à 0ms de VOT) (Hoonhorst et al., 2009b). Les données collectées chez les nouveau-nés indiquent un apport essentiel de l'exposition à la langue maternelle dans les premiers mois de vie pour acquérir des frontières phonémiques spécifiques à sa langue maternelle. Par ailleurs, des frontières universelles ont été mises en évidence par l'étude de marqueurs neurologiques spécifiques à la reconnaissance de certains traits. On sait par exemple que le N100, un pic tardif de potentiels évoqués, est un indicateur de reconnaissance du trait de voisement (Hoonhorst et al., 2009a). La catégorisation phonémique serait donc le résultat d'interactions entre des phénomènes physiologiques universels et l'exposition à un matériel langagier.

Néanmoins, catégoriser les phonèmes selon une frontière phonémique ne suffit pas à percevoir les sons de parole de manière catégorielle. Il est également nécessaire de préciser les frontières entre catégories, afin de limiter le nombre de confusions entre catégories. En effet, on observe un effet de l'âge sur la précision catégorielle. Medina et al. (2010) constatent une augmentation de la précision de la frontière sur un continuum de voisement depuis l'âge de 9 ans jusqu'à l'âge adulte. Les auteurs interprètent ce développement tardif de la précision catégorielle à l'adolescence comme une perte progressive de la sensibilité inter-catégories au profit d'une reconnaissance des phonèmes par catégorie, avec l'exposition à un matériel langagier.

Considérés ensemble, ces résultats mettent en évidence deux phénomènes importants : d'une part, les capacités à catégoriser les phonèmes interviendraient de manière précoce dans le développement des catégories phonémiques, et d'autre part, la perception des catégories phonémiques subirait des remaniements continus tout au long de la vie pour préciser les phonèmes en fonction de la langue d'exposition. La capacité à préciser les catégories s'acquiert en parallèle avec le développement des habiletés auditives et l'exposition continue au langage parlé. Elle repose sur des habiletés précoces à traiter les informations acoustiques qui évoluent dans l'enfance.

### **3.3. Perception et précision catégorielle : intégration des indices acoustiques**

Plusieurs études concordent en effet à montrer l'évolution de la perception des informations d'enveloppe porteuse et de structure fine au cours de l'enfance. Bertoncini, Nazzi, Cabrera et Lorenzi (2011) ont montré dans leur étude que des enfants, pouvaient, dès 6 mois, discriminer le trait articulatoire de voisement en se basant uniquement sur les indices portés par l'enveloppe sonore. De plus, selon Bertoncini, Serniclaes et Lorenzi (2009), les habiletés auditives et les capacités de perception de la parole sont matures vers

---

10-12 ans. Plus précisément, les capacités d'utilisation des indices issues de l'enveloppe sonore deviendraient identiques à celles de l'adulte vers 10 ans (Eisenberg et al., 2000 ; cités par Bertoncini et al., 2009). Actuellement il n'y a pas d'études indiquant l'âge de maturation des capacités d'utilisation des indices issus de la structure fine mais on sait que l'utilisation des informations d'enveloppe porteuse arrive à une forme de maturation avant l'adolescence.

Le phonème est une unité de référence pour décrire, segmenter et identifier le signal de parole. Il a une réalité linguistique (Dufour et Frauenfelder, 2007) et physiologique (Hoonhorst et al., 2009a). Mais percevoir la parole ne consiste pas uniquement en la capacité à identifier des phonèmes. Le système perceptif intègre également des informations lexicales qui influencent la perception des mots parlés et des phonèmes qui les composent.

#### **4. Le traitement lexical de la parole**

La reconnaissance des mots est une activité nécessaire et essentielle à une communication réussie. Dans le langage parlé, le mot constitue la plus petite unité porteuse de sens. Acquérir une langue, c'est donc développer un lexique mental qui comprend les représentations phonémiques, phonologiques, sémantiques et syntaxiques des mots connus. Dans cette perspective, la reconnaissance des mots parlés s'effectuerait en comparant l'entrée sensorielle (ce que l'on entend) et les représentations des mots en mémoire.

Depuis les années 1970, les chercheurs se sont intéressés à l'influence des représentations sémantiques et lexicales dans la reconnaissance des mots. Cutler, Mehler, Norris et Segui (1987) citent de nombreux protocoles qui mettent en évidence l'importance des connaissances lexicales dans l'identification du mot : soit en mettant en opposition des mots et des pseudo-mots (chaîne de caractères ressemblant à un mot réel et respectant les règles orthographiques de la langue, mais n'ayant pas de signification) (Rubin, Turvey et Vangelder, 1976), ou des mots de fréquence d'apparition dans la langue faible et des mots de fréquence d'apparition élevée (Morton et Long, 1976 ; Segui et Frauenfelder, 1986), soit en jouant avec le contexte d'apparition de la phrase (Cutler et Fodor, 1979 ; Dell et Newman, 1980). L'ensemble de ces recherches montre que les indices sémantiques améliorent la perception des mots.

##### **4.1. Les principaux modèles de reconnaissance de la parole**

L'analyse des indices acoustiques pour déterminer les phonèmes, puis l'association des phonèmes pour sélectionner des candidats mots dans le stock lexical correspondent à des mécanismes ascendants d'analyse appelés « bottom-up ».

Marslen-Wilson et Welsh, en 1978 ont développé le modèle de la Cohorte, qui se base uniquement sur un processus « bottom-up » de perception de la parole. L'information initiale du mot est l'information acoustique, et permet la sélection mentale d'un ensemble de mots candidats (une cohorte). Ces candidats ne sont activés que s'ils correspondent en tout point aux premiers phonèmes du mot-cible. Par exemple, la séquence de phonèmes /ta/ permet de former une cohorte initiale qui comporterait les mots "tableaux", "tapis",

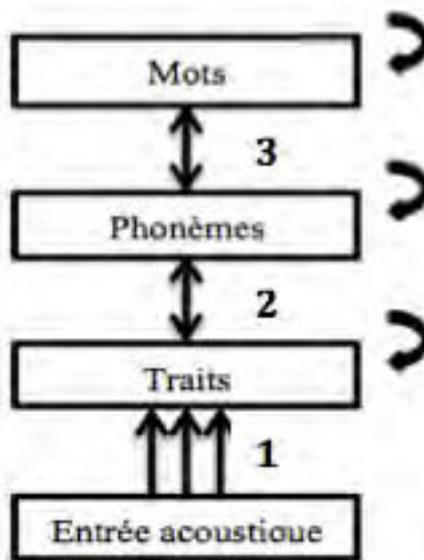


Figure 5 : Représentation du modèle TRACE

Ce schéma permet de représenter les interactions entre le traitement lexical et le traitement phonémique. Lors de la reconnaissance de mots dans le signal de parole, différents processus entrent en jeu.

Le premier processus est un processus croissant appelé l'effet « bottom-up », il est détaillé ici en 3 étapes :

1. Tout d'abord, le sujet perçoit des informations acoustiques dont il extrait des caractéristiques (traits acoustiques). Ce sont les indices acoustiques spécifiques de ces traits qui permettent de les identifier.
2. Puis ces traits acoustiques permettent de sélectionner différents phonèmes dans les représentations phonémiques du sujet.
3. La sélection de certains phonèmes va activer dans le lexique mental du sujet, différents mots pouvant correspondre aux indices acoustiques apportés par le signal de parole. La particularité du modèle TRACE est de décrire un second processus, appelé effet « top-down ».

Le modèle TRACE fait l'hypothèse d'un phénomène descendant, dans lequel les mots activés par les informations acoustiques et phonémiques auraient un effet rétroactif sur l'activation des phonèmes. Ainsi les mots sélectionnés grâce à l'activation de certains phonèmes en représentation mentale pourraient renforcer le niveau d'activation de ces mêmes phonèmes et inhiberaient d'autant plus les autres phonèmes non activés.

---

"talus", etc. Dans ce modèle, l'accès au lexique est donc basé sur les représentations phonémiques.

Norris (1994) propose le modèle « Short-list », également uniquement basé sur des processus « bottom-up » de perception de la parole, qui prédit qu'un mot est reconnu par reconnaissance des unités pré-lexicales. Lors d'une première étape, un nombre limité de représentations lexicales ("Short-list") est sélectionné par correspondance entre le nombre de phonèmes du signal d'entrée et de chaque unité lexicale. Durant la seconde étape, la « Short-list » est importée dans un réseau de compétition lexicale où le mot cible sera sélectionné par inhibition des autres mots pré-sélectionnés. Le modèle « Short-list » apparie donc une entrée sensorielle et une représentation lexicale sur la base du phonème.

Le modèle TRACE (McClelland et Elman, 1986) postule l'existence de trois niveaux de traitement successif lors de l'accès au lexique : les traits phonétiques, les phonèmes et les mots. Ces niveaux sont organisés hiérarchiquement avec des connexions interactives représentant des processus « bottom-up » et « top-down » (Figure 5). L'activation des candidats lexicaux est renforcée en fonction de leur correspondance avec le mot cible, ce qui amplifie également l'activation des phonèmes et des traits acoustiques qui les composent au détriment des autres éléments qui seront davantage inhibés. Ce modèle permet d'expliquer le fait qu'un phonème est plus précisément reconnu dans un mot que dans un pseudo-mot (Frauenfelder et Peeters, 1990; Ganong, 1980; Samuel, 1981). Le modèle TRACE postule que lorsqu'un phonème est situé dans un mot, les unités phonémiques qui le composent reçoivent de l'activation ascendante mais également de l'activation descendante. En revanche, lorsqu'un phonème est situé dans un pseudo-mot, elle ne peut recevoir que de l'activation ascendante puisque les pseudo-mots ne disposent pas de représentations lexicales qui leur permettraient de renforcer l'activation des unités phonémiques qui les composent.

Les modèles qui postulent que la reconnaissance repose sur des effets « bottom-up » uniquement, comme Cohorte et « Short-list », sont dits autonomes. Les modèles qui intègrent des effets « top-down », comme TRACE, sont dits interactionnistes. La nécessité d'un effet rétroactif des unités les plus grandes (mot) sur les unités les plus petites (phonèmes et traits phonétiques) est débattue. Une des questions qui se posent encore est celle de savoir si les mécanismes descendants sont nécessaires à la reconnaissance de la parole (Norris, McQueen et Cutler, 2000). Une autre question est de savoir s'ils sont actifs durant la reconnaissance (online) ou a posteriori, dans les processus d'apprentissage notamment (Van Linden, Stekelendurg, Tuomainen et Vroomen, 2007).

## **4.2. Effet du lexique sur le traitement phonémique**

Les effets du lexique pour compenser une altération du signal sonore dans des tâches de reconnaissance de la parole ont été mis en évidence à plusieurs reprises depuis les années 1970. Warren, en 1970, en remplaçant un phonème par un autre bruit dans une tâche de reconnaissance de mot, a mis en évidence la capacité du système de perception à accéder malgré tout à la représentation lexicale du mot dégradé, produisant une illusion de perception. Samuel reprend ce paradigme expérimental en 1981, confirmant les résultats de Warren et la prédictibilité du signal de parole en fonction des connaissances lexicales. En accord avec le modèle interactionniste, Ganong (1980) montre, dans une tâche de décision phonémique, que les connaissances préalables de l'auditeur peuvent influencer la

---

perception des frontières catégorielles des phonèmes ambigus. Ces trois études princeps mettent en avant un effet du lexique sur l'unité phonème, corroborant ainsi les théories interactionnistes.

VanLinden et al. ont repris en 2007 le paradigme expérimental de Ganong (1980) et y ont ajouté une mesure de l'activation cérébrale dans cette tâche, grâce à la MMN (Mismatch Negativity), qui permet de voir les activations automatiques dues à la perception d'un changement acoustique. Ils ont ainsi trouvé que les changements acoustiques étaient mieux perçus quand le résultat de la déviation phonétique générait un pseudo-mot que quand elle générait un mot. Ce résultat confirme non seulement l'existence de mécanismes d'activation « top-down » du lexique sur le phonème, mais il montre également que ces mécanismes interviennent de façon automatique et ont un corrélât neurophysiologique.

Dans la vie de tous les jours, les auditeurs perçoivent le signal de parole de manière fiable malgré une grande variabilité de réalisations (débit de parole, voix du locuteur, accent), si bien que l'on peut considérer que la reconnaissance a toujours lieu dans des conditions dégradées (Mattys, Davis, Bradlow et Scott, 2012). Une large variété de formes acoustiques est perçue comme un même phonème ou un même mot. L'influence du lexique sur la fiabilité de la perception est encore mal définie. Dans cette étude, nous nous intéressons à l'influence du lexique dans une condition de dégradation particulière : la parole vocodée.

### **III. La perception de la parole en contexte de signal dégradé**

#### **1. La perception de la parole de l'enfant implanté**

Malgré la réception d'informations auditives partielles, les personnes porteuses d'implant cochléaire réussissent à percevoir et comprendre la parole. Les enfants implantés avant l'âge de deux ans présentent une plasticité cérébrale propice au développement d'un langage oral efficient (Kral et Sharma, 2012). Ces enfants sont capables de se baser sur les indices portés par l'enveloppe sonore pour développer leur perception de la parole. Percevoir les sons de parole semble en effet possible en se basant uniquement sur l'enveloppe temporelle du signal. Les enfants normo-entendants exposés à un signal dégradé peuvent dès 6 mois discriminer des syllabes, en l'absence d'information spectrale ou de structure temporelle fine (Bertoncini et al., 2011).

Néanmoins, les enfants munis d'un implant cochléaire ont des compétences de perception des sons de parole inférieures à celles d'enfants normo-entendants. Medina, Loundon, Busquet, Petroff et Serniclaes (2009), montrent que des enfants implantés âgés de 5;9 ans à 11 ans, ayant en moyenne une période d'implantation de 3;3 ans, ont des résultats inférieurs à ceux d'enfants normo-entendants appariés en âge chronologique pour des tâches de discrimination de phonèmes. Ainsi, pour les voyelles, la nasalité est moins bien perçue que l'aperture, le lieu ou la labialité. Pour les consonnes, le déficit touche le lieu articulaire et la nasalité plutôt que le mode ou le voisement.

Une autre étude a mis en évidence que les enfants implantés ont des scores d'identification et de discrimination de paires minimales inférieurs à ceux d'enfants

---

normo-entendants de même âge audio-perceptif (Bouton, Serniclaes et al., 2012). Les difficultés mises en évidence seraient associées aux capacités de précision catégorielle, mais le développement de la perception catégorielle des enfants implantés serait équivalent à celui d'enfants normo-entendants de même âge auditivo-perceptif.

Les enfants implantés auraient donc des représentations phonémiques et phonologiques moins bien définies que les enfants normo-entendants. Selon les modèles précédemment cités (partie II.3 et II.4), il semble que le niveau phonémique est essentiel à la perception des mots. Les enfants implantés pourraient développer des traitements lexicaux, associés à la perception des mots, différemment de ceux des enfants normo-entendants. L'étude de Bouton, Colé et al. (2012) propose de comparer la perception de mots et de pseudo-mots d'enfants implantés et d'enfants normo-entendants. Pour les deux groupes, les scores de discrimination de paires minimales sont supérieurs pour les mots que pour les pseudo-mots. Ce résultat indique une influence des connaissances lexicales des enfants implantés et normo-entendants lorsqu'ils perçoivent des phonèmes, et cette influence est de même ampleur dans les deux groupes. Le traitement lexical des mots parlés seraient donc équivalent entre enfants implantés et normo-entendants. L'étude de Kirk, Pisoni et Osberger (1995) permet de compléter cette approche puisqu'ils mesurent l'effet de la fréquence lexicale, de la longueur des mots et de la densité lexicale (la quantité de mots-voisins phonologiquement proches du mot-cible) sur la reconnaissance des mots chez des enfants implantés et normo-entendants. Les résultats démontrent que les enfants implantés reconnaissent les mots en recourant aux mêmes traitements que les enfants normo-entendants. Ils procèdent par comparaison au sein de leur stock lexical entre les mots-voisins du mot-cible et le mot-cible. La haute fréquence, ainsi que la faible densité lexicale seraient une aide à la reconnaissance du mot aussi bien chez les enfants implantés que chez les normo-entendants, et le nombre de voisins orthographiques influence la perception des mots chez les deux groupes.

Le développement de la précision catégorielle ne peut se faire sur la seule base des informations transmises par l'implant. Les sujets implantés ont donc besoin de suivre un entraînement complémentaire de leur propre perception afin d'affiner et de développer leur précision catégorielle. Contrairement aux études précédemment citées qui se basent sur les compétences des enfants implantés, la perception de la parole chez l'adulte implanté présente des caractéristiques différentes. En effet, les adultes implantés sont, à l'heure actuelle, généralement des sujets présentant une surdité post-linguale, alors que les enfants des études précédentes ont développé une surdité pré-linguale. Les compétences en perception de la parole se voient alors modifiées du fait de l'existence d'un stock lexical avant l'apparition de la surdité.

## **2. La perception de la parole de l'adulte implanté**

Les performances en perception de la parole chez les adultes sourds post-linguaux implantés sont variables. Elles vont de sujets capables d'entretenir une conversation téléphonique en l'absence de toute aide de la lecture labiale, à des sujets pour qui l'implant n'offre qu'un complément à la lecture labiale. Cette différence de résultats peut être la conséquence de la variabilité du nombre de neurones auditifs encore viables et de facteurs comme la durée de la surdité avant implantation. Cependant la façon dont les informations sonores sont codées par l'implant et l'analyse que peut en faire l'auditeur, en fonction de ses connaissances phonémiques et lexicales, sont aussi essentielles.

---

Meyer, Frisch, Pisoni, Miyamoto et Svirsky (2003) ont réalisé une étude sur la perception de la parole et la reconnaissance des mots chez des adultes implantés présentant une surdité post-linguale. Pour se faire les auteurs ont réalisé un modèle computationnel configuré selon le « Neighborhood Activation Model » (Luce et Pisoni, 1998). Ce modèle suppose que les mots de la chaîne parlée sont reconnus par l'individu par comparaison avec d'autres mots aux consonances communes au sein de son stock lexical. Les mots aux consonances communes à celles du mots-cible forment son voisinage lexical. Un mot avec peu de mots-voisins sera plus facile à reconnaître qu'un mot avec un voisinage lexical important. De plus, pour deux mots ayant le même nombre de voisins lexicaux, le mot ayant des voisins lexicaux dont la fréquence d'apparition dans la langue est élevée, voire plus élevée que sa propre fréquence, sera plus difficile à reconnaître que le mot dont les voisins lexicaux sont de plus faible fréquence. Le modèle intègre des traitements « bottom-up » et « top-down » proposés par le modèle TRACE. Les résultats démontrent que les adultes implantés identifient les mots entendus sur la base des caractéristiques acoustiques transmises par leur implant, mais utilisent aussi leurs connaissances lexicales et phonémiques, ainsi que la fréquence d'apparition des mots dans leur langue. La fréquence d'apparition des mots dans la langue serait un indice particulièrement important pour les sujets implantés post-linguaux. L'utilisation de ces différents traitements est fondamentalement similaire à celle des adultes normo-entendants. Par conséquent, il semble que les adultes implantés devenus sourds après le développement du langage gardent une représentation lexicale identique à celle qu'ils avaient lorsqu'ils étaient entendants. De plus, les informations transmises par l'implant permettraient aux patients d'accéder de la même manière que les adultes normo-entendants à leur stock lexical.

Kaiser, Kirk, Lachs et Pisoni (2003) démontrent aussi un effet de la lexicalité dans les tâches de perception de la parole chez des adultes implantés suite à une surdité post-linguale, comme chez des adultes normo-entendants. Les deux groupes détectent plus aisément les mots ayant peu de voisins lexicaux et étant plus fréquents que leurs voisins lexicaux que les mots ayant un grand nombre de voisins lexicaux et ayant une moindre fréquence d'apparition que leurs voisins lexicaux. Cependant les adultes sourds post-linguaux implantés présentent un score inférieur aux normo-entendants pour la reconnaissance de mots ayant un grand nombre de voisins lexicaux et ayant une moindre fréquence d'apparition que leurs voisins lexicaux. Cette difficulté pourrait révéler un traitement lexical différent entre adultes implantés et normo-entendant lors de la perception des sons de parole.

### **3. Le vocodeur : principe et application**

Le vocodeur est un dispositif électronique de traitement du signal sonore, habituellement utilisé pour simuler la perception sonore fournie par l'implant cochléaire. Ce mécanisme permet le découpage et l'analyse des différents éléments spectraux de l'ensemble des messages sonores (la voix comme la musique). Grâce à ce traitement, une modification de l'ensemble du signal sonore peut être réalisée, et ainsi une reproduction de l'implant cochléaire (Gnansia, 2009). Pour simuler la perception obtenue en situation d'implantation, une séparation de l'enveloppe sonore et de la structure fine est nécessaire. Les informations portées par la structure fine sont alors dégradées et celles transmises par les modulations d'amplitudes de l'enveloppe serviront de trame.

---

Les sujets normo-entendants utilisent un mécanisme pour discriminer le signal de parole des bruits environnants. Ce mécanisme s'appelle le « speech masking release » ou démasquage de la parole, en français (Lorenzi, Husson, Ardoit et Debruille, 2006 ; Gnansia, Péan, Meyer et Lorenzi, 2009). Ce processus se base sur les différences spectrales et les variations de structure fine de l'ensemble des sources sonores. Il est spécialement utilisé pour la discrimination des fréquences correspondant au fondamental de la voix humaine et à ses harmoniques graves. Ainsi le démasquage de la parole permet, chez les normo-entendants, une augmentation de la perception des informations portées par certains traits articulatoires (le voisement, le lieu et le mode articulatoire). L'implant cochléaire, en transmettant les informations issues de l'enveloppe sonore et non de la structure fine, ne permettrait pas cette sélectivité de la voix humaine. Cela expliquerait la plainte de perception de la parole dans le bruit chez ces patients. Dans le cadre d'une expérimentation avec des participants normo-entendants, la dégradation de la structure fine par un vocodeur ne suffirait pas à reproduire cette situation de perte de sélectivité de la voix humaine dans le bruit. Ainsi une perturbation des informations issues de la fréquence est nécessaire. Pour ce faire un procédé appelé « vocoder tone » est utilisé. Il permet une dégradation des contours de la hauteur de la voix et une altération de la perception des formants, tout en ajoutant des sons perturbateurs (sons purs).

De plus, comme nous l'avons vu précédemment, l'un des problèmes majeurs de l'implantation est le risque de chevauchement des informations portées par les électrodes. En effet, les électrodes baignent dans la périlymphe, une substance très conductrice. Elles ont tendance à diffuser leur excitation électrique à leurs voisines, et cela malgré les différents procédés de transmission des impulsions électriques. La principale conséquence de ce phénomène est une perte de sélectivité en fréquence et une réduction de la fréquence de résolution. D'après Tyler, Hall, Glasberg, Moore et Patterson (1984), ce phénomène s'observerait chez l'ensemble des déficiences auditives d'origine cochléaire et impliquerait majoritairement une diminution de la résolution des extrémités basse et haute de la gamme de fréquence perçue par l'homme. Afin de reproduire ce phénomène de chevauchement des informations, un mécanisme appelé « smearing » est employé. Ce procédé a pour but de produire un élargissement du découpage en bandes de fréquence réalisé par la cochlée. Ainsi la largeur des bandes est multipliée par deux ou quatre, produisant une imprécision fréquentielle (Annexe II avec les Figures 21, 22 et 23, ainsi que le Tableau IV).

Le vocodeur permet de modéliser le fonctionnement de l'implant cochléaire. Il est ainsi un outil particulièrement utile pour les recherches en perception de la parole, car il permet de mettre des sujets normo-entendants n'ayant jamais subi de période de privation auditive et présentant un stock lexical ainsi que des représentations phonémiques matures en situation de perception auditive dégradée. Cette condition particulière offre la possibilité d'étudier les conséquences de cette dégradation du signal sur traitement de la parole et en particulier sur les niveaux de traitement phonémique et lexical.

---

**Chapitre II**  
**PROBLEMATIQUE ET HYPOTHESES**

---

## I. Problématique

Les différentes études citées dans la partie théorique (Bouton, Colé et al., 2012 ; Bouton, Serniclaes, et al., 2012 ; Médina et al., 2009) démontrent que l'implant ne transmet pas de la même manière et avec la même qualité tous les indices phonémiques. Certains traits sont mieux perçus que d'autres par les sujets implantés. Or, lors de la reconnaissance des mots de la chaîne parlée, les informations issues des données phonémiques sont indispensables. L'effet « bottom-up » est une des hypothèses tentant d'expliquer comment les sujets traitent le signal de parole. C'est un processus qui mène à la sélection de candidats lexicaux pouvant correspondre aux mots du message, sur la seule base des informations acoustiques perçues par le sujet. En fonction des phonèmes reconnus dans le message, les sujets activent des représentations lexicales contenant ces phonèmes et inhibent les autres. En complément de cette hypothèse, un autre modèle de traitement de la parole a été développé : le modèle TRACE (McClelland et Elman, 1986). Il expose le fonctionnement du processus « top-down », qui permet une action du traitement lexical sur le traitement phonémique. En lien avec cette théorie, des études prouvent que les sujets implantés s'aident, tout comme les sujets normo-entendants, de leurs connaissances lexicales et phonémiques (Bouton, Colé et al., 2012 ; Kaiser et al., 2003 ; Kirk et al., 1995 ; Meyer et al., 2003).

Actuellement, peu d'études en perception de la parole chez les sujets implantés sont menées à partir de populations adultes. La majorité de ces études se concentre sur la perception d'enfants implantés en comparaison à celle d'enfants normo-entendants (Bertoncini et al., 2011 ; Bouton, Colé et al., 2012 ; Bouton, Serniclaes et al., 2012 ; Kirk et al., 1995 ; Kral et al., 2012, Medina et al., 2009). Si ces études nous éclairent sur les processus mis en œuvre par les enfants implantés, elles ne donnent aucun indice sur les conséquences respectives de la composante développementale et de la composante acoustique. En effet, ces études ne permettent pas de séparer les observations découlant d'une influence développementale du lexique de celles résultant des informations transmises par l'implant. Ainsi, les conséquences dues aux représentations lexicales immatures et celles résultant des informations acoustiques dégradées ne peuvent être correctement identifiées.

Dans l'objectif de départager ces deux types d'influences, nous avons réalisé une étude sur des adultes normo-entendants en recréant artificiellement ce que transmet l'implant grâce au vocodage. En effet, à l'heure actuelle, il est possible de modéliser le fonctionnement de l'implant à l'aide d'un vocodeur. Ce programme informatique permet de simuler le signal de l'implant cochléaire et de le faire entendre à des sujets normo-entendants. Grâce à l'utilisation de ce procédé dans notre étude, les conséquences de l'aspect développemental, lié aux capacités de représentations lexicales, sont exclues. Les observations ne peuvent être attribuées qu'à l'influence de la composante acoustique, liée aux informations portées par l'implant.

Ainsi pour cette étude, nous avons cherché à observer les conséquences de la dégradation du signal auditif sur le traitement de la parole, et en particulier sur les interactions entre les niveaux de traitement phonémique et lexical.

---

## **II. Hypothèses**

### **1. Hypothèse générale**

Notre hypothèse de travail se base sur le fait que durant l'identification des éléments d'un signal de parole, il y aurait un effet du lexique sur les perceptions phonémiques. De plus, cette influence du traitement lexical augmenterait dans le cas d'une perception dégradée au vocodeur. Pour nous permettre d'observer l'influence du traitement lexical, nous avons défini différentes variables (Annexe III pour un tableau récapitulatif du matériel) :

#### **1.1. Les variables indépendantes (VI)**

- VI1 : le type de matériel (mots vs pseudo-mots). Cette variable nous permet d'observer l'effet de lexicalité.
- VI2 : le niveau de dégradation (8 bandes vs 16 bandes vs non vocodé). Cette variable nous permet de mesurer l'importance de la dégradation dans la mise en place des stratégies compensatoires.
- VI3 : la tâche (discrimination vs identification). Cette variable nous permet de calculer des scores en perception et en précision catégorielle.

#### **1.2. Les variables contrôles (VC)**

- VC1 : la structure syllabique du matériel (composé de consonnes (C) et voyelles (V) peut être de forme CVC vs CVCV). Cette variable nous permet de contrôler le rôle de la structure des stimuli avec des mots et pseudo-mots (monosyllabiques et bisyllabiques).
- VC2 : les traits articulatoires (voisement/nasalité/lieu/mode). Cette variable nous permet de proposer des stimuli présentant l'ensemble des caractéristiques phonémiques des consonnes et ainsi de ne pas privilégier des informations phonémiques mieux transmises que d'autres par l'implant.
- VC3 : la fréquence des mots (haute fréquence vs basse fréquence). Cette variable nous permet de présenter des résultats non biaisés par un effet de fréquence d'apparition dans la langue française.

#### **1.3. Les variables dépendantes**

- VD1 : le pourcentage de réponses correctes correspondant au nombre de réponses correctes sur le nombre de réponses totales. A partir du pourcentage de réponses correctes, nous pouvons calculer : la perception catégorielle (le rapport entre le score de discrimination et le score d'identification) ; la précision catégorielle (la moyenne des scores aux tâches de discrimination et d'identification).
- VD2 : le temps de réponse (score t) : correspondant au temps de latence entre l'affichage de l'écran de réponse et la réponse du participant (nous avons différencié ce phénomène pour la tâche de discrimination et d'identification).

---

## **2. Hypothèses opérationnelles**

Afin de répondre à notre hypothèse générale, nous formulons cinq hypothèses opérationnelles :

### **Hypothèse 1 : l'effet de lexicalité**

Il y a plus de réponses correctes et le temps de réponse est moins long pour les tâches comprenant des mots par rapport à celles avec des pseudo-mots.

### **Hypothèse 2 : l'effet de dégradation du signal**

- a) Il y a plus de réponses correctes et le temps de réponse est moins long pour les tâches avec la condition non vocodée que pour celles avec la condition vocodée (à 8 ou 16 bandes de fréquence).
- b) Il y a plus de réponses correctes et le temps de réponse est moins long pour les tâches avec la condition vocodée à 16 bandes de fréquence que pour celles avec la condition vocodée à 8 bandes de fréquence.

### **Hypothèse 3 : l'effet du lexique sur la perception et la précision catégorielle**

- a) Le score en perception catégorielle restera le même que les stimuli soient des mots ou des pseudo-mots.
- b) Le score en précision catégorielle sera meilleur pour des mots que pour des pseudo-mots.

### **Hypothèse 4 : l'effet du lexique selon la dégradation du signal**

- a) Il y a plus de réponses correctes et le temps de réponse est moins long pour les mots par rapport aux pseudo-mots dans les conditions avec vocodage. L'effet de lexicalité est plus important quand le signal est dégradé, c'est-à-dire pour les conditions vocodées à 16 et 8 bandes par rapport à la condition non vocodée.
- b) Il y a plus de réponses correctes et le temps de réponse est moins long pour les mots par rapport aux pseudo-mots dans la condition la plus dégradée (8 bandes). L'effet de lexicalité augmente avec le niveau de dégradation du signal, c'est-à-dire pour la condition vocodée à 8 bandes par rapport à la condition vocodée à 16 bandes.

### **Hypothèses 5 : l'effet de la dégradation du signal sur la perception catégorielle**

- a) La perception catégorielle est affectée par la dégradation du signal : le rapport entre les scores de discrimination et d'identification est meilleur dans la condition non vocodée par rapport aux conditions vocodées.
- b) La perception catégorielle est d'autant plus affectée que le niveau de dégradation est important : le rapport entre les scores de discrimination et d'identification est meilleur dans la condition vocodée à 16 bandes par rapport à la condition vocodée à 8 bandes.

---

**Chapitre III**  
**PARTIE EXPERIMENTALE**

---

Notre travail de recherche a pour but d'étudier l'effet du lexique sur la perception de la parole dans un contexte de signal dégradé. L'étude de groupe est apparue comme un protocole adéquat pour étudier la perception de la parole comme un phénomène global dont nous essayons de déterminer un certain nombre d'invariants.

## I. Population

50 adultes normo-entendants, de langue maternelle française, ne souffrant d'aucun trouble langagier acquis ou développemental ont été recrutés pour notre expérience.

Nous avons déterminé des critères de sélection pour les participants de notre étude :

- Les critères d'inclusion :
  - être normo-entendant
  - être âgé entre 20 et 35 ans
  - être de langue maternelle française
  
- Les critères d'exclusion :
  - présenter des antécédents de problèmes auditifs
  - présenter des difficultés associées au langage (troubles spécifiques du langage)
  - présenter des troubles neurologiques (antécédents médicaux)
  - présenter des troubles attentionnels
  - présenter des troubles visuels sans correction optique adéquate

Ces 25 femmes et 25 hommes, âgés de 20 à 34 ans ont été recrutés sur la base du volontariat grâce à un mail (Annexe IV) expliquant notre expérimentation et les critères de sélection des candidats. Ce mail a été transmis via la liste de diffusion des étudiants de Lyon 1 (etudiants@univ-lyon1.fr) et celle des étudiants en orthophonie de Lyon (ortholyon@yahoogroupes.fr). Grâce à ces mails, nous avons reçu une soixantaine de réponses. L'âge moyen des candidats ( $M$ ) est de 24,12 ans, la déviation standard ( $DS$ ) est de 2,74 ans.

L'ensemble des participants sont ou ont été étudiants dans les universités de Lyon 1 et de Lyon 2, ou sont des conjoints d'étudiants. 27 candidats sur 50 ont réalisé un audiogramme n'objectivant aucun trouble auditif dans une période allant de six mois à cinq ans avant notre expérimentation. De plus, avant de commencer chaque passation, tous les participants devaient régler le volume sonore auquel les stimuli de l'expérience leur seraient présentés. Pour l'ensemble des candidats, ce réglage du volume sonore a été compris entre 70 et 85 dB SPL. Nous n'avons pas réalisé d'évaluation du niveau lexical de nos participants. Cependant 86% des participants ont réalisé des études universitaires ou sont en cours de formation. Les 14% restant ont suivi une formation type CAP, BEP, BTS ou IUT. Nous avons donc fait l'hypothèse que nos participants présentaient un niveau lexical suffisant pour notre expérimentation.

---

Nous avons divisé notre population en deux groupes:

- 1 groupe qui a réalisé la tâche à 16 bandes (signal dégradé avec 16 bandes de vocodage) et la tâche contrôle (signal non dégradé).
- 1 groupe qui a réalisé la tâche à 8 bandes (signal dégradé avec 8 bandes de vocodage) et la tâche contrôle (signal non dégradé).

Le groupe ayant réalisé la tâche à 16 bandes compte 25 participants dont 12 femmes et 13 hommes. Le groupe ayant réalisé la tâche à 8 bandes compte aussi 25 participants dont 13 femmes et 12 hommes. La partie contrôle n'a pas pu être proposée à l'ensemble des participants : 41 candidats ont passé cette tâche ; parmi eux 20 femmes et 21 hommes.

La problématique de notre étude nous a amenées à tester les conséquences de la dégradation du signal auditif sur les niveaux de traitement lexical et phonémique. Dans la plupart des études que nous avons pu consulter, la population expérimentale était composée d'enfants implantés et d'enfants normo-entendants. Dans les conclusions de ces recherches, il est difficile de différencier l'influence de la composante développementale du lexique et l'influence de la composante acoustique des informations transmises par l'implant. Il est difficile d'évaluer les compétences d'enfants implantés du fait de l'hétérogénéité de la population. En effet, l'histoire de la surdité des sujets comporte de nombreux éléments à prendre en compte : l'âge d'apparition de la surdité, la durée de la privation auditive avant l'implantation, l'âge d'implantation, la durée d'utilisation de l'implant, mais aussi le niveau de développement du langage. Pour limiter une trop grande variabilité des connaissances lexicales et phonémiques chez les enfants, il est possible d'évaluer le lexique avec des tests orthophoniques. Cependant, les tests actuels sont principalement centrés sur du lexique concret et imageable, ce qui ne permet pas de faire un inventaire complet des capacités des participants en réception. Afin de préciser l'influence de la composante acoustique sur la perception de mots, nous nous sommes intéressées à une population ayant des représentations phonémiques et un stock lexical matures : des adultes. Face à la difficulté de réunir un groupe homogène de participants adultes implantés sourds post-linguaux, et du fait de la diversité des compétences perceptives et de l'historique des soins, nous avons constitué une population d'adultes normo-entendants en éliminant tout antécédent de privation auditive. Nous avons alors placé ces participants en situation de perception dégradée reproduisant celle transmise par l'implant cochléaire. Notre population d'adultes normo-entendants s'est trouvée dans la situation simulée de sujets implantés présentant une surdité post-linguale.

Nos participants ont été sélectionnés parmi la tranche d'âge 20-35 ans. Plusieurs résultats nous ont permis de délimiter cet intervalle d'âge pour la population de notre expérimentation. Les potentiels évoqués auditifs (PEA) chez les humains évoluent jusqu'à la fin de l'adolescence. En effet Sussman, Steinschneider, Gumenyuk, Grushko et Lawson (2008) démontrent que la morphologie des PEA présente des différences entre un groupe de participants âgés de 16 ans et un autre comptant des individus compris entre 22 et 40 ans, alors qu'aucune évolution n'est repérée au sein du groupe d'adultes. De plus, Medina et al. (2010), constatent une augmentation de la précision de la frontière pour le trait de voisement entre l'âge de 9 ans et la majorité. Enfin, la limite d'âge supérieure nous a permis d'éliminer des participants pouvant présenter une presbycusie (une perte d'audition, liée à l'âge, bilatérale et symétrique, affectant surtout les fréquences élevées). Elle apparaît généralement à partir de 60 ans, mais peut survenir dès 40 ans pour les cas les plus précoces (Menner, 2005).

---

## II. Matériel

Le matériel expérimental a été construit à partir de la base de données Lexique3. Nous avons constitué un ensemble de paires minimales de mots et de pseudo-mots nous permettant de tester les capacités de perception phonémiques de nos participants (exemple de paire minimale de mots : talon-salon ; exemple de paire minimale de pseudo-mots : tadon-sadon). Un tableau récapitulatif de la construction du matériel est disponible en Annexe V.

Les études récentes dans le domaine de la perception de la parole ont utilisé le trait articulatoire comme différence minimale impliquée dans la perception des mots et des phonèmes (Bertoncini et al., 2011 ; Bouton, Colé, et al., 2012 ; Hoonhorst et al., 2009a et 2009b; Medina et al., 2010).

Dans la lignée de ces études, nous avons choisi de faire varier les paires minimales de mots et de pseudo-mots d'un trait articulatoire seulement. Les paires minimales sont ainsi construites en suivant les quatre caractéristiques articulatoires des consonnes (voisement, nasalité, mode articulatoire et lieu articulatoire). Par exemple :

- « poule », dont la consonne initiale est non voisée a été apparié à « boule », dont la consonne initiale est voisée,
- « motte », dont la consonne initiale est nasalisée a été apparié à « botte », dont la consonne initiale n'est pas nasalisée,
- « fer » dont la consonne initiale est une constrictive a été apparié à « père, dont la consonne initiale est une occlusive,
- « cœur », dont la consonne initiale est postérieure a été apparié à « peur », dont la consonne initiale est antérieure.

Dans certaines études (Bouton, Serniclaes et al, 2012), la perception phonémique pour l'ensemble des traits consonantiques et vocaliques, a été testée. D'autres études se sont focalisées sur un trait articulatoire uniquement (le voisement, souvent) (Bertoncini et al., 2011 ; Hoonhorst et al., 2009a et 2009b ; Medina et al., 2010). Nous avons ici choisi de nous restreindre à l'étude de la perception des traits articulatoires consonantiques uniquement, mais en présentant l'ensemble des traits consonantiques. Tous les traits consonantiques sont ainsi représentés : voisement, nasalité, mode et lieu d'articulation. Les traits vocaliques (voisement, nasalité, aperture et lieu d'articulation) ne sont pas étudiés.

Par ailleurs, les mots choisis ne varient que sur leur premier phonème. Certains modèles de reconnaissance de la parole, comme le modèle Cohorte, (Marslen-Wilson et Welsch, 1978) et certains protocoles expérimentaux (Jakimik, 1980, cité par Spinelli et Ferrand, 2005) insistent en effet sur l'importance des premiers phonèmes dans la reconnaissance des mots. En plaçant l'ambiguïté en début de mot, nous avons souhaité réduire l'effet d'indigence potentiel qu'aurait pu amener le début du mot. Nous avons également voulu privilégier un mécanisme de reconnaissance immédiat, qui impliquerait des temps de latence très courts (Jakimik, 1980, cité par Spinelli et Ferrand, 2005). Pour cette raison, les mots sont courts. Ils comptent peu de phonèmes ( $M = 3,40$  ;  $DS = 0,49$ ) et peu de

---

syllabes ( $M = 1,40$  ;  $DS = 0,42$ ). Ils sont tous de structure phonologique CVC (unissyllabique), ou CVCV (bisyllabique).

Le contrôle de la structure s'est fait au niveau phonologique et orthographique. En effet, pour une même structure phonologique, il est possible d'avoir différentes structures orthographiques. Par exemple : « poule » a pour structure phonologique CVC, mais sa structure orthographique est CVVCV. De la même manière, « barre » a pour structure phonologique CVC et sa structure orthographique CVCCV.

La forme orthographique des mots a également été contrôlée. En effet, les mots sont présentés en modalité écrite sur l'écran de réponse de l'épreuve d'identification. Ils comptent en moyenne 5,02 lettres ( $DS = 0,81$ ) et ont une moyenne de 7,28 voisins orthographiques ( $DS = 5,50$ ). La rapidité de la lecture dépend de la structure orthographique des mots. On considère qu'autour de 5 lettres, un mot peut être lu en un seul mouvement oculaire par un normo-lecteur adulte sans trouble visuel (Legge, Mansfield, et Chung, 2001 ; Pelli, 2010). Nous avons respecté cette norme pour le choix de la structure orthographique des mots.

Nous savons également que la taille et la fréquence des mots influencent leur reconnaissance (« Neighborhood Activation Model », Luce et Pisoni, 1998). Les mots longs et ayant peu de voisins phonologiques sont plus aisément reconnus que les mots courts avec de nombreux voisins phonologiques. Les mots choisis sont tous courts et possèdent en moyenne 20,50 voisins phonologiques ( $DS = 7,65$ ). Ils ont tous au moins un voisin phonologique, utilisé pour construire la paire minimale.

Les mots ont été contrôlés en fréquence, de sorte que les deux éléments d'une même paire soient de fréquence similaire. Nous avons ainsi défini deux niveaux de fréquence : des mots de haute fréquence d'une part (fréquence supérieure à 20 apparitions par millions d'occurrences) et des mots de basse fréquence d'autre part (fréquence d'apparition comprise entre 1 et 15 apparitions par millions d'occurrences). La base de données Lexique3 intègre en effet les informations sur la fréquence d'apparition des mots à partir d'un large corpus oral (films) et écrit (livres). Une moyenne des fréquences films et livres a été faite pour déterminer le seuil. Les fréquences d'apparition film (orales) ont été privilégiées en cas de conflit, puisque le protocole expérimental porte sur la reconnaissance des mots parlés. Ainsi par exemple, des mots de haute fréquence (exemple : « poule » et « boule ») ont pu être appariés et des mots de basse fréquence (exemple : « mâche » et « bâche ») ont pu être appariés (Tableau II).

Nous avons limité les éléments de variabilité des occurrences en choisissant des noms singuliers ou invariables sans préférence de genre. Les homophones adjectivaux ou verbaux ont été éliminés. Les homophones nominaux ont quant à eux été pris en compte pour corriger la fréquence le cas échéant. Dans ce cas, l'occurrence la plus fréquente a été sélectionnée. Par exemple, « calot » et « galop » ont chacun un homophone nominal. Cet homophone nominal a été pris en compte dans le calcul de la fréquence d'apparition.

**Tableau II: Tableau représentant les critères de sélection des paires lexicales**

Les informations ont été extraites à partir de la banque de données Lexique3. Ce tableau représente les différentes informations que nous avons prises en compte dans notre choix de matériel. Nous avons sélectionné les mots en fonction de leur catégorie grammaticale, de leur genre, de leur nombre, de leur fréquence d'apparition dans la langue et de leur structure phonémique. Pour la fréquence d'apparition dans la langue nous avons pris en compte : la langue orale « films » et la langue écrite « livres ». Nous avons fait une moyenne de ces deux éléments. Nous avons ensuite calculé le niveau de fréquence de chaque mot : haute fréquence (fréquence supérieure à 20 apparitions par millions d'occurrences) et basse fréquence d'autre part (fréquence d'apparition comprise entre 1 et 15 apparitions par millions d'occurrences).

ortho- graphe	phoné- tique	informations grammaticales			fréquence d'apparition				structure orthographique	structure phonémique
		catégorie	genre	nombre	films	livres	moyenne	niveau de fréquence		
boule	bul	NOM	f	s	19.29	38.31	37.375	haute	cvcev	cvc
poule	pul	NOM	f	s	23.5	16.69	27.1775	haute	cvcev	cvc
bâche	baʃ	NOM	f	s	2.3	10.07	7.0375	basse	cvcev	cvc
mâche	maʃ	NOM	f	s	0.81	1.08	0.945	basse	cvcev	cvc

Parallèlement, un set de paires de pseudo-mots a été dérivé des paires minimales de mots de façon à préserver les phonèmes initiaux, la rime, le nombre de phonèmes, le nombre de syllabes et la structure syllabique (CVC ou CVCV). Par exemple, la paire de pseudo-mots CVC « bête-mêpe » a été dérivée de la paire de mots CVC « bêche-mèche »; « tadon-sadon » (CVCV) a été dérivée de « talon-salon » (CVCV). Nous avons réalisé ces échantillons de paires de mots et de paires de pseudo-mots parce que nous souhaitons comparer les stratégies utilisées pour percevoir les mots et celles utilisées pour percevoir les pseudo-mots. Le matériel comprend 40 paires minimales de mots et 40 paires minimales de pseudo-mots (Annexes III et V pour visualiser le matériel complet).

Un enregistrement audio des 80 mots et des 80 pseudo-mots a été effectué en studio d'enregistrement professionnel (Laboratoire de Mécanique et d'Acoustique, Marseille), par une voix féminine. Plusieurs échantillons de chaque item ont été enregistrés de façon à pouvoir sélectionner les meilleures réalisations. Nous avons découpé l'enregistrement à l'aide du logiciel Audacity (logiciel libre pour la manipulation de données audionumériques, <http://audacity.sourceforge.net/>) et chaque échantillon a été égalisé en temps. Les blancs initiaux ont aussi été égalisés. Chaque item dure ainsi 2 secondes avec un blanc initial de 70ms. La sélection des 160 items s'est ensuite faite à l'oreille dans un premier temps. Les échantillons trop rapides, trop lents et les « couacs » dans la voix ont été éliminés. Puis nous avons vérifié l'appariement des items à l'aide du logiciel Praat (logiciel libre pour la manipulation, le traitement et la synthèse de sons vocaux, <http://www.praat.org/>). Pour chaque item d'une paire, la durée générale de l'item, la durée de chacun des phonèmes, la courbe intonative, la position des formants et les délais de voisement étaient similaires, de sorte que les mots appariés diffèrent le moins possible au-delà du trait articulatoire que nous avons choisi de faire varier (Figure 6).

« Bonnet »

« Monnaie »

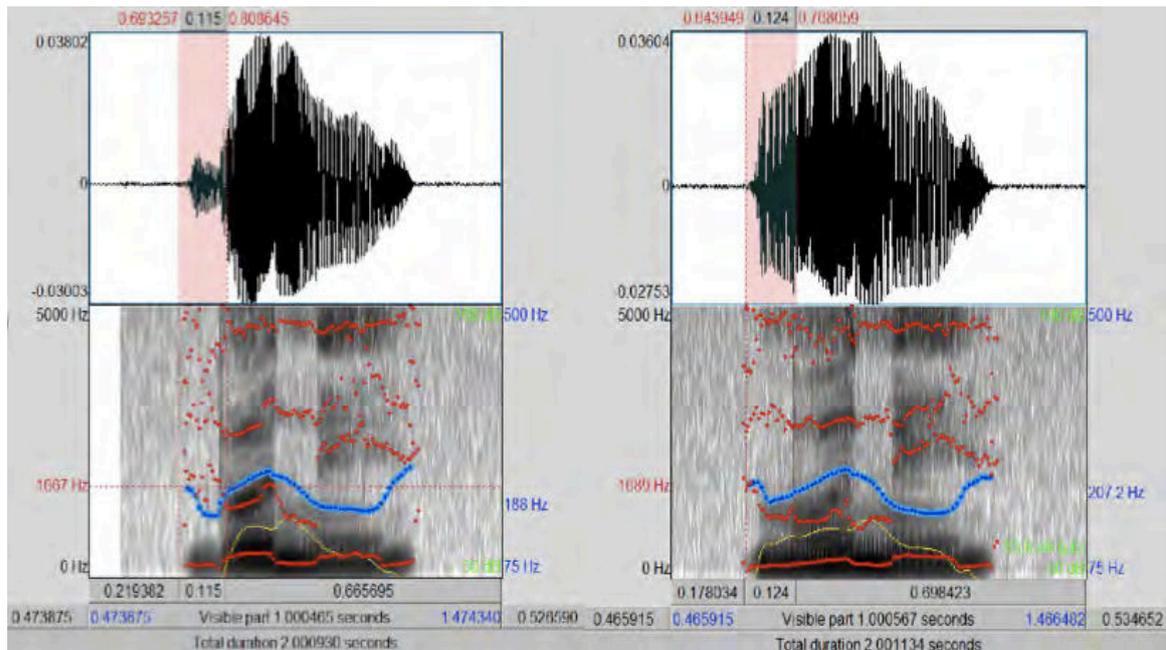


Figure 6 : Analyse spectrale d'une paire minimale avec le logiciel Praat

Cette représentation permet de visualiser les oscillogrammes et les spectrogrammes des mots « bonnet » (à gauche) et « monnaie » (à droite). Les oscillogrammes (amplitudes (Hz) en fonction du temps (s)) sont situés au-dessus des spectrogrammes (fréquence (Hz) en fonction du temps (s)). Sur les oscillogrammes, les bandes roses délimitent la représentation spectrale du premier phonème de chaque mot : /b/ pour « bonnet » et /m/ pour « monnaie ». Sur les spectrogrammes la courbe bleue représente la fréquence fondamentale, la courbe jaune représente la courbe d'intensité et les courbes rouges représentent les formants.

Nous avons contrôlé la longueur des items (2 secondes) et la longueur de chacun des phonèmes : ici /b/ (à gauche) et /m/ (à droite) apparaissent en rose, ils durent environ 0,12 secondes tous les deux. Nous nous sommes également assurées que la courbe de la fréquence fondamentale (points bleus) et les courbes d'intensité (ligne jaune), ainsi que la concordance des formants (points rouges) étaient similaires pour les deux items de la paire (Annexe I pour une analyse spectrale complète des paires pour chacun des traits consonantiques).

Les 160 items sélectionnés ont été égalisés en fréquence fondamentale et en intensité, puis ils ont été vocodés avec le logiciel Matlab (logiciel de calcul numérique permettant entre autres, la manipulation de matrices, <http://www.mathworks.fr/products/matlab/>). Deux modalités de vocodage ont été réalisées : les items ont été découpés en 16 bandes de fréquence de largeur équivalente, mais aussi en 8 bandes de fréquence de largeur équivalente (Annexe II); en ne conservant dans les deux cas que les informations d'enveloppe porteuse (de 2 à 50 Hz environ). Pour chaque bande, la fréquence a été réduite à une moyenne. Un « vocoder tone » a ensuite été utilisé : il ajoute des sons purs à l'échantillon et perturbe donc la perception de la hauteur de la voix et des formants. Enfin, un effet de recouvrement a été appliqué. Il permet de reproduire les chevauchements fréquentiels qui se produisent dans la périlymphe de la cochlée avec l'implant cochléaire.

---

Nous avons donc au total un matériel de 160 stimuli auditifs (mots et pseudo-mots), tous disponibles en trois modalités différentes : non vocodés, vocodés à 16 bandes et vocodés à 8 bandes. Le vocodage ainsi que la présentation des mots aux participants ont été programmés sous Matlab (psychtoolbox, boîte à outils permettant la programmation d'expériences).

### **III. Tâches**

Les tâches de discrimination et d'identification sont les tâches qui permettent de déterminer le degré de perception catégorielle selon la définition classique de Liberman et al., (1957).

Dans la tâche de discrimination, les participants sont soumis à deux stimuli auditifs présentés séquentiellement et doivent ensuite dire si les deux items entendus sont « pareils » ou « différents ». Ces deux stimuli auditifs sont soit les deux éléments d'une paire, soit deux fois le même élément d'une paire.

Dans la tâche d'identification, les participants sont soumis à un seul stimulus auditif. Ils doivent ensuite choisir parmi deux propositions écrites à l'écran laquelle a été entendue. Pour cette tâche les réponses sont données en modalité visuelle, le participant doit donc avoir une représentation orthographique du mot accessible rapidement.

Les scores de chacune de ces deux tâches peuvent être comparés, pour déterminer d'une part un indice de perception catégorielle (le rapport entre le score de discrimination et le score d'identification) et d'autre part un indice de précision catégorielle (la moyenne des scores aux tâches de discrimination et d'identification).

### **IV. Passation**

La passation a été réalisée à l'aide d'un ordinateur portable (Dell), muni du logiciel Matlab et d'un casque audio (Sennheiser). L'expérimentation s'est déroulée dans un environnement calme, pour qu'il n'y ait pas de bruits environnants qui puissent gêner les perceptions auditives des participants. Ainsi les différentes passations ont été réalisées selon les participants : au domicile des participants (40%), à notre domicile (18%), dans une salle de travail vide au sein de l'université (30%) ou dans des bibliothèques universitaires (12%).

L'expérimentation se décomposait en 4 ou 8 sessions pour une durée de 45 min ou 1h20-1h30. La différence de temps et du nombre de sessions à effectuer dépend du temps que les participants pouvaient nous accorder. Sur les 50 participants, 41 ont accepté de réaliser 4 tâches vocodées (identification de mots, identification de pseudo-mots, discrimination de mots, discrimination de pseudo-mots) et les 4 tâches correspondantes non vocodées. Cela correspond respectivement aux tâches expérimentales et contrôle. Les 9 autres candidats, faute de temps, n'ont réalisé que les 4 tâches expérimentales vocodées. L'ordre de présentation des tâches a été modifié à chaque passation. Ainsi aucun participant n'a effectué les tâches dans le même ordre. Cependant la partie vocodée a toujours été présentée avant la partie contrôle et l'ordre des tâches de la partie contrôle était le même que celui de la partie vocodée pour un même participant. Les participants

---

réalisaient les tâches expérimentales soit avec un niveau de vocodage de 16 bandes, soit avec un niveau de vocodage de 8 bandes. Aucun participant n'a entendu les deux niveaux de dégradation. Le niveau de dégradation proposé et l'ordre de présentation des tâches ont été déterminés par l'ordre de passage des candidats (Annexe VI pour l'ordre de présentation des épreuves pour tous les candidats).

Pour toutes les épreuves, les consignes ont été présentées à l'oral par l'examineur avant chaque tâche. Un énoncé a été préalablement établi pour cette présentation afin d'éviter toute variation entre expérimentateur (Annexe VII). De plus, avant de commencer chaque tâche, le participant se voyait rappeler les consignes sur l'écran d'ordinateur. Il pouvait les faire défiler au rythme souhaité en appuyant sur la barre d'espace du clavier.

Les participants étaient installés à une table, face à l'écran d'ordinateur. Nous leur expliquions les conditions générales de la passation (durée de la passation et nature des tâches), puis ils plaçaient le casque sur leur tête. La passation débutait par un réglage du volume sonore. Une tâche d'écoute d'un stimulus répété était proposée, afin de permettre aux participants de régler le volume sonore. Les participants étaient invités à choisir un niveau sonore qui corresponde à un volume d'écoute confortable. Ce volume sonore était enregistré et s'appliquait ensuite à l'ensemble des tâches proposées. Après le réglage du niveau sonore, l'expérimentation à proprement parler pouvait commencer. Chaque épreuve était lancée individuellement. Les consignes étaient rappelées avant chaque tâche.

Dans la tâche de discrimination, après avoir entendu les deux stimuli, le participant devait dire le plus rapidement possible si ces deux stimuli étaient identiques ou non. Pour cela, le participant devait appuyer sur les touches « a » ou « p » du clavier de l'ordinateur. En effet, pendant que la stimulation sonore était diffusée, les propositions « pareils » et « différents » s'affichaient de chaque côté de l'écran. Lorsque le candidat choisissait la réponse de gauche, il appuyait sur la touche « a » et s'il choisissait la réponse de droite il appuyait sur « p ». Les réponses pouvaient apparaître indifféremment à droite ou à gauche de l'écran. Le participant devait ainsi systématiquement lire les propositions sur l'écran. Cette randomisation a été réalisée dans le but de maintenir l'attention du candidat et pour éviter une habitude du geste de réponse.

Pour les épreuves d'identification, la passation se déroulait de la même façon avec le matériel mots ou pseudo-mots. Pour les tâches expérimentales, le son était dégradé selon le niveau de dégradation correspondant à l'ordre de passage du candidat (16 bandes ou 8 bandes). Pour les tâches contrôles, le son n'était pas dégradé. Après avoir entendu le stimulus, le participant devait déterminer le plus rapidement possible ce qu'il venait d'entendre en choisissant parmi deux propositions affichées sur l'écran de l'ordinateur. Pour cela, le participant devait appuyer sur les touches « a » ou « p » du clavier de l'ordinateur. En effet, pendant que la stimulation sonore était diffusée, deux propositions s'affichaient sur l'écran de l'ordinateur. Il s'agissait du stimulus que venait d'entendre le participant et de l'autre élément de la paire minimale. Par exemple pour les pseudo-mots, le candidat entendait « gapé » et devait choisir entre « gapé » et « kapé ». Pour les mots, le candidat entendait « cale » et devait choisir entre « cale » et « châte ». Les deux propositions étaient écrites de chaque côté de l'écran. Lorsque le candidat choisissait la réponse de gauche, il appuyait sur la touche « a » et s'il choisissait la réponse de droite il appuyait sur la touche « p ».

---

**Chapitre IV**  
**PRESENTATION DES RESULTATS**

---

## I. Préambule

Lors de notre expérimentation, nous avons mesuré deux types de variables:

- le pourcentage de réponses correctes correspondant au nombre de réponses correctes sur le nombre de réponses totales.
- le temps de réponse correspondant au temps qui s'écoule entre l'affichage de l'écran de réponse et la réponse du participant.

Pour chaque variable, une moyenne des scores ainsi que la déviation standard ont été calculées, afin de déterminer la distribution des résultats pour chacune des conditions testées. Les conditions étaient les suivantes :

- la condition mot et la condition pseudo-mots, qui permettent de mesurer l'effet de lexicalité
- la condition sans vocodage, la condition vocodage à 16 bandes de fréquence et la condition vocodage à 8 bandes de fréquence, qui permettent de mesurer l'effet de vocodage
- la condition discrimination et la condition identification, qui permettent de mesurer l'effet de la tâche

Une analyse de variance (ANOVA) a été appliquée afin de déterminer si la distribution des résultats dans les conditions définies différait de manière significative. Elle a été appliquée indépendamment pour le pourcentage de réponses correctes et pour les temps de réponse. Elle permet de mettre en évidence trois effets principaux :

- effet de lexicalité
- effet de vocodage
- effet de la tâche

Et des interactions entre ces trois effets principaux:

- l'interaction entre l'effet de lexicalité et l'effet de la tâche
- l'interaction entre l'effet de lexicalité et l'effet de vocodage
- l'interaction entre l'effet de vocodage et l'effet de la tâche

L'analyse de variance permet de mettre à l'épreuve l'hypothèse nulle : « l'écart à la moyenne est égal dans toutes les conditions définies », afin de déterminer avec quel degré de probabilité les résultats diffèrent de cette hypothèse nulle. A cette hypothèse nulle est opposée une hypothèse adverse : « l'écart à la moyenne est différent dans chacune des conditions ». Cette comparaison donne un indice de probabilité compris entre 0 et 1. On considère que plus l'indice de probabilité est proche de 0, plus l'hypothèse nulle est écartée, et plus les écarts à la moyenne sont significativement différents dans les différentes conditions. Au contraire, plus l'indice de probabilité est proche de 1, plus les résultats coïncident avec l'hypothèse nulle, et plus la distribution de ces résultats est semblable dans les différentes conditions. On considère que la variabilité due aux

---

différentes conditions est significative quand l'indice de probabilité est inférieur à 0.05 ( $p < .05$ ) (Howell, Yzerbyt et Bestgen, 2008). Pour tous les effets significatifs à plus de deux conditions, une comparaison planifiée avec correction de Boole-Bonferroni a été appliquée. La comparaison planifiée avec correction de Boole-Bonferroni permet en effet de corriger le seuil de significativité en cas de comparaisons multiples.

## II. Analyse pour les scores de réponses correctes

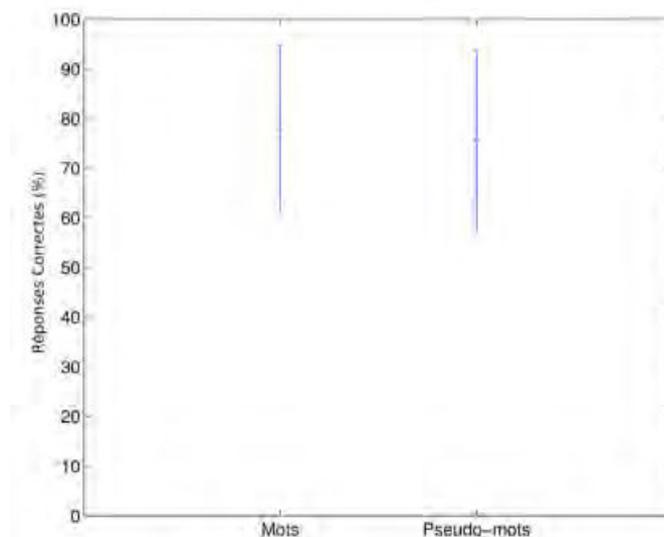
### 1. Résultats des effets principaux

#### 1.1. L'effet de lexicalité

Notre hypothèse 1 postule qu'il y a plus de réponses correctes et que le temps de réponse est moins long pour les tâches comprenant des mots par rapport à celles avec des pseudo-mots.

L'ANOVA montre un effet significatif de la lexicalité ( $F(1,26) = 8,03 ; p < .01$ ). Les résultats montrent qu'il y a plus de réponses correctes pour les tâches comprenant des mots ( $M = 77,75 ; DS = 16,88$ ) par rapport à celles avec des pseudo-mots ( $M = 75,45 ; DS = 18,29$ ) (Figure 7).

Notre hypothèse 1 est validée pour ce qui est du pourcentage de réponses correctes. Il y a plus de réponses correctes pour les tâches comprenant des mots par rapport à celles avec des pseudo-mots, aussi bien pour les tâches d'identification que pour celles de discrimination, qu'il y ait vocodage ou non et quel que soit le niveau de dégradation.



**Figure 7 : Représentation du pourcentage de réponses correctes (moyenne et déviation standard) dans les conditions mots et pseudo-mots**

**N.B :** Pour l'ensemble des graphiques de cette partie, la moyenne (M) est représentée par la marque au milieu des barres verticales. Les extrémités des barres correspondent aux déviations standard (DS).

---

## 1.2. L'effet de la dégradation du signal

Notre hypothèse 2 postule que:

- a) Il y a plus de réponses correctes et le temps de réponse est moins long pour les tâches avec la condition non vocodée que pour celles avec la condition vocodée (à 16 ou 8 bandes de fréquence).
- b) Il y a plus de réponses correctes et le temps de réponse est moins long pour les tâches avec la condition vocodée à 16 bandes de fréquence que pour celles avec la condition vocodée à 8 bandes de fréquence.

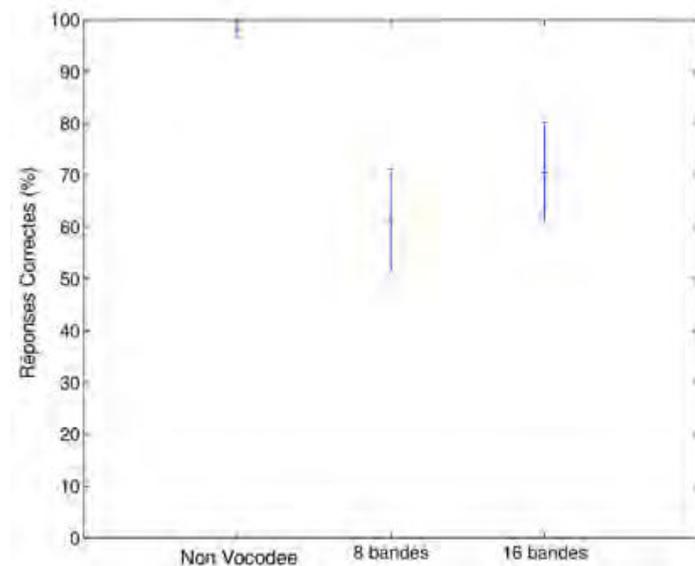
L'ANOVA met en évidence un effet significatif de la dégradation du signal ( $F(2,26) = 743,43$  ;  $p < .001$ ). Les trois conditions de dégradation (16 bandes de fréquence, 8 bandes de fréquence et aucun vocodage) ne sont pas traitées de la même façon. La distribution des résultats a été soumise à une comparaison planifiée avec correction de Boole-Bonferroni. Cette comparaison planifiée montre que les participants réussissent mieux les tâches non vocodées que celles avec un vocodage, qu'elles soient à 16 bandes ( $t(91) = 26,85$  ;  $p < .001$ ) ou à 8 bandes ( $t(91) = 36,27$  ;  $p < .001$ ). En revanche, la différence de traitement entre les deux niveaux de vocodage (16 et 8 bandes de fréquence) n'est pas significative ( $t(91) < 1$  ;  $p > .20$ ). Par conséquent, les participants réussissent mieux les tâches non vocodées que celles avec vocodage, mais la différence de traitement entre les deux niveaux de vocodage (16 et 8 bandes) n'est pas significative.

Les résultats montrent que (Figure 8) :

- a) Il y a plus de réponses correctes dans les tâches avec la condition non vocodée ( $M = 98,08$  ;  $DS = 1,46$ ) que pour celles avec la condition vocodée, qu'elles soient à 16 bandes de fréquence ( $M = 70,50$  ;  $DS = 9,69$ ) ou à 8 bandes de fréquence ( $M = 61,22$  ;  $DS = 9,88$ ).
- b) Il n'y a pas plus de réponses correctes pour les tâches avec la condition vocodée à 16 bandes de fréquence ( $M = 70,50$  ;  $DS = 9,69$ ) que pour celles avec la condition vocodée à 8 bandes de fréquence ( $M = 61,22$  ;  $DS = 9,88$ ).

L'hypothèse 2a est validée pour ce qui est du pourcentage de réponses correctes. Il y a plus de réponses correctes dans les tâches avec la condition non vocodée que pour celles avec la condition vocodée, qu'elles soient à 16 bandes de fréquence ou à 8 bandes de fréquence.

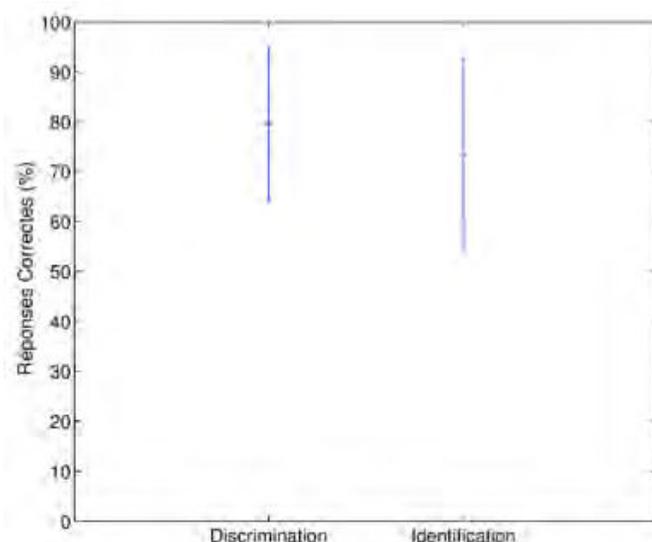
L'hypothèse 2b n'est pas validée pour ce qui est du pourcentage de réponses correctes. Il n'y a pas plus de réponses correctes pour les tâches avec la condition vocodée à 16 bandes de fréquence que pour celles avec la condition vocodée à 8 bandes de fréquence.



**Figure 8 : Représentation du pourcentage de réponses correctes (moyenne et déviation standard) dans les conditions non vocodée, vocodage à 8 bandes et vocodage à 16 bandes de fréquence**

### 1.3. L'effet du type de tâche

Il n'y a pas d'hypothèse pour cet effet. Nous utiliserons les résultats de réponses correctes en identification et en discrimination pour calculer les scores de perception catégorielle et de précision catégorielle dans les effets d'interaction. En effet, la perception catégorielle correspond au rapport entre le score de discrimination et le score d'identification. La précision catégorielle correspond à la moyenne des scores aux tâches de discrimination et d'identification. L'ANOVA montre un effet significatif du type de tâche ( $F(1,26) = 55,53 ; p < .001$ ). Les résultats montrent qu'il y a plus de réponses correctes pour les tâches de discrimination ( $M = 79,63 ; DS = 15,40$ ) que pour celles d'identification ( $M = 73,58 ; DS = 19,15$ ) (Figure 9). Les participants réussissent mieux les tâches de discrimination que celles d'identification.



**Figure 9 : Représentation du pourcentage de réponses correctes (moyenne et déviation standard) dans les conditions discrimination et identification**

---

## 2. Résultats des interactions

### 2.1. L'interaction entre la lexicalité et le type de tâche

Notre hypothèse 3 postule que :

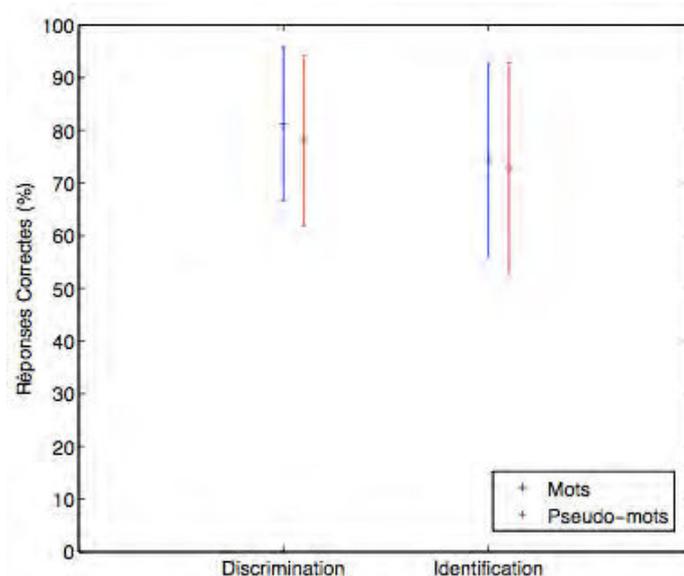
- a) Le score en perception catégorielle restera le même que les stimuli soient des mots ou des pseudo-mots.
- b) Le score en précision catégorielle sera meilleur si les stimuli sont des mots par rapport à des pseudo-mots.

L'ANOVA ne met pas en évidence d'effet significatif de l'implication du lexique sur le type de tâche ( $F(1,26) = 1,15 ; p > .20$ ). L'étude de l'interaction entre la lexicalité et le type de tâche est un indicateur de l'influence de la lexicalité sur la perception catégorielle. L'ANOVA montre qu'il n'y a pas d'interaction significative entre la lexicalité et la tâche : les informations lexicales n'influencent pas la perception catégorielle des phonèmes (Figure 10).

Notre hypothèse 3a concernant les scores de perception catégorielle est validée. Le score en perception catégorielle reste le même si les stimuli sont des mots ou des pseudo-mots.

Par ailleurs, nous avons procédé à une comparaison planifiée des résultats de mots et de pseudo-mots selon le type de tâches. Nous savons que la précision catégorielle est indiquée par la moyenne des réponses correctes en discrimination et en identification. La comparaison planifiée nous permet de déterminer le poids relatif du lexique dans la précision catégorielle des participants. Cette comparaison met en évidence un effet du lexique sur la précision catégorielle. Cet effet est significatif en discrimination ( $t(68) = 3,55 ; p < .01$ ) et en identification ( $t(68) = 6,10 ; p < .001$ ). La précision catégorielle est améliorée par les informations lexicales de manière significative dans les deux tâches. Les résultats pour les tâches de discrimination et d'identification avec les mots (respectivement  $M = 81,21 ; DS = 14,55$  et  $M = 74,29 ; DS = 18,39$ ) sont supérieurs aux résultats pour les tâches de discrimination et d'identification avec les pseudo-mots (respectivement  $M = 78,04 ; DS = 16,16$  et  $M = 72,86 ; DS = 19,99$ ).

L'hypothèse 3b est validée pour ce qui est de l'influence du lexique sur la précision catégorielle : le score en précision catégorielle est meilleur si les stimuli sont des mots par rapport à des pseudo-mots.



**Figure 10 : Représentation du pourcentage de réponses correctes (moyenne et déviation standard) pour les épreuves de discrimination et d'identification des mots et pseudo-mots**

## 2.2. L'interaction entre la lexicalité et la dégradation du signal

Notre hypothèse 4 postule que :

- Il y a plus de réponses correctes et le temps de réponse est moins long pour les mots par rapport aux pseudo-mots dans les conditions avec vocodage. L'effet de lexicalité est plus important quand le signal est dégradé, c'est-à-dire pour les conditions vocodées à 16 et 8 bandes par rapport à la condition non vocodée.
- Il y a plus de réponses correctes et le temps de réponse est moins long pour les mots que pour les pseudo-mots dans la condition la plus dégradée. L'effet de lexicalité augmente avec le niveau de dégradation du signal.

L'ANOVA met en évidence un effet d'interaction significatif entre la lexicalité et le niveau de dégradation : ( $F(2,26) = 3,68 ; p < .05$ ). Nous avons effectué une comparaison planifiée des résultats pour les mots et pseudo-mots selon les trois niveaux de dégradation. Cette comparaison met en évidence que les différences de résultats entre mots et pseudo-mots sont significatives pour le niveau de vocodage à 16 bandes de fréquence, ( $t(45) = 3,76 ; p < .001$ ) et à 8 bandes de fréquence ( $t(45) = 2,50 ; p < .05$ ), mais non significatives en l'absence de vocodage ( $t(45) < 1 ; p > .20$ ).

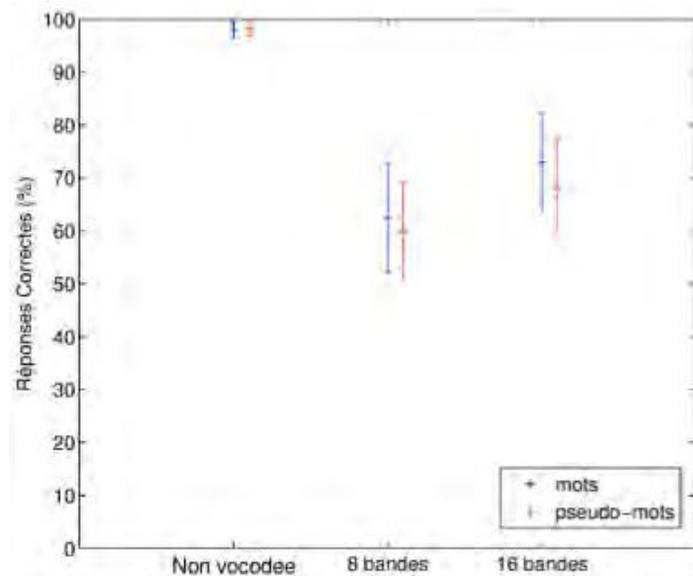
Les résultats montrent que (Figure 11) :

- Dans les conditions vocodage à 16 et à 8 bandes, il y a plus de réponses correctes pour les mots (respectivement  $M = 72,93 ; DS = 9,31$  et  $M = 62,50 ; DS = 10,28$ ) que pour les pseudo-mots (respectivement  $M = 68,07 ; DS = 9,54$  et  $M = 59,95 ; DS = 9,41$ ). En l'absence de vocodage, il n'y a pas plus de réponses correctes en mots ( $M = 97,82 ; DS = 1,33$ ) qu'en pseudo-mots ( $M = 98,34 ; DS = 1,56$ ).
- Dans la condition vocodage à 16 bandes, il y a plus de réponses correctes pour les mots ( $M = 72,93 ; DS = 9,31$ ) que pour les pseudo-mots ( $M = 68,07 ; DS = 9,54$ ).

- 
- c) Dans la condition vocodage à 8 bandes, il y a plus de réponses correctes pour les mots ( $M = 62,50$  ;  $DS = 10,28$ ) que pour les pseudo-mots ( $M = 59,95$  ;  $DS = 9,41$ ). La différence entre mots et pseudo-mots est équivalente dans les deux conditions.

L'hypothèse 4a est partiellement validée pour ce qui est du pourcentage de réponses correctes. Il y a plus de réponses correctes pour les mots par rapport aux pseudo-mots dans les conditions avec vocodage. En revanche, cet effet de lexicalité est absent lorsque le signal n'est pas vocodé.

L'hypothèse 4b n'est pas validée pour ce qui est du pourcentage de réponses correctes. Il n'y a pas plus de réponses correctes pour les mots par rapport aux pseudo-mots dans la condition la plus dégradée. L'effet de lexicalité n'augmente pas avec la dégradation du signal.



**Figure 11 : Représentation du pourcentage de réponses correctes (moyenne et déviation standard) pour les mots et les pseudo-mots en condition sans vocodage, avec vocodage à 8 bandes et avec vocodage à 16 bandes**

### 2.3. L'interaction entre la dégradation du signal et le type de tâche

L'hypothèse 5 postule que :

- La perception catégorielle est affectée par la dégradation du signal : le rapport entre les scores de discrimination et d'identification est meilleur dans la condition non vocodée par rapport aux conditions vocodées.
- La perception catégorielle est d'autant plus affectée que le niveau de dégradation est important : le rapport entre les scores de discrimination et d'identification est meilleur dans la condition vocodée à 16 bandes par rapport à la condition vocodée à 8 bandes.

L'ANOVA montre un effet significatif de l'implication de la dégradation sur le type de tâche ( $F(2,26) = 24,7$  ;  $p < .001$ ). Nous avons procédé à une comparaison planifiée avec correction de Boole-Bonferroni pour évaluer le niveau d'interaction relatif des différents degrés de dégradation sur la différence entre les résultats de discrimination et d'identification. Nous savons que la différence entre les scores de discrimination et d'identification nous donne un indice de perception catégorielle.

La comparaison des différences de scores dans les deux tâches pour la condition non vocodée et pour la condition de vocodage à 16 bandes de fréquence donne un résultat significatif ( $F(1,18) = 7,30$  ;  $p < .01$ ). De même, les différences de scores sont significatives lorsqu'on compare la condition non vocodée à la condition de vocodage à 8 bandes de fréquence ( $F(1,18) = 8,20$  ;  $p < .01$ ). En revanche la comparaison des deux modalités de vocodage entre elles ne donne pas de résultat significatif ( $F(1,18) < 1$  ;  $p > .20$ ). La perception catégorielle des participants n'est pas significativement différente dans les deux modalités de vocodage. Elle l'est en revanche dans les deux conditions de vocodage par rapport à la condition vocodée.

Les résultats montrent que (Figure 12) :

- La différence entre les résultats en identification (16 bandes :  $M = 68,78$  ;  $DS = 8,81$  et non vocodée  $M = 97,72$  ;  $DS = 1,43$ ) et les résultats en discrimination (16 bandes :  $M = 72,23$  ;  $DS = 10,30$  et non vocodée  $M = 68,21$  ;  $DS = 7,90$ ) n'est pas équivalente pour les conditions de vocodage à 16 bandes et sans vocodage. La perception catégorielle n'est pas équivalente dans les deux conditions.
- La différence entre les résultats en identification (8 bandes :  $M = 54,24$  ;  $DS = 6,03$  et non vocodée  $M = 97,72$  ;  $DS = 1,43$ ) et les résultats en discrimination (8 bandes :  $M = 68,21$  ;  $DS = 7,90$  et non vocodée  $M = 68,21$  ;  $DS = 7,90$ ) n'est pas équivalente pour les conditions de vocodage à 8 bandes et sans vocodage. La perception catégorielle n'est pas équivalente dans les deux conditions (Tableau III).
- La différence entre les résultats en identification (16 bandes :  $M = 68,78$  ;  $DS = 8,81$  et 8 bandes :  $M = 54,24$  ;  $DS = 6,03$ ) et les résultats en discrimination (16 bandes :  $M = 72,23$  ;  $DS = 10,30$  et 8 bandes :  $M = 68,21$  ;  $DS = 7,90$ ) est équivalente pour les conditions 16 et 8 bandes. La perception catégorielle est équivalente dans les deux conditions de vocodage.

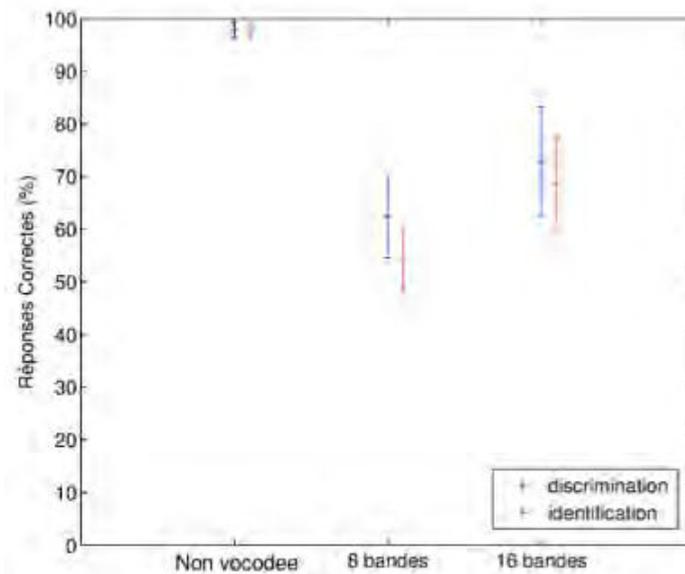
**Tableau III : Présentation des moyennes et déviation standard des pourcentages de réponses correctes pour les tâches de discrimination d'identification selon les différents niveaux de vocodage**

Epreuve	Niveau de vocodage	Moyenne	Déviatiion Standard
Discrimination	Aucun	98,45	1,42
	16 bandes de fréquence	72,23	10,30
	8 bandes de fréquence	68,21	7,90
Identification	Aucun	97,72	1,43
	16 bandes de fréquence	68,78	8,81
	8 bandes de fréquence	54,24	6,03

---

L'hypothèse 5a est validée : la perception catégorielle des participants est affectée par la dégradation du signal. Le rapport entre les scores de discrimination et d'identification est meilleur dans la condition non vocodée par rapport aux conditions vocodées.

L'hypothèse 5b n'est pas validée : la perception catégorielle des participants n'est pas d'autant plus affectée que le niveau de dégradation est important. Le rapport entre les scores de discrimination et d'identification est équivalent dans la condition vocodée à 16 bandes de fréquence et dans la condition vocodée à 8 bandes de fréquence.



**Figure 12 : Représentation des pourcentages de réponses correctes (moyenne et déviation standard) pour les tâches d'identification et de discrimination en condition sans vocodage, avec vocodage à 8 bandes et avec vocodage à 16 bandes**

### III. Analyse pour les scores de temps de réponse

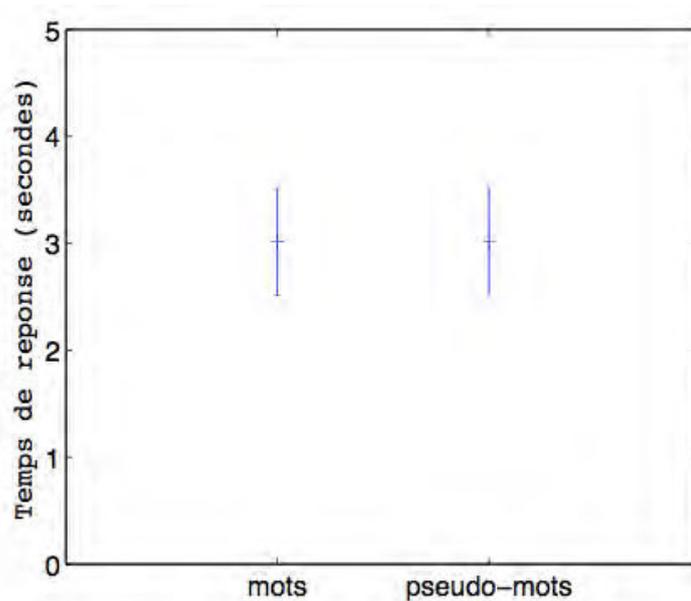
#### 1. Résultats des effets principaux

##### 1.1. L'effet de lexicalité

L'hypothèse 1 postule que le temps de réponse est moins long pour les tâches comprenant des mots par rapport à celles avec des pseudo-mots.

L'ANOVA ne met pas en évidence un effet significatif de la lexicalité sur les temps de réponse ( $F(1, 26) = 0,49$  ;  $p > .05$ ). Les résultats montrent que les participants n'ont pas été plus rapides pour traiter les mots ( $M = 3,07$  ;  $DS = 0,63$ ) par rapport aux pseudo-mots ( $M = 3,11$  ;  $DS = 0,60$ ) (Figure 13).

L'hypothèse 1 n'est pas validée pour ce qui concerne les temps de réponse : le temps de réponse n'est pas moins long pour les tâches comprenant des mots que pour celles avec des pseudo-mots.



**Figure 13 : Représentation des temps de réponse (moyenne et déviation standard) pour les conditions mots et pseudo-mots**

## 1.2. L'effet de la dégradation du signal

L'hypothèse 2 postule que :

- Le temps de réponse est moins long pour les tâches avec la condition non vocodée que pour celles avec la condition vocodée (à 16 ou 8 bandes de fréquence).
- Le temps de réponse est moins long pour les tâches dans la condition vocodée à 16 bandes de fréquence que pour celles dans la condition vocodée à 8 bandes de fréquence.

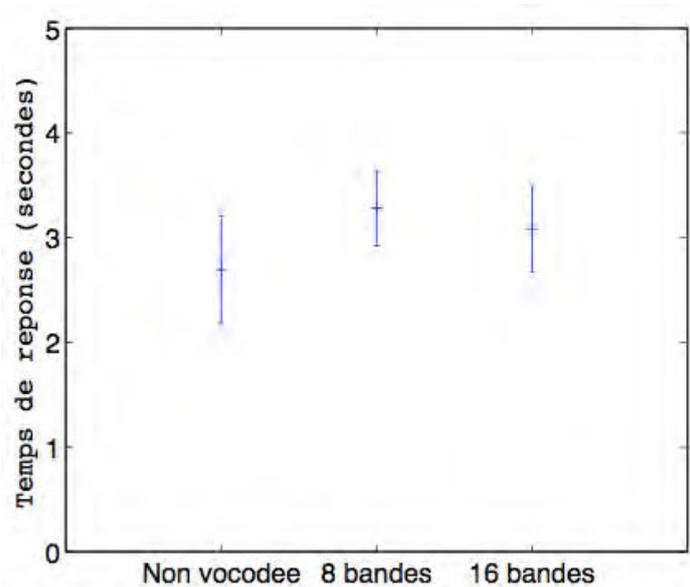
L'ANOVA met en évidence un effet de dégradation significatif ( $F(2,26) = 52,59 ; p < .001$ ). Les participants traitent différemment les tâches selon les niveaux de dégradation. Nous avons procédé à une comparaison planifiée des résultats pour les différents niveaux de dégradation deux à deux. Toutes les comparaisons donnent des résultats significatifs (Figure 14) :

- La modalité non vocodée comparée au vocodage à 16 bandes de fréquence ( $t(91) = 8,10 ; p < .001$ ) Les participants répondent significativement plus vite en l'absence de vocodage ( $M = 2,69 ; DS = 0,51$ ) qu'en présence de la condition à 16 bandes ( $M = 3,23 ; DS = 0,59$ )
- La modalité non vocodée comparée au vocodage à 8 bandes de fréquence ( $t(91) = 10,40 ; p < .001$ ). Les participants répondent significativement plus vite en l'absence de vocodage ( $M = 2,69 ; DS = 0,51$ ) qu'en présence de la condition vocodée à 8 bandes ( $M = 3,36 ; DS = 0,53$ )
- Les deux modalités de vocodage comparées entre elles ( $t(91) = 4,42 ; p < .001$ ). Les participants répondent plus vite dans la condition de dégradation à 16 bandes que dans la condition de dégradation à 8 bandes.

---

Plus le signal est dégradé et plus les participants sont lents. La différence entre les résultats dans les différentes modalités de dégradation est significative.

L'hypothèse 2a est validée pour ce qui est des temps de réponse : le temps de réponse est moins long pour les tâches sans vocodage par rapport à celles avec vocodage. L'hypothèse 2b est validée pour ce qui est des temps de réponse : le temps de réponse est moins long chez les participants ayant entendu la dégradation à 16 bandes de fréquence que chez ceux ayant entendu la dégradation à 8 bandes de fréquence.

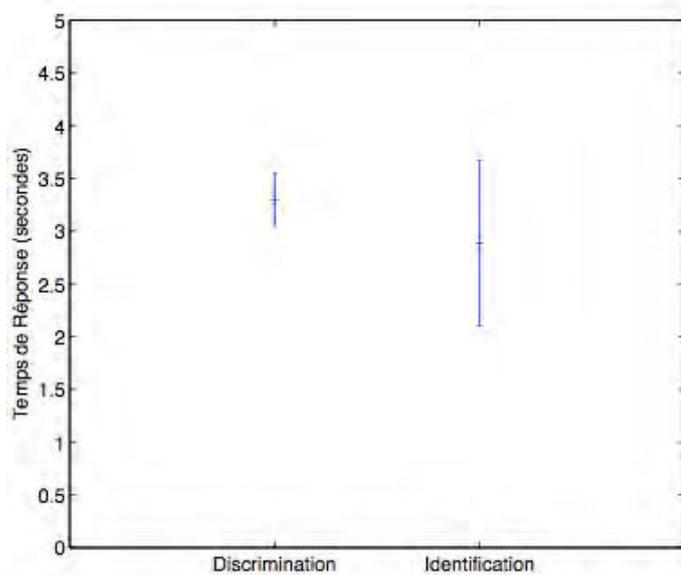


**Figure 14 : Représentation des temps de réponse (moyenne et déviation standard) pour les différents niveaux de dégradation : pas de vocodage, vocodage à 8 bandes de fréquence et vocodage à 16 bandes de fréquence**

### 1.3. L'effet du type de tâche

Nous n'avons pas fait d'hypothèse pour cet effet. Seuls les résultats des réponses correctes pour les tâches de discrimination et d'identification nous permettent de calculer les scores de perception catégorielle et de précision catégorielle.

Cependant nous notons que l'effet de la tâche est significatif ( $F(1,26) = 53,31 ; p < .001$ ). Les participants répondent significativement plus vite dans les tâches d'identification ( $M = 2,89 ; DS = 0,78$ ) que dans les tâches de discrimination ( $M = 3,30 ; DS = 0,25$ ) (Figure 15).



**Figure 15 : Représentation des temps de réponse (moyenne et déviation standard) pour les tâches de discrimination et d'identification**

## 2. Résultats des interactions

### 2.1. L'interaction entre la lexicalité et le type de tâche

Nous n'avons pas fait d'hypothèse pour cet effet. Seuls les résultats des réponses correctes pour les tâches de discrimination et d'identification nous permettent de calculer les scores de perception catégorielle et de précision catégorielle.

L'ANOVA ne met pas en évidence d'effet d'interaction entre la lexicalité et la tâche ( $F(1,26) = 0,54$  ;  $p > .05$ ). La lexicalité et le type de tâche n'interagissent pas sur les temps de réponse (Figure 16).

Les résultats montrent que les temps de réponse pour la condition discrimination sont similaires pour les mots ( $M = 3,30$  ;  $DS = 0,24$ ) et les pseudo-mots ( $M = 3,30$  ;  $DS = 0,26$ ). De la même manière, les temps de réponse pour la condition identification, sont similaires pour les pseudo-mots ( $M = 2,93$  ;  $DS = 0,77$ ) par rapport aux mots ( $M = 2,85$  ;  $DS = 0,80$ ).

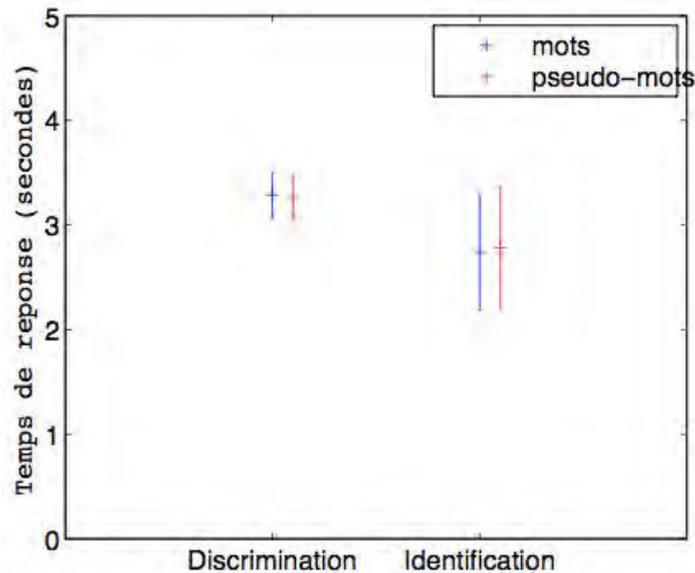


Figure 16 : Représentation des temps de réponse (moyenne et déviation standard) selon le type de tâche (discrimination ou identification) et le type de matériel (mots ou pseudo-mots)

## 2.2. L'interaction entre la lexicalité et la dégradation du signal

L'hypothèse 4 postule que :

- Le temps de réponse est moins long pour les mots par rapport aux pseudo-mots dans les conditions avec vocodage. L'effet de lexicalité est plus important quand le signal est dégradé, c'est-à-dire pour les conditions vocodées à 16 et 8 bandes par rapport à la condition non vocodée.
- Le temps de réponse est moins long pour les mots par rapport aux pseudo-mots dans la condition la plus dégradée. L'effet de lexicalité augmente avec le niveau de dégradation du signal, c'est-à-dire pour la condition vocodée à 8 bandes par rapport à la condition vocodée à 16 bandes.

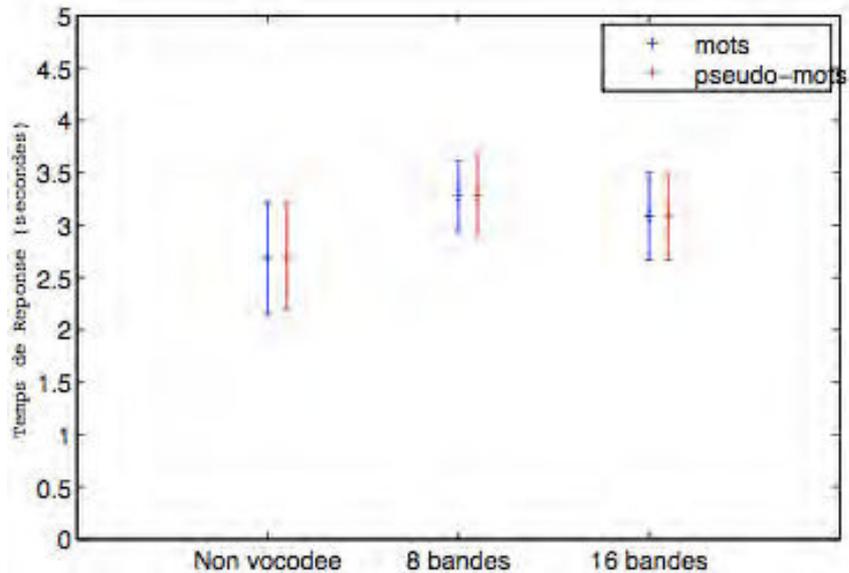
L'ANOVA ne montre pas d'effet d'interaction entre la lexicalité et le vocodage ( $F(2,266) = 0,35$  ;  $p > .05$ ). Les temps de réponse des participants ne sont pas significativement affectés par les différents niveaux de dégradation. Les résultats des moyennes et des déviations standard sont (Figure 17) :

- pour la condition vocodée à 16 bandes ( $M = 3,18$  ;  $DS = 0,57$ ) pour les mots et ( $M = 3,35$  ;  $DS = 0,61$ ) pour les pseudo-mots,
- pour la condition vocodée à 8 bandes ( $M = 3,36$  ;  $DS = 0,60$ ) pour les mots et ( $M = 3,29$  ;  $DS = 0,45$ ) pour les pseudo-mots,
- pour la condition non vocodée ( $M = 2,68$  ;  $DS = 0,53$ ) pour les mots et ( $M = 2,71$  ;  $DS = 0,50$ ) pour les pseudo-mots.

L'hypothèse 4a n'est pas validée pour ce qui est des temps de réponse. Les temps de réponse ne sont pas significativement moins longs pour les mots par rapport aux pseudo-mots dans les conditions avec vocodage par rapport à la condition sans vocodage. L'hypothèse 4b n'est pas validée : les temps de réponse ne sont pas significativement moins longs pour les mots par rapport aux pseudo-mots pour la condition vocodée à 16

---

bandes de fréquences par rapport à la condition vocodée à 8 bandes. L'effet de lexicalité n'augmente pas avec le niveau de dégradation du signal.



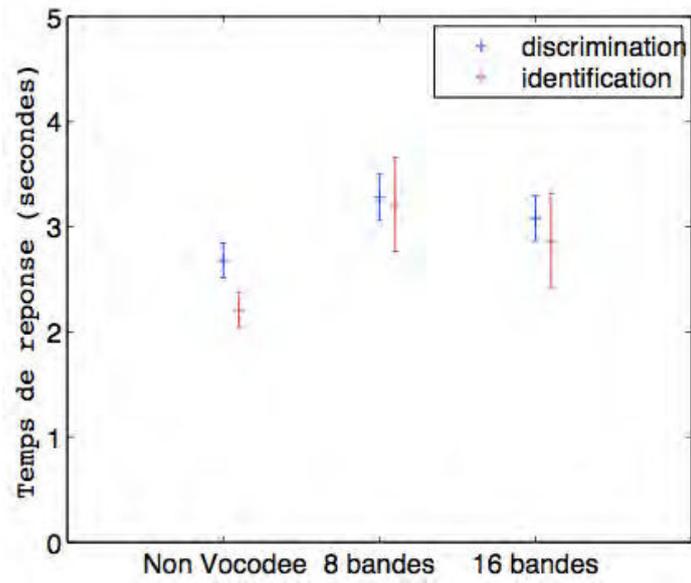
**Figure 17 : Représentation des temps de réponse (moyenne et déviation standard) selon les différents niveaux de dégradation (pas de vocodage, vocodage à 8 bandes de fréquence et vocodage à 16 bandes de fréquence) et le type de matériel (mots et pseudo-mots)**

### 2.3. L'interaction entre la dégradation du signal et le type de tâche

Nous n'avons pas fait d'hypothèse pour cet effet. Seuls les résultats des réponses correctes pour les tâches de discrimination et d'identification nous permettent de calculer les scores de perception catégorielle et de précision catégorielle.

L'ANOVA montre un effet significatif de l'implication de la dégradation sur le type de tâche ( $F(2,26) = 26,12 ; p < .001$ ). L'effet du type de tâche et l'effet de dégradation montrent une interaction significative. En effet, l'effet du type de tâche est significatif dans les conditions non vocodée ( $t(180) = 6,52 ; p < 0.05$ ) et vocodage 8 bandes ( $t(180) = 8,61 ; p < 0.05$ ) mais pas pour la condition vocodage 16 bandes ( $t(180) = 1,59 ; p = 0.11$ ).

Les résultats montrent que les temps de réponse en discrimination ( $M = 3,18 ; DS = 0,16$ ) sont supérieurs à ceux en identification ( $M = 2,21 ; DS = 0,17$ ) en l'absence de vocodage. De la même manière, les temps de réponse en discrimination ( $M = 3,37 ; DS = 0,25$ ) sont supérieurs à ceux en identification ( $M = 3,35 ; DS = 0,71$ ) dans la modalité de vocodage à 8 bandes. Par contre, les temps de réponse en discrimination ( $M = 3,35 ; DS = 0,28$ ) ne sont pas supérieurs aux temps de réponse en identification ( $M = 3,10 ; DS = 0,77$ ) dans la modalité de vocodage à 16 bandes (Figure 18).



**Figure 18 : Représentation des temps de réponse (moyenne et déviation standard) selon le niveau de dégradation (pas de vocodage, vocodage à 8 bandes de fréquence, vocodage à 16 bandes de fréquence) et le type de tâche (discrimination ou identification)**

---

**Chapitre V**  
**DISCUSSION DES RESULTATS**

---

## I. Préambule

L'objectif de notre travail était d'examiner l'influence du lexique sur la perception phonémique dans un contexte de signal dégradé. En effet, depuis les années 1970, les études psycholinguistiques se sont intéressées à l'influence des connaissances lexicales sur la perception de la parole. Cet effet a été largement exploré (Ganong, 1980 ; Warren, 1970 ; ou plus récemment Van Linden et al., 2007 ; Fort, Spinelli, Savariaux et Kandel, 2010). Parallèlement, les aptitudes de catégorisation des sons de la parole ont été mises en évidence (Serniclaes, 1987). Elles permettent une perception catégorielle des phonèmes (Liberman et al., 1957 ; Liberman et al., 1967), différenciée selon les langues (Hoonhorst et al., 2009a). Elles se mettent en place dans les premières années de la vie (Hoonhorst et al., 2009b) et évoluent ensuite avec l'exposition continue à la langue (Medina et al., 2010). Il n'y a pas de consensus sur les effets de ces différents niveaux de traitement les uns sur les autres (Norris et al., 2000 ; McClelland et al., 2006).

L'implant cochléaire permet à des sujets présentant une surdité sévère à profonde de percevoir toutes les fréquences entre 125 Hz et 8000 Hz pour des intensités comprises entre 30 et 40 dB. Les informations transmises par l'implant sont partielles, mais restent suffisantes pour analyser la parole (Loundon et Busquet, 2009). Contrairement aux informations transmises par une cochlée normale, l'implant ne fournit que les éléments portés par l'enveloppe temporelle et dégrade ceux donnés par la structure fine. Ainsi les éléments prosodiques, le timbre, mais aussi certaines caractéristiques articulatoires (comme la nasalité) ne sont pas communiquées. Face à ces difficultés sensorielles, les personnes implantées doivent compléter leurs capacités perceptives afin de décrypter la parole.

Nous supposons que la dégradation du signal par l'implant cochléaire pourrait avoir un impact sur ces niveaux de traitement dans la perception. Les informations lexicales prendraient d'autant plus d'importance quand le signal est dégradé (Kirk et al., 1995 ; Meyer et al., 2003). Nous nous sommes basés sur le modèle TRACE (McClelland et Elman, 1986) qui postule un effet des niveaux de traitement lexicaux sur la reconnaissance des unités phonémiques, et nous avons fait l'hypothèse d'une influence du lexique sur le traitement phonémique de la parole plus importante dans un contexte de signal dégradé.

Pour cela nous avons utilisé le vocodeur qui nous a permis de simuler le signal auditif transmis par l'implant cochléaire. Nous avons fait écouter à des adultes normo-entendants des paires minimales de mots et de pseudo-mots dégradées à l'aide du vocodeur. Nous souhaitons observer les stratégies mises en place par nos candidats face à cette dégradation du signal. Nous leur avons proposé deux tâches : la discrimination et l'identification pour les conditions mots et pseudo-mots. Ces deux types de tâches nous ont permis de calculer des scores en perception catégorielle et en précision catégorielle.

Nous avons aussi fait varier le niveau de dégradation. La moitié de notre population a entendu le matériel avec une dégradation à 16 bandes de fréquence et l'autre moitié avec une dégradation à 8 bandes de fréquence. De plus, 41 candidats sur 50 ont passé une seconde fois notre expérimentation sans aucun vocodage afin de constituer le groupe contrôle.

---

Nous avons mesuré deux variables : le pourcentage de réponses correctes et le temps de réponse de chaque candidat. Nous avons évalué pour ces deux variables : l'effet du lexique, l'effet de la dégradation du signal, l'effet de la tâche, l'effet de la lexicalité sur la dégradation du signal, l'effet de la lexicalité sur le type de tâche et l'effet de la dégradation du signal sur le type de tâche.

Les résultats ont été traités selon une analyse de variance (ANOVA). Pour les effets significatifs comptant plus de deux conditions, nous avons réalisé des comparaisons planifiées avec correction de Boole-Bonferroni.

Les résultats confirment l'existence d'un effet du lexique sur la perception phonémique des participants. Cet effet existe en identification comme en discrimination et il est équivalent dans les deux tâches. La lexicalité du matériel affecte la précision des participants, mais pas leur perception catégorielle, ni leurs temps de réponse. Cet effet lexical est présent dans le contexte dégradé et est équivalent pour les deux conditions de dégradation, mais il n'a pas été mis en évidence en situation d'écoute ordinaire (signal non dégradé). Les résultats mettent également en évidence un effet de la dégradation du signal. Le fait que le signal soit dégradé perturbe la perception, la précision catégorielle et les temps de réponse des participants. Elle conditionne par ailleurs l'apparition de l'effet de lexicalité. Cependant, le fait que le signal soit dégradé à 8 plutôt qu'à 16 bandes de fréquence n'affecte pas les traitements phonémiques, mais provoque des temps de réponse plus longs. Par ailleurs, les participants discriminent moins vite qu'ils n'identifient dans les conditions sans vocodage et vocodage à 8 bandes de fréquence, alors que dans la condition vocodage à 16 bandes il y a une égalisation des temps de réponse dans les deux tâches.

Nous allons à présent discuter l'ensemble des résultats obtenus, puis nous réfléchirons aux limites de notre étude, avant de voir les apports qu'elle peut présenter pour l'orthophonie.

## **II. Interprétation des résultats**

### **1. L'effet du lexique**

#### **1.1. L'influence du lexique**

L'hypothèse principale de cette étude était que la lexicalité avait un impact sur les capacités des participants à percevoir les phonèmes, à la fois sur le pourcentage de réponses correctes et sur la vitesse de réponse.

Contrairement à de nombreuses études mesurant l'effet de lexicalité sur la perception phonémique, nous n'avons pas cherché à établir de frontière phonémique et à montrer le glissement de cette frontière vers le candidat mot, dans un choix opposant le mot et le pseudo-mot d'une paire minimale (Ganong, 1980, Vanlinden et al., 2010). Toutes nos tâches ont mis les candidats face à une alternative de paires minimales composées de mots (mots/mots), ou face à une alternative de paires minimales composées de pseudo-mots (pseudo-mots/pseudo-mots). Nous avons pu déduire des résultats en discrimination

---

et en identification un score de perception catégorielle d'une part, et un score de précision catégorielle d'autre part. Ainsi, nous avons cherché à montrer une plus grande rapidité et une plus grande précision des réponses dans les choix mot/mot une préférence des participants pour les candidats mots, mais des temps plus courts et plus de réponses correctes pour les paires minimales composées de mots. Les résultats confirment l'existence d'un effet du lexique sur la perception phonémique des participants. La lexicalité du matériel affecte la précision des participants, mais pas leur perception catégorielle ni leurs temps de réponse.

## **1.2. L'influence du lexique sur la perception catégorielle et la précision catégorielle**

Nous avons également fait l'hypothèse que, dans ce paradigme expérimental, la perception catégorielle des mots ne différerait pas de celle des pseudo-mots, dans la lignée des résultats de Bouton, Serniclaes et al., 2012. Nous nous attendions aussi à des scores supérieurs pour la discrimination que pour l'identification. Mais la différence de scores entre les deux tâches devait rester équivalente pour les mots et pour les pseudo-mots.

Les résultats montrent en effet que la discrimination est toujours mieux réussie que l'identification pour ce qui est du pourcentage de réponses correctes. Cependant, les participants sont au contraire plus rapides pour effectuer les tâches d'identification.

Une explication possible consisterait à considérer la tâche de discrimination comme une tâche donnant à comparer deux à deux un ensemble d'indices acoustiques. Tandis que la tâche d'identification consisterait à détecter dans un ensemble d'indices acoustiques, l'indice pertinent pour la reconnaissance de l'item donné. Nous pouvons donc supposer que la discrimination est un processus plus long, mais donnant plus d'indices et plus de réponses correctes, tandis que l'identification s'apparenterait plutôt à un mécanisme de détection, plus court, mais donnant moins d'indices et moins de réponses correctes. Toutefois, l'intérêt de notre étude n'est pas dans la différence entre les scores de discrimination et d'identification. Elle se trouve dans le fait que cette différence est similaire pour les mots et pour les pseudo-mots. Elle indique ainsi que la perception catégorielle est affectée de la même manière lorsque les participants sont exposés à des mots et à des pseudo-mots. Bouton, Serniclaes et al., 2012, ont montré que la perception catégorielle de sujets implantés cochléaires était semblable à celle de sujets normo-entendants. Nous nous attendions donc à des résultats similaires.

Nous savons en effet que les catégories phonémiques se forment dans la première année de vie, sur des bases neurosensorielles universelles et sur la base du matériel langagier auquel le sujet est exposé (Hoonhorst et al. 2009b). Or, une fois ces catégories formées, elles entrent en jeu dans la perception de l'ensemble du matériel langagier entendu : que les mots soient réels ou fictifs, qu'ils appartiennent au système phonologique de la langue d'exposition ou non, et quelle que soit la tâche proposée. L'auditeur perçoit le matériel langagier sous le spectre du système phonologique de sa langue (Hoonhorst et al., 2009b). Ce spectre n'est pas figé, et nous savons que la précision catégorielle peut évoluer jusqu'à la fin de l'adolescence (Medina et al., 2010), mais nous savons aussi que des sujets monolingues qui acquièrent une seconde langue la perçoivent sur la base de leur système

---

phonologique d'origine, au moins dans un premier temps (Flege, MacKay et Meador, 1999).

Les participants de notre étude ont été sélectionnés pour avoir un stock lexical et phonémique mature. Par ailleurs, les pseudo-mots de notre étude ont été construits en respectant les règles phonotactiques du Français. Nous pensons donc qu'ils ont été traités sur la base du système phonologique Français par les participants, pour les tâches de discrimination comme pour celles d'identification. Ainsi, si les pseudo-mots ont été traités différemment des mots, cette différence de traitement est identique pour la discrimination et pour l'identification. C'est effectivement ce que nos résultats montrent. Les mots sont mieux traités. La précision catégorielle est donc meilleure pour les mots. En revanche, la différence de traitement est équivalente pour la discrimination et pour l'identification. Les tâches de discrimination et d'identification sont affectées de façon identique par le fait que le matériel soit non lexical. La perception catégorielle est similaire pour les mots et les pseudo-mots.

### **1.3. L'influence du lexique en présence de signal dégradé**

L'analyse statistique des résultats a montré qu'il y a plus de réponses correctes pour les mots que pour les pseudo-mots dans les conditions avec vocodage, mais que cet effet de lexicalité est absent lorsque le signal n'est pas vocodé. Les connaissances lexicales des participants leur permettent de mieux réussir les tâches contenant des mots en situation de signal dégradé. Cependant cette aide n'intervient que lorsque le signal est dégradé.

En l'absence de dégradation du signal, les participants ne sont pas plus en difficulté avec un matériel de pseudo-mots qu'avec un matériel de mots. Nous nous attendions à ce que l'effet de lexicalité soit majoré par la dégradation du signal, avec des résultats progressivement chutés face à l'augmentation de la dégradation. Nos résultats suggèrent plutôt un traitement en « tout ou rien ». Lorsque le signal n'est pas dégradé, l'effet de lexicalité est absent. Lorsqu'il est dégradé, l'effet de lexicalité apparaît, et il reste équivalent quel que soit le niveau de dégradation.

Ce résultat confirme l'hypothèse d'un effet du lexique que Bouton, Colé et al. (2012) avaient mis à l'épreuve chez un public d'enfants porteurs d'implants cochléaires. Ils avaient pu mettre en évidence un effet du lexique chez les implantés cochléaires, mais celui-ci était équivalent à celui des normo-entendants, évacuant l'hypothèse d'une compensation possible par le lexique, mais montrant au contraire des mécanismes de reconnaissance similaires chez les normo-entendants et les implantés cochléaires. La configuration de notre expérience diffère de celle-ci par le fait que d'une part le stock lexical de nos participants s'est construit sur la base d'un signal intact, et que d'autre part il est mature. Cela entraîne deux suppositions pour expliquer les résultats de la précédente étude :

- Nous pouvons supposer que la construction d'un stock lexical fiable permet une compensation par le lexique impossible lorsque celui-ci s'est construit sur la base d'un signal dégradé,
- Nous pouvons également nous demander si le stock lexical d'enfants implantés cochléaires âgés de 7;11 ans à 11;6 ans est suffisamment mature pour mettre en évidence un tel effet de compensation.

---

Notre étude sur des participants adultes donne des résultats en faveur d'un effet de compensation : les participants mettraient en jeu des mécanismes de perception différents lorsqu'ils sont soumis à un signal dégradé et lorsqu'ils sont soumis à un signal non dégradé. Ces mécanismes comporteraient des activations « top-down » compensatrices permettant de pallier les manques du message auditif dégradé. Soumis à un message non dégradé, les participants mettent en œuvre des stratégies dont certaines seraient basées sur les mécanismes « bottom-up », similaires pour les mots et les pseudo-mots. Cette analyse n'exclut pas que des mécanismes « top-down » entrent également en jeu dans la perception de la parole en condition d'écoute normale, mais ils n'ont pas été mis en évidence dans notre étude. Ces résultats sont en accord avec le modèle TRACE (McClelland et Elman, 1986) et les données comportementales et neurologiques récentes (Davis et Johnsrude, 2007 ; Davis, Johnsrude, Hervais-Adelman, 2005 ; Vanlinden et al., 2007 ; voir Mattys et al., 2012 pour une revue de la littérature à ce sujet), qui objectivent l'existence de mécanismes compensatoires liés aux connaissances lexicales lorsque le locuteur est soumis à des conditions de perception défavorables.

#### **1.4. L'influence du lexique selon le niveau de dégradation du signal**

L'analyse statistique a montré que la différence entre le pourcentage de réponses correctes pour les mots et les pseudo-mots n'augmentait pas dans la condition vocodée à 8 bandes de fréquence par rapport à la condition vocodée à 16 bandes de fréquence. Par conséquent, l'aide apportée par les connaissances lexicales en situation de signal dégradé n'augmente pas avec la dégradation du signal. L'effet de lexicalité reste le même quelle que soit l'importance de la dégradation. Ainsi les connaissances lexicales sont utilisées en situation de signal dégradé, mais cette utilisation n'augmente pas en efficacité avec le niveau de dégradation.

Nous pouvons nous demander s'il n'y a pas un niveau « seuil » de dégradation au-delà duquel la perception n'est plus possible comme dans les conditions « normales ». Ce résultat évoque en effet une compensation par le lexique qui fonctionnerait en « tout ou rien ». Lorsque le signal est dégradé, il semble que des mécanismes de compensation se mettent en jeu avec des effets « top-down ». Lorsqu'il ne l'est pas, ce serait un autre type de mécanismes qui serait en jeu, plus axé sur les effets « bottom-up ». Ce résultat est cohérent avec un argument donné par Davis et Johnsrude (2007). Ils mettent en parallèle les résultats de perception face à deux types de dégradation très différents : le signal sans enveloppe (« sine-wave » en anglais), qui garde des mouvements de haute fréquence sur les formants sans en garder l'enveloppe, et le signal vocodé avec du bruit, qui garde au contraire des informations de basse fréquence, auxquelles sont superposées du bruit blanc. Dans les deux cas, les auditeurs arrivent après une exposition supérieure à une demi-heure à des résultats de perception similaires. Ainsi, dans le cadre d'un mécanisme de compensation, la qualité des connaissances préalables de l'auditeur entrerait en jeu dans la perception de la parole d'une façon importante. Nos résultats à 16 et 8 bandes de fréquence semblent donner des arguments en ce sens. Si nous considérons que des mécanismes de compensation se mettent en place sur la base d'un savoir préalable, nous pouvons penser que la qualité de ce savoir préalable est déterminante dans la performance de l'auditeur. Pour un niveau de savoir préalable similaire, les performances semblent similaires. Pour un niveau lexical similaire, nous faisons l'hypothèse que nos participants ont mis en place des mécanismes de compensation similaires et obtenu des performances similaires face aux différents niveaux de dégradation auxquels ils ont été soumis. Au vu

---

des résultats de Bouton, Colé et al., 2012 nous pouvons nous questionner sur le niveau de connaissance « minimal » à avoir pour pouvoir mettre en place ces mécanismes de compensations, et sur la quantité et la qualité « minimale » d'informations acoustiques nécessaires pour que ces mécanismes puissent se mettre en place.

### **1.5. L'influence du lexique sur les temps de réponse**

L'analyse des résultats ne montre pas que les participants soient plus rapides pour traiter les mots par rapport aux pseudo-mots. Pour ce qui est du temps de réponse, les participants réagissent de la même manière pour traiter les mots et les pseudo-mots. La différence de temps de traitement entre mots et pseudo-mots n'est pas non plus affectée par la dégradation du signal, ni par le type de tâche.

Cutler et al. (1987) ont montré un effet de supériorité du mot (« word superiority effect » en anglais). Ils ont en effet mesuré des temps de réponse plus courts pour détecter un phonème dans un mot que dans un pseudo-mot. Ce protocole a été repris par Fort et al. en 2010 avec des résultats similaires. Dans la lignée de ces études, nous nous attendions à obtenir des temps de réponse plus courts pour les mots dans les tâches d'identification et de discrimination proposées aux participants. Or, nous n'observons pas cet effet. Il est à noter que dans les expériences citées, le paradigme expérimental mettait l'auditeur face à une tâche de détection : une fois exposé à un mot ou un pseudo-mot, il devait dire le plus rapidement possible si ce mot ou pseudo-mot contenait ou non un phonème-cible. Notre expérience ne demandait pas de détecter, mais d'estimer si oui ou non des paires de mots ou de pseudo-mots étaient identiques (discrimination) et de choisir parmi une alternative de deux mots ou de deux pseudo-mots l'item entendu (identification). Face à un choix, nous pouvons supposer que les participants se sont concentrés davantage sur l'exactitude de leur réponse, la présence d'une réponse adverse impliquant des mécanismes de comparaison absents dans une tâche de détection.

### **1.6. Conclusion sur l'influence du lexique**

Si les participants n'ont pas été plus rapides dans les tâches comprenant des mots, leurs performances en précision ont toutefois été augmentées par rapport aux tâches comprenant des pseudo-mots. Nous observons également avec une analyse par niveau de dégradation que cet effet de lexicalité n'apparaît en fait que dans la condition dégradée, ce qui argumente en faveur d'un effet de compensation, avec deux mécanismes de perception bien distincts : un mécanisme ascendant (« bottom-up ») lorsque les conditions auditives sont optimales et un mécanisme descendant (« top-down ») lorsque le signal est dégradé. Ce mécanisme descendant de compensation serait largement dépendant des connaissances lexicales de l'auditeur.

---

## **2. L'effet de dégradation du signal**

### **2.1. L'influence de la dégradation sur le pourcentage de réponses correctes**

L'analyse statistique des résultats a montré que les participants réussissaient mieux les tâches avec la condition non vocodée que celles avec la condition vocodée, qu'elles soient à 16 bandes de fréquence ou à 8 bandes de fréquence. Les candidats sont plus performants lorsque le signal n'est pas perturbé. Cependant ils ne réussissent pas mieux les tâches présentées dans la condition la moins dégradée (16 bandes de fréquence). Le niveau de dégradation n'influence pas significativement les compétences des participants. Ce résultat démontre que la dégradation du signal perturbe les performances des participants, mais cette perturbation ne s'accroît pas avec le niveau de dégradation.

L'implant cochléaire réalise un filtrage du message auditif et ne transmet que les informations correspondant à l'enveloppe sonore. Ces informations permettent la compréhension des données principales de la parole, mais les éléments prosodiques et le timbre ne sont pas transmis. De plus, la transmission des informations acoustiques pour les différents phonèmes de la langue française est inégale. En effet, les caractéristiques du voisement et du mode articulaire sont portées par l'enveloppe sonore alors que les indices de la nasalité et du lieu sont transmis par la structure fine. Ainsi les sujets implantés perçoivent mieux les traits de voisement et de lieu pour les consonnes. Les différences de perception des traits articulatoires vocaliques sont liées à une variation de discrimination des formants et au codage de ces fréquences pour les électrodes (Gallégo et Collet, 2008 ; Gnansia, 2009 ; Micheyl, 2005). Les informations apportées par le signal vocodé sont parcellaires. Nous nous attendions donc à obtenir de moindres résultats dans les conditions vocodées que dans la condition non vocodée.

Nous nous attendions toutefois, comme Shannon, Zeng, Kamath, Wygonski et Ekelid, (1995), à des performances croissantes avec l'augmentation du nombre de bandes de fréquence. Cet effet n'est pas observé sur la précision des réponses. Il est difficile de déterminer pourquoi, mais nous pouvons donner des hypothèses de réponses. La première hypothèse supposerait l'intervention de mécanismes « top-down ». Les compensations prendraient le pas sur l'analyse acoustique du signal et le niveau de dégradation aurait alors une moindre importance. Une seconde explication peut être trouvée dans la qualité des informations transmises : pour une même stratégie de codage, le nombre de bandes de fréquences n'affecte pas le type d'informations acoustiques transmises. Une étude des performances des participants en fonction des indices acoustiques transmis et du nombre de bandes de fréquence permettrait de déterminer la part des mécanismes « bottom-up » impliqués dans la perception et de qualifier les éléments acoustiques donnés par chacun des niveaux de dégradation. Une troisième explication se baserait sur le fait que le protocole de l'étude de Shannon et al. 1995, comprenait un entraînement des participants au signal dégradé. Les participants auraient eu le temps de se familiariser avec le signal dégradé et ils auraient ainsi amélioré leurs performances pour la condition la moins dégradée par rapport à celle avec un niveau de dégradation supérieur.

---

## 2.2. L'influence de la dégradation sur les temps de réponse

L'analyse des temps de réponse montre que les participants sont plus rapides pour les tâches présentées dans la condition non vocodée. Ils sont aussi plus rapides pour les tâches présentées dans la condition vocodée à 16 bandes par rapport à la condition vocodée à 8 bandes. La présence d'une dégradation du signal ralentit le traitement des candidats et ce ralentissement augmente avec le niveau de dégradation. Ce résultat est cohérent avec des données récentes de la littérature qui explorent les mécanismes d'apprentissage. Davis et Gaskell (2009) proposent un modèle d'apprentissage de nouveaux mots basés sur deux étapes distinctes :

- une première étape rapide, avec des activations de l'hippocampe, qui semble correspondre à un mécanisme de perception immédiat « bottom-up ».
- une seconde étape de renforcement plus lente, liée à des activations corticales a posteriori, qui semblent correspondre à des mécanismes de rétroaction « top-down ».

Il aurait été intéressant de pouvoir mesurer l'activation cérébrale de nos participants, afin de voir si les temps de réponse plus longs observés correspondent à des mécanismes cérébraux distincts. Il est impossible de conclure à deux mécanismes distincts seulement à partir de données comportementales. Nous nous demandons toutefois si les temps de réponse accrus ne correspondent pas à des mécanismes d'apprentissages impliquant des processus « top-down ». Cependant, nous ne pouvons pas exclure des facteurs extérieurs non contrôlables comme les variations du niveau d'attention, la fatigabilité et les stratégies privilégiées par les participants (le fait de favoriser l'exactitude des réponses au détriment de la rapidité).

## 2.3. L'influence de la dégradation du signal sur la perception catégorielle

L'analyse statistique des résultats a montré que la perception catégorielle des participants est affectée par la dégradation du signal. Le rapport entre les scores de discrimination et d'identification est meilleur dans la condition non vocodée par rapport aux conditions vocodées. La dégradation du signal affecte les capacités de perception phonémique. Cependant, que le vocodage soit à 16 ou à 8 bandes de fréquence, la perception catégorielle est affectée de manière identique. Ce résultat n'est pas cohérent avec l'étude de Bouton, Colé et al., 2012, qui montrait que des enfants implantés avaient des performances en perception catégorielle similaires à celles d'enfants normo-entendants appariés en âge auditif. Nous savons que les sujets implantés de l'étude de Bouton, Colé et al., 2012 ont toujours été exposés à un signal dégradé. Leur perception phonémique s'est alors construite sur la base de ce signal. Nos participants au contraire, ont construit leurs catégories phonémiques sur la base d'un signal non dégradé. Nous pouvons supposer que l'exposition à un signal dégradé perturbe les repères pris par les sujets normo-entendants et qu'il faille une exposition prolongée à ce signal pour développer une perception catégorielle cohérente. Nous savons que le développement de la perception catégorielle dépend du repérage d'un certain nombre d'invariants par les sujets dans un matériel langagier donné (Hoonhorst et al., 2009b). Aussi, face à un matériel langagier différent (langue étrangère), les sujets ne réussissent à percevoir les indices pertinents dans la langue nouvelle qu'après une exposition prolongée (Flege et al., 1999). Ainsi nous

---

pouvons supposer que les adultes soumis au signal dégradé réagissent comme s'ils étaient face à une langue inconnue et qu'ils sont incapables d'en retirer les invariants nécessaires au repérage des différents phonèmes. Cependant, nous pouvons faire l'hypothèse qu'avec une exposition prolongée à ce signal, les sujets normo-entendants pourraient peut-être finir par atteindre des scores de perception similaires à ceux de sujets implantés ayant toujours perçu un signal dégradé.

### **III. Critiques et limites de notre expérimentation**

#### **1. Pour notre population**

Notre population est quasiment composée exclusivement d'étudiants, plus particulièrement d'étudiants en orthophonie. Cette caractéristique nous permet de faire l'hypothèse d'un groupe ayant un niveau de connaissances lexicales homogène. Cependant, cette composition n'est pas réellement représentative de la population francophone. Les stratégies mises en place par les candidats sont-elles alors généralisables ? La réalisation du même protocole expérimental avec une population d'adultes normo-entendants plus hétérogène serait sans doute plus compatible avec la réalité observée chez les sujets implantés.

#### **2. Pour notre passation**

Le temps d'expérimentation était long (environ 1h15). Le coup cognitif des épreuves pour les candidats est une donnée non négligeable dans l'interprétation des résultats. Il était important d'encourager les sujets à rester vigilants et rapides. Les résultats ne montrent pas d'effet significatif des temps de réponse. Nous pouvons en conclure que les participants ont favorisé la qualité de leurs réponses à la rapidité.

L'expérimentation devait se dérouler dans un environnement calme afin de ne pas perturber l'écoute des participants. Afin de présenter des résultats parfaitement comparables, il serait possible de réaliser les passations dans un lieu insonorisé. Nous avons réalisé les passations dans des endroits silencieux, comme des salles de travail à l'université ou dans des bibliothèques universitaires, ainsi qu'au domicile des participants ou à notre domicile. Ces différents lieux apportent une certaine variabilité à nos conditions de relevés.

Nous avons randomisé l'ordre de présentation des différentes tâches de notre expérimentation pour que les résultats des premières tâches ne soient pas perturbés par un effet de nouveauté et que les dernières tâches ne soient pas systématiquement mieux réussies grâce à un effet d'apprentissage. Les candidats ont en effet été surpris lors de leur première écoute d'un signal dégradé, mais ils nous ont fait aussi remarquer qu'après deux tâches, ils parvenaient à mieux percevoir les stimuli. Ils semblent qu'ils se soient adaptés au signal dégradé et qu'ils soient parvenus à améliorer leurs capacités de traitement face à la dégradation. Il serait intéressant de vérifier statistiquement cette impression. Le fait que des sujets normo-entendants sans antécédent de trouble auditif arrivent à améliorer leur performance en situation de signal dégradé au bout d'un temps d'exposition de 45 minutes - 1 heure est très intéressant pour la clinique. Nous pouvons alors faire un

---

parallèle avec les sujets sourds nouvellement implantés et l'importance du temps d'utilisation de l'appareil. En effet, en clinique, nous observons souvent que les patients nouvellement implantés sont gênés dans un premier temps par les informations qu'ils perçoivent avec l'implant. En premier lieu, il est nécessaire de les encourager à porter leur implant seul (sans la prothèse contro-latérale), à certains moments de la journée, afin qu'ils affinent leurs capacités de traitement du signal. Si un jour il est possible, grâce au vocodeur, de modéliser avec précision les perceptions que le patient recevra avec son implant, nous pouvons imaginer que des simulations d'écoute permettraient au patient tout au moins de se faire une idée du signal auquel il va être soumis et pourquoi pas de s'entraîner à sa reconnaissance avant son implantation.

### **3. Pour l'analyse des données**

Dans la continuité de notre étude, une analyse des résultats (pourcentage de réponses correctes et temps de réponse) par trait articulatoire pourrait être réalisée, afin d'apporter des informations sur le traitement des différents composants acoustiques de la parole. Lors de la constitution de notre matériel nous avons choisi de créer des paires minimales à partir des quatre caractéristiques acoustiques des consonnes (le voisement, la nasalité, le lieu articulatoire et le mode articulatoire). Il serait intéressant de comparer les résultats pour les différents traits acoustiques, car l'implant cochléaire ne transmet pas de la même manière les informations acoustiques. L'aide apportée par les connaissances lexicales en fonction de la dégradation de différentes informations acoustiques serait une donnée importante à développer. Il serait intéressant de croiser des données quantitatives issues des résultats pour les différents traits acoustiques avec les informations qualitatives relevées auprès des candidats. De nombreux candidats nous ont fait part de leur ressenti sur le matériel de notre expérimentation : ils avaient le sentiment de mieux percevoir les différences avec le matériel composé de premiers phonèmes marqués par le trait de voisement que pour le matériel caractérisé par le trait de nasalité. Ces réactions corroborent les résultats obtenus par certaines études. En effet, d'après l'étude de Medina et al., 2009, pour les voyelles, la nasalité est moins bien perçue que l'aperture, le lieu ou la labialité. Pour les consonnes, le déficit touche le lieu articulatoire et la nasalité plutôt que le mode ou le voisement. D'autres études concluent à des résultats sensiblement différents, en particulier pour les voyelles (Tyler, Preece, Lansing et Gantz (1992), Fitzgerald et al. (2007), Tye-Murray, Spencer et Gilbert-Bedia (1995), cités par Bouton, Serniclaes et al. (2012)). Il est difficile de déterminer la part du signal auditif qui entre en jeu dans la perception de chacun des traits. Cela est d'autant plus difficile que les indices acoustiques ne sont pas uniques et se chevauchent dans le temps. Cependant, l'étude des informations acoustiques impliquées dans la perception des sons de parole est nécessaire à l'amélioration des traitements effectués par l'implant cochléaire.

De la même manière, nous avons sélectionné nos paires de mots en fonction de leur fréquence d'apparition dans la langue française et nous avons comptabilisé le nombre d'homophones nominaux. Une autre piste d'extension de notre étude serait de calculer l'importance de l'aide lexicale en fonction du niveau de fréquence d'apparition des mots dans la langue et du nombre de voisins lexicaux. Nous savons en effet que les manipulations des indices acoustiques peuvent affecter l'activation lexicale (McMurray et al., 2003 ; Andruski, Blumstein et Burton, 1994). Ces auteurs postulent que les manipulations des indices acoustiques entraînent l'activation de voisins phonologiques qui peut perturber l'activation de la cible. Nous pouvons donc nous demander si dans

---

notre protocole expérimental, les activations lexicales des voisins phonologiques n'ont pas été déclenchées avec la dégradation du signal, et ce pour les mots comme pour les pseudo-mots. Il serait intéressant de pondérer les effets potentiels de ces activations adverses avec un calcul de probabilité en fonction de la fréquence d'apparition des paires de mots sélectionnées et du nombre de voisins phonologiques, dans la lignée du « Neighborhood Activation Model » de Luce et Pisoni (1998). Ce modèle présente en effet un intérêt pour la compréhension des mécanismes de perception de la parole dans des conditions dégradées, comme l'a montré Meyer et al. (2003), sur un public d'adultes implantés cochléaires.

#### **IV. Apports de l'étude pour la pratique orthophonique**

L'implant cochléaire a pour objectif de réhabiliter la fonction auditive chez des sujets présentant une déficience sévère ou profonde. Cependant, les informations transmises par cet appareil sont incomplètes et variables selon les stimuli. En effet, seuls les indices acoustiques portés par l'enveloppe temporelle sont transmis. Les informations apportées par la structure fine sont dégradées. Ces caractéristiques de codage engendrent une différence de perception pour les traits acoustiques des phonèmes du français. Comme nous l'avons vu précédemment, l'implant ne transmet pas les indices acoustiques de la même manière que la cochlée. Pour les consonnes, les informations portant sur les traits de nasalité et de lieu sont réduites par rapport aux informations portant sur les traits de voisement et de mode. Pour les voyelles, les différences se situent au niveau du codage des différences de fréquences des formants. De ces inégalités de transmission découlent des difficultés de perception (Gallégo et Collet, 2008 ; Gnansia, 2009 ; Micheyl, 2005).

Notre étude a permis de mettre en évidence un effet de la dégradation du signal sur la perception catégorielle. En situation de signal dégradé, les participants présentent de moins bonnes capacités pour différencier deux sons lorsqu'ils appartiennent à deux catégories phonémiques différentes. Leurs compétences en discrimination et en identification de mots ou de pseudo-mots sont ainsi réduites. Face aux difficultés perceptives des sujets implantés, une éducation auditive spécifique est nécessaire. L'objectif est de permettre aux patients de développer leurs capacités perceptives en s'ancrant dans le monde sonore pour structurer et catégoriser les informations, mais aussi pour leur donner du sens (Le Calvez, 2009).

Pour les enfants implantés, l'éducation auditive comporte plusieurs étapes (Loundon et Busquet, 2009). Dans un premier temps, elle s'appuie sur l'écoute des bruits environnants : les stimuli sonores sont variés et non spécifiques au langage. L'objectif est de différencier la présence et l'absence de sons. Des jeux sonores peuvent être utilisés pour travailler la discrimination auditive, puis l'identification. La musique est également un média utile pour appréhender les variations de rythme, de prosodie, d'intensité, de hauteur et de durée. Dans un deuxième temps, l'éducation perceptive se base sur du matériel verbal. Il faut d'abord placer l'enfant dans des conditions optimales d'écoute avant de proposer des situations en présence de bruits parasitant le travail d'écoute. Le patient travaille sur des phonèmes, puis des mots et ensuite des phrases (approche analytique), mais une partie du protocole veille également à favoriser les interactions et à développer le plaisir des échanges en proposant des activités de reconnaissance de certains mots courts et usuels au sein de la chaîne de parole (approche globale).

---

Pour les adultes implantés sourds post-linguaux, la réhabilitation auditive est plus courte. Contrairement aux enfants implantés sourd pré-linguaux, l'éducation auditive des adultes a pour objectif de se réapproprier le monde sonore et de s'adapter au nouveau signal auditif. Ces patients ont déjà un système phonémique mature, il s'agit donc d'adapter les capacités perceptives et non de les construire. Comme pour les enfants, le travail perceptif est composé d'exercices de détection, de discrimination, de reconnaissance et d'identification basés sur des stimuli sonores de l'environnement, qu'ils soient verbaux ou non (Cardon et Collet, 2010). La différence dans le protocole se situe dans l'ordre de présentation du matériel verbal. Les adultes travaillent d'abord sur des phrases, puis sur des mots ou des logatomes et ensuite sur des oppositions phonémiques de plus en plus fines (Laborde, 2005). Par ailleurs, ils s'entraînent dans des conditions perceptives de difficultés croissantes avec de la compréhension de phrases, de textes et de conversations dans des conditions optimales dans un premier temps, puis dans des conditions plus difficiles comme un échange téléphonique ou dans un environnement bruyant.

Notre travail donne à réfléchir sur le type de matériel que nous pouvons proposer en rééducation chez l'adulte. Nous venons de donner les progressions respectives pour les enfants et les adultes. Nous pouvons également réfléchir plus précisément au type de matériel que nous pouvons utiliser en rééducation adulte. Cardon et Collet (2010) préconisent le choix du matériel sur des critères phonétiques, mais aussi linguistiques. Il faut choisir d'abord des mots peu fréquents et ayant peu d'homophones ou de voisins phonologiques, puis nous proposerons au patient des mots de plus en plus fréquents dans la langue et ayant de nombreux homophones et voisins phonologiques. Meyer et al. (2003), en prédisant les performances de sujets implantés cochléaires post-linguaux à partir de ces deux facteurs (fréquence et nombres de voisins phonologiques) donnait déjà des arguments en cette faveur. Si notre protocole expérimental met les adultes face à des tâches qui demandent de résoudre une ambiguïté fine à partir de paires minimales, il montre toutefois que même dans ces tâches, les participants réussissent mieux pour les mots que pour les pseudo-mots. Nous n'avons pas vérifié l'effet de la fréquence et du nombre de voisins phonologiques sur les réponses. Nous pensons toutefois qu'une telle analyse pourrait montrer des différences de performances. Nous pensons donc qu'il est important de réfléchir au choix du matériel proposé aux adultes selon ces critères : proposer des mots plutôt que pseudo-mots pour s'appuyer sur les connaissances de l'auditeur et choisir des mots de difficulté croissante, c'est-à-dire avec de plus en plus de voisins phonologiques et avec une fréquence croissante.

Notre travail, à la suite de l'étude de Bouton, Colé et al. (2012), apporte une justification théorique au fait que l'éducation auditive présentée ici soit différenciée pour les adultes et pour les enfants. Dans le cas de l'éducation auditive des adultes sourds post-linguaux, le travail auditif va des unités les plus grandes (énoncés) aux unités les plus fines (phonèmes), afin de développer les stratégies d'adaptation à l'implant en se basant sur les compensations que nous avons explorées dans notre étude. Pour les enfants sourds pré-linguaux, le va-et-vient entre informations acoustiques et informations lexicales est permanent. Il s'agit de construire le lexique car, de la même manière que les enfants entendants, les enfants implantés doivent se construire des représentations (phonémiques et lexicales) du monde sonore. S'ils semblent mettre en place la perception catégorielle de manière similaire aux enfants normo-entendants (Bouton, Colé et al., 2012), cela n'est possible que grâce à des apprentissages explicites, à une exposition privilégiée au matériel langagier et à un étayage permettant la construction des représentations acoustiques et lexicales.

---

Nous ne pouvons pas dissocier un travail perceptif d'un travail sur l'accès au sens. Les traitements acoustiques ne sont pas suffisants pour traiter la parole. En effet, la perception en situation naturelle n'est jamais optimale (Mattys et al, 2012). Dans les conditions de perceptions naturelles, le signal de parole est perturbé entre autres par les bruits environnants. La gêne occasionnée est d'autant plus importante chez les sujets implantés. En effet, les sujets normo-entendants utilisent un mécanisme pour discriminer le signal de parole des bruits environnants. Ce mécanisme s'appelle le « speech masking release » ou démasquage de la parole, en français (Lorenzi et al., 2006 ; Gnansia et al., 2009). Ce processus permet la discrimination des fréquences correspondant au fondamental de la voix humaine. L'implant cochléaire, en ne transmettant que les informations issues de l'enveloppe temporelle, ne permet pas cette sélectivité de la voix humaine. De plus, lorsqu'un auditeur entend un signal de parole, c'est un flot continu d'informations qui lui parvient. Un contexte linguistique naturel ne permet jamais la perception de mots isolés. Nos résultats montrent que les déficits acoustiques sont compensés par un recours aux connaissances lexicales chez les sujets adultes. Ces résultats confirment que nous pouvons nous reposer sur des stratégies de compensation en thérapie. De plus, si les sujets implantés développent des stratégies de compensation pour décoder la chaîne parlée, il faut alors veiller à ce que leur compréhension ne se limite pas à quelques mots de la chaîne parlée. « *On remarque au cours des rééducations que le dernier mot d'une phrase est reconnu avant le premier mot de cette même phrase. Les confusions de l'enfant implanté seront d'autant plus nombreuses que le stock lexical ne sera pas important.* » (Le Calvez, 2009). Il va falloir permettre au patient de compléter ces informations grâce à des connaissances lexicales et morphosyntaxiques (Sadek-Khalil, 1997). Les sujets doivent pouvoir se servir de leurs connaissances lexicales pour étayer leur perception phonémique, mais aussi de leurs connaissances morphosyntaxiques pour comprendre les liens entre les différents éléments d'un énoncé.

En montrant l'importance du lexique dans la compréhension de la chaîne parlée, notre étude soutient les pratiques orthophoniques visant à permettre à l'enfant de développer à la fois son lexique et sa morphosyntaxe. Contrairement aux adultes, le travail de compréhension de la parole chez les enfants nécessite un apprentissage des structures lexicales et morphosyntaxiques de la langue. Cet apprentissage du lexique nécessite une implication aussi bien du patient que de son entourage. L'accompagnement familial prend alors toute son importance. Les enfants implantés présentent souvent de grandes disparités dans la construction de leur stock lexical : ils peuvent connaître des mots correspondant à leur tranche d'âge et ne pas maîtriser des mots normalement acquis bien plus tôt. Un travail basé sur la morphologie lexicale permet une généralisation des concepts de même ordre. Il faut également aider à la mise en place des processus de catégorisation des éléments pour permettre une automatisation des stratégies d'apprentissages (Dumont, 2008). Pour ce qui est de la morphosyntaxe, un travail explicite sur les marqueurs grammaticaux est indispensable, car leur perception est particulièrement difficile. En effet au sein de la chaîne parlée, ces éléments sont rapides, faiblement accentués et courts (Le Normand, 2004). Ce travail explicite peut s'appuyer sur toutes les informations accessibles à l'enfant. Un étayage multimodal permet à l'enfant de pallier les manques du signal sonore et de repérer les éléments signifiants dans ce signal. Les aides visuelles sont particulièrement intéressantes, car cette modalité est bien investie par les patients sourds (Le Calvez, 2009 ; Loundon et Busquet, 2005). La lecture labiale est un étayage visuel universel (Fort et al., 2010). Des outils spécifiques ont été développés pour favoriser les aides visuelles, comme le français signé, et le langage parlé complété. Chez les enfants sourds, ces outils favorisent la compréhension

---

du langage oral, l'acquisition de la morphosyntaxe et des représentations phonémiques, ainsi que l'acquisition du langage écrit (Alegria, Hage, Charlier et Leybaert, 2007). D'autres méthodes multimodales comme la Dynamique Naturelle de la Parole (DNP) et la méthode verbo-tonale peuvent favoriser la perception de l'ensemble des éléments de la chaîne parlée. Nous supposons que ces méthodes peuvent également favoriser la mise en place des représentations phonémiques et lexicales, car elles permettent le repérage de tous les éléments signifiants du signal de parole.

Notre étude démontre un effet du lexique sur la dégradation du signal. Lorsque le signal est dégradé, les participants réussissent mieux les tâches comportant un matériel lexical. Ils s'aident de leurs connaissances lexicales pour compenser le manque d'informations auditives. Ce résultat est un exemple de suppléance mentale. Les informations acoustiques étant parcellaires, les auditeurs s'aident de leurs connaissances lexicales pour combler les manques. Les informations lexicales permettent une activation d'indices acoustiques dans le processus de perception de la parole. En clinique, les confusions sont d'autant plus importantes que le stock lexical est faible. Dans ce cas, la suppléance mentale ne permet pas de rectifier les erreurs ou de compléter les informations non perçues (Le Calvez, 2009). Il est donc important de tester entre autres les connaissances lexicales des sujets sourds post-linguaux avant l'implantation. Aussi, une évaluation régulière des connaissances lexicales et morphosyntaxiques des sujets implantés sourds pré-linguaux permet d'adapter l'intervention orthophonique à leurs capacités de suppléance. Les autres capacités de compensation, comme les aides visuelles, doivent également être prises en compte. Enfin les spécificités du signal perçu doivent être connues afin de savoir sur quel signal se base la perception auditive. En effet, chez les adultes, comme chez les enfants implantés, ce sont les connaissances linguistiques aussi bien lexicales que morphosyntaxiques qui vont permettre au sujet de décrypter le message et de le structurer face aux perturbations perceptives, qu'elles viennent de l'implant ou de l'environnement.

---

## CONCLUSION

---

Nous avons mis en place un protocole expérimental basé sur des tâches d'identification et de discrimination de mots et de pseudo-mots. Ce protocole permet de tester l'influence des représentations lexicales sur la perception et la précision catégorielle. Nous avons pour cela développé un matériel permettant de tester l'ensemble des traits acoustiques des consonnes du Français. Nous avons aussi contrôlé ce matériel selon la fréquence d'apparition des mots dans la langue. Nous avons soumis une population d'adultes normo-entendants à ce matériel que nous avons dégradé grâce à un vocodeur, de façon à reproduire le signal fourni par l'implant cochléaire.

Notre objectif était de mettre en évidence les mécanismes compensatoires mis en place par les sujets pour percevoir le signal dégradé. Nous pensions en effet pouvoir montrer que les connaissances lexicales des sujets influençaient positivement leurs performances. L'hypothèse de tels mécanismes avait déjà été faite sur un public d'enfants implantés cochléaires par Bouton, Colé et al. (2012). Les auteurs n'avaient pas pu montrer de compensation par le lexique chez les enfants implantés cochléaires. Nous avons souhaité tester à nouveau cette hypothèse sur un public d'adultes normo-entendants, n'ayant jamais subi de privation auditive et présentant un stock lexical mature, en les soumettant au signal dégradé au vocodeur. Ce paradigme nous a permis d'isoler l'effet de la dégradation des autres facteurs ayant pu affecter les performances des sujets implantés cochléaires.

Nous avons pu mettre en évidence un effet de compensation du lexique chez les sujets soumis au signal vocodé. Ce résultat ouvre des perspectives pour une meilleure compréhension des mécanismes de perception de la parole chez les implantés cochléaires. Il est en effet nécessaire de mieux comprendre ces mécanismes pour le développement technique de l'implant, comme pour la pratique orthophonique. Le lexique peut être utilisé d'une part comme base du travail pour l'éducation auditive chez l'adulte, et d'autre part comme unité de référence pour mettre en place les représentations phonémiques et lexicales chez l'enfant. De manière générale, ce résultat nous incite particulièrement à être attentif au niveau lexical de nos patients implantés, adultes ou enfants.

Une analyse détaillée des indices acoustiques en jeu dans la perception des différents traits acoustiques que nous avons testés pourrait présenter un intérêt pour la compréhension des conséquences directes de la dégradation du signal sur la perception des différents indices acoustiques. Une analyse des effets de compensation en fonction de la fréquence d'apparition des mots dans la langue et de leurs voisins phonologiques aurait également un intérêt, pour une meilleure compréhension des mécanismes d'accès au lexique. La question se pose également de savoir si les mécanismes de compensation mis en évidence dans notre étude sont activés de la même manière dans d'autres conditions de dégradation et jusqu'à quel niveau de dégradation ils peuvent l'être. Enfin, il est également possible que ces mécanismes de compensation aient un rôle dans des processus d'apprentissage, qui sont encore à explorer.

---

## BIBLIOGRAPHIE

---

- Alegria, J., Hage, C., Charlier, B. et Leybaert, J. (2007). Phonologie audiovisuelle : lecture, lecture labiale et lecture labiale complétée. In *Surdité et langage: prothèses, LPC et implants cochléaires*. Saint-Denis: Presses universitaires de Vincennes.
- Andruski, J. E., Blumstein, S. E. et Burton, M. (1994). The effect of subphonetic differences on lexical access. *Cognition*, 52(3), 163–187.
- Bertoncini, J., Nazzi, T., Cabrera, L. et Lorenzi, C. (2011). Six-month-old infants discriminate voicing on the basis of temporal envelope cues (L). *The Journal of the Acoustical Society of America*, 129, 2761. doi:10.1121=1.3571424.
- Bertoncini, J., Serniclaes, W. et Lorenzi, C. (2009). Discrimination of speech sounds based upon temporal envelope versus fine structure cues in 5-to 7-year-old children. *Journal of Speech, Language and Hearing Research: JSLHR*, 52(3), 682–695. doi:10.1044/1092-4388(2008/07-0273).
- Bouton, S., Colé, P. et Serniclaes, W. (2012). The influence of lexical knowledge on phoneme discrimination in deaf children with cochlear implants. *Speech Communication*, 54(2), 189–198. doi:10.1016/j.specom.2011.08.002.
- Bouton, S., Serniclaes, W., Bertoncini, J. et Colé, P. (2012). Perception of speech features by French-speaking children with cochlear implants. *Journal of speech, language, and hearing research: JSLHR*, 55(1), 139–153. doi:10.1044/1092-4388(2011/10-0330).
- Cardon, M. et Collet, C. (2010). Manuel d'entraînement à l'éducation auditivo-verbale de l'adulte sourd implanté cochléaire. Mémoire d'orthophonie. Récupéré du site de la bibliothèque universitaire de Lille 2 en avril 2013 : [http://www.scd.univ-lille2.fr/fileadmin/user\\_upload/memoires\\_ortho/2010/LIL2\\_SMOR\\_2010\\_002.pdf](http://www.scd.univ-lille2.fr/fileadmin/user_upload/memoires_ortho/2010/LIL2_SMOR_2010_002.pdf)
- Cutler, A., Mehler, J., Norris, D. et Segui, J. (1987). Phoneme identification and the lexicon. *Cognitive psychology*, 19, 141–177.
- Cutler, A. et Fodor, J. A. (1979). Semantic focus and sentence comprehension. *Cognition*, 7(1), 49–59.
- Davis, M. H. et Gaskell, M. G. (2009). A complementary systems account of word learning: neural and behavioural evidence. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, 364(1536), 3773–3800. doi:10.1098/rstb.2009.0111.
- Davis, M. H. et Johnsrude, I. S. (2007). Hearing speech sounds: top-down influences on the interface between audition and speech perception. *Hearing research*, 229(1-2), 132–147. doi:10.1016/j.heares.2007.01.014.

- 
- Davis, M. H., Johnsrude, I. S., Hervais-Adelman, A., Taylor, K. et McGettigan, C. (2005). Lexical information drives perceptual learning of distorted speech: evidence from the comprehension of noise-vocoded sentences. *Journal of experimental psychology. General*, 134(2), 222–241. doi:10.1037/0096-3445.134.2.222.
- Dell, G. S. et Newman, J. E. (1980). Detecting phonemes in fluent speech. *Journal of Verbal Learning and Verbal Behavior*, 19(5), 608–623. doi:10.1016/S0022-5371(80)90661-1.
- Dufour, S. et Frauenfelder, U. H. (2007). L'activation et la sélection lexicales lors de la reconnaissance des mots parlés: modèles théoriques et données expérimentales. *L'Année psychologique*, 107(01), 87–111.
- Dumont, A. (2008). *Orthophonie et surdité: communiquer, comprendre, parler*. Issy-les-Moulineaux: Elsevier Masson.
- Flege, J. E., MacKay, I. R. et Meador, D. (1999). Native Italian speakers' perception and production of English vowels. *The Journal of the Acoustical Society of America*, 106(5), 2973–2987.
- Fort, M., Spinelli, E., Savariaux, C. et Kandel, S. (2010). The word superiority effect in audiovisual speech perception. *Speech Communication*, 52(6), 525–532. doi:10.1016/j.specom.2010.02.005.
- Frauenfelder, U. H. et Peeters, G. (1990). On lexical segmentation in TRACE: An exercise in Simulation. In *Cognitive models of speech processing: psycholinguistic and computational perspectives*. Cambridge, Mass: MIT Press.
- Gallégo, S. et Collet, L. (2008). L'implant cochléaire : comment et pourquoi cela marche? *ORL autrement*, 9, 9-23.
- Ganong, W. F. (1980). Phonetic categorization in auditory word perception. *Journal of Experimental Psychology: Human Perception and Performance*, 6(1), 110–125. doi:10.1037/0096-1523.6.1.110.
- Gnansia, D. (2009). *Intelligibilité dans le bruit et démasquage de la parole chez les sujets normo-entendants, malentendants et implantés cochléaires*. Thèse de doctorat, Université Pierre et Marie Curie, Paris, France. Numéro national de thèse : 2009PA066052.
- Gnansia, D., Péan, V., Meyer, B. et Lorenzi, C. (2009). Effects of spectral smearing and temporal fine structure degradation on speech masking release. *The Journal of the Acoustical Society of America*, 125(6), 4023–4033. doi:10.1121/1.3126344
- Haute Autorité de Santé (HAS). (2007). Traitement de la surdité par pose d'implants cochléaires ou d'implants du tronc cérébral. Récupéré du site internet: [http://www.has-sante.fr/portail/jcms/c\\_559051/traitement-de-la-surdite-par-pose-dimplants-cochleaires-ou-dimplants-du-tronc-cerebral](http://www.has-sante.fr/portail/jcms/c_559051/traitement-de-la-surdite-par-pose-dimplants-cochleaires-ou-dimplants-du-tronc-cerebral)
-

- 
- Haute Autorité de Santé (HAS). (2009). Recommandation de bonne pratique. Surdit  de l'enfant : accompagnement des familles et suivi de l'enfant de 0   6 ans, hors accompagnement scolaire. R cup r  du site internet: [http://www.has-sante.fr/portail/upload/docs/application/pdf/2010-03/surdite\\_de\\_lenfant\\_-\\_0\\_a\\_6\\_ans\\_-\\_recommandations.pdf](http://www.has-sante.fr/portail/upload/docs/application/pdf/2010-03/surdite_de_lenfant_-_0_a_6_ans_-_recommandations.pdf)
- Hoonhorst, I., Colin, C., Markessis, E., Radeau, M., Deltenre, P. et Serniclaes, W. (2009a). N100 component : an electrophysiological cue for voicing perception. In *Some aspects of speech and the brain*. Frankfurt am Main ; New York: Peter Lang.
- Hoonhorst, I., Colin, C., Markessis, E., Radeau, M., Deltenre, P. et Serniclaes, W. (2009b). French native speakers in the making: From language-general to language-specific voicing boundaries. *Journal of Experimental Child Psychology*, 104(4), 353–366. doi:10.1016/j.jecp.2009.07.005.
- Howell, D. C., Yzerbyt, V. et Bestgen, Y. (2008). *M thodes statistiques en sciences humaines*. Bruxelles: De Boeck.
- Kaiser, A. R., Kirk, K. I., Lachs, L. et Pisoni, D. B. (2003). Talker and lexical effects on audiovisual word recognition by adults with cochlear implants. *Journal of speech, language, and hearing research: JSLHR*, 46(2), 390–404.
- Katagiri, S. (2000). *Handbook of neural networks for speech processing*. Boston: Artech House.
- Kirk, K. I., Pisoni, D. B. et Osberger, M. J. (1995). Lexical effects on spoken word recognition by pediatric cochlear implant users. *Ear and hearing*, 16(5), 470–481.
- Knudsen, E. I. (2004). Sensitive periods in the development of the brain and behavior. *Journal of Cognitive Neuroscience*, 16(8), 1412–1425.
- Kral, A. et Sharma, A. (2012). Developmental neuroplasticity after cochlear implantation. *Trends in Neurosciences*, 35(2), 111–122. doi:10.1016/j.tins.2011.09.004.
- Laborde, M-L. (2005). Prise en charge orthophonique de l'adulte post-lingual porteur d'un implant cochl aire : Proposition d'une r ducation. R cup r  du site des journ es GEORRIC, Toulouse en avril 2013 : [http://jeanmarc.chanal.free.fr/marie\\_laborde.pdf](http://jeanmarc.chanal.free.fr/marie_laborde.pdf)
- Le Calvez, V. (2009). La m thode verbo-tonale   l'heure de l'implant cochl aire. *Connaissances Surdit s. La revue ACFOS*, (29), 13–17. R cup r  du site internet de l'ACFOS en avril 2013 : [http://www.acfos.org/publication/acfos/sommaire\\_28\\_v2.pdf](http://www.acfos.org/publication/acfos/sommaire_28_v2.pdf)
- Legge, G. E., Mansfield, J. S., et Chung, S. T. (2001). Psychophysics of reading. XX. Linking letter recognition to reading speed in central and peripheral vision. *Vision research*, 41(6), 725–743.
-

- 
- Le Normand, M. T. (2004). Evaluation du lexique de production chez des enfants sourds profonds munis d'un implant cochléaire sur un suivi de trois ans. *Rééducation orthophonique*, 217, 125–140. Récupéré en ligne en avril 2013 : <http://s3.amazonaws.com/publicationslist.org/data/marielenormand/ref-21/RO217%20article%20LeNormand.pdf>
- Liberman, A. M., Harris, K., Hoffman, H. S. et Griffiths, B. C. (1957). The discrimination of speech sounds across and within phoneme boundaries. *Journal of Experimental Psychology*, 54(5), 358–368.
- Liberman, A. M., Cooper, F. S., Shankweiler, D. P. et Studdert-Kennedy, M. (1967). Perception of the speech code. *Psychological Review*, 74(6), 431–461.
- Loizou, P. C., Dorman, M. et Tu, Z. (1999). On the number of channels needed to understand speech. *The Journal of the Acoustical Society of America*, 106, 2097.
- Lorenzi, C., Husson, M., Ardoint, M. et Debrulle, X. (2006). Speech masking release in listeners with flat hearing loss: effects of masker fluctuation rate on identification scores and phonetic feature reception. *International journal of audiology*, 45(9), 487–495. doi:10.1080/14992020600753213.
- Loundon, N. et Busquet, D. (2009). *Implant cochléaire pédiatrique et rééducation orthophonique : comment adapter les pratiques ?* Paris : Médecine-Sciences Flammarion.
- Luce, P. A. et Pisoni, D. B. (1998). Recognizing spoken words: The neighborhood activation model. *Ear and hearing*, 19(1), 1–36.
- Marslen-Wilson, W. D. et Welsh, A. (1978). Processing interactions and lexical access during word recognition in continuous speech. *Cognitive Psychology*, 10(1), 29–63. doi:10.1016/0010-0285(78)90018-X.
- Mattys, S. L., Davis, M. H., Bradlow, A. R. et Scott, S. K. (2012). Speech recognition in adverse conditions: A review. *Language and Cognitive Processes*, 27(7-8), 953–978. doi:10.1080/01690965.2012.705006.
- McClelland, J. L. et Elman, J. L. (1986). The TRACE model of speech perception. *Cognitive Psychology*, 18(1), 1–86. doi:10.1016/0010-0285(86)90015-0.
- McMurray, B., Tanenhaus, M. K., Aslin, R. N. et Spivey, M. J. (2003). Probabilistic constraint satisfaction at the lexical/phonetic interface: Evidence for gradient effects of within-category VOT on lexical access. *Journal of psycholinguistic research*, 32(1), 77–97.
- Medina, V., Hoonhorst, I., Bogliotti, C. et Serniclaes, W. (2010). Development of voicing perception in French: Comparing adults, adolescents, and children. *Journal of Phonetics*, 38(4), 493–503. doi:10.1016/j.wocn.2010.06.002.
- Medina, V., Loundon, N., Busquet, D., Petroff, N. et Serniclaes, W. (2009). Perception catégorielle des sons de parole chez des enfants avec implant cochléaire.
-

---

Communication présentée au 15e Journées d'étude sur la parole [JEP-TALN-RECITAL], Fès, Maroc. Récupéré en janvier 2012 de <http://www.lpl.univ-aix.fr/jep-taln04/proceed/actes/jep2004/Medina.pdf>

- Menner, A. L. (2005). *Guide de poche des affections de l'oreille*. Paris: Flammarion médecine-sciences.
- Meyer, T. A., Frisch, S. A., Pisoni, D. B., Miyamoto, R. T. et Svirsky, M. A. (2003). Modeling open-set spoken word recognition in postlingually deafened adults after cochlear implantation: some preliminary results with the neighborhood activation model. *Otology & neurotology: official publication of the American Otological Society, American Neurotology Society [and] European Academy of Otology and Neurotology*, 24(4), 612–620.
- Micheyl, C. (2005). Influence du nombre de canaux et de la cadence de stimulation sur les performances de reconnaissance de la parole des patients implantés cochléaires : Le point sur les données scientifiques. Récupéré du site des journées GEORRIC, Toulouse en janvier 2013 : [http://jeanmarc.chanal.free.fr/com\\_micheyl.pdf](http://jeanmarc.chanal.free.fr/com_micheyl.pdf)
- Morton, J. et Long, J. (1976). Effect of word transitional probability on phoneme identification. *Journal of Verbal Learning and Verbal Behavior*, 15(1), 43–51. doi:10.1016/S0022-5371(76)90005-0.
- Norris, D., McQueen, J. M. et Cutler, A. (2000). Merging information in speech recognition: feedback is never necessary. *The Behavioral and brain sciences*, 23(3), 299–325; discussion 325–370.
- Norris, D. (1994). Shortlist: A connectionist model of continuous speech recognition. *Cognition*, 52(3), 189–234.
- Pelli, D. G. (2010). Towards an easier way to measure the visual span. *Journal of Vision*, 9(8), 1002–1002. doi:10.1167/9.8.1002.
- Rubin, P., Turvey, M. T. et VanGelder, P. (1976). Initial Phonemes are detected faster in spoken words than in spoken nonwords. *Perception & Psychophysics*, 19(5), 394–398.
- Sadek-Khalil, D. (1997). *L'enfant sourd et la construction de la langue*. Montreuil [France]: Editions du Papyrus .
- Samuel, A. G. (1981). Phonemic restoration: insights from a new methodology. *Journal of experimental psychology. General*, 110(4), 474–494.
- Segui, J. et Frauenfelder, U.H. (1986). The effects of lexical constraints upon speech perception. In F. Klix et H. Hagendorf (Eds.) *Human memory and cognitive capabilities: Symposium in memoriam Hermann Ebbinghaus*, Amsterdam: North-Holland.

- 
- Serniclaes, W. (1987). *Etude expérimentale de la perception du trait de voisement des occlusives du français*. Thèse de doctorat, Université Libre de Bruxelles, Bruxelles, Belgique. Récupérée du site du laboratoire de Psychologie de la Perception (LPP), Université Descartes Paris 5 en avril 2013: <http://lpp.psych.univ-paris5.fr/pdf/2554.pdf>
- Serniclaes, W., Medina, V., Schepers, F. et Simon, P. (2007). Le développement de la communication parlée avec implant cochléaire. In *Surdit  et langage : prothèses, LPC et implants cochléaires*. Saint-Denis: Presses universitaires de Vincennes.
- Shannon, R. V., Zeng, F. G., Kamath, V., Wygonski, J. et Ekelid, M. (1995). Speech recognition with primarily temporal cues. *Science (New York, N.Y.)*, 270(5234), 303–304.
- Sharma, A., Dorman, M. F. et Kral, A. (2005). The influence of a sensitive period on central auditory development in children with unilateral and bilateral cochlear implants. *Hearing research*, 203(1-2), 134–143. doi:10.1016/j.heares.2004.12.010.
- Spinelli, E. et Ferrand, L. (2005). *Psychologie du langage : l'écrit et le parlé, du signal à la signification*. Paris: A. Colin.
- Sussman, E., Steinschneider, M., Gumenyuk, V., Grushko, J. et Lawson, K. (2008). The maturation of human evoked brain potentials to sounds presented at different stimulus rates. *Hearing research*, 236(1-2), 61–79. doi:10.1016/j.heares.2007.12.001.
- Tyler, R. S., Hall, J. W., Glasberg, B. R., Moore, B. C. et Patterson, R. D. (1984). Auditory filter asymmetry in the hearing impaired. *The Journal of the Acoustical Society of America*, 76(5), 1363–1368.
- Van Linden, S., Stekelenburg, J. J., Tuomainen, J. et Vroomen, J. (2007). Lexical effects on auditory speech perception: an electrophysiological study. *Neuroscience letters*, 420(1), 49–52. doi:10.1016/j.neulet.2007.04.006.
- Warren, R. M. (1970). Perceptual restoration of missing speech sounds. *Science (New York, N.Y.)*, 167(3917), 392–393.
- Werker, J. F. et Tees, R. C. (1984). Cross-Language Speech Perception : Evidence for Reorganization During the First Year of Life. *Infant Behaviour and Development*, 7, 49–63.
- Wilson, B. S. et Dorman, M. F. (2008). Cochlear implants: current designs and future.

---

## **ANNEXES**

---

## LISTE DES ANNEXES

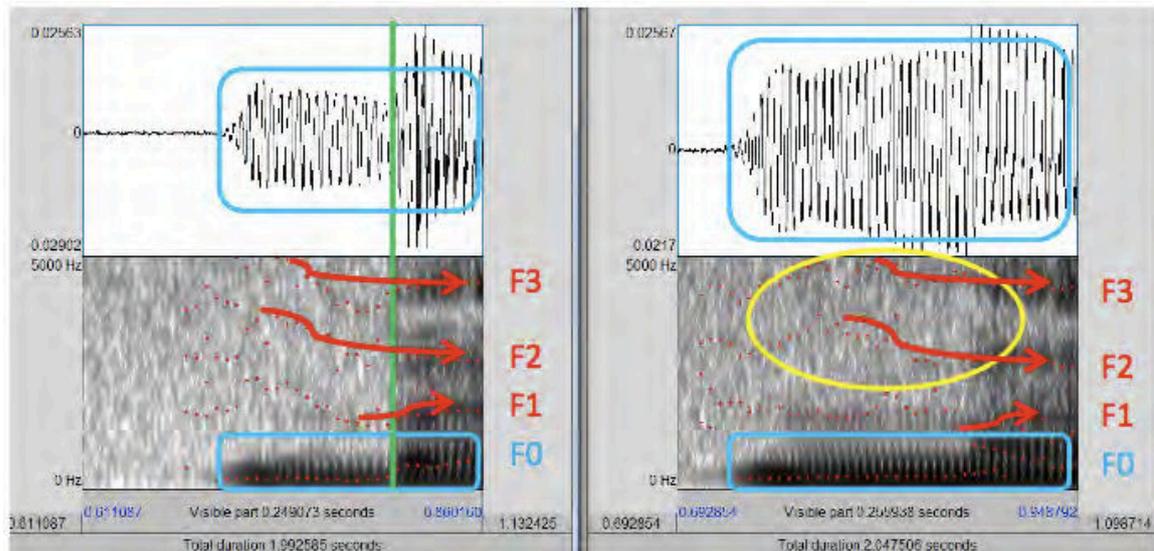
---

Annexe I : Analyse spectrale de deux paires minimales issues de notre matériel .....	78
Annexe II : Représentations du mot « marron » en conditions non vocodée et vocodée .....	80
Annexe III : Tableau récapitulatif du matériel expérimental .....	83
Annexe IV : Document d'information pour notre recherche de population .....	85
Annexe V : Tableau de construction des paires .....	86
1. Pour la nasalité.....	86
2. Pour le lieu articulatoire .....	87
3. Pour le mode articulatoire.....	88
4. Pour le voisement .....	89
Annexe VI : Ordre de présentation des tâches .....	90
Annexe VII : Consignes présentées lors de la passation.....	92
1. Consignes écrites (pour Matlab) .....	92
2. Consignes orales (trame des explications).....	92
2.1. Discrimination.....	92
2.2. Identification .....	93

---

## Annexe I : Analyse spectrale de deux paires minimales issues de notre matériel

Cette annexe a pour but de mettre en évidence les indices spectraux principaux correspondant aux traits articulatoires des consonnes.



**Figure 19 : Comparaison des représentations acoustiques de /b/ (à gauche) et /m/ (à droite), dans la paire minimale « bâche »/« mâche », issue de notre matériel.**

La figure comprend deux oscillogrammes (en haut) et deux spectrogrammes (en bas) issus de Praat. La barre de voisement est encadrée en bleu ; les transitions formantiques sont représentées par des flèches rouges ; la barre d'explosion apparaît en vert et les anti-formants sont entourés en jaune.

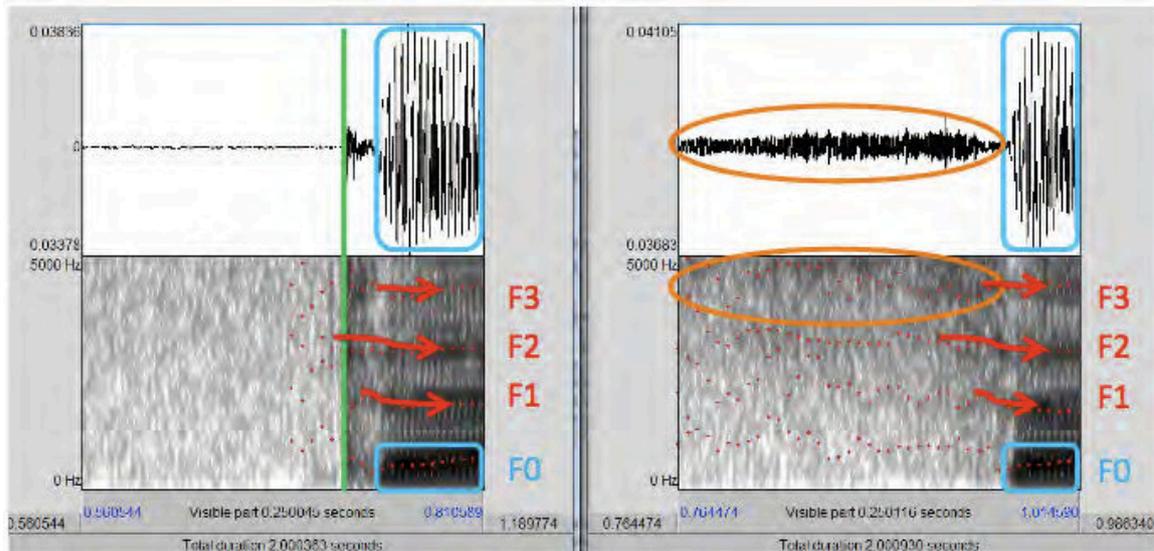
La barre de voisement sur le spectrogramme et la zone de périodicité sur l'oscillogramme sont encadrées en bleu. Elles correspondent à la fréquence fondamentale (F0) de la vibration laryngée. Nous notons que le début de la barre de voisement coïncide avec le début de la production du /b/ pour /ba/, et de la production du /m/ pour /ma/. Cela nous donne une indication sur le trait articulatoire de voisement : /b/ et /m/ sont toutes deux des consonnes voisées.

Les transitions formantiques pour les formants F1, F2 et F3 sont représentées par des flèches rouges. Elles montrent les changements articulatoires lors du passage des consonnes /b/ ou /m/ à la voyelle /a/. Les transitions formantiques ont des points de départ similaires pour les deux consonnes. Cela indique que les deux consonnes ont le même lieu articulatoire. Elles sont toutes deux produites en avant de la bouche, au niveau des lèvres. Ce sont des consonnes labiales.

Une barre d'explosion apparaît en vert sur la figure de gauche. Elle correspond au relâchement d'air consécutif à l'ouverture des lèvres lors de la production du /b/. C'est un indicateur du mode articulatoire, car il indique le relâchement de l'occlusion : /b/ est une consonne occlusive. /m/ est également une occlusive, mais la barre d'explosion n'apparaît

pas à droite, car le relâchement d'air n'est pas assez conséquent pour générer une perturbation acoustique visible au spectrogramme.

/m/ est un phonème nasal, qui présente des anti-formants sur les hautes fréquences. Ces anti-formants apparaissent sur le spectrogramme comme des zones noircies de façon parcellaire. Ils indiquent la résonance de la parole dans la cavité nasale. Ces anti-formants ont été entourés en jaune. Ils sont des indices spectraux du trait de nasalité.



**Figure 20 : Comparaison des représentations acoustiques de /t/ (à gauche) et /s/ (à droite) dans la paire minimale « talon »-« salon », issue de notre matériel.**

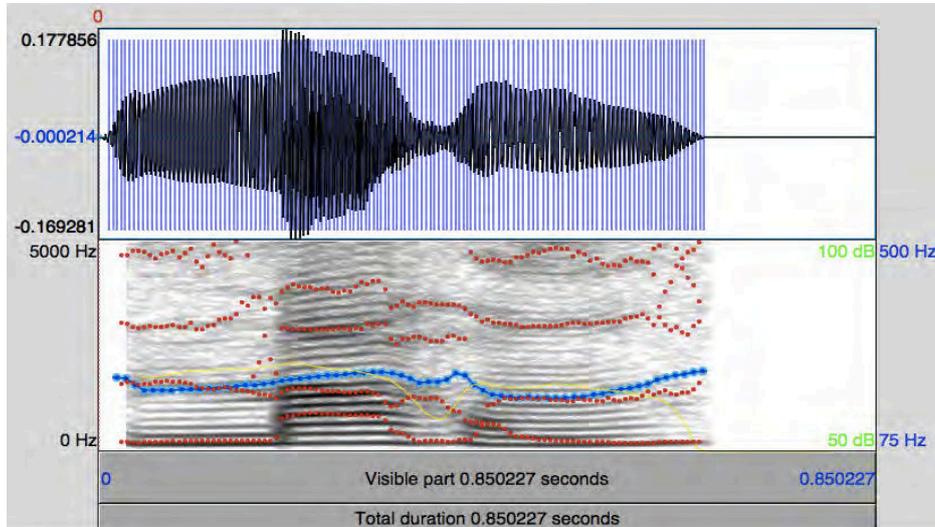
La figure comprend deux oscillogrammes (en haut) et deux spectrogrammes (en bas) issus de Praat. La barre de voisement est encadrée en bleu ; les transitions formantiques sont représentées par des flèches rouges ; la barre d'explosion apparaît en vert et la zone de friction est entourée en orange.

La barre de voisement sur le spectrogramme et la zone de périodicité sur l'oscillogramme sont encadrées en bleu. Elles correspondent à la fréquence fondamentale (F0) de la vibration laryngée. Nous notons que le début de la barre de voisement coïncide avec le début de la production de la voyelle /a/. Il n'y a pas de barre de voisement lors de la production de /t/ et de /s/. Cela nous donne une indication sur le trait articulaire de voisement : /t/ et /s/ sont toutes deux des consonnes non voisées ou sourdes.

Les transitions formantiques pour les formants F1, F2 et F3 sont représentées par des flèches rouges. Elles montrent les changements articulaires lors du passage des consonnes /t/ ou /s/ à la voyelle /a/. Les transitions formantiques ont des points de départ similaires pour les deux consonnes. Cela indique un lieu articulaire similaire : /t/ et /s/ sont toutes deux produites en arrière des dents. Ce sont des consonnes alvéolaires.

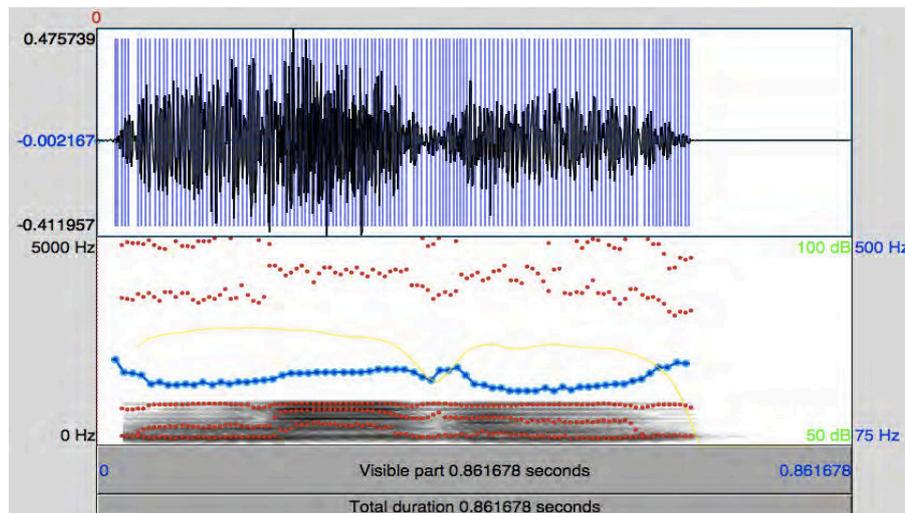
La barre d'explosion du /t/ apparaît en vert à gauche. Elle marque le relâchement d'air consécutif à l'occlusion du conduit buccal. La zone de friction du /s/ est entourée en orange à droite. Elle correspond à l'écoulement d'air qui a lieu tout au long de la constriction du conduit buccal lors de la production d'un /s/. Cet écoulement d'air produit un bruit, visible comme une onde apériodique à l'oscillogramme, et comme une zone de plus forte intensité sur les hautes fréquences au spectrogramme. Ces éléments spectraux donnent des indices sur le mode articulaire. La barre d'explosion indique l'occlusion pour le /t/ et la zone de friction marque la constriction pour le /s/.

## Annexe II : Représentations du mot « marron » en conditions non vocodée et vocodée



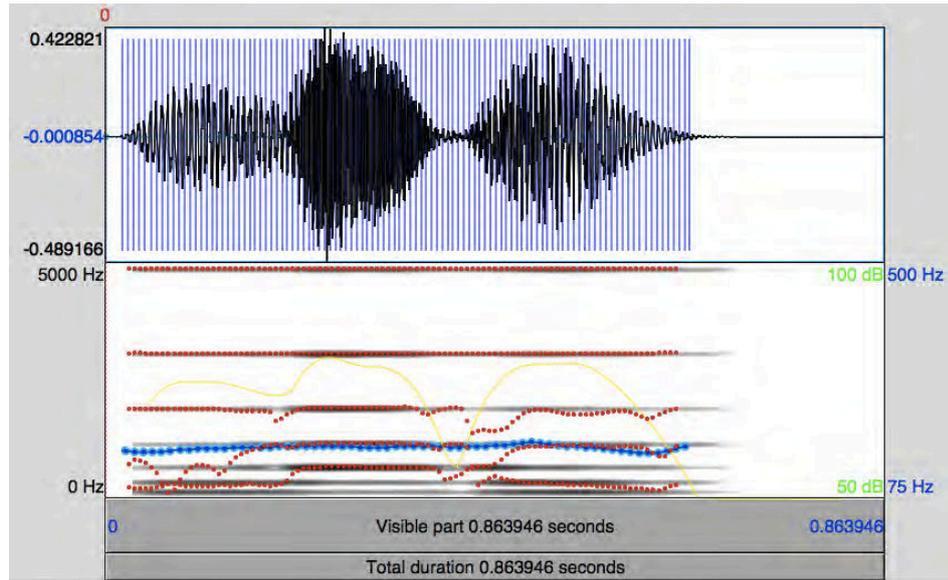
**Figure 21 : Représentation de l'oscillogramme et du spectrogramme du mot « marron » dans la condition non vocodée**

Cette représentation a été réalisée à l'aide du logiciel Praat. L'oscillogramme est situé au-dessus du spectrogramme. L'oscillogramme (amplitude de la courbe (Hz) en fonction du temps (s)) représente l'ensemble des indices spectraux (enveloppe porteuse et structure temporelle fine). Le spectrogramme (fréquence (Hz) en fonction du temps (s)) permet de visualiser la courbe de fréquence (en bleu), la courbe d'intensité (en jaune) et les courbes des formants (en rouge).



**Figure 22 : Représentation de l'oscillogramme et du spectrogramme du mot « marron » dans la condition vocodée à 16 bandes**

Cette représentation a été réalisée à l'aide du logiciel Praat. L'oscillogramme est situé au-dessus du spectrogramme. L'oscillogramme (amplitude de la courbe (Hz) en fonction du temps (s)) représente l'ensemble des indices spectraux. Le spectrogramme (fréquence (Hz) en fonction du temps (s)) permet de visualiser la courbe de fréquence (en bleu), la courbe d'intensité (en jaune) et les courbes des formants (en rouge).



**Figure 23 : Représentation de l'oscillogramme et du spectrogramme du mot « marron » dans la condition vocodé à 8 bandes**

Cette représentation a été réalisée à l'aide du logiciel Praat. L'oscillogramme est situé au dessus du spectrogramme. L'oscillogramme (amplitude de la courbe (Hz) en fonction du temps (s)) représente l'ensemble des indices spectraux. Le spectrogramme (fréquence (Hz) en fonction du temps (s)) permet de visualiser la courbe de fréquence (en bleu), la courbe d'intensité (en jaune) et les courbes des formants (en rouge).

Les spectrogrammes font apparaître une diminution des informations acoustiques entre les messages sonores non vocodé et vocodés. Pour le vocodage à 16 bandes, les informations sont condensées entre 0 et 1000Hz. Au contraire, pour le vocodage à 8 bandes, les informations sont très espacées les unes par rapport aux autres sur une gamme de fréquences allant de 0Hz à 5000Hz.

Sur les spectrogrammes, la courbe jaune représente la courbe d'intensité. Un creux apparaît vers 0,38 secondes sur l'ensemble des spectrogrammes. Cependant, la courbe d'intensité du signal vocodé à 8 bandes est beaucoup plus perturbée que celles des signaux non vocodé et vocodé à 16 bandes.

La courbe d'évolution de la fréquence est représentée en bleu. Contrairement à la courbe d'intensité, la courbe de fréquence du signal vocodé à 8 bandes est aplanie par rapport à celles des signaux non vocodé et vocodé 16 bandes.

De la même manière, les courbes des formants (en rouge) sont plus aplanies pour le vocodage à 8 bandes. Néanmoins, ces courbes sont espacées entre elles, alors que pour le vocodage à 16 bandes, elles sont condensées entre 0 Hz et 1000 Hz pour les trois premiers formants. Pour les formants 4 et 5, le vocodage à 16 bandes présente des courbes perturbées à partir d'environ 3500Hz.

**Tableau IV : Relevé des résultats des différentes caractéristiques du signal acoustique (non vocodé, vocodé à 16 bandes et vocodé à 8 bandes)**

L'ensemble des données est calculé à l'aide du logiciel Praat à partir des trois représentations des oscillogrammes et des spectrogrammes du mot « marron » dans les conditions non vocodée, vocodée à 16 bandes et vocodée à 8 bandes.

Caractéristiques observées	Éléments observés	Non vocodé	Vocodage à 16 bandes	Vocodage à 8 bandes
Durée (s)		0,85	0,86	0,86
Intensité (dB)	moyenne	67,17	75,15	75,34
	maximum	70,82	78,37	80,10
	minimum	43,57	50,07	46,03
	max-min	27,25	28,30	34,07
Fréquence (Hz)	moyenne	210,47	211,87	167,13
	maximum	237,58	245,35	176,48
	minimum	184,14	186,87	157,12
	max-mini	53,44	58,48	19,36
Formants (Hz)	F1	399,04	286,57	397,09
	F2	1343,98	676,56	999,00
	F3	2927,87	1295,13	1843,43
	F4	4221,10	3973,62	3092,96

Grâce au logiciel Praat, nous avons pu extraire des données chiffrées concernant l'intensité, la fréquence et les formants des différents types de signal acoustique (non vocodé, vocodé à 16 bandes et vocodé à 8 bandes).

Nous remarquons que l'intensité moyenne des signaux vocodés est supérieure à celle du signal non vocodé. Cette différence se situe aux alentours de 30dB.

La fréquence moyenne est similaire entre le signal non vocodé et le vocodage à 16 bandes. Elle diminue d'une trentaine d'Hertz pour le signal vocodé 8 à bandes. Comme nous avons pu le remarquer en observant la courbe de fréquence sur les spectrogrammes, les variations d'amplitude de la fréquence pour le signal vocodé à 8 bandes sont beaucoup plus faibles que pour les deux autres enregistrements, ce qui confirme l'aplanissement de la courbe de fréquence du signal vocodé à 8 bandes.

La fréquence des formants est nettement moindre pour les signaux vocodés par rapport à celle des formants du signal non vocodé. Seul le premier formant du signal vocodé à 8 bandes est identique à celui du signal non vocodé. De plus, nous notons que les trois premiers formants du vocodage à 16 bandes sont inférieurs en fréquence à ceux du signal vocodé à 8 bandes. Au contraire, la fréquence du quatrième formant du signal vocodé à 16 bandes est supérieure de 900Hz par rapport à celle du signal vocodé à 8 bandes.

## Annexe III : Tableau récapitulatif du matériel expérimental

Les paires minimales de mots ont été contrôlées en fonction du trait articulatoire de variation (nasalité, mode, voisement ou lieu), de la structure syllabique répartie en consonnes (C) ou voyelles (V) (CVCV ou CVC) et de la fréquence d'usage (haute ou basse) des mots utilisés.

Paire	Mots	Pseudo-mots	Trait articulatoire	Structure syllabique	Fréquence	Correspondances Matlab			
						Stimulus	Trait (Feature)	Structure syllabique	Fréquence
1	baron	bavon	nasalité	CVCV	haute	S1	F1	CV2	Fr1
	marron	mavon	nasalité			S2	F1	CV2	Fr1
2	bonté	bonlé	nasalité	CVCV	haute	S3	F1	CV2	Fr1
	montée	monlé	nasalité			S4	F1	CV2	Fr1
3	bonnet	bojet	nasalité	CVCV	haute	S5	F1	CV2	Fr1
	monnaie	mojet	nasalité			S6	F1	CV2	Fr1
4	bouton	boudon	nasalité	CVCV	haute	S7	F1	CV2	Fr1
	mouton	moudon	nasalité			S8	F1	CV2	Fr1
5	banque	banle	nasalité	CVC	haute	S9	F1	CV1	Fr1
	manque	manle	nasalité			S10	F1	CV1	Fr1
6	bâche	bame	nasalité	CVC	basse	S11	F1	CV1	Fr2
	mâche	mame	nasalité			S12	F1	CV1	Fr2
7	bêche	bèpe	nasalité	CVC	basse	S13	F1	CV1	Fr2
	mèche	mèpe	nasalité			S14	F1	CV1	Fr2
8	biche	bige	nasalité	CVC	basse	S15	F1	CV1	Fr2
	miche	mige	nasalité			S16	F1	CV1	Fr2
9	botte	bope	nasalité	CVC	basse	S17	F1	CV1	Fr2
	motte	mope	nasalité			S18	F1	CV1	Fr2
10	buse	bude	nasalité	CVC	basse	S19	F1	CV1	Fr2
	muse	mude	nasalité			S20	F1	CV1	Fr2
11	fer	fède	mode	CVC	haute	S21	F2	CV1	Fr1
	père	pède	mode			S22	F2	CV1	Fr1
12	bague	bage	mode	CVC	haute	S23	F2	CV1	Fr1
	vague	vagè	mode			S24	F2	CV1	Fr1
13	photo	fosso	mode	CVCV	haute	S25	F2	CV2	Fr1
	poteau	posso	mode			S26	F2	CV2	Fr1
14	talon	tadon	mode	CVCV	haute	S27	F2	CV2	Fr1
	salon	sadon	mode			S28	F2	CV2	Fr1
15	gâteau	gapo	mode	CVCV	haute	S29	F2	CV2	Fr1
	râteau	rapo	mode			S30	F2	CV2	Fr1
16	boeuf	beude	mode	CVC	basse	S31	F2	CV1	Fr2
	veuf	veude	mode			S32	F2	CV1	Fr2
17	galon	gasson	mode	CVCV	basse	S33	F2	CV2	Fr2
	jalon	jasson	mode			S34	F2	CV2	Fr2
18	benne	beipe	mode	CVC	basse	S35	F2	CV1	Fr2
	veine	veipe	mode			S36	F2	CV1	Fr2

Paire	Mots	Pseudo-mots	Trait articulatoire	Structure syllabique	Fréquence	Correspondances Matlab			
						Stimulus	Trait (Feature)	Structure syllabique	Fréquence
19	feinte	finle	mode	CVC	basse	S37	F2	CV1	Fr2
	pinte	pinle	mode			S38	F2	CV1	Fr2
20	cale	cabe	mode	CVC	basse	S39	F2	CV1	Fr2
	châle	chabe	mode			S40	F2	CV1	Fr2
21	poule	poute	voisement	CVC	haute	S41	F3	CV1	Fr1
	boule	boute	voisement			S42	F3	CV1	Fr1
22	balai	bajai	voisement	CVCV	haute	S43	F3	CV2	Fr1
	palais	pajai	voisement			S44	F3	CV2	Fr1
23	boulet	bounet	voisement	CVCV	haute	S45	F3	CV2	Fr1
	poulet	pounet	voisement			S46	F3	CV2	Fr1
24	fer	fède	voisement	CVC	haute	S47	F3	CV1	Fr1
	verre	vède	voisement			S48	F3	CV1	Fr1
25	phase	faje	voisement	CVC	haute	S49	F3	CV1	Fr1
	vase	vaje	voisement			S50	F3	CV1	Fr1
26	cale	cabe	voisement	CVC	basse	S51	F3	CV1	Fr2
	gale	gabe	voisement			S52	F3	CV1	Fr2
27	char	chabe	voisement	CVC	basse	S53	F3	CV1	Fr2
	jarre	jabe	voisement			S54	F3	CV1	Fr2
28	calot	casso	voisement	CVCV	bas	S55	F3	CV1	Fr2
	galop	gasso	voisement			S56	F3	CV1	Fr2
29	bottin	bolin	voisement	CVCV	basse	S57	F3	CV1	Fr2
	potin	polin	voisement			S58	F3	CV1	Fr2
30	bagne	bane	voisement	CVC	basse	S59	F3	CV1	Fr2
	pagne	pane	voisement			S60	F3	CV1	Fr2
31	coeur	coeuke	lieu	CVC	haute	S61	F4	CV1	Fr1
	peur	peuke	lieu			S62	F4	CV1	Fr1
32	gâteau	gano	lieu	CVCV	haute	S63	F4	CV2	Fr1
	bateau	bano	lieu			S64	F4	CV2	Fr1
33	galet	gadet	lieu	CVCV	haute	S65	F4	CV2	Fr1
	balai	badet	lieu			S66	F4	CV2	Fr1
34	bar	bane	lieu	CVC	haute	S67	F4	CV1	Fr1
	gare	gane	lieu			S68	F4	CV1	Fr1
35	cage	cabe	lieu	CVC	haute	S69	F4	CV1	Fr1
	page	pabe	lieu			S70	F4	CV1	Fr1
36	quinte	kinfe	lieu	CVC	basse	S71	F4	CV1	Fr2
	pinte	pinfe	lieu			S72	F4	CV1	Fr2
37	peigne	peive	lieu	CVC	basse	S73	F4	CV1	Fr2
	teigne	teive	lieu			S74	F4	CV1	Fr2
38	baume	baute	lieu	CVC	low	S75	F4	CV1	Fr2
	dôme	daute	lieu			S76	F4	CV1	Fr2
39	pari	pabi	lieu	CVCV	basse	S77	F4	CV2	Fr2
	carie	cabi	lieu			S78	F4	CV2	Fr2
40	vallon	vamon	lieu	CVCV	basse	S79	F4	CV2	Fr2
	jalon	jamon	lieu			S80	F4	CV2	Fr2

---

## Annexe IV : Document d'information pour notre recherche de population

### UNIVERSITE CLAUDE-BERNARD LYON I

#### Institut Sciences et Techniques de Réadaptation : Formation Orthophonie

Adresse postale : 8, avenue Rockefeller 69373 Lyon cedex 08

Téléphone standard : 04 78 77 70 00

Année universitaire 2012-2013

#### **JOBIT Pauline**

36 rue de la Thibaudière 69007 Lyon

06.63.90.77.12

paulinejobit@gmail.com

#### **JOYE Nelly**

14 rue Duroc 69001 Lyon

06.95.73.06.50

nelly.joye@wanadoo.fr

<b>Mémoire de recherche en orthophonie</b>
--

Madame, Monsieur,

Dans le cadre de notre mémoire de fin de formation, nous sommes amenées à réaliser une étude sur la conséquence de la dégradation du signal acoustique sur les traitements perceptifs de la parole, en particulier sur le niveau de traitement lexical. Nous nous intéresserons aux traitements lexicaux mis en œuvre par l'individu lorsque celui-ci reçoit un message auditif dégradé. Pour tenter de répondre à cette question, nous nous sommes intéressées au cas de l'implant cochléaire et au signal dégradé apporté par cet appareil. A l'aide d'un logiciel informatique : le vocodeur, nous allons tenter de reproduire les paramètres acoustiques caractéristiques de l'implant cochléaire. Nous cherchons à savoir si des adultes normo-entendants, n'ayant jamais subi de période de privation auditive comme les personnes sourdes implantées, vont développer des stratégies compensatoires, associées à leurs connaissances lexicales, lorsqu'ils sont soumis à un signal de parole dégradé au vocodeur.

Nous souhaiterions soumettre des adultes normo-entendants à un message auditif dégradé au vocodeur lors d'une tâche de décision lexicale (discrimination et identification de mots existants et de mots inventés). Les sujets entendront à travers un casque un signal de parole à un seuil d'audition confortable (60-65 dB SPL), qu'ils pourront régler au préalable. Les réponses ainsi que le temps de latence avant chaque réponse seront enregistrés.

Pour les mois de novembre à décembre 2012, nous sommes à la recherche d'une population pour notre étude, comprenant des adultes **masculins et féminins dont l'âge est compris entre 20 et 35 ans. Les sujets ne présenteront pas d'antécédents de troubles auditifs et seront de langue maternelle française.** Le lieu de passation, qui doit être calme et sans bruit ambiant, sera déterminé en fonction des disponibilités de chaque personne. Nous pourrions nous retrouver à l'université, à notre domicile ou chez la personne participant à l'étude. Le temps de passation est estimé à une heure. Les réponses données lors de cette étude seront enregistrées, mais ne feront pas l'objet d'une diffusion individuelle, puisque notre objectif est de réaliser une analyse statistique. De plus, l'anonymat des participants sera respecté.

Si vous souhaitez participer à cette étude et ainsi nous aider à réaliser notre projet, contactez-nous par téléphone ou par mail aux coordonnées mentionnées plus haut.

Cordialement,

Pauline JOBIT et Nelly JOYE

## Annexe V : Tableau de construction des paires

### 1. Pour la nasalité

**1** : Forme orthographique ; **2** : Forme phonologique ; **3** : Lemme ; **4** : Catégorie grammaticale ; **5** : Fréquence d'apparition du lemme dans les films (par million d'occurrences) ; **6** : Fréquence d'apparition du lemme dans les livres (par million d'occurrence) ; **7** : Fréquence d'apparition du mot dans les films (par million d'occurrences) ; **8** : Fréquence d'apparition du mot dans les livres (par million d'occurrences) ; **9** : Moyenne de toutes les fréquences (lemme et mot, livres et films), par million d'occurrences ; **10** : Niveau de fréquence (haute si >20 par million d'occurrences, basse si <20 par million d'occurrences) ; **11** : Nombre d'homographes ; **12** : Nombre de voisins orthographiques ; **13** : Nombre d'homophones ; **14** : Nombre de voisins phonologiques ; **15** : Nombre de lettres ; **16** : Structure orthographique ; **17** : Nombre de phonèmes ; **18** : Structure phonémique ; **19** : Nombre de syllabes ; **20** : Structure syllabique

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
baron	baRɔ̃	baron	NOM	20.82	28.24	16.07	17.64	20.69	Haute	1	3	3	27	5	CVCVC	4	CVCV	2	CV-CV
marron	maRɔ̃	marron	NOM	4.89	7.36	2.09	3.45	4.44		2	2	4	27	6	CVCCVC	4	CVCV	2	CV-CV
bonté	bɔ̃te	bonté	NOM	17.88	19.8	17.31	18.65	18.41	Basse	1	7	2	11	5	CVCCV	4	CVCV	2	CV-CV
montée	mɔ̃te	montée	NOM	2.67	15.61	2.48	14.39	8.78		3	4	12	15	6	CVCCVV	4	CVCV	2	CV-CV
monnaie	monɛ	monnaie	NOM	27.31	28.99	26.82	25.34	27.11	Haute	2	1	3	15	7	CVCCVVV	4	CVCV	2	CV-CV
bonnet	bonɛ	bonnet	NOM	8.37	18.58	6.62	14.66	12.05		1	3	2	16	6	CVCCVC	4	CVCV	2	CV-CV
bouton	butɔ̃	bouton	NOM	32.44	44.46	21.29	21.55	29.93	Haute	1	4	3	21	6	CVVCVC	4	CVCV	2	CV-CV
mouton	mutɔ̃	mouton	NOM	15.20	30.47	6.080	14.12	16.46		1	2	2	14	6	CVVCVC	4	CVCV	2	CV-CV
banque	bɑ̃k	banque	NOM	79.35	30.88	70.79	25.54	51.64	Haute	1	5	2	10	6	CVCCVV	3	CVC	1	CVC
manque	mɑ̃k	manque	NOM	24.77	37.23	24.48	36.28	30.69		2	6	5	10	6	CVCCVV	3	CVC	1	CVC
bâche	baʃ	bâche	NOM	2.40	13.38	2.30	10.07	7.03	Basse	2	11	3	31	5	CVCCV	3	CVC	1	CVC
mâche	maʃ	mâche	NOM	0.81	1.08	0.81	1.08	0.94		2	11	4	28	5	CVCCV	3	CVC	1	CVC
bêche	bɛʃ	bêche	NOM	0.46	3.72	0.42	3.51	2.02	Basse	2	7	4	22	5	CVCCV	3	CVC	1	CVC
mèche	mɛʃ	mèche	NOM	9.87	31.01	8.21	19.12	17.05		1	8	2	21	5	CVCCV	3	CVC	1	CVC
bièche	biʃ	bièche	NOM	5.80	13.92	5.29	7.3	8.07	Basse	2	10	4	30	5	CVCCV	3	CVC	1	CVC
miche	miʃ	miche	NOM	2.20	9.80	0.70	4.59	4.32		1	10	2	23	5	CVCCV	3	CVC	1	CVC
botte	bɔt	botte	NOM	32.19	44.93	6.38	8.51	23.00	Basse	2	14	6	28	5	CVCCV	3	CVC	1	CVC
motte	mɔt	motte	NOM	0.51	9.19	0.26	2.97	3.2325		1	10	2	28	5	CVCCV	3	CVC	1	CVC
buse	byz	buse	NOM	1.28	2.84	0.82	2.09	1.7575	Basse	1	9	2	16	4	CVCV	3	CVC	1	CVC
muse	myz	muse	NOM	2.42	1.35	1.82	0.81	1.6		2	11	3	12	4	CVCV	3	CVC	1	CVC

## 2. Pour le lieu articulatoire

**1** : Forme orthographique ; **2** : Forme phonologique ; **3** : Lemme ; **4** : Catégorie grammaticale ; **5** : Fréquence d'apparition du lemme dans les films (par million d'occurrences) ; **6** : Fréquence d'apparition du lemme dans les livres (par million d'occurrence) ; **7** : Fréquence d'apparition du mot dans les films (par million d'occurrences) ; **8** : Fréquence d'apparition du mot dans les livres (par million d'occurrences) ; **9** : Moyenne de toutes les fréquences (lemme et mot, livres et films), par million d'occurrences ; **10** : Niveau de fréquence (haute si >20 par million d'occurrences, basse si <20 par million d'occurrences) ; **11** : Nombre d'homographes ; **12** : Nombre de voisins orthographiques ; **13** : Nombre d'homophones ; **14** : Nombre de voisins phonologiques ; **15** : Nombre de lettres ; **16** : Structure orthographique ; **17** : Nombre de phonèmes ; **18** : Structure phonémique ; **19** : Nombre de syllabes ; **20** : Structure syllabique

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
coeur	kœR	coeur	NOM	239.97	400.74	224.98	380.07	311.44	Haute	1	1	4	13	5	CVVVC	3	CVC	1	CVC
peur	pœR	peur	NOM	557.21	311.69	551.83	307.23	431.99		1	8	1	8	4	CVVC	3	CVC	1	CVC
gâteau	gato	gâteau	NOM	55.19	32.16	42.33	13.92	35.90	Haute	1	1	2	13	6	CVCVVV	4	CVCV	2	CV-CV
bateau	bato	bateau	NOM	124.82	82.36	106.55	61.22	93.73		1	0	2	17	6	CVCVVV	4	CVCV	2	CV-CV
galet	galE	galet	NOM	1.18	15.74	0.54	2.70	5.04	Basse	1	4	2	30	5	CVCVC	4	CVCV	2	CV-CV
balai	balE	balai	NOM	10.91	17.70	8.24	11.96	12.20		1	2	9	26	5	CVCV	4	CVCV	2	CV-CV
gare	gaR	gare	NOM	42.15	84.53	40.28	78.58	61.38	Haute	3	23	6	34	4	CVCV	3	CVC	1	CVC
bar	baR	bar	NOM	67.57	61.35	60.17	52.57	60.41		1	15	7	38	3	CVC	3	CVC	1	CVC
cage	kaʒ	cage	NOM	18.30	41.82	16.61	34.86	27.89	Haute	1	15	2	24	4	CVCV	3	CVC	1	CVC
page	paʒ	page	NOM	39.41	115.54	25.16	55.88	58.99		1	15	2	20	4	CVCV	3	CVC	1	CVC
quinte	kɛ̃t	quinte	NOM	1.42	4.86	1.25	3.85	2.84	Basse	1	6	2	15	6	CVVCCV	3	CVC	1	CVC
pinte	pɛ̃t	pinte	NOM	1.34	0.61	0.90	0.34	0.79		2	10	7	17	5	CVCCV	3	CVC	1	CVC
peigne	pɛjɛ̃	peigne	NOM	6.81	10.81	6.07	8.85	8.13	Basse	2	6	6	20	6	CVVCCV	3	CVC	1	CVC
teigne	tɛjɛ̃	teigne	NOM	1.22	1.15	1.18	0.95	1.12		2	3	4	14	6	CVVCCV	3	CVC	1	CVC
baume	bom	baume	NOM	1.94	3.65	1.79	2.91	2.57	Basse	1	4	3	10	5	CVVCV	3	CVC	1	CVC
dôme	dom	dôme	NOM	1.93	4.93	1.81	3.78	3.11		1	5	2	11	4	CVCV	3	CVC	1	CVC
carie	kaRi	carie	NOM	0.89	1.15	0.38	0.61	0.75	Basse	1	13	7	31	5	CVCV	4	CVCV	2	CV-CV
pari	paRi	pari	NOM	26.61	12.57	13.93	4.59	14.42		1	16	5	29	4	CVCV	4	CVCV	2	CV-CV
vallon	valɔ̃	vallon	NOM	2.50	6.89	2.11	5.74	4.31	Basse	1	3	3	18	6	CVCCVC	4	CVCV	2	CV-CV
jalon	ʒalɔ̃	jalon	NOM	0.17	2.70	0.03	1.08	0.99		1	5	2	10	5	CVCVC	4	CVCV	2	CV-CV

### 3. Pour le mode articulatoire

**1**: Forme orthographique ; **2**: Forme phonologique ; **3**: Lemme ; **4**: Catégorie grammaticale ; **5**: Fréquence d'apparition du lemme dans les films (par million d'occurrences) ; **6**: Fréquence d'apparition du lemme dans les livres (par million d'occurrence) ; **7**: Fréquence d'apparition du mot dans les films (par million d'occurrences) ; **8**: Fréquence d'apparition du mot dans les livres (par million d'occurrences) ; **9**: Moyenne de toutes les fréquences (lemme et mot, livres et films), par million d'occurrences ; **10**: Niveau de fréquence (haute si >20 par million d'occurrences, basse si <20 par million d'occurrences) ; **11**: Nombre d'homographes ; **12**: Nombre de voisins orthographiques ; **13**: Nombre d'homophones ; **14**: Nombre de voisins phonologiques ; **15**: Nombre de lettres ; **16**: Structure orthographique ; **17**: Nombre de phonèmes ; **18**: Structure phonémique ; **19**: Nombre de syllabes ; **20**: Structure syllabique

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
fer	fɛR	fer	NOM	37.08	114.19	33.65	106.28	72.80	Haute	1	9	5	28	3	CVC	3	CVC	1	CVC
père	pɛR	père	NOM	895.55	723.51	879.31	708.11	801.62		1	13	13	32	4	CVCV	3	CVC	1	CVC
baguette	bag	baguette	NOM	30.32	22.36	26.14	16.08	23.72	Haute	2	9	3	26	5	CVCVV	3	CVC	1	CVC
vague	vag	vague	NOM	21.59	73.58	12.80	38.18	36.53		3	8	6	15	5	CVCVV	3	CVC	1	CVC
photo	fɔto	photo	NOM	209.1	94.59	122.47	54.66	120.20	Haute	1	1	2	10	5	CCVCV	4	CVCV	2	CV-CV
poteau	poto	poteau	NOM	5.17	13.04	3.88	7.70	7.44		1	1	3	11	6	CVCVVV	4	CVCV	2	CV-CV
talon	talɔ̃	talon	NOM	11.72	49.26	4.03	12.36	19.34	Haute	1	5	4	19	5	CVCVC	4	CVCV	2	CV-CV
salon	salɔ̃	salon	NOM	39.24	101.22	37.06	84.12	65.41		1	10	3	21	5	CVCVC	4	CVCV	2	CV-CV
gâtesu	gato	gâtesu	NOM	55.19	32.16	42.33	13.92	35.90	Haute	1	1	2	13	6	CVCVVV	4	CVCV	2	CV-CV
râteau	Rato	râteau	NOM	0.83	2.50	0.77	1.62	1.43		1	1	2	16	6	CVCVVV	4	CVCV	2	CV-CV
boeuf	bœf	boeuf	NOM	10.39	21.01	8.58	14.32	13.57	Basse	2	0	2	12	5	CVVVC	3	CVC	1	CVC
veuf	vœf	veuf	NOM	18.41	32.70	0.82	1.62	13.38		2	6	4	9	4	CVVC	3	CVC	1	CVC
galon	galɔ̃	galon	NOM	3.16	14.73	0.89	4.12	5.72	Basse	1	6	4	21	5	CVCVC	4	CVCV	2	CV-CV
jalon	ʒalɔ̃	jalon	NOM	0.17	2.70	0.03	1.08	0.99		1	5	2	10	5	CVCVC	4	CVCV	2	CV-CV
benne	bɛn	benne	NOM	2.38	3.11	2.12	1.76	2.34	Basse	1	6	2	25	5	CVCCV	3	CVC	1	CVC
veine	vɛn	veine	NOM	19.74	35.41	10.75	15.27	20.29		1	6	5	20	5	CVVCV	3	CVC	1	CVC
feinte	fɛ̃t	feinte	NOM	1.04	2.36	0.86	1.15	1.3525	Basse	3	5	6	15	6	CVVCCV	3	CVC	1	CVC
pinte	pɛ̃t	pinte	NOM	1.34	0.61	0.9	0.34	0.7975		2	10	7	17	5	CVCCV	3	CVC	1	CVC
cale	kal	cale	NOM	4.08	6.15	3	4.8	4.5075	Basse	2	24	9	34	4	CVCV	3	CVC	1	CVC
châle	ʃal	châle	NOM	2.23	12.09	2.08	9.32	6.43		1	1	3	18	5	CCVCV	3	CVC	1	CVC

## 4. Pour le voisement

**1** : Forme orthographique ; **2** : Forme phonologique ; **3** : Lemme ; **4** : Catégorie grammaticale ; **5** : Fréquence d'apparition du lemme dans les films (par million d'occurrences) ; **6** : Fréquence d'apparition du lemme dans les livres (par million d'occurrence) ; **7** : Fréquence d'apparition du mot dans les films (par million d'occurrences) ; **8** : Fréquence d'apparition du mot dans les livres (par million d'occurrences) ; **9** : Moyenne de toutes les fréquences (lemme et mot, livres et films), par million d'occurrences ; **10** : Niveau de fréquence (haute si >20 par million d'occurrences, basse si <20 par million d'occurrences) ; **11** : Nombre d'homographes ; **12** : Nombre de voisins orthographiques ; **13** : Nombre d'homophones ; **14** : Nombre de voisins phonologiques ; **15** : Nombre de lettres ; **16** : Structure orthographique ; **17** : Nombre de phonèmes ; **18** : Structure phonémique ; **19** : Nombre de syllabes ; **20** : Structure syllabique

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
poule	pul	poule	NOM	36.7	31.82	23.50	16.69	27.17	Haute	1	13	3	21	5	CVVCV	3	CVC	1	CVC
boule	bul	boule	NOM	30.68	61.22	19.29	38.31	37.37		2	17	6	26	5	CVVCV	3	CVC	1	CVC
palais	palɛ	palais	NOM	29.55	58.72	29.55	58.72	44.13	Haute	1	8	3	30	6	CVCCVC	4	CVCV	2	CV-CV
balai	balɛ	balai	NOM	10.91	17.7	8.24	11.96	12.20		1	2	9	26	5	CVCVV	4	CVCV	2	CV-CV
poulet	pulɛ	poulet	NOM	41.25	25.34	32.33	14.53	28.36	Haute	1	5	2	19	6	CVVCVC	4	CVCV	2	CV-CV
boulet	bulɛ	boulet	NOM	2.95	8.11	1.99	3.78	4.20		1	6	3	28	6	CVVCVC	4	CVCV	2	CV-CV
fer	fɛR	fer	NOM	37.08	114.19	33.65	106.28	72.80	Haute	1	9	5	28	3	CVC	3	CVC	1	CVC
verre	vɛR	verre	NOM	176.57	230.07	154.13	175.2	183.99		1	10	11	22	5	CVCCV	3	CVC	1	CVC
vase	vaz	vase	NOM	10.74	32.97	9.83	26.76	20.07	Haute	1	9	2	17	4	CVCV	3	CVC	1	CVC
phase	faz	phase	NOM	13.22	9.32	12.42	6.76	10.43		1	1	2	21	5	CCVCV	3	CVC	1	CVC
cale	kal	cale	NOM	4.08	6.15	3.00	4.80	4.50	Basse	2	24	9	34	4	CVCV	3	CVC	1	CVC
gale	gal	gale	NOM	3.52	0.47	3.50	0.47	1.99		1	18	6	28	4	CVCV	3	CVC	1	CVC
char	ʃaR	char	NOM	11.86	27.57	8.60	7.91	13.98	Basse	1	7	4	25	4	CCVC	3	CVC	1	CVC
jarre	ʒaR	jarre	NOM	1.25	3.92	1.17	1.49	1.957		1	4	4	28	5	CVCCV	3	CVC	1	CVC
calot	kalo	calot	NOM	0.23	5.81	0.12	4.19	2.58	Basse	1	8	3	32	5	CVCCV	4	CVCV	2	CV-CV
galop	galo	galop	NOM	2.83	13.72	2.83	12.77	8.03		1	2	6	16	5	CVCVC	4	CVCV	2	CV-CV
potin	potɛ	potin	NOM	2.02	3.85	0.61	1.55	2.00	Basse	1	2	2	8	5	CVCVC	4	CVCV	2	CV-CV
bottin	botɛ	bottin	NOM	1.05	1.35	1.05	0.68	1.03		1	0	2	9	6	CVCCVC	4	CVCV	2	CV-CV
pagne	paʃl	pagne	NOM	0.40	1.69	0.37	1.42	0.97	Basse	1	9	2	22	5	CVCCV	3	CVC	1	CVC
bagne	baʃl	bagne	NOM	2.50	3.85	2.49	3.31	3.03		1	9	2	25	5	CVCCV	3	CVC	1	CVC

## Annexe VI : Ordre de présentation des tâches

Ordre de passation des différentes tâches pour chaque participant. La plupart des participants ont passé à la fois la tâche Expérimentale « E » (avec vocodage) et sa tâche Contrôle « C » (sans vocodage). L'ordre de passation des épreuves pour chacune des deux tâches est strictement identique. Les épreuves varient en type de vocodage : non vocodé (No), vocodé à 16 bandes de fréquence (16) ou vocodé à 8 bandes de fréquence (8) ; en type de tâche : discrimination (dis) ou identification (id) ; en matériel utilisé : mots (m) ou pseudo-mots (pm).

Participant	Tâche	Epreuve 1			Epreuve 2			Epreuve 3			Epreuve 4		
1	E	8	dis	pm	8	dis	m	8	id	pm	8	id	m
	C	No	dis	pm	No	dis	m	No	id	pm	No	id	m
2	E	16	dis	pm	16	dis	m	16	id	pm	16	id	m
	C	No	dis	pm	No	dis	m	No	id	pm	No	id	m
3	E	8	dis	pm	8	dis	m	8	id	m	8	id	pm
	C	No	dis	pm	No	dis	m	No	id	m	No	id	pm
4	E	16	dis	pm	16	dis	m	16	id	m	16	id	pm
	C	No	dis	pm	No	dis	m	No	id	m	No	id	pm
5	E	8	dis	pm	8	id	m	8	id	m	8	dis	pm
	C	No	dis	pm	No	id	m	No	id	m	No	dis	pm
6	E	16	dis	pm	16	id	m	16	id	m	16	dis	pm
	C	No	dis	pm	No	id	m	No	id	m	No	dis	pm
7	E	8	dis	pm	8	id	m	8	dis	m	8	id	pm
	C	No	dis	pm	No	id	m	No	dis	m	No	id	pm
8	E	16	dis	pm	16	id	m	16	dis	m	16	id	pm
	C	No	dis	pm	No	id	m	No	dis	m	No	id	pm
9	E	8	dis	pm	8	id	pm	8	dis	m	8	id	m
	C	No	dis	pm	No	id	pm	No	dis	m	No	id	m
10	E	16	dis	pm	16	id	pm	16	dis	m	16	id	m
	C	No	dis	pm	No	id	pm	No	dis	m	No	id	m
11	E	8	dis	pm	8	id	pm	8	id	m	8	dis	m
	C	No	dis	pm	No	id	pm	No	id	m	No	dis	m
12	E	16	dis	pm	16	id	pm	16	id	m	16	dis	m
	C	No	dis	pm	No	id	pm	No	id	m	No	dis	m
13	E	8	dis	m	8	dis	pm	8	id	m	8	id	pm
	C	No	dis	m	No	dis	pm	No	id	m	No	id	pm
14	E	16	dis	m	16	dis	pm	16	id	m	16	id	pm
	C	No	dis	m	No	dis	pm	No	id	m	No	id	pm
15	E	8	dis	m	8	dis	pm	8	id	pm	8	id	m
	C	No	dis	m	No	dis	pm	No	id	pm	No	id	m
16	E	16	dis	m	16	dis	pm	16	id	pm	16	id	m
	C	No	dis	m	No	dis	pm	No	id	pm	No	id	m
17	E	8	dis	m	8	id	m	8	id	pm	8	dis	pm
	C	No	dis	m	No	id	m	No	id	pm	No	dis	pm
18	E	16	dis	m	16	id	m	16	id	pm	16	dis	pm
	C	No	dis	m	No	id	m	No	id	pm	No	dis	pm
19	E	8	dis	m	8	id	m	8	dis	pm	8	id	pm
	C	No	dis	m	No	id	m	No	dis	pm	No	id	pm
20	E	16	dis	m	16	id	m	16	dis	pm	16	id	pm
	C	No	dis	m	No	id	m	No	dis	pm	No	id	pm
21	E	8	dis	m	8	id	pm	8	dis	pm	8	id	m
	C	No	dis	m	No	id	pm	No	dis	pm	No	id	m
22	E	16	dis	m	16	id	pm	16	dis	pm	16	id	m
	C	No	dis	m	No	id	pm	No	dis	pm	No	id	m
23	E	8	dis	m	8	id	pm	8	id	m	8	dis	pm
	C	No	dis	m	No	id	pm	No	id	m	No	dis	pm

Participant	Tâche	Epreuve 1			Epreuve 2			Epreuve 3			Epreuve 4		
24	E	16	dis	m	16	id	pm	16	id	m	16	dis	pm
	C	No	dis	m	No	id	pm	No	id	m	No	dis	pm
25	E	8	id	m	8	id	pm	8	dis	m	8	dis	pm
	C	No	id	m	No	id	pm	No	dis	m	No	dis	pm
26	E	16	id	m	16	id	pm	16	dis	m	16	dis	pm
	C	No	id	m	No	id	pm	No	dis	m	No	dis	pm
27	E	8	id	m	8	id	pm	8	dis	pm	8	dis	m
	C	No	id	m	No	id	pm	No	dis	pm	No	dis	m
28	E	16	id	m	16	id	pm	16	dis	pm	16	dis	m
	C	No	id	m	No	id	pm	No	dis	pm	No	dis	m
29	E	8	di	m	8	dis	pm	8	id	pm	8	dis	m
	C	No	di	m	No	dis	pm	No	id	pm	No	dis	m
30	E	16	di	m	16	dis	pm	16	id	pm	16	dis	m
	C	No	di	m	No	dis	pm	No	id	pm	No	dis	m
31	E	8	id	m	8	dis	pm	8	dis	m	8	id	pm
	C	No	id	m	no	dis	pm	No	dis	m	No	id	pm
32	E	16	id	m	16	dis	pm	16	dis	m	16	id	pm
	C	No	id	m	no	dis	pm	No	dis	m	No	id	pm
33	E	8	id	m	8	dis	m	8	dis	pm	8	id	pm
	C	No	id	m	No	dis	m	No	dis	pm	No	id	pm
34	E	16	id	m	16	dis	m	16	dis	pm	16	id	pm
	C	No	id	m	No	dis	m	No	dis	pm	No	id	pm
35	E	8	id	m	8	dis	m	8	id	pm	8	dis	pm
	C	No	id	m	No	dis	m	No	id	pm	No	dis	pm
36	E	16	id	m	16	dis	m	16	id	pm	16	dis	pm
	C	No	id	m	No	dis	m	No	id	pm	No	dis	pm
37	E	8	id	pm	8	id	m	8	dis	m	8	dis	pm
	C	No	id	pm	No	id	m	No	dis	m	No	dis	pm
38	E	16	id	pm	16	id	m	16	dis	m	16	dis	pm
	C	No	id	pm	No	id	m	No	dis	m	No	dis	pm
39	E	8	id	pm	8	id	m	8	dis	pm	8	dis	m
	C	No	id	pm	No	id	m	No	dis	pm	No	dis	m
40	E	16	id	pm	16	id	m	16	dis	pm	16	dis	m
	C	No	id	pm	No	id	m	No	dis	pm	No	dis	m
41	E	8	id	pm	8	dis	pm	8	id	m	8	dis	m
	C	No	id	pm	No	dis	pm	No	id	m	No	dis	m
42	E	16	id	pm	16	dis	pm	16	id	m	16	dis	m
	C	No	id	pm	No	dis	pm	No	id	m	No	dis	m
43	E	8	id	pm	8	dis	pm	8	dis	m	8	id	m
	C	No	id	pm	No	dis	pm	No	dis	m	No	id	m
44	E	16	id	pm	16	dis	pm	16	dis	m	16	id	m
	C	No	id	pm	No	dis	pm	No	dis	m	No	id	m
45	E	8	id	pm	8	dis	m	8	dis	pm	8	id	m
	C	No	id	pm	No	dis	m	No	dis	pm	No	id	m
46	E	16	id	pm	16	dis	m	16	dis	pm	16	id	m
	C	No	id	pm	No	dis	m	No	dis	pm	No	id	m
47	E	8	id	pm	8	dis	m	8	id	m	8	dis	pm
	C	No	id	pm	No	dis	m	No	id	m	No	dis	pm
48	E	16	id	pm	16	dis	m	16	id	m	16	dis	pm
	C	No	id	pm	No	dis	m	No	id	m	No	dis	pm
49	E	8	dis	pm	8	dis	m	8	id	pm	8	id	m
	C	No	dis	pm	No	dis	m	No	id	pm	No	id	m
50	E	16	dis	pm	16	dis	m	16	id	pm	16	id	m
	C	No	dis	pm	No	dis	m	No	id	pm	No	id	m

---

## Annexe VII : Consignes présentées lors de la passation

### 1. Consignes écrites (pour Matlab)

<b>Discrimination</b>	<b>Identification</b>
Veillez suivre les instructions qui vont vous être présentées.	Veillez suivre les instructions qui vont vous être présentées.
Vous allez entendre 2 mots. Ecoutez-les attentivement.	Vous allez entendre 1 mot. Ecoutez-le attentivement.
Attention, le son est dégradé (à ne mettre que pour la tâche vocodeur).	Attention, le son est dégradé (à ne mettre que pour la tâche vocodeur).
Dites si les mots sont pareils ou différents.	Sélectionnez le mot que vous aurez entendu parmi les 2 propositions.
Cliquez sur la touche « a » pour sélectionner la réponse de gauche.	Cliquez sur la touche « a » pour sélectionner la réponse de gauche.
Cliquez sur la touche « p » pour sélectionner la réponse de droite.	Cliquez sur la touche « p » pour sélectionner la réponse de droite.
Répondez le plus spontanément possible.	Répondez le plus spontanément possible.
Restez bien concentré jusqu'à la fin.	Restez bien concentré jusqu'à la fin.

### 2. Consignes orales (trame des explications)

Installez-vous confortablement et mettez le casque. Nous allons vous proposer deux tâches :

#### 2.1. Discrimination

Dans cette tâche, nous allons vous faire écouter deux mots et vous devrez dire si ce sont les mêmes ou non.

Attention, les mots n'existent pas toujours dans la langue française. Par exemple : vous allez entendre « gapé »/« kapé ».

A l'écran vont apparaître les mots « pareils » et « différents ».

Pour choisir la réponse de gauche appuyez sur « a », pour choisir la réponse de droite, appuyez sur « p ».

---

Attention, les mots « pareils » et « différents » seront présentés à droite ou à gauche de l'écran de façon aléatoire.

Attention, le son que vous allez entendre est altéré (uniquement pour le groupe vocodeur).

Soyez bien attentif tout le long de la tâche et répondez le plus spontanément possible.

Ce n'est pas une tâche de rapidité, mais il ne faut pas réfléchir trop longtemps.

## **2.2. Identification**

Dans cette tâche, nous allons vous faire écouter un seul mot, puis vous proposer un choix de deux mots sur l'écran. Choisissez celui qui correspond à ce que vous avez entendu.

Attention, les mots n'existent pas toujours dans la langue française. Par exemple : vous allez entendre « gapé ». Vous devrez choisir entre « gapé » et « kapé ».

Pour le choix de réponses tapez sur « a » pour la réponse de gauche et sur « p » pour la réponse de droite.

Attention, le son que vous allez entendre est altéré (uniquement pour le groupe vocodeur).

Soyez bien attentif jusqu'à la fin de la tâche et répondez le plus spontanément possible.

Ce n'est pas une tâche de rapidité, mais il ne faut pas réfléchir trop longtemps.

---

# TABLE DES ILLUSTRATIONS

---

## I. Liste des figures

Figure 1 : Eléments anatomiques et indices acoustiques mis en jeu dans la production de la parole.....	10
a. Représentation schématique des cavités de résonance principales pour la production de la parole.....	10
b. Représentation à l’oscillogramme (haut) et au spectrogramme (bas) de la syllabe /ba/ dans le mot «bâche ».....	10
Figure 2 : Oscillogrammes de la production /aba/ (amplitude (en Hz) en fonction du temps en (ms)).....	10
Figure 3 : Représentation du système auditif.....	11
a. Représentation schématique du système auditif dans sa totalité.....	11
b. Schéma représentant la cochlée.....	11
c. Schéma de l’organe de Corti sur une coupe de la cochlée.....	11
Figure 4 : Représentation de l’installation d’un implant cochléaire.....	12
Figure 5 : Représentation du modèle TRACE.....	20
Figure 6 : Analyse spectrale d’une paire minimale avec le logiciel Praat .....	35
Figure 7 : Représentation du pourcentage de réponses correctes (moyenne et déviation standard) dans les conditions mots et pseudo-mots .....	40
Figure 8 : Représentation du pourcentage de réponses correctes (moyenne et déviation standard) dans les conditions non vocodée, vocodage à 8 bandes et vocodage à 16 bandes de fréquence ....	42
Figure 9 : Représentation du pourcentage de réponses correctes (moyenne et déviation standard) dans les conditions discrimination et identification .....	42
Figure 10 : Représentation du pourcentage de réponses correctes (moyenne et déviation standard) pour les épreuves de discrimination et d’identification des mots et pseudo-mots.....	44
Figure 11 : Représentation du pourcentage de réponses correctes (moyenne et déviation standard) pour les mots et les pseudo-mots en condition sans vocodage, avec vocodage à 8 bandes et avec vocodage à 16 bandes.....	45
Figure 12 : Représentation des pourcentages de réponses correctes (moyenne et déviation standard) pour les tâches d’identification et de discrimination en condition sans vocodage, avec vocodage à 8 bandes et avec vocodage à 16 bandes .....	47

---

Figure 13 : Représentation des temps de réponse (moyenne et déviation standard) pour les conditions mots et pseudo-mots .....	48
Figure 14 : Représentation des temps de réponse (moyenne et déviation standard) pour les différents niveaux de dégradation : pas de vocodage, vocodage à 8 bandes de fréquence et vocodage à 16 bandes de fréquence .....	49
Figure 15 : Représentation des temps de réponse (moyenne et déviation standard) pour les tâches de discrimination et d'identification .....	50
Figure 16 : Représentation des temps de réponse (moyenne et déviation standard) selon le type de tâche (discrimination ou identification) et le type de matériel (mots ou pseudo-mots).....	51
Figure 17 : Représentation des temps de réponse (moyenne et déviation standard) selon les différents niveaux de dégradation (pas de vocodage, vocodage à 8 bandes de fréquence et vocodage à 16 bandes de fréquence) et le type de matériel (mots et pseudo-mots) .....	52
Figure 18 : Représentation des temps de réponse (moyenne et déviation standard) selon le niveau de dégradation (pas de vocodage, vocodage à 8 bandes de fréquence, vocodage à 16 bandes de fréquence) et le type de tâche (discrimination ou identification).....	53
Figure 19 : Comparaison des représentations acoustiques de /b/ (à gauche) et /m/ (à droite), dans la paire minimale « bâche »/« mâche », issue de notre matériel.....	78
Figure 20 : Comparaison des représentations acoustiques de /t/ (à gauche) et /s/ (à droite) dans la paire minimale « talon »/« salon », issue de notre matériel.....	79
Figure 21 : Représentation de l'oscillogramme et du spectrogramme du mot « marron » dans la condition non vocodée .....	80
Figure 22 : Représentation de l'oscillogramme et du spectrogramme du mot « marron » dans la condition vocodée à 16 bandes .....	80
Figure 23 : Représentation de l'oscillogramme et du spectrogramme du mot « marron » dans la condition vocodé à 8 bandes .....	81

---

---

## II. Liste des tableaux

Tableau I : Correspondance entre traits articulatoires et indices acoustiques principaux.....	15
Tableau II: Tableau représentant les critères de sélection des paires lexicales .....	34
Tableau III : Présentation des moyennes et déviation standard des pourcentages de réponses correctes pour les tâches de discrimination et d'identification selon les différents niveaux de vocodage .....	46
Tableau IV : Relevé des résultats des différentes caractéristiques du signal acoustique (non vocodé, vocodé à 16 bandes et vocodé à 8 bandes) .....	82

---

# TABLE DES MATIERES

---

<b>ORGANIGRAMMES</b> .....	<b>2</b>
1- UNIVERSITE CLAUDE BERNARD LYON1 .....	<b>ERREUR ! SIGNET NON DEFINI.</b>
1.1. Secteur Santé : .....	<i>Erreur ! Signet non défini.</i>
1.2. Secteur Sciences et Technologies : .....	<i>Erreur ! Signet non défini.</i>
2- INSTITUT SCIENCES ET TECHNIQUES DE READAPTATION FORMATION ORTHOPHONIE	<b>ERREUR ! SIGNET NON DEFINI.</b>
<b>REMERCIEMENTS</b> .....	<b>4</b>
<b>SOMMAIRE</b> .....	<b>5</b>
<b>INTRODUCTION</b> .....	<b>7</b>
<b>PARTIE THEORIQUE</b> .....	<b>9</b>
I. L'IMPLANT COCHLEAIRE : ENJEUX ACTUELS.....	10
1. Le signal sonore.....	10
2. L'audition normale .....	11
3. Le fonctionnement de l'implant cochléaire .....	12
4. L'implant cochléaire : indications et limites .....	13
II. LA PERCEPTION DE LA PAROLE .....	14
1. Perception de la parole, phonèmes et traits articulatoires .....	14
2. Description acoustique et articulatoire des phonèmes du français.....	15
3. Le traitement phonémique de la parole .....	16
3.1. Perception et précision catégorielle : définition.....	17
3.2. Perception et précision catégorielle : décours développemental.....	17
3.3. Perception et précision catégorielle : intégration des indices acoustiques.....	18
4. Le traitement lexical de la parole .....	19
4.1. Les principaux modèles de reconnaissance de la parole .....	19
4.2. Effet du lexique sur le traitement phonémique .....	20
III. LA PERCEPTION DE LA PAROLE EN CONTEXTE DE SIGNAL DEGRADE .....	21
1. La perception de la parole de l'enfant implanté .....	21
2. La perception de la parole de l'adulte implanté .....	22
3. Le vocodeur : principe et application.....	23
<b>PROBLEMATIQUE ET HYPOTHESES</b> .....	<b>25</b>
I. PROBLEMATIQUE .....	26
II. HYPOTHESES .....	27
1. Hypothèse générale .....	27
1.1. Les variables indépendantes (VI).....	27
1.2. Les variables contrôles (VC).....	27
1.3. Les variables dépendantes.....	27
2. Hypothèses opérationnelles .....	28
<b>PARTIE EXPERIMENTALE</b> .....	<b>29</b>
I. POPULATION .....	30
II. MATERIEL.....	32
III. TACHES .....	36
IV. PASSATION.....	36
<b>PRESENTATION DES RESULTATS</b> .....	<b>38</b>
I. PREAMBULE .....	39
II. ANALYSE POUR LES SCORES DE REPONSES CORRECTES .....	40
1. Résultats des effets principaux.....	40
1.1. L'effet de lexicalité .....	40
1.2. L'effet de la dégradation du signal.....	41
1.3. L'effet du type de tâche.....	42
2. Résultats des interactions .....	43
2.1. L'interaction entre la lexicalité et le type de tâche.....	43
2.2. L'interaction entre la lexicalité et la dégradation du signal .....	44

---

2.3. L'interaction entre la dégradation du signal et le type de tâche.....	45
III. ANALYSE POUR LES SCORES DE TEMPS DE REPONSE.....	47
1. <i>Résultats des effets principaux</i> .....	47
1.1. L'effet de lexicalité.....	47
1.2. L'effet de la dégradation du signal.....	48
1.3. L'effet du type de tâche.....	49
2. <i>Résultats des interactions</i> .....	50
2.1. L'interaction entre la lexicalité et le type de tâche.....	50
2.2. L'interaction entre la lexicalité et la dégradation du signal.....	51
2.3. L'interaction entre la dégradation du signal et le type de tâche.....	52
<b>DISCUSSION DES RESULTATS.....</b>	<b>54</b>
I. PREAMBULE.....	55
II. INTERPRETATION DES RESULTATS.....	56
1. <i>L'effet du lexique</i> .....	56
1.1. L'influence du lexique.....	56
1.2. L'influence du lexique sur la perception catégorielle et la précision catégorielle.....	57
1.3. L'influence du lexique en présence de signal dégradé.....	58
1.4. L'influence du lexique selon le niveau de dégradation du signal.....	59
1.5. L'influence du lexique sur les temps de réponse.....	60
1.6. Conclusion sur l'influence du lexique.....	60
2. <i>L'effet de dégradation du signal</i> .....	61
2.1. L'influence de la dégradation sur le pourcentage de réponses correctes.....	61
2.2. L'influence de la dégradation sur les temps de réponse.....	62
2.3. L'influence de la dégradation du signal sur la perception catégorielle.....	62
III. CRITIQUES ET LIMITES DE NOTRE EXPERIMENTATION.....	63
1. <i>Pour notre population</i> .....	63
2. <i>Pour notre passation</i> .....	63
3. <i>Pour l'analyse des données</i> .....	64
IV. APPORTS DE L'ETUDE POUR LA PRATIQUE ORTHOPHONIQUE.....	65
<b>CONCLUSION.....</b>	<b>69</b>
<b>BIBLIOGRAPHIE.....</b>	<b>70</b>
<b>ANNEXES.....</b>	<b>76</b>
<b>LISTE DES ANNEXES.....</b>	<b>77</b>
ANNEXE I : ANALYSE SPECTRALE DE DEUX PAIRES MINIMALES ISSUES DE NOTRE MATERIEL.....	78
ANNEXE II : REPRESENTATIONS DU MOT « MARRON » EN CONDITIONS NON VOCODEE ET VOCODEE.....	80
ANNEXE III : TABLEAU RECAPITULATIF DU MATERIEL EXPERIMENTAL.....	83
ANNEXE IV : DOCUMENT D'INFORMATION POUR NOTRE RECHERCHE DE POPULATION.....	85
ANNEXE V : TABLEAU DE CONSTRUCTION DES PAIRES.....	86
1. <i>Pour la nasalité</i> .....	86
2. <i>Pour le lieu articulatoire</i> .....	87
3. <i>Pour le mode articulatoire</i> .....	88
4. <i>Pour le voisement</i> .....	89
ANNEXE VI : ORDRE DE PRESENTATION DES TACHES.....	90
ANNEXE VII : CONSIGNES PRESENTEES LORS DE LA PASSATION.....	92
1. <i>Consignes écrites (pour Matlab)</i> .....	92
2. <i>Consignes orales (trame des explications)</i> .....	92
2.1. Discrimination.....	92
2.2. Identification.....	93
<b>TABLE DES ILLUSTRATIONS.....</b>	<b>94</b>
I. LISTE DES FIGURES.....	94
II. LISTE DES TABLEAUX.....	96
<b>TABLE DES MATIERES.....</b>	<b>97</b>

---

Pauline Jobit et Nelly Joye

## **INFLUENCE DU LEXIQUE SUR LA PERCEPTION PHONEMIQUE DANS UN CONTEXTE DE SIGNAL DEGRADE**

98 Pages

Mémoire d'Orthophonie n°1691 -UCBL-ISTR- Lyon 2013

---

### **RESUME**

---

La perception de la parole est résistante aux distorsions acoustiques (accent, bruits environnants), reflétant ainsi les compétences de l'adulte à ajuster ces traitements perceptifs en fonction du contexte auditif. Cet ajustement peut accentuer certains traitements automatiques de la parole, comme l'influence du lexique sur la perception phonémique.

Notre projet s'inscrit dans une perspective de compréhension des mécanismes de perception de la parole. Il s'intéresse au signal fourni par l'implant cochléaire avec pour objectif une meilleure compréhension des mécanismes de perception mis en place par les sujets implantés. Il se base sur le modèle TRACE (McClelland et Elman, 1986) qui postule l'existence d'effets « top-down » du lexique sur les phonèmes.

Nous avons évalué l'influence des représentations lexicales sur l'identification et la discrimination de paires minimales de mots (exemple : talon/salon) et de pseudo-mots (exemple : tadon/sadon) vocodés. Le vocodage est une manipulation qui supprime les détails spectraux du signal de parole, dégradant ainsi les compétences des adultes à percevoir la parole. Les manipulations de vocodage utilisées dans ce mémoire correspondent à des simulations de ce qu'un adulte sourd pourrait entendre avec un implant cochléaire.

Les résultats n'indiquent pas de différence de perception catégorielle puisque la différence de scores entre discrimination et identification est similaire lorsque les participants entendent un signal vocodé ou ordinaire. Mais les résultats montrent une différence de précision catégorielle puisque les participants ont de moins bonnes performances d'identification et de discrimination de paires minimales en condition où le signal est vocodé par rapport à la condition où le signal n'est pas vocodé. De plus, l'effet du lexique, absent en condition contrôle, apparaît en condition où le signal est dégradé (vocodage). Nos résultats suggèrent l'existence de mécanismes de compensation basés sur le lexique, mis en jeu dans la perception de la parole en condition d'écoute dégradée.

---

### **MOTS-CLES**

---

Perception de la parole – Perception Phonémique – Représentations Lexicales – Perception catégorielle – Précision catégorielle – Implant cochléaire – Vocodeur

---

### **MEMBRES DU JURY**

---

Stéphanie Colin – Marie Ozil – Eric Truy

---

### **MAITRE DE MEMOIRE**

---

Sophie Bouton – Anne-Dominique Lodeho

---

### **DATE DE SOUTENANCE**

---

27 juin 2013

---