Communauté UNIVERSITÉ Grenoble Alpes

THÈSE

Pour obtenir le grade de

DOCTEUR DE L'UNIVERSITÉ GRENOBLE ALPES

Spécialité : **Sciences du Langage – Linguistique, sociolinguistique et acquisition du langage**

Arrêté ministériel : 7 août 2006

Présentée par

Bénédicte GRANDON

Thèse dirigée par **Anne VILAIN**

Préparée au sein du **Laboratoire Grenoble, Images, Parole, Signal & Automatique (GIPSA-Lab, UMR 5216)**
et de **l'École Doctorale Lettres, Langues et Sciences Humaines**

Développement typique et atypique de la production de parole :

caractéristiques segmentales et intelligibilité de la parole d'enfants porteurs d'un implant cochléaire et d'enfants normo-entendants de 5 à 11 ans

Thèse soutenue publiquement le **9 décembre 2016**,
devant le jury composé de :

Madame Elisabetta CARPITELLI
Professeur, Université Grenoble Alpes, Présidente
Monsieur Steven GILLIS
Professeur, Université d'Anvers, Examinateur
Madame Sophie KERN
Chargée de Recherche, HDR, Université de Lyon, Rapporteur
Madame Christine MEUNIER
Chargée de Recherche, HDR, Aix-Marseille Université, Rapporteur
Monsieur Eric TRUY
Professeur des Universités – Praticien Hospitalier, Groupement Hospitalier Edouard Herriot (Lyon), Examinateur
Madame Anne VILAIN
Maître de Conférences, Université Grenoble Alpes, Directrice de thèse

Cette thèse a été financée
par une allocation doctorale ARC2
de la Région Rhône-Alpes

RÉSUMÉ

L'implant cochléaire est l'un des dispositifs de remédiation proposés dans le cadre de la surdité de l'adulte et de l'enfant, dans le cadre d'une surdité congénitale ou acquise. En 2010, plus de 200 000 personnes dans le monde en étaient équipées, dont plus de 10 000 en France (adultes et enfants). La technologie utilisée pour les premiers implants commercialisés au milieu des années 1980 a beaucoup évolué et l'implant cochléaire permet désormais à son utilisateur d'avoir accès à des caractéristiques acoustiques de plus en plus précises des sons de son environnement et notamment des sons de parole. Cependant, l'information auditive fournie par l'implant reste limitée, ce qui a pour conséquence des difficultés persistantes de production de certains sons de parole par l'utilisateur d'implant cochléaire, même après plusieurs années d'utilisation. Ces difficultés de production peuvent se traduire également par une intelligibilité moindre, et peuvent avoir des répercussions sur les relations familiales et sociales, en particulier chez l'enfant.

Les études disponibles dans la littérature se concentrent sur les effets et les bénéfices à court-terme de l'implant cochléaire chez l'enfant, et il existe relativement peu d'études de production de parole chez l'enfant en âge scolaire, en particulier chez l'enfant francophone. L'objectif de ce travail est donc de proposer une évaluation des difficultés de production de plusieurs contrastes phonologiques chez l'enfant sourd, porteur d'implant cochléaire, plusieurs années après l'implantation cochléaire, et des facteurs qui influencent son intelligibilité.

Dans le cadre de notre thèse, nous avons constitué un corpus de productions de parole de 13 enfants âgés de 6;6 à 10;7 ans, atteints de surdité pré- ou périlinguistique, ayant reçu un implant cochléaire entre 1;1 et 6;6 ans, et l'utilisant depuis plus d'un an et de 20 enfants normo-entendants appariés en âge chronologique (de 5;7 à 10;6 ans).

Dans un premier temps, nous avons comparé les caractéristiques acoustiques des voyelles orales, des occlusives et des fricatives du français, ainsi que la réalisation de la coarticulation dans des séquences occlusive-voyelle par ces deux groupes d'enfants. Les résultats montrent une grande proximité entre les productions des enfants typiques et implantés, et certaines différences, par exemple sur les voyelles antérieures arrondies, les fricatives alvéolaires et les occlusives vélaires. Ces différences peuvent s'expliquer par les caractéristiques technologiques de l'implant et son usage par l'enfant, et révèlent également le poids des différents facteurs du développement phonologique : contraintes articulatoires, contraintes perceptives, caractéristiques de l'input langagier et de la langue maternelle.

Dans un deuxième temps, nous avons élaboré une méthode d'évaluation perceptive de l'intelligibilité de la parole, que nous avons soumise à 9 auditeurs experts en parole pathologique et à 17 auditeurs naïfs, tous francophones. Notre étude perceptive d'intelligibilité met en évidence 1) une absence d'effet d'expertise sur le jugement d'intelligibilité puisque les notes données par les auditeurs experts et naïfs sont corrélées, 2) un effet de l'audition sur l'intelligibilité, puisque les enfants sourds porteurs d'implant cochléaire sont jugés moins intelligibles que les enfants normo-entendants, et 3) une meilleure intelligibilité chez les enfants implantés précocement (avant 20 mois) mais pas d'effet de la durée d'utilisation de l'implant cochléaire sur l'intelligibilité.

Ce travail montre donc le bénéfice apporté par l'implant pour la communication orale, mais aussi l'existence de difficultés persistantes, qui doivent être prises en compte dans la rééducation et l'accompagnement familial, scolaire et social des enfants.

Notre étude fournit en outre un ensemble de données de référence sur le développement phonologique tardif des enfants francophones, et un corpus de parole utilisable pour d'autres travaux de recherche sur le développement typique et pathologique.

ABSTRACT

Cochlear implantation has developed as one of the options available in case of congenital or acquired deafness in adults and children. As of 2010, cochlear implant has been used by over 200 000 persons (adults and children) worldwide and by 10 000 persons in France. Technology has largely improved since the first devices were put on the market in the 1980's, and it now provides its user with more and more detailed acoustical information about the sounds of their environment in general, and about speech sounds in particular. However, perception with a cochlear implant remains limited, which leads to persisting difficulties in producing some speech sounds, even after several years of implant use. These difficulties in speech production can translate into a lower intelligibility, with effects on interactions with family and community members for its users, especially for children.

A large body of studies in the literature focusses on short-term effects and gains of cochlear implant for young children and infants, but fewer studies in speech production in school-age children are available, especially in French-speaking children. The goal of this project is to assess the difficulties in the production of selected phonological contrasts in cochlear-implanted children, several years after cochlear implantation, and the factors influencing their intelligibility.

For this thesis, we recorded thirteen 6;6-to-10;7 year old, pre- or perilinguistically deaf children wearing cochlear implants since they were 1;1 to 6;6 years old, with more than a year of use, and twenty normal-hearing age-matched children (chronological age ranging from 5;7 to 10;6 years).

In a first experiment, we compared the acoustical characteristics of ten oral vowels, six stops and three fricatives of French. We also studied the realization of coarticulation in plosive-vowel sequences in these two groups of children.

Our results show a high degree of similarity in typical and implant using children, and several differences, for example, for front rounded vowels, alveolar fricatives and velar stops. These differences can be explained by the implant's technology and its use by the child. They also show the importance of several factors in phonological development: articulatory constraints, perceptual constraints, characteristics of language input and mother tongue.

In a second experiment, we designed a method to perceptually assess the intelligibility of speech: 9 expert and 17 naïve French-speaking listeners participated in this experiment.

Our perceptual assessment of intelligibility shows 1) no effect of expertise since grades given by expert and naïve listeners are correlated, 2) an effect of hearing level on intelligibility since cochlear-implanted children are less intelligible than their normal-hearing peers, and 3) a better intelligibility in early implanted children (before 20 months) but no effect of duration of implant use on intelligibility.

This work emphasizes the benefits for oral communication provided by cochlear implants, but also the remaining difficulties, that need to be taken into acount for rehabilitation and support from the child's family, school and society.

Our study also provides both reference data on late stages of phonological development in French-speaking children, and a corpus of speech production which can be used for further research on typical and pathological speech development.

REMERCIEMENTS

Je tiens tout d'abord à remercier ma directrice de thèse, Anne Vilain. Merci de m'avoir confié ce projet et de m'avoir permis de le mener à bien. Merci de ton encadrement et de tes conseils. Merci de m'avoir appris à mettre les résultats de mon travail en valeur et de m'avoir encouragée à aller les présenter dans des conférences à Londres, Halifax, Paris, Lyon... Enfin, merci de ta disponibilité dès ton retour au labo après quelques mois d'absence et de ton investissement dans la lecture et la correction rapide et efficace des quelque 300 pages que je t'ai envoyées à la fin de l'été !

Je remercie ensuite les rapporteurs, Sophie Kern et Christine Meunier d'avoir accepté de lire et d'évaluer mon travail. Merci également à Steven Gillis et Eric Truy d'avoir participé au jury de ma soutenance en tant qu'examinateurs, ainsi qu'à Elisabetta Carpitelli de l'avoir présidé. Merci à tous pour vos commentaires, questions et conseils lors de la soutenance, qui me permettent de continuer à réfléchir aux implications de mon travail et d'envisager de nouvelles pistes de travail.

Merci à Michel Hoen et Hélène Loevenbruck d'avoir participé à l'élaboration de ce projet. Merci au Professeur Sébastien Schmerber ainsi qu'aux orthophonistes Anne-Marie Fluttaz et Catherine Rebière, du CHU de Grenoble et au Professeur Eric Truy et à Florence Dambelin de l'Hôpital Edouard Herriot de Lyon, de m'avoir aidée à recruter les enfants. Merci à Adeline Leclercq-Samson, Marie-José Martinez et Silvain Gerber pour leurs conseils statistiques. Merci aux stagiaires Corinne, Marine, Myriam, Noémie et Vanessa pour leur aide, à différents stades du projet.

Un grand merci à tous les enfants qui ont participé à l'étude, ainsi qu'à leurs parents qui m'ont permis de les enregistrer. Merci également à tous les adultes qui ont participé aux deux études.

Merci à toutes les personnes que j'ai côtoyées lors de ces trois années au Gipsa, en particulier la fine équipe du bâtiment Stendhal, par ordre alphabétique et en essayant de n'oublier personne : Avril, Carole, Christine, Christophe, Coriandre, Diane, Elisabetta, Florie, Geneviève, Giovanni, Hien, Jean-Pierre, Julien, Katja, Laurence, Leslie, Lucie, Marine, Marjorie, NathalieH, NathalieV, Silvia, Yanchen. Merci pour les pauses, les expériences gastronomiques… Merci aussi aux filles du bureau E004 d'avoir supporté ma « bonne » humeur pendant ces trois ans !

Un merci tout particulier à Elisabetta. Tu as été ma tutrice de monitorat pendant mes deuxième et troisième années, je ne te remercierai jamais assez pour les conseils et les discussions sur l'enseignement, la rédaction, les relations professionnelles… et ta porte toujours ouverte.

Un gros, gros merci à Florie, Geneviève et Silvia, pour les bavardages, les sorties ski, les restos, etc. on s'est bien amusées quand même et j'espère que ça va continuer ! Merci à Hien pour les bons moments passés au labo.

Merci également à Jérôme Mars, directeur du Gipsa-lab et Denis Beautemps, responsable du Département Parole et Cognition, ainsi qu'à Christian Commault et Cécilia Mendes pour leur aide et pour avoir fait en sorte que je puisse terminer ma thèse dans de bonnes conditions.

Merci à Hiyon Yoo, Elisabeth Delais-Roussarie et Jacqueline Vaissière de m'avoir permis d'assister à la conférence ISSP2014 à Cologne et de pouvoir y présenter mon travail de Master, grâce à l'aide financière du labex EFL.

Merci à mes amies de toujours, Anne-Laure, Marie-Anne et Dorothée, ainsi qu'à la petite bande de Paris 7, Xin, Lucie, Qianwen et Zeke, pour votre amitié longue-distance.

Mille mercis à Hiyon, grâce à qui j'ai découvert la phonétique, la gastronomie coréenne et la vie à Séoul. Merci d'avoir dirigé mes premières tentatives de recherche et de m'avoir encouragée à continuer dans cette voie. Juste une question : maintenant que cette thèse est finie, quand est-ce qu'on retravaille ensemble ?

Merci à mes parents, à mes frères et à ma sœur… Merci de m'avoir soutenue et supportée chacun à votre façon dans les difficultés, en particulier ces dix-sept dernières années. Par contre, je suis désolée de vous décevoir mais je crois que je ne finirai probablement jamais cette fichue maîtrise d'allemand !

Enfin, merci à toutes les personnes qui sont venues, de Blois, Paris, Thônes, Strasbourg… et même de Grenoble (!) assister à ma soutenance malgré le froid et le changement de salle de dernière minute.

TABLE DES MATIÈRES

INTRODUCTION 9

PREMIERE PARTIE : Approches théoriques 13
CHAPITRE 1 : Théories d'acquisition de la langue maternelle et développement phonologique 15
1. Introduction 15
2. Quelques théories de l'acquisition de la langue maternelle 15
2.1. Behaviourisme 15
2.2. Nativisme/cognitivisme 16
2.3. (Socio-)constructivisme 17
3. Développement de la parole : perception 20
3.1. Rétrécissement perceptif 20
3.2. Perception multimodale 22
4. Développement de la parole : production 24
4.1. Etapes du développement de la production de parole 24
4.2. Ordre d'apparition des phonèmes 26
5. Développement phonologique : quelques modèles 28
5.1. Modèles descendants (ou « top-down ») du développement phonologique 28
5.2. Modèles ascendants (ou « bottom-up ») du développement phonologique 29
5.3. Modèles de développement plus tardif 32
6. Conclusion 33
CHAPITRE 2 : Audition, surdité et implant cochléaire 34
1. Introduction 34
2. Fonctionnement non pathologique de l'oreille 34
2.1. Oreille externe 35
2.2. Oreille moyenne 35
2.3. Oreille interne 35
2.3.1. Composition de l'oreille interne 35
2.3.2. Fonctionnement de la cochlée 35
3. Surdité 37
3.1. Caractéristiques de la surdité 37
3.1.1. Surdités unilatérales et bilatérales, âge de survenue de la surdité, étiologies 37
3.1.2. Degré de surdité 38
3.1.3. Types de surdité 38
3.1.3.1. Surdité de transmission 38
3.1.3.2. Surdité de perception 38
3.1.3.3. Surdité centrale 38
3.2. Types de dispositifs de remédiation auditive disponibles 39
4. Implant cochléaire 40
4.1. Composition de l'implant cochléaire 40
4.2. Fonctionnement de l'implant cochléaire et stratégies de codage 41
4.2.1. Fonctionnement général de l'implant cochléaire 41
4.2.2. Stratégies de codage des fréquences des sons de parole par l'implant cochléaire 42
5. Conséquence de l'implant sur les capacités de communication et la qualité de vie 44
5.1. Capacités perceptives 44
5.2. Capacités de communication et qualité de vie 45
5.2.1. Conséquences de l'implantation sur les capacités langagières de l'enfant 46
5.2.2. Conséquences de l'implantation sur la qualité de vie de l'enfant et sa famille 46
5.2.3. Modes de communication des personnes sourdes 47
5.2.4. Scolarité et intégration sociale de l'enfant porteur d'implant cochléaire 48
5.2.5. Problèmes éthiques soulevés par l'implant cochléaire 49
6. Conclusion 50
CHAPITRE 3 : production de parole et intelligibilité de la parole des enfants porteurs d'implants cochléaires : état de l'art 51
1. Caractéristiques segmentales de la parole d'enfants porteurs d'implants cochléaires 51
1.1. Voyelles 51
1.1.1. Introduction 51
1.1.2. Etudes perceptives / inventaires de voyelles 51
1.1.3. Etudes acoustiques 53

1.1.3.1. Etudes longitudinales 53
- Premier et deuxième formants (F1 et F2) 53
- Espace vocalique 54

1.1.3.2. Implants ON vs OFF vs enfants NH 56
- Implant ON vs OFF 56
- Implant ON vs enfants NH 56

1.1.3.3. NH vs CI vs HA 56
- Méthodologies 58
- Premier formant (F1) 59
- Deuxième formant (F2) 60
- Espace vocalique 61
- Autres résultats 64

1.1.4. Conclusion 64
1.2. Consonnes 65
1.2.1. Introduction 65
1.2.2. Etudes perceptives : inventaires de consonnes produites par les enfants porteurs d'implants cochléaires 65
1.2.3. Etudes acoustiques 69
1.2.3.1. Occlusives 69
1.2.3.2. Fricatives 71
1.2.4. Conclusion 75
1.3. Coarticulation 75
1.3.1. Introduction 75
1.3.2. Etude perceptive : types d'erreurs et évolution de la production d'enfants CI 76
1.3.3. Etudes acoustiques de la coarticulation 77
1.3.4. Conclusion 79
2. Intelligibilité de la parole d'enfants porteurs d'implants cochléaires : évaluation et facteurs explicatifs 80
2.1. Facteurs explicatifs du niveau d'intelligibilité 80
2.1.1. Corrélats de l'intelligibilité dans la parole non-pathologique 80
2.1.2. Facteurs explicatifs de l'intelligibilité dans la parole d'enfants porteurs d'implants cochléaires 81
2.1.2.1. Facteurs individuels ou environnementaux 81
2.1.2.2. Corrélats phonétiques 81
2.2. Résultats des études d'intelligibilité 82
2.2.1. Facteurs individuels et environnementaux de l'intelligibilité 82
2.2.1.1. Durée d'implantation, âge d'implantation et âge chronologique 82
2.2.1.2. Caractéristiques de l'implant 83
2.2.1.3. Capacités perceptives et niveau d'audition 84
2.2.1.4. Caractéristiques familiales, modes de communication et nombre de L1 84
2.2.2. Corrélats acoustiques de l'intelligibilité 84
2.2.2.1. Caractéristiques segmentales 85
2.2.2.2. Caractéristiques suprasegmentales et qualité de voix 87
2.3. Conclusion 88
3. Notre projet 89
3.1. Problématique et hypothèses de l'étude de production 89
3.1.1. Problématique 89
3.1.2. Hypothèses 90
3.1.2.1. Effet du groupe sur la réalisation des segments et de la coarticulation par les enfants CI et NH 90
3.1.2.2. Effet de l'âge chronologique sur la réalisation des segments et de la coarticulation par les enfants NH 90
3.1.2.3. Effet de l'âge à l'implantation et de la durée d'utilisation de l'implant sur la réalisation des segments et de la coarticulation par les enfants CI 91
3.1.2.4. Effet de l'input auditif immédiat sur la réalisation des segments (consonnes et voyelles) et de la coarticulation CV 91
3.1.2.5. Effet de la complexité du mot sur la réalisation des segments (consonnes et voyelles) 91
3.2. Problématique et hypothèses de l'étude d'intelligibilité 92

DEUXIÈME PARTIE : Production de parole (1ère étude expérimentale) 93
CHAPITRE 4 : Étude de production (1ère étude expérimentale) : méthodologie 95
1. Recueil des données 95
1.1. Participants 95

1.1.1. Enfants normo-entendants 95
1.1.1.1. Recrutement des enfants normo-entendants 95
1.1.1.2. Age et origine géographique des enfants normo-entendants 96
1.1.1.3. Audition des enfants normo-entendants 96
1.1.1.4. Développement langagier des enfants normo-entendants 96
1.1.2. Enfants sourds, porteurs d'implant cochléaire 96
1.1.2.1. Recrutement des enfants sourds, porteurs d'implant cochléaire 96
1.1.2.2. Age et origine géographique des enfants sourds, porteurs d'implant cochléaire 97
1.1.2.3. Audition des enfants sourds, porteurs d'implant cochléaire 97
1.1.2.4. Modes de communication des enfants sourds, porteurs d'implant cochléaire 98
1.2. Protocole 99
1.2.1. Tâches 99
1.2.1.1. Motivation du choix des tâches 99
1.2.1.2. Répétition 100
1.2.1.3. Dénomination 101
1.2.1.4. Narration 101
1.2.2. Corpus 102
1.2.2.1. Critères de choix des mots utilisés dans les tâches de répétition et de dénomination 102
1.2.2.2. Choix des mots du corpus pour la tâche de répétition 103
1.2.2.3. Choix des mots du corpus pour la tâche de dénomination 103
1.2.2.4. Listes des mots du corpus pour la tâche de répétition et la tâche de dénomination 103
1.3. Procédure d'enregistrement 106
1.3.1. Matériel, logiciels et conditions de passation 106
1.3.2. Elaboration des stimuli de la tâche de répétition 106
1.3.3. Lieux d'enregistrement 107
1.4. Traitement des enregistrements 107
1.4.1. Organisation du corpus 107
1.4.2. Annotations 107
2. Analyses acoustiques 109
2.1. Voyelles orales : formants (F1, F2, F3), espaces vocaliques F1-F2 et F2-F3 109
2.1.1. Description du système vocalique du français 109
2.1.1.1. Caractéristiques articulatoires des voyelles du français 109
2.1.1.2. Caractéristiques acoustiques des voyelles du français 111
2.1.2. Mesures des formants 114
2.1.3. Normalisation des valeurs de formants : étude préliminaire 115
2.1.3.1. Introduction 115
2.1.3.2. Variation dans la production de voyelles chez des adultes et chez des enfants 116
2.1.3.3. Etudes de systèmes vocaliques chez des enfants et jeunes adultes sourds 118
2.1.3.4. Comparaisons des méthodes de normalisation les plus fréquemment utilisées dans la littérature 121
2.1.3.5. Notre étude : comparaison de six méthodes de normalisation et application à notre problématique 123
- Corpus 123
- Méthodes de normalisation testées 124
- Comparaison : analyses et résultats 126
2.1.3.6. Conclusion et limites 127
2.2. Consonnes : VOT, moments spectraux 127
2.2.1. Description du système consonantique du français : occlusives 128
2.2.2. Description du système consonantique du français : fricatives 129
2.2.3. Caractéristiques acoustiques des consonnes occlusives 130
2.2.3.1. Définition du VOT dans la littérature 130
2.2.3.2. Caractéristiques du VOT en français 131
2.2.3.3. VOT dans les études développementales 132
2.2.3.4. Mesure du VOT 134
2.2.4. Moments spectraux 135
2.2.4.1. Caractérisation acoustique des fricatives 135
2.2.4.2. Moments spectraux dans les études développementales 135
2.2.4.3. Mesure des moments spectraux 139
2.3. Coarticulation 139
2.3.1. Définition de l'équation de locus dans la littérature 140
2.3.2. Utilisation de la pente de l'équation de locus dans des études développementales 143
2.3.3. Calcul de la pente de l'équation de locus 146

3. Analyses statistiques 148
CHAPITRE 5 : Étude de production (1ère étude expérimentale) : résultats 151
1. Voyelles 151
1.1. Données analysées 151
1.2. Effets du groupe et de la catégorie vocalique 154
1.2.1. F1 155
1.2.2. F2 157
1.2.3. F3 159
1.2.4. Dispersion F1-F2 160
1.2.5. Dispersion F2-F3 161
1.2.6. Focalisation : différences F2-F1 et F3-F2 162
1.2.7. Résumé : effets du groupe et de la catégorie vocalique 164
1.3. Enfants NH : effet de l'âge chronologique 165
1.3.1. F1 165
1.3.2. F2 166
1.3.3. F3 167
1.3.4. Dispersion F1-F2 167
1.3.5. Dispersion F2-F3 167
1.3.6. Résumé : effets de l'âge chronologique (enfants NH) 168
1.4. Enfants CI : effet de l'âge à l'implantation et de la durée d'utilisation de l'implant 168
1.4.1. F1 168
1.4.2. F2 169
1.4.3. F3 170
1.4.4. Dispersion F1-F2 171
1.4.5. Dispersion F2-F3 172
1.4.6. Résumé : effets de l'âge à l'implantation et de la durée d'utilisation de l'implant (enfants CI) 173
1.5. Effet de l'input auditif immédiat 173
1.5.1. F1 173
1.5.2. F2 174
1.5.3. F3 175
1.5.4. Dispersion F1-F2 176
1.5.5. Dispersion F2-F3 177
1.5.6. Résumé : effet de l'input auditif immédiat 177
1.6. Effet de la complexité du mot 178
1.6.1. F1 178
1.6.2. F2 179
1.6.3. F3 180
1.6.4. Dispersion F1-F2 181
1.6.5. Dispersion F2-F3 182
1.6.6. Résumé : effet de la complexité du mot 183
1.7. Discussion 184
1.7.1. Production de voyelles chez les enfants NH 184
1.7.2. Production de voyelles chez les enfants CI 185
1.7.2.1. Effet de l'implant : groupes CI vs NH 185
1.7.2.2. Effet de l'âge à l'implantation et de l'âge auditif 186
1.7.2.3. Effet de la tâche et de la complexité du mot 186
1.7.3. Focalisation F1-F2 et F2-F3 chez les deux groupes d'enfants 187
1.8. Conclusion 187
2. Occlusives 189
2.1. Données analysées 189
2.2. Effets du groupe et du lieu d'articulation 190
2.2.1. VOT 190
2.2.2. Distance entre VOT des occlusives voisées et non-voisées 192
2.2.3. Résumé : effets du groupe et du lieu d'articulation 193
2.3. Effet du contexte vocalique 194
2.3.1. VOT 194
2.3.2. Distance entre VOT des occlusives voisées et non-voisées 195
2.3.3. Résumé : effet du contexte vocalique 196
2.4. Enfants NH : effet de l'âge chronologique 197
2.4.1. VOT 197
2.4.2. Distance entre VOT des occlusives voisées et non-voisées 198

2.4.3. Résumé : effet de l'âge chronologique (enfants NH) ... 199
2.5. Enfants CI : effet de l'âge à l'implantation et de la durée d'utilisation de l'implant ... 200
2.5.1. VOT ... 200
2.5.2. Distance entre VOT des occlusives voisées et non-voisées ... 201
2.5.3. Résumé : effets de l'âge à l'implantation et de la durée d'utilisation de l'implant (enfants CI) 203
2.6. Effet de l'input auditif immédiat ... 203
2.6.1. VOT ... 203
2.6.2. Distance entre VOT des occlusives voisées et non-voisées ... 204
2.6.3. Résumé : effet de l'input audio immédiat ... 205
2.7. Effet de la complexité du mot ... 206
2.7.1. VOT ... 206
2.7.2. Distance entre VOT des occlusives voisées et non-voisées ... 208
2.7.3. Résumé : effet de la complexité du mot ... 208
2.8. Discussion ... 210
2.8.1. Production d'occlusives chez les enfants NH ... 210
2.8.2. Production d'occlusives chez les enfants CI ... 212
3. Fricatives ... 214
3.1. Données analysées ... 214
3.2. Effets du groupe et du lieu d'articulation ... 215
3.2.1. Centre de gravité ... 215
3.2.2. Distance entre centres de gravité de /s/ et /ʃ/ ... 217
3.2.3. Ecart-type ... 217
3.2.4. Skewness (coefficient de dissymétrie) ... 219
3.2.5. Kurtosis (coefficient d'aplatissement) ... 220
3.2.6. Résumé : effets du groupe et du lieu d'articulation ... 221
3.3. Effet du contexte vocalique ... 223
3.3.1. Centre de gravité ... 223
3.3.2. Distance entre centres de gravité de /s/ et /ʃ/ ... 225
3.3.3. Ecart-type ... 226
3.3.4. Skewness ... 227
3.3.5. Kurtosis ... 228
3.3.6. Résumé : effet du contexte vocalique ... 229
3.4. Enfants NH : effet de l'âge chronologique ... 230
3.4.1. Centre de gravité ... 230
3.4.2. Distance entre centres de gravité de /s/ et /ʃ/ ... 231
3.4.3. Ecart-type ... 232
3.4.4. Skewness ... 233
3.4.5. Kurtosis ... 234
3.4.6. Résumé : effets de l'âge chronologique (enfants NH) ... 235
3.5. Enfants CI : effet de l'âge à l'implantation et de la durée d'utilisation de l'implant ... 236
3.5.1. Centre de gravité ... 236
3.5.2. Distance entre centres de gravité de /s/ et /ʃ/ ... 238
3.5.3. Ecart-type ... 239
3.5.4. Skewness ... 240
3.5.5. Kurtosis ... 242
3.5.6. Résumé : effets de l'âge à l'implantation et de la durée d'utilisation de l'implant (enfants CI) 243
3.6. Effet de l'input auditif immédiat ... 244
3.6.1. Centre de gravité ... 244
3.6.2. Distance entre centres de gravité de /s/ et /ʃ/ ... 245
3.6.3. Ecart-type ... 245
3.6.4. Skewness ... 246
3.6.5. Kurtosis ... 246
3.6.6. Résumé : effet de l'input auditif immédiat ... 247
3.7. Effet de la complexité du mot ... 247
3.7.1. Centre de gravité ... 247
3.7.2. Distance entre centres de gravité de /s/ et /ʃ/ ... 249
3.7.3. Ecart-type ... 249
3.7.4. Skewness ... 250
3.7.5. Kurtosis ... 251
3.7.6. Résumé : effet de la complexité du mot ... 251
3.8. Discussion ... 253

3.8.1. Production de fricatives par les enfants NH ... 253
3.8.2. Production de fricatives par les enfants CI ... 255
4. Coarticulation ... 258
4.1. Données analysées ... 258
4.2. Effets du groupe et du lieu d'articulation ... 259
4.2.1. Résumé : effets du groupe et du lieu d'articulation ... 260
4.3. Enfants NH : effet de l'âge chronologique ... 261
4.3.1. Résumé : effet de l'âge chronologique (enfants NH) ... 261
4.4. Enfants CI : effet de l'âge à l'implantation et de la durée d'utilisation de l'implant ... 262
4.4.1. Résumé : effets de l'âge à l'implantation et de la durée d'utilisation de l'implant (enfants CI) 263
4.5. Effet de l'input auditif immédiat ... 263
4.5.1. Résumé : effet de l'input auditif immédiat ... 264
4.6. Discussion ... 265
4.6.1. Coarticulation chez les enfants NH ... 265
4.6.2. Coarticulation chez les enfants CI ... 267
5. Discussion générale des résultats de notre étude de production de parole ... 269

TROISIÈME PARTIE : Intelligibilité de la parole (2ème étude expérimentale) ... 273
CHAPITRE 6 : Étude d'intelligibilité (2ème étude expérimentale) ... 275
1. Méthodologie d'études d'intelligibilité dans la littérature ... 275
1.1. Introduction ... 275
1.2. Revue de méthodologies disponibles ... 275
1.2.1. Méthodes d'élicitation de la parole ... 278
1.2.1.1. Enregistrement des productions de parole ... 278
Répétition de mots ou de phrases ... 278
Elicitation de parole spontanée ... 278
1.2.1.2. Populations ... 279
1.2.2. Méthodes de jugement de l'intelligibilité ... 280
1.2.2.1. Juges d'intelligibilité ... 282
1.2.2.2. Méthodes d'évaluation de l'intelligibilité ... 282
Evaluations objectives : transcriptions ... 284
Evaluations subjectives : notes d'intelligibilité ... 284
Combinaison d'évaluations objectives et subjectives : transcriptions et notes ... 284
Autres méthodes ... 285
1.3. Conclusion ... 285
2. Méthodologie de notre étude d'intelligibilité ... 286
2.1. Etude pilote ... 286
2.2. Notre étude ... 287
2.2.1. Protocole ... 287
2.2.1.1. Question ... 287
2.2.1.2. Echelle ... 288
2.2.1.3. Interface ... 288
2.2.1.4. Conditions de passation ... 289
2.2.2. Stimuli ... 289
2.2.2.1. Groupes d'enfants ... 289
2.2.2.2. Choix des énoncés ... 290
Choix des énoncés pour la phase d'entrainement ... 290
Choix des énoncés pour la phase de test ... 291
2.2.2.3. Traitement des stimuli ... 291
2.2.3. Participants ... 291
2.3. Analyses statistiques ... 292
2.3.1. Analyses préliminaires : différence entre auditeurs naïfs et auditeurs experts et différences intra-sujet ... 292
2.3.2. Analyses statistiques ... 293
3. Résultats ... 294
3.1. Effet du groupe ... 294
3.2. Enfants NH : effet de l'âge chronologique ... 296
3.3. Enfants CI : effet de l'âge à l'implantation ... 296
3.4. Enfants CI : effet de la durée d'utilisation de l'implant ... 297
3.5. Enfants CI : effet de l'âge chronologique ... 298
4. Discussion ... 300

4.1. Liens entre caractéristiques de l'implantation cochléaire et intelligibilité 300
4.2. Liens entre caractéristiques acoustiques et intelligibilité 300

CONCLUSION ET PERSPECTIVES 305
Références 313
Liste des figures 329
Liste des tables 333
Annexes 339
Annexe 1 : Formulaire de consentement pour l'étude de production – enfants NH 339
Annexe 2 : Formulaire de consentement pour l'étude de production – enfants CI 340
Annexe 3 : Questionnaire – enfants CI 341
Annexe 4 : Fréquence des mots utilisés dans les tâches de répétition et de dénomination 342
Annexe 5 : Correspondance entre mots et codes utilisés pour l'organisation du corpus 343
Annexe 6 : Scénario du test d'intelligibilité réalisé sur Praat (Boersma & Weenink, 2015) – phase d'entraînement 344
Annexe 7 : Scénario du test d'intelligibilité réalisé sur Praat (Boersma & Weenink, 2015) – phase de test 345

INTRODUCTION

Ce projet de thèse est né d'un double constat sur les descriptions de l'acquisition de la langue maternelle par l'enfant, proposées dans la littérature : d'une part, peu d'études s'intéressent aux effets à long terme d'un accès partiel à l'audition sur la production de parole par l'enfant sourd, porteur d'un implant cochléaire ou de prothèses auditives traditionnelles et d'autre part, les descriptions du développement phonologique chez l'enfant typique s'arrêtent la plupart du temps après ses deux premières années de vie. Or ces deux questions sont cruciales : d'une part, l'accès au langage oral et l'intelligibilité de la parole représentent un enjeu majeur pour les interactions de l'enfant sourd appareillé avec ses pairs et la société dans laquelle il évolue et d'autre part, le manque de données développementales acoustiques chez l'enfant typique francophone de plus de deux ans ne permet pas de situer et d'identifier la parole dite « pathologique » par rapport à une trajectoire attendue du développement phonologique. Notre thèse propose donc de contribuer à combler ces deux manques.

Depuis le milieu des années 1980, l'implant cochléaire s'est développé comme l'une des options proposées dans le cadre d'une surdité sévère et profonde, acquise ou congénitale, de l'adulte et de l'enfant. Les caractéristiques techniques de l'implant n'ont cessé d'évoluer depuis cette période, permettant à son utilisateur d'avoir accès à des caractéristiques acoustiques de plus en plus précises des sons de l'environnement et de (re)-construire des représentations des phonèmes de sa langue maternelle pour avoir une parole la plus intelligible possible.

Cependant, les effets de l'utilisation à long terme de l'implant cochléaire sur la production de parole et sur l'intelligibilité des enfants restent, à ce jour, très peu étudiés. Cette thèse est une description et une proposition d'analyse du développement de la parole d'enfants sourds, porteurs d'implant cochléaire, âgés de cinq à onze ans. Ce projet a été conçu dans une démarche qui utilise à la fois les techniques et méthodes d'analyse de la phonétique acoustique, de la phonétique clinique et de la psycholinguistique. Nous souhaitons proposer à travers cette étude une analyse des liens entre production de parole et intelligibilité des enfants qui utilisent un implant cochléaire depuis plusieurs années, de façon à évaluer les difficultés persistantes et résiduelles de l'enfant à produire des sons de parole, malgré une audition partiellement restaurée et un accompagnement orthophonique.

Dans une première partie théorique, nous allons présenter les enjeux de la thèse du point de vue des théories de l'acquisition du langage et du développement phonologique chez l'enfant typique et normo-entendant (Première partie, chapitre 1), nous allons ensuite présenter le fonctionnement du système auditif, les différents types de surdité et techniques de remédiation disponibles et proposées dans le cadre de surdités chez l'adulte et l'enfant (Première partie, chapitre 2), et enfin les différentes études disponibles dans la littérature sur la production de parole et l'intelligibilité des enfants porteurs d'implant cochléaire (Première partie, chapitre 3). Cette première partie a pour but de présenter une revue de la littérature qui nous a permis d'élaborer ce travail de thèse, dont les enjeux sont aussi bien théoriques que méthodologiques et même cliniques.

La deuxième partie de ce volume sera consacrée à l'étude de production de parole comparée chez des enfants sourds porteurs d'implant cochléaire et chez des enfants normo-entendants appariés en âge chronologique, réalisée dans le cadre de cette thèse. Dans un premier temps (Deuxième partie, chapitre 4), nous présenterons notre démarche expérimentale, qui prend en compte à la fois les spécificités physiologiques des enfants, mais également le système phonologique de leur

langue maternelle, dans un contexte clinique. Nous proposerons une réflexion sur la collecte des données chez une population d'enfants particulière, sur les méthodologies d'analyse de la production de parole en phonétique acoustique ainsi que sur les techniques de traitement des données que nous allons utiliser. Le deuxième volet de notre étude expérimentale (Deuxième partie, chapitre 5) sera consacré à l'analyse des productions de trois types de sons de parole (voyelles, consonnes occlusives, consonnes fricatives) et de la coarticulation de séquences occlusive-voyelle chez les enfants sourds porteurs d'implant cochléaire.

Le troisième volet de cette thèse présente la deuxième étude expérimentale (Troisième partie, chapitre 6) que nous avons réalisée dans le cadre de ce projet, destinée à caractériser perceptivement l'intelligibilité de la parole de deux groupes d'enfants de cinq à onze ans, sourds porteurs d'implant cochléaire et normo-entendants. Après un aperçu de méthodes de jugements de l'intelligibilité disponibles dans la littérature, nous proposons notre propre méthode d'analyse de l'intelligibilité, à partir d'un corpus de parole semi-spontanée, puis les résultats que nous obtenons, en les confrontant aux résultats de notre étude de production.

Enfin, dans la conclusion générale de cette thèse, nous synthétiserons les résultats, ainsi que les limites de notre travail. Nous proposerons également des pistes d'analyses complémentaires qui pourraient permettre d'explorer d'autres aspects de la parole des enfants, qu'ils soient normo-entendants ou utilisateurs d'implant cochléaire.

PREMIERE PARTIE :
Approches théoriques

CHAPITRE 1 : Théories d'acquisition de la langue maternelle et développement phonologique

1. Introduction

Notre travail cherche à comprendre les mécanismes de l'acquisition de la langue maternelle par l'enfant, dans le contexte particulier d'une surdité, et nous proposons donc dans un premier temps une revue de théories d'acquisition de cette langue maternelle par l'enfant au développement typique. Au cours du vingtième siècle notamment, plusieurs courants de la psychologie et de la linguistique se sont intéressés aux mécanismes de l'acquisition du langage en général, et du développement phonologique en particulier, en relation avec les capacités de l'enfant à percevoir et à produire les sons de parole. Ces études, dont le but est de décrire et comprendre les étapes de l'émergence du langage et du développement phonologique chez l'enfant, donnent lieu à des analyses différentes selon l'approche théorique et le cadre formel adoptés par les différents auteurs. Ce chapitre propose donc dans un premier temps une revue de quelques théories de l'acquisition du langage, puis une sélection d'études développementales en perception et en production, et enfin une présentation des principaux modèles du développement phonologique de l'enfant.

2. Quelques théories de l'acquisition de la langue maternelle

Dans ce paragraphe, nous avons choisi de présenter trois courants qui ont décrit et analysé l'acquisition de la langue maternelle selon différentes perspectives: le behaviourisme, le nativisme (associé au cognitivisme), et le constructivisme (et le socio-constructivisme). Ce choix d'approches théoriques n'est pas exhaustif, mais il nous permet de définir quelques unes des problématiques liées à l'étude de l'acquisition du langage et du développement phonologique.

2.1. Behaviourisme

Le behaviourisme est une théorie du début du vingtième siècle, notamment développée par Watson (1925) qui analyse tous les comportements humains (parmi lesquels la parole) comme un ensemble de stimuli et réponses. Le but de la psychologie est donc pour les behaviouristes la « prédiction et le contrôle du comportement » (Bronckart, 1977), et le langage n'est vu que comme l'un des comportements humains qui peuvent être décrits et prédits sous la forme d'une réponse à un stimulus. Dans le cadre behaviouriste, les comportements ne sont pas liés à des facultés innées du nouveau-né, mais sont acquis par l'enfant à partir de la naissance par le mécanisme de réponse à un stimulus. Pour Watson (1925) cependant, la seule spécificité du langage par rapport à un autre comportement est qu'il s'agit d'un comportement verbal qui peut déclencher un type de réponse qui n'est pas systématiquement verbale. A la suite de Watson (1925), la théorie behaviouriste va proposer une théorie de l'acquisition du langage par l'enfant, et notamment Skinner présente dans son ouvrage *Verbal Behavior* (1957) une approche qui repose sur trois mécanismes décrits dans Bronckart (1977) pour expliquer le fonctionnement des interactions verbales et l'acquisition du langage dans la théorie behaviouriste : 1) un stimulus est émis par un locuteur, et 2) peut provoquer une réponse de l'auditeur, qui 3) à son tour peut être renforcée positivement. L'analyse behaviouriste consiste donc à décrire ce processus de stimulus-réponse, et les conditions d'apparition du comportement, ce qui va permettre de pouvoir prédire la réponse en fonction du stimulus émis.

Skinner (1957) propose trois types de comportements de l'enfant, résumés par Bronckart (1977) : 1) les « mands » (demandes, ordres, interdictions) qui sont des réponses verbales à des stimuli internes qui sont notamment liés à des besoins et qui apparaissent avec le renforcement chez les très jeunes enfants, 2) les comportements « échoïques, textuels et intraverbaux » qui sont des réponses à des stimuli externes (l'enfant va par exemple imiter des productions de parole de l'adulte), et 3) les « tacts », qui sont des réponses verbales à un événement ou un objet référentiel dans le cadre d'un récit ou d'un discours, la présence du stimulus accroissant alors la probabilité d'apparition de la réponse.

Dans le cas particulier de l'acquisition du langage, le stimulus émis est par exemple un input maternel, et la réponse peut être la répétition ou l'imitation par l'enfant : celui-ci utilise le retour de sa mère (par exemple ses encouragements et réactions) comme renforcement, pour développer une réponse adaptée au stimulus. Le babillage est ainsi vu par les behaviouristes comme un essai de production en réponse à un stimulus, qui va se préciser au fur et à mesure du développement de l'enfant, en fonction du renforcement positif qu'il reçoit de son entourage : une absence de réaction positive de l'adulte à une réponse de l'enfant va conduire celui-ci à l'abandonner petit à petit puisqu'elle ne convient pas au stimulus, pour ne conserver que les réponses appropriées.

Dans ce cadre, l'interaction verbale est non-spécifique, et le comportement verbal est donc vu comme un simple échange entre un locuteur et un auditeur ou un enchaînement de séquences « stimulus-réponse », mais cette approche ne prend pas en compte la situation et le contexte de la communication, qui incluent notamment l'intention du locuteur ou tout aspect sémantique du langage et de la parole.

L'approche behaviouriste a été remise en cause, notamment par Chomsky (1959), qui s'oppose à l'aspect purement fonctionnel de l'analyse de la parole comme une succession de réponses à des stimuli mais qui ne prend en compte ni processus cognitifs sous-jacents à la réponse ni les spécificités du langage, qui pour Chomsky (1959) n'est pas un comportement parmi d'autres. En outre, Chomsky (1959) conteste l'idée que le renforcement est rendu possible selon Skinner (1957) par la seule action d'une personne extérieure mais que le locuteur n'a qu'un rôle secondaire dans l'échange verbal. L'approche behaviouriste est également critiquée par Chomsky (1959) en raison de son aspect exclusivement expérimental, ce qui permet à Skinner (1957) d'observer des phénomènes à l'aide de situations extrêmement contrôlées. Ces situations ne reflètent cependant pas la diversité de la production de parole en conditions naturelles, et l'analyse de l'acquisition du langage par les enfants issue de ces conditions expérimentales très contrôlées peut notamment être remise en cause. Enfin, la simple relation stimulus-réponse ne permet pas non plus de rendre compte selon Chomsky (1959) de la créativité de l'enfant et du locuteur en général ainsi que de sa capacité à produire un nombre infini d'énoncés.

2.2. Nativisme/cognitivisme

Les approches nativistes et cognitivistes se distinguent donc de l'approche behaviouriste. Ces deux courants sont notamment liés aux travaux de Chomsky en linguistique et de Pinker en psychologie cognitive, et partagent de nombreuses similitudes puisqu'ils cherchent à expliquer les mécanismes sous-jacents de la production de parole et de l'acquisition du langage.

L'approche de Chomsky (entre autres 1967, 1969, 1975) s'inscrit dans le courant nativiste (ou innéiste) qui postule que l'enfant vient au monde en ayant des capacités cognitives qui lui

permettent d'acquérir le langage : selon Chomsky, l'enfant dispose à la naissance d'un module abstrait, le « Language Acquisition Device » (ou LAD), qui lui permet d'acquérir le langage. Celui-ci est lié à une capacité abstraite universelle (aussi appelée « compétence linguistique ») qui se réalise dans les productions de la langue maternelle (aussi appelée « performance »). Ainsi, la « compétence linguistique » permet à l'enfant de disposer à la naissance de ce que Chomsky nomme une « Grammaire Universelle », qui lui permet en théorie de produire une infinité de phrases syntaxiquement bien formées indépendamment de sa langue maternelle. Dans ce cadre théorique aussi appelé « Principes et paramètres », la « Grammaire Universelle » est donc un ensemble de « principes » préexistants à toute langue, et la production de parole dans une langue donnée est le résultat de la mise en place des « paramètres » propres à celle-ci au fur et à mesure de l'exposition à cette seule langue. Le processus d'acquisition de la langue maternelle va ainsi consister pour l'enfant à acquérir les « paramètres » de cette langue-cible pour pouvoir produire des énoncés dans celle-ci. Si la notion de « Grammaire Universelle » implique que l'enfant possède une capacité innée à produire a priori n'importe quel énoncé dans n'importe quelle langue, l'acquisition de la langue maternelle va consister à se spécialiser vers cette langue maternelle progressivement, à partir de la naissance. Ce processus d'acquisition du langage dépend donc à la fois de structures préexistantes, spécifiques au langage, ainsi que de l'input langagier reçu par l'enfant dès la naissance.

Le cognitivisme dont Pinker (1994, 2002) est l'un des théoriciens défend des thèmes similaires, mais si l'approche de Chomsky est avant tout liée à un formalisme linguistique, Pinker adopte la perspective de la psychologie expérimentale et des liens entre acquisition et processus physiologiques et cérébraux : il défend l'idée selon laquelle l'acquisition est un processus rendu possible par les caractéristiques biologiques de l'être humain mais auquel l'environnement familial et social participe également. Ainsi, Pinker (2002) montre que l'héritage génétique contribue à hauteur de 40 à 50% du développement de l'enfant, l'environnement familial entre 0 et 10% et les interactions sociales de l'enfant en dehors de sa famille à hauteur de 50% : l'acquisition du langage est donc rendue possible grâce à des capacités cognitives disponibles à la naissance et liées au patrimoine génétique de l'enfant, mais ne peut être réalisée que grâce aux interactions de l'enfant avec son environnement familial et social (Pinker, 1994, 2002).

Ces liens entre environnement (familial et social) et développement de l'enfant ne sont pas sans rappeler le constructivisme, théorisé notamment par Piaget et Vygotski, pour lesquels le développement et notamment l'accès au langage se font avant tout par interaction de l'enfant avec son environnement.

2.3. (Socio-)constructivisme

Dans la première moitié du vingtième siècle se développe le courant constructiviste sous l'impulsion de Piaget : le développement de l'enfant se fait par une succession de stades dont l'ordre est prédéterminé et qui correspondent à la présence de ce que Piaget (2008) appelle des « conduites » ou « comportements ». Le langage n'est pour Piaget que l'une de ces « conduites ».

Le premier stade de développement est la période sensori-motrice (de 0 à 18 mois), pendant laquelle l'enfant interagit avec son environnement pour arriver à coordonner des gestes puis des actions et à acquérir une organisation spatio-temporelle des objets et à comprendre la relation de causalité entre action et résultat. Au cours de cette période, l'enfant commence à produire des sons, puis à babiller pour arriver aux premières vocalisations assimilables à des mots.

Puis vient le stade pré-opératoire (de 2 à 7 ans) qui correspond également à l'âge d'apparition et de stabilisation des premiers mots, puis de production des premiers énoncés qui deviennent de plus en plus complexes et qui permettent à l'enfant de représenter son environnement en le nommant ; c'est le stade où se développent peu à peu ce que Piaget (1989) appelle « langage égocentrique » qui consiste pour l'enfant à se parler à lui-même, ce qui peut être vu comme un entraînement à la parole ou une forme de réflexion à haute voix, et « langage socialisé » qui correspond au langage que l'enfant utilise dans ses interactions avec d'autres enfants ou avec des adultes. Au cours de ce stade, l'enfant va devenir capable de nommer les objets qui l'entourent, ce qui pour Piaget contribue à l'apparition de la fonction symbolique ou sémiotique (l'enfant assigne un signifiant à un signifié).

Le stade des opérations concrètes (de 7 à 11 ans) correspond ensuite à la période où l'enfant acquiert la logique, devient capable de raisonner et d'effectuer des opérations arithmétiques. Enfin, lors du stade des opérations formelles (de 11 à 16 ans), l'enfant est capable d'avoir un raisonnement abstrait.

L'acquisition du langage s'inscrit donc pour Piaget et le courant constructiviste dans un contexte plus large de développement de l'enfant, et se fait à partir de l'enfant qui construit, par le langage, sa représentation du monde dont il est le centre.

Le socio-constructivisme est un courant dont Vygotski est l'un des théoriciens et qui, à la différence du constructivisme piagetien, voit l'enfant comme un être social en premier lieu et dont le développement est guidé par le rapport à la société dans laquelle il évolue.

Vygotski (1997) s'intéresse aux fonctions du langage et à son évolution au cours du développement de l'enfant : le langage est en premier lieu pour l'enfant un outil de communication avec l'adulte. Alors que pour Piaget, le langage dit « égocentrique », qui permet à l'enfant de se situer par rapport à l'environnement, se développe avant le langage « socialisé » qui lui permet de communiquer avec son entourage, il est analysé à l'inverse par Vygotski (1997) pour qui le langage de l'enfant est d'abord externe (il permet à l'enfant de communiquer avec son environnement), pour ensuite devenir interne (le langage devient alors la pensée). Il justifie notamment cette analyse par le fait que dans un premier temps, l'enfant utilise le langage pendant ou après une action, ce qui peut être vu comme un commentaire de l'action, pour ensuite accéder à un stade où il utilise le langage avant de commencer l'action, dans une démarche de planification de cette action, ce qui pour Vygotski s'apparente à l'expression de sa pensée.

Dans cette approche socio-constructiviste, l'enfant est donc d'abord envisagé comme un être social qui utilise le langage pour communiquer avec son environnement et en particulier avec les adultes qui l'entourent, et dont le langage se développe par les interactions sociales, alors que dans l'approche constructiviste de Piaget, le langage se développe chez l'enfant d'abord pour représenter le monde qui l'entoure, puis pour communiquer. Ces deux approches sont par ailleurs similaires car elles envisagent le développement de l'enfant et l'acquisition du langage comme une série d'étapes préprogrammées et qui vont se réaliser à partir de la naissance et au fur et à mesure que l'enfant grandit. Dans cette approche, l'acquisition du langage fait partie du développement cognitif plus général et est rendu possible par les interactions sociales de l'enfant.

Cette présentation non-exhaustive de quelques théories a permis de mettre en évidence que l'acquisition du langage est un processus complexe qui s'inscrit dans le cadre global du

développement de l'enfant, qui est possible à la fois grâce aux capacités cognitives de l'enfant, mais également grâce à ses interactions avec l'environnement.

Le but de notre travail de thèse étant d'étudier les difficultés de production de parole chez des enfants qui accèdent à l'oral plus tardivement et partiellement par l'intermédiaire de l'implant cochléaire, il nous semblait important de proposer en premier lieu cet aperçu théorique de l'acquisition du langage et du développement phonologique chez l'enfant typique : l'enjeu de notre travail est à la fois de comprendre comment l'enfant sourd, porteur d'implant cochléaire accède au langage, mais également d'identifier plus précisément les mécanismes de cette acquisition par l'enfant normo-entendant. Dans la suite de ce chapitre, nous allons donc nous intéresser plus précisément aux modalités nécessaires à l'enfant pour percevoir et produire la parole, puis nous présenterons quelques modèles de développement phonologique chez l'enfant typique.

3. Développement de la parole : perception

Dès ses premiers mois de vie, la perception de son environnement va jouer un rôle important dans le développement de l'enfant et dans son accès à la parole. Nous allons donc dans un premier temps proposer une revue de quelques études qui contribuent à la compréhension des mécanismes auxquels l'enfant a recours pour percevoir le monde qui l'entoure.

3.1. Rétrécissement perceptif

Les capacités perceptives de l'enfant au cours de la première année de vie sont caractérisées par une grande plasticité qui peut être envisagée comme le reflet d'une période critique du développement ou comme le résultat d'une exposition à l'input, notamment langagier.

Le rétrécissement perceptif est un mécanisme qui se met en place relativement tôt chez l'enfant et qui va lui permettre de construire le système phonologique propre à sa langue maternelle, nécessaire ensuite à la production de parole. Ce rétrécissement perceptif repose sur une évolution de la perception catégorielle. Nazzi (2006) propose une définition de la perception catégorielle des sons de parole :

> « […] deux sons séparés par une même distance acoustique seront discriminés s'ils se trouvent de part et d'autre d'une frontière phonétique, mais ne le seront pas s'ils sont situés du même côté de cette frontière. »

Plusieurs expériences ont mis en évidence la capacité précoce de l'enfant à percevoir des contrastes segmentaux, même absents de la langue maternelle, entre un et quatre mois (Eimas et al., 1971, Kuhl et al., 1992), puis la perte progressive de la perception de contrastes non-natifs entre huit et dix mois, qui ne seront finalement plus discriminés par les enfants vers la fin de la première année de vie (Werker & Tees, 1984). Une étude de Maye et al. (2002) met en évidence les liens entre perception catégorielle d'un contraste phonologique et distribution statistique des occurrences participant à ce contraste dans l'input : dans cette étude, des enfants de six et huit mois sont capables de discriminer un contraste consonantique dont la distribution est bimodale (c'est-à-dire, une distribution localisée principalement aux deux extrémités du continuum) s'ils y ont été exposés auparavant lors d'une phase d'entrainement mais pas s'ils ont été exposés à une distribution unimodale (c'est-à-dire une distribution localisée au centre du continuum) . Cette étude de Maye et al. (2002) est un argument supplémentaire en faveur de la mise en place progressive d'un rétrécissement perceptif, qui a lieu vers la fin de la première année de vie de l'enfant et qui dépend de l'exposition à sa langue maternelle.

Différentes propositions de modélisations du rétrécissement perceptif ont été faites, parmi lesquelles le modèle d'assimilation perceptive ou PAM (Best, 1994), le modèle perceptif proposé par Kuhl (2000), issu de la théorie des prototypes, ou le modèle perceptif issu de la théorie des exemplaires (Jusczyk, 1992, Bybee 2001).

Le modèle d'assimilation perceptive (Best, 1994) postule notamment que lorsqu'un contraste entre deux catégories phonologiques ne fait pas partie de l'inventaire de la langue maternelle de l'enfant, celui-ci continuera à être discriminé si les sons de ce contraste ne correspondent pas à des catégories de la langue maternelle (par exemple les clicks sont discriminés par des locuteurs d'une langue qui n'en possède pas) mais ne sera au contraire plus discriminé si les sons sont similaires ou proches d'une catégorie de sons existants dans la langue maternelle : dans ce cas, les

sons non-natifs seront perçus comme faisant partie de cette catégorie proche (par exemple, le [θ] présent en anglais pourra être assimilé à un [s] ou un [f] par un locuteur francophone).

Le modèle perceptif proposé par Kuhl (2000) postule que des « prototypes » de sons de parole sont répartis dans l'espace acoustique : le « prototype » fonctionne comme un aimant (« magnet effect », Kuhl, 1991 et Iverson & Kuhl, 1995) pour chaque son de parole de la langue maternelle : chaque son est perçu comme une réalisation du « prototype » qui en est le plus proche dans l'espace acoustique. Plus l'enfant va acquérir de l'expérience en grandissant, plus le son de parole va être assimilé à l'un des « prototypes » de la langue maternelle, ce qui conduit au déclin de la perception des catégories qui ne font pas partie de l'inventaire de la langue maternelle et ce dès l'âge de six mois.

A l'inverse, dans le modèle perceptif issu de la théorie des « exemplaires » (Jusczyk, 1992, Bybee, 2001) décrit notamment dans Pierrehumbert (2001), chaque occurrence d'un son de parole est classé perceptivement en fonction de ses propriétés acoustiques puis stocké par la mémoire dans une zone acoustique. Du point de vue perceptif, une catégorie correspond ainsi à une zone où un certain nombre d'occurrences du son sont regroupées car elles partagent des propriétés acoustiques, et la perception se fait par identification de l'occurrence au sein de cette zone. La fréquence d'apparition de l'occurrence participe ainsi à l'émergence et à la distinction des catégories : il s'agit d'une approche statistique de la perception.

Il est en outre intéressant de noter que les capacités de l'enfant puis de l'adulte à percevoir un contraste phonologique évoluent au cours de la vie, notamment si l'exposition à ce contraste diminue voire disparaît. L'étude de Ventureyra et al. (2004) a montré que des adultes qui avaient été adoptés en Corée du Sud par des familles francophones entre trois et neuf ans ne sont plus capables de discriminer le contraste des occlusives du coréen (qui se distinguent acoustiquement par la durée du VOT), qui pourtant était leur première langue, avant l'adoption. La perception catégorielle des consonnes par les participants à cette étude a donc été modifiée, puisqu'ils ne sont plus capables de discriminer le contraste des trois types de consonnes de leur première langue, qu'ils ont remplacé par le contraste de voisement des consonnes du français : la plasticité de la perception ne serait donc pas exclusivement limitée à la première année de vie, ce qui pourrait remettre en question la notion de « période critique » dans l'acquisition du langage.

Un grand nombre d'études sur le rétrécissement perceptif chez le bébé et le jeune enfant se concentre sur les caractéristiques segmentales de la parole, et quelques-unes s'intéressent au rétrécissement perceptif sur la base de l'utilisation d'indices suprasegmentaux ou prosodiques. Ainsi, les études de Dehaene-Lambertz & Houston (1998), de Ramus et al. (2000) et de Nazzi & Ramus (2003) montrent que les enfants, dès le plus jeune âge, peuvent discriminer des langues grâce à leur prosodie puis se spécialisent vers les structures prosodiques de leur langue maternelle : dans l'étude de Dehaene-Lambertz & Houston (1998), des bébés de deux mois sont capables de distinguer leur langue maternelle d'une langue étrangère en utilisant les caractéristiques prosodiques de la phrase ; dans l'étude de Ramus et al. (2000), des bébés francophones âgés d'une semaine environ arrivent à discriminer le japonais du néerlandais par le rythme ; et dans l'étude de Nazzi & Ramus (2003), des bébés francophones de cinq mois sont capables de discriminer deux langues appartenant à des classes rythmiques différentes, mais ne sont plus capables de discriminer deux langues appartenant à la même classe rythmique, ce qui

pourrait indiquer que le rétrécissement perceptif n'agit pas seulement au niveau segmental, mais également au niveau suprasegmental.

La modalité auditive joue donc un grand rôle dans le rétrécissement perceptif au cours des premiers mois et des premières années de la vie de l'enfant. Nous pouvons donc nous interroger sur l'impact de la surdité sur le rétrécissement perceptif qui a lieu lors de cette première période de la vie de l'enfant : dans quelle mesure l'accès tardif à l'oral et l'utilisation de l'implant cochléaire gênent-ils ce rétrécissement perceptif et quelles sont les conséquences du recours à une communication utilisant d'autres modalités que la modalité auditive sur le rétrécissement perceptif ?

3.2. Perception multimodale

La perception par l'enfant n'est pas liée qu'à la modalité auditive, mais à plusieurs types de modalités, qui lui permettent de construire une représentation du monde qui l'entoure. En effet, plusieurs auteurs ont montré que la perception est un processus qui évolue, parfois très rapidement, dès la naissance et se fait par plusieurs biais : la perception peut être entre autres haptique, auditive, visuelle, ou multimodale, lorsqu'elle utilise plusieurs modalités conjointement. Pour comprendre les liens entre ces différentes modalités qui entrent en jeu lors de la perception, nous allons présenter ici une sélection non-exhaustive de plusieurs études, qui s'intéressent à la perception multimodale chez le très jeune enfant.

Meltzoff & Borton (1979) ont par exemple montré qu'à un mois, l'enfant est capable d'identifier visuellement une tétine qu'il a eue en bouche auparavant, ce qui témoigne d'une capacité à associer plusieurs modalités pour identifier un objet. De même, Aronson & Rosenbloom (1971) mettent en évidence la préférence de bébés du même âge pour la cohérence spatiale entre information auditive et visuelle lors d'une expérience où était présentée conjointement une personne et sa voix : les bébés ont une préférence pour la présentation cohérente de cette personne en face de lui et du son de sa voix en stéréo, plutôt qu'avec un seul haut-parleur latéral, ce qui ne correspond pas à la localisation de la personne qui est en face de lui.

Kuhl & Meltzoff (1982) et Legerstee (1990) s'intéressent à la perception bi-modale de la parole : en mesurant le temps de fixation du regard de l'enfant, Kuhl & Meltzoff (1982) montrent que des bébés de 18 à 20 semaines ont une préférence pour des stimuli congruents, dans lesquels le son de la voyelle /i/ ou /a/ testée correspond à l'information visuelle proposée, et en comptant le nombre de productions de voyelles /a/ ou /u/ produites par des bébés après qu'on leur a présenté des voyelles sous forme audio ou visuelle, Legerstee (1990) montre de la même façon qu'à l'âge de trois-quatre mois, ceux-ci sont capables de percevoir et d'imiter les voyelles pour lesquelles l'information auditive et visuelle du stimulus proposé est congruente. Ceci suppose qu'il a déjà acquis une représentation abstraite des gestes articulatoires nécessaire à la production de parole.

Ces quelques expériences permettent de comprendre que le développement des capacités perceptives de l'enfant, notamment pour la parole, est relativement précoce puisque dès la naissance et dans ses premiers mois de vie, celui-ci fait appel à plusieurs modalités pour percevoir à la fois le monde qui l'entoure et également la parole.

L'enfant normo-entendant va ainsi mettre à profit plusieurs modalités conjointement, dont la modalité visuelle et la modalité auditive, pour percevoir les sons de parole et construire son

système phonologique. Mais qu'en est-il de l'enfant sourd qui a eu peu ou pas d'accès à l'oral dans sa première année de vie ? Existe-t-il un processus d'acquisition différent pour les informations visuelles ou pour les information auditives ? Si l'enfant entendant est capable d'intégrer plusieurs types d'informations pour percevoir les sons de parole, sur quelles informations l'enfant sourd va-t-il pouvoir s'appuyer pour développer ses capacités à percevoir le monde qui l'entoure ?

Cette présentation non exhaustive de quelques études a permis de mettre en lumière le processus d'acquisition par l'enfant de la perception des caractéristiques et contrastes phonologiques de sa langue maternelle, au détriment des indices acoustiques qui ne sont pas pertinents dans ce système. Ce processus d'acquisition repose sur plusieurs modalités conjointes, et permet à l'enfant de développer sa représentation du monde et la catégorisation des sons de sa langue maternelle. Cette spécialisation vers la perception des caractéristiques de la langue maternelle se fait en parallèle du développement de la production de parole par l'enfant que nous allons présenter dans le paragraphe suivant.

4. Développement de la parole : production

Le développement de la production de la parole par l'enfant suit une trajectoire inverse à celle de la perception. En effet, si l'enfant est tout d'abord capable de percevoir les caractéristiques acoustiques fines des sons de parole, pour ensuite perdre cette capacité au cours de son développement, et ne plus percevoir que les sons dont la catégorisation est pertinente dans le système phonologique de sa langue maternelle, sa production de parole va suivre une trajectoire qui dépend à la fois de cette perception mais également de ses capacités notamment articulatoires à produire les sons de parole. Ce paragraphe propose de retracer les différentes étapes qui se succèdent dans le développement de la production de parole par l'enfant.

4.1. Etapes du développement de la production de parole

Boysson-Bardies (1996) propose une description parallèle de l'évolution de la perception et de production de parole chez l'enfant à partir de la naissance dont une synthèse est présentée dans la table 1.1 ci-dessous, tirée de Wauquier-Gravelines (2005).

PERCEPTION	PRODUCTION
0-1 mois perception catégorielle préference voix de la mère/ langue maternelle sensibilité voix (timbre, chant) et prosodie	**0-1 mois** Sons végétatifs, pleurs, Expression malaise / confort
1-5 mois catégorisation contrastes phonétiques discrimination schèmes intonatifs reconnaissance de syllabes	**1-5 mois** rires, cris, expressions vocales émotions jeux vocaux, arreuh début contrôle de la phonation
5-7 mois catégorisation voyelles détection des clauses prosodiques grands constituants début perception intermodale	**5-7 mois** vocalisation variation, imitation intonative trills
8-10 mois détection frontières syntagmes sensibilité à l'accent début reconnaissance mots connus en contexte	**8-10 mois** voyelles langue cible babillage canonique (bababa) puis babillage varié orienté langue cible contours intonatifs langue cible
10-12 mois détection frontières mots surdité phonologique reconnaissance mots connus hors contexte compréhension mots (+/- 30 mots) avec référents	**10-12 mois** syllabes langue cible simplifiées CV/V babillage varié et intoné production formes stabilisées en contexte premiers mots (maman, dodo ...)
12-16 mois compréhension +/- 100 mots	**12-16 mois** stabilisation des formes de babillage varié complexification des structures intonatives 16 mois : stade dit des « 50 mots », noms début acquisition phonologique (reduplications, troncations, harmonies, substitutions)
16-20 mois compréhension +/- 200 mots début sensibilité classes morpho-syntaxiques	**16-20 mois** 50 / 150 mots premiers verbes acquisition phonologique : pleine période
20-26 mois compréhension mots début parsing syntaxique	**20-26 mois** explosion lexicale accélération de l'acquisition phonologique et début stabilisation des formes lexicales
26-48 mois compréhension tous mots concrets même hors contextes + mots nouveaux en contexte traitement syntaxique	**26-48 mois** stabilisation des formes lexicales fin acquisition de la phonologie acquisition morpho-phonologie

Table 1.1 : chronologie de l'évolution de la perception et de la production de la parole chez l'enfant de 0 à 48 mois (Boysson-Bardies, 1996, cité dans Wauquier-Gravelines, 2005)

Comme l'indique la table 1.1, la production de parole se fait par étapes : les premières productions assimilables à des sons de parole contrôlés arrivent vers la fin de la première année, avec l'apparition du babillage canonique vers 8-10 mois (productions de type CVCV, avec une seule consonne et une seule voyelle), suivi du babillage varié (structures CVCV dans lesquelles

alternent différentes consonnes et différentes voyelles), des premiers mots, puis de mots plus complexes, organisés en énoncés eux-mêmes de plus en plus complexes.

Koopmans van Beinum & Van der Stelt (1986) décrivent les mécanismes qui entrent en jeu dans cette acquisition progressive du contrôle articulatoire, qui permet à l'enfant de produire des sons assimilables à des sons de parole à partir de l'apparition du babillage. Ils se placent dans la théorie source-filtre distinguant phonation (la source) et articulation (le filtre) et proposent six stades de développement entre 0 et 27 semaines : 1) phonation ininterrompue, de type voyelle [a] tenue, 2) phonation interrompue, de type [aʔaʔa], 3) phonation accompagnée d'un mouvement articulatoire supraglottique, de type [awa] ou [aʁə], 4) enchaînement de variations phonatoires avec ou sans geste supraglottique (qualités de voix différentes), 5) enchaînement de gestes articulatoires rédupliqués ou babillage canonique, de type [bababa], 6) premiers mots identifiables.

Les liens entre babillage canonique et premiers mots ont été étudiés par plusieurs auteurs qui les analysent cependant différemment. Pour Jakobson (1941, 1949) le babillage et les premiers mots sont deux étapes indépendantes, et le babillage est une phase d'entraînement articulatoire qui lui sera nécessaire à la production de la parole et dans laquelle l'enfant utilise la perception des sons de son environnement pour s'entrainer à identifier et à reproduire les contrastes acoustiques et phonologiques de sa langue maternelle de plus en plus précisément et fréquemment.

A l'inverse, pour Stoel-Gammon (1992) par exemple, babillage et premiers mots forment un continuum, le babillage étant une étape indispensable du développement qui permet de faire des prédictions sur la production ultérieure de la parole par l'enfant :

> « the baby who produces more canonical syllables generally exhibits more advanced speech and language skills than his peer who babbles less […] (1) more canonical vocalizations provide more *practice* in producing sounds and syllable types that can serve as the basis for words and (2) more vocalizations provide the opportunity for the establishment of a *feedback loop* that allows babies to determine the relationship between their own oral-motor movements and the sounds that result from these movements »[1].

Ainsi, d'après Stoel-Gammon (1992), l'émergence précoce du babillage favorise la mise en place de la boucle audio-phonatoire, donc l'émergence du système phonologique, grâce à un aller-retour permanent entre perception et production de sons de parole par l'enfant. L'étude de Oller & Eilers (1988) chez des enfants normo-entendants et sourds appareillés va dans ce sens puisqu'elle montre que les enfants sourds appareillés produisent dans un premier temps des vocalisations similaires à celles des enfants normo-entendants, mais ont un babillage canonique plus tardif ainsi que des difficultés et un retard dans la production de sons de parole assimilables à des phonèmes et similaires à ceux des enfants normo-entendants.

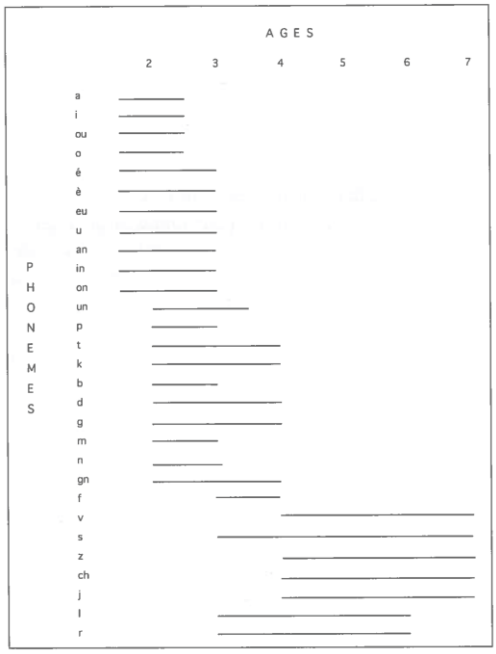
[1] « le bébé qui produit un plus grand nombre de syllabes canoniques fait généralement preuve de capacités de parole et de langage plus avancées que celui qui babille moins […] (1) un plus grand nombre de vocalisations canoniques conduit à une *pratique* plus importante de la production des sons et des types de syllabes qui peuvent servir de base pour les mots, et (2) un plus grand nombre de vocalisations permet d'établir une *boucle de feedback* qui permet aux bébés de déterminer la relation entre leurs propres gestes oro-moteurs et les sons qui résultent de ces gestes » (notre traduction)

Parallèlement à l'apparition de segments qui vont s'approcher de plus en plus de ceux de la langue maternelle, l'enfant va petit à petit développer un système prosodique dont les caractéristiques rythmiques et intonatives sont celles de sa langue maternelle. L'étude de Konopczynski (1990a, b) de productions de séquences CV dans des énoncés multisyllabiques par des enfants francophones montre qu'entre 10 et 24 mois, l'enfant va acquérir peu à peu la prosodie de sa langue maternelle : la production de l'enfant va se caractériser d'une part par un allongement progressif de la durée des syllabes finales et la diminution de celle des syllabes non-finales, qui se stabilisent vers l'âge de 16 mois, et d'autre part par la modulation de la F0 par l'enfant dès l'âge de 9 mois, pour produire notamment des contours intonatifs de questions (Konopczynski, 1991). Ces résultats sont en accord avec l'étude de Hallé et al. (1991) chez des enfants francophones et japonophones de 18 mois, qui indique que l'enfant a acquis les caractéristiques prosodiques de sa langue maternelle : les enfants francophones produisent une proportion importante de mots bisyllabiques caractérisés par un allongement final et un contour montant, alors que les enfants japonophones ont tendance à produire des mots avec un contour descendant et des durées de syllabes équivalentes en position non-finale et finale.

4.2. Ordre d'apparition des phonèmes

L'émergence de la production de parole par l'enfant consiste donc en une suite d'étapes bien précises, liées à la fois à l'acquisition du contrôle articulatoire et à l'évolution des capacités perceptives qui se spécialisent peu à peu vers la langue maternelle. Ainsi, si les étapes de l'acquisition sont communes à tous les enfants, l'émergence des sons de parole est influencée par l'environnement linguistique auquel ils sont exposés.

Rondal (1999) propose une chronologie de l'ordre d'acquisition des phonèmes chez des enfants francophones (Figure 1.1 ci-dessous), en distinguant « émergence » (qui correspond à la première apparition du phonème dans l'inventaire de l'enfant) et « acquisition » du phonème (qui correspond à son utilisation systématique et appropriée dans des mots, par l'enfant).



Note. Les traits correspondent en leur point de départ à l'âge auquel environ 50% des enfants prononcent les phonèmes correctement et en leur point d'arrivée à l'âge auquel le phonème est acquis par la très grande majorité des enfants.

Figure 1.1 : ordre d'apparition des phonèmes chez des enfants francophones (tiré de Rondal, 1999)

L'ordre d'acquisition, c'est-à-dire l'utilisation adéquate et systématique du phonème par l'enfant, proposé par Rondal (1999) est le suivant : voyelles orales /a/ et occlusives et nasales labiales /p, b, m/ sont les premiers phonèmes que l'enfant utilise, suivis des voyelles /i, u, o/, des voyelles intermédiaires et centrales /e, ɛ, ø, y/ puis des voyelles nasales /ã, ɔ̃, ɛ̃/, et des occlusives alvéolaires et vélaires /t,d,k,g/ et nasales /n, ɲ/, les fricatives /f, v/ et liquides /l, ʀ/, enfin fricatives /s, z, ʃ, ʒ/ sont les sons de parole que l'enfant acquiert le plus tardivement (jusqu'à l'âge de sept ans d'après Rondal, 1999).

Cette chronologie de l'émergence des phonèmes chez l'enfant francophone est donnée à titre indicatif, elle ne prend pas en compte le contexte dans lequel ceux-ci sont réalisés, et qui résulte notamment de la position du phonème dans le mot ou du statut accentuel de la syllabe ou du mot dans lequel il est placé.

L'étude de MacLeod et al. (2011) s'intéresse à la production des consonnes chez 156 enfants francophones québécois de 18 à 54 mois et met en évidence que les consonnes /t, m, n, z/ sont acquises avant 36 mois, que les consonnes /p, b, d, k, g, ɲ, f, v, ʀ, l, w, ɥ/ sont acquises entre 36 et 53 mois et que les consonnes /s, ʒ, ʃ, j/ sont les consonnes les plus tardivement acquises, après 53 mois. Elles notent également des différences selon la position de la consonne dans le mot (initiale, médiane, finale). MacLeod et al. (2011) montrent également que seuls cinq clusters consonantiques, /bl, fl, pw, kʁ, bw/ sur les dix qu'elles étudient sont acquis par les enfants à 53 mois. Elles notent également que les consonnes qui émergent (c'est-à-dire qui apparaissent pour la première fois dans l'inventaire de l'enfant) le plus tôt sont également celles qui sont acquises (ou utilisées de façon adéquate) le plus tôt par les enfants. A la différence des consonnes, il n'existe pas à notre connaissance d'étude comparable sur les voyelles, permettant d'établir une chronologie précise de l'émergence et de l'acquisition des voyelles par les enfants francophones.

Les étapes de l'acquisition du langage et de l'apparition des sons de parole restent relativement peu étudiées, en particulier chez l'enfant de plus de deux ans et chez l'enfant francophone : quel est le rôle des interactions entre perception et production dans ces étapes d'acquisition plus tardive ? Plusieurs modèles phonétiques et phonologiques, ont été proposés pour expliquer la chronologie de l'émergence des sons de parole et l'évolution de la production de la parole par l'enfant au cours de ses premières années. Nous avons choisi de présenter certains de ces modèles du développement phonologique dans le paragraphe §5 ci-après.

5. Développement phonologique : quelques modèles

Plusieurs modèles ont été proposés pour rendre compte du développement phonologique chez l'enfant. Ils sont soit descendants (ou « top-down ») soit ascendants (ou « bottom-up ») et s'inscrivent dans un courant théorique particulier de la phonologie : les modèles descendants postulent l'existence d'une structure phonologique sous-jacente qui va émerger progressivement au cours du développement de l'enfant, et à l'inverse, les modèles ascendants envisagent le développement phonologique comme une organisation progressive des sons de parole produits par l'enfant en un système de catégories phonologiques distinctes sous l'effet de contraintes articulatoires ou de processus production-perception et de l'exposition à la langue maternelle.

5.1. Modèles descendants (ou « top-down ») du développement phonologique

Nous avons choisi de présenter plusieurs modèles descendants : un modèle issu de la phonologie structuraliste (Jakobson, 1941, 1949), un modèle issu de la phonologie générative post-SPE (Wauquier-Gravelines, 2003) et un modèle proposé dans la théorie de l'optimalité (Smolensky, 1996).

La phonologie structuraliste associe fréquence des phonèmes dans la langue maternelle et ordre d'apparition dans le système phonologique de l'enfant. Jakobson (1941, 1949) par exemple s'intéresse aux liens entre développement phonologique et caractéristiques segmentales de la langue maternelle en se basant sur les caractéristiques acoustiques des segments. Sa modélisation s'inscrit dans une démarche structuraliste, qui à la suite de Saussure (1916) notamment voit la phonologie de la langue comme un système, dont les éléments constitutifs se distinguent grâce à une opposition de traits binaires. Dans cette approche théorique, les phonèmes sont hiérarchisés au sein du système selon leur marque. La marque est à la fois le reflet de la fréquence des sons dans les langues du monde et dans la langue considérée. Pour Jakobson (1941, 1949), le système phonologique de l'enfant se construit par opposition des traits qui caractérisent les phonèmes, et en fonction de leur « marque », ce qui va permettre de prédire leur ordre d'apparition dans la production de l'enfant, du moins marqué au plus marqué.

Les sons du système sont dits « marqués » ou « non-marqués » selon leur position dans la hiérarchie : les voyelles sont « non-marquées » et sont donc présentes dans les productions de l'enfant avant les consonnes qui sont quant à elles « marquées », les consonnes non-voisées par exemple sont « non-marquées » et vont apparaître avant les sons voisés qui sont « marqués » ; les occlusives sont « non-marquées » et apparaissent avant les fricatives, « marquées » ; les voyelles nasales sont « marquées » et apparaissent après les voyelles orales, « non-marquées ». L'acquisition suit donc un ordre précis et rigide, et il existe, selon Jakobson (1941, 1949), une chronologie universelle de l'acquisition, selon la fréquence des sons dans les langues du monde puis dans la langue maternelle (par exemple, le son /a/ est le premier à apparaître), « les oppositions relativement rares dans les langues du monde sont parmi les dernières acquisitions de l'enfant ».

Les approches nativiste et cognitiviste de l'acquisition sont à l'origine des modélisations du développement dans le cadre de la phonologie générative post-SPE (entre autres Wauquier-Gravelines, 2003) et de la théorie de l'optimalité (entre autres Smolensky, 1996). Ces deux courants s'accordent sur une modélisation comportant à la fois le niveau abstrait de la « compétence linguistique » qui correspond à une structure sous-jacente et celui de la « performance » qui est la réalisation de surface de la production de parole.

Dans le cadre de la phonologie générative post-SPE, (Wauquier-Gravelines, 2005), à partir du babillage, la production de l'enfant va passer par plusieurs étapes pour arriver une production similaire à celle de l'adulte. En effet, cette approche postule que le mot a une forme sous-jacente (ou gabarit) qui est universelle, de type CV(CV)CV (Wauquier-Gravelines, 2003) et comme l'enfant n'est pas capable de produire tous les sons de sa langue maternelle ni les clusters consonantiques avant un certain âge, sa production va s'adapter en propageant les sons les plus saillants à leurs positions adjacentes de façon à combler toutes les positions disponibles sur le gabarit, selon les règles phonologiques et phonotactiques de la langue. Au fur et à mesure de l'avancée du processus d'acquisition, l'enfant va avoir de plus en plus de sons à disposition, et les mots qu'il produit seront de plus en plus proches de la cible adulte.

La modélisation de Smolensky (1996) dans le cadre de la théorie de l'optimalité (OT) postule également l'existence d'une forme du mot sous-jacente dont la réalisation de surface résulte d'une succession de contraintes ordonnées qui autorise ou interdit certaines combinatoires de phonèmes, et qui correspond à une grammaire de la langue maternelle. Ainsi, Smolensky (1996) définit cette grammaire ou succession de contraintes de la façon suivante :

> « An OT grammar is to be understood now as a means of evaluating the relative Harmony (unmarkedness) of structural descriptions via a language-particular ranking of universal constraints ».[2]

D'après Smolensky (1996), comme la production de parole par l'enfant se fait en appliquant une série de contraintes à partir de la forme sous-jacente, il serait envisageable qu'à l'inverse, la compréhension nécessite une deuxième série de contraintes (et donc une deuxième grammaire) qui part de la forme de surface pour arriver à la forme sous-jacente du mot. Cependant Smolensky (1996) postule l'existence d'une grammaire unique qui permet à l'enfant d'avoir une représentation phonologique sous-jacente proche de celle de l'adulte, mais avec laquelle il a une production de surface limitée par des contraintes articulatoires.

5.2. Modèles ascendants (ou « bottom-up ») du développement phonologique

A la différence des modèles descendants (ou « top-down »), les modèles ascendants (ou « bottom-up ») postulent que le développement de la production de parole chez l'enfant se fait par acquisition progressive des gestes articulatoires, pour arriver à une production similaire à celle de l'adulte. Plusieurs types de modélisations ont été proposés pour rendre compte de cette acquisition « bottom-up » de la parole, parmi lesquels des modèles articulatoires (MacNeilage, 1998 et Davis & MacNeilage, 2004, Vihman, 1996 et Vihman et Croft, 2007), des modèles statistiques (Bybee, 2001, Pierrehumbert, 2001, 2003, Kern & Dos Santos, 2012 et Stokes et al., 2012).

A la différence de l'approche phonologique structuraliste de Jakobson (1941, 1949) qui propose un modèle qui lie caractéristiques acoustiques des sons de parole et développement phonologique, l'approche théorique « Frame, then content » (MacNeilage, 1998) développe une

[2] « Dans la théorie de l'optimalité (OT), une grammaire doit être comprise désormais comme un moyen d'évaluer l'Harmonie relative (la minimisation de la marque) des descriptions structurelles, à travers un ensemble de contraintes universelles ordonnées d'une façon spécifique : la langue. » (notre traduction)

modélisation du développement phonologique en lien avec les caractéristiques articulatoires des sons de parole.

L'approche articulatoire « Frame, then content » postule la coexistence de deux mécanismes articulatoires lors de la production de parole : l'oscillation de la mandibule (le cadre, « frame ») et les mouvements des articulateurs (le contenu, « content »). Ainsi, la production de parole chez l'adulte se décompose en oscillations de la mandibule accompagnées du contrôle de la position de la langue et du velum. Dans cette approche théorique, le développement de la production de parole est vu comme un processus de développement successif de ces deux mécanismes : le cadre puis le contenu (« frame then content »). Davis & MacNeilage (2004) proposent un modèle articulatoire dans lequel les premières structures proto-syllabiques vont être produites par l'oscillation de la mandibule, c'est-à-dire des cycles d'ouverture et de fermeture de la bouche, sans contrôle des autres articulateurs parmi lesquels la langue ou le velum : l'enfant produit dans un premier temps des sons grâce aux gestes articulatoires utilisés depuis la naissance pour se nourrir.

Ainsi, ce modèle prédit que les premiers sons produits lors du babillage ou des premiers mots sont le reflet des capacités articulatoires des enfants et de la plus grande facilité à produire des sons de parole dont les mouvements sont similaires à ceux qu'ils utilisent pour se nourrir (gestes de succion, mastication…). Ce modèle de développement phonologique à partir de gestes disponibles par ailleurs est résumé de la façon suivante par Davis & MacNeilage (2004) :

> « Speech ontogeny entails an initial frame stage involving basic capabilities available to the production system. Earliest frames provide evidence of self-organizing processes in a complex system rather than learning. »[3]

Ce modèle postule en outre une préférence, dans le babillage et les premiers mots, pour trois types de syllabes CV : [ba], [di] et [ku], pour lesquelles la voyelle est produite avec une inertie de la langue, dont le mouvement suit celui de la mandibule. Davis & MacNeilage (2004) notent également que l'enfant a une préférence pour les consonnes les plus antérieures (labiales et coronales) qui vont être produites avant les dorsales. De même, les séquences [ba] vont être produites avant les séquences [di] qui nécessitent une mise en place plus complexe des articulateurs et les occlusives seront produites avant les fricatives qui demandent un contrôle plus fin de la position des articulateurs.

Dans ce modèle, l'acquisition du langage et le développement phonologique en production sont liés en premier lieu aux contraintes physiologiques et articulatoires de l'enfant (qui sont des caractéristiques universelles, biologiques) et en deuxième lieu aux caractéristiques de la langue ambiante.

Le modèle de développement phonologique proposé par Vihman (1996) et Vihman & Croft (2007) prend en compte également les contraintes du développement articulatoire du bébé et son interaction avec l'input. Ainsi, le babillage et les premiers mots sont des tentatives de l'enfant de produire des mots similaires à ceux des adultes qui l'entourent sans pouvoir y parvenir. Ces tentatives sont des patrons idiosyncrasiques qui lui permettent de s'approcher des cibles adultes.

[3] « L'ontogénèse de la parole comprend un stade avec un cadre initial utilisant des capacités basiques disponibles dans le système de production. Les cadres les plus précoces témoignent de l'existence de processus auto-organisés en un système complexe plutôt que d'apprentissage. » (notre traduction)

Les premiers mots sont donc des formes relativement simplifiées des productions adultes, qui vont peu à peu se diversifier et se complexifier, en fonction de la maturation physiologique des différents articulateurs nécessaires à la production, mais également de l'expérience de l'enfant et de sa familiarité avec la langue maternelle. L'enfant produit donc des patrons de mots entiers, plutôt que des enchaînements de segments ou de syllabes, qui deviennent de plus en plus élaborés en fonction de l'exposition au lexique et de la stimulation par l'input proposé par les adultes de son environnement.

Bybee (2001) et Pierrehumbert (2001, 2003) proposent une modélisation dans le cadre de la théorie des exemplaires, qui lie le développement phonologique et la distribution statistique des phonèmes de la langue maternelle.

Pour Bybee (2001), le développement phonologique en production se fait par la répétition par l'enfant du geste articulatoire nécessaire pour produire un son, qui va être caractérisé tout d'abord par une variation initiale importante. Cependant, cette variation va avoir tendance à se réduire au fur et à mesure que la pratique de ce geste articulatoire augmente. Ainsi, d'après Bybee (2001), la variation va se réduire de plus en plus dans les mots les plus fréquemment utilisés par l'enfant, qui va être ensuite capable de généraliser ces gestes devenus systématiques aux nouveaux mots qu'il est amené à produire.

Dans une démarche similaire d'étude des relations entre développement phonologique et acquisition statistique de la langue maternelle, Pierrehumbert (2003) postule l'utilisation par l'enfant de deux niveaux (input et output) qui interagissent lors de l'acquisition : celle-ci est donc liée à l'input ainsi qu'à deux types de feedback, interne et environnemental. Dans ce modèle, l'acquisition consiste à utiliser à la fois l'input et le feedback pour créer des catégories phonologiques et à organiser le système, à partir d'une analyse statistique et acoustique des sons de parole. Le développement phonologique est vu par Pierrehumbert (2003) comme une émergence bottom-up des phonèmes à partir de modes statistiques utilisés dans l'espace phonétique. Cette émergence bottom-up est rendue possible par les boucles perception-production qui entrent en jeu dans la parole et se base sur deux postulats : 1) les propriétés statistiques superficielles de la parole disponibles chez le bébé reflètent les éléments contrastifs et distinctifs de la parole de l'adulte et 2) le système se développe grâce au feedback interne (qui consiste pour l'enfant à élaborer une hiérarchie des sons les plus fréquents) statistique plutôt que grâce au feedback lexical (qui est lié à l'exposition au lexique).

Kern & Dos Santos (2012) et Stokes et al. (2012) proposent également une analyse des relations entre « densité de voisinage » qui correspond au « nombre de voisins phonologiques que possède le mot à acquérir », « fréquence d'occurrence » des mots dans la langue (calculée à partir de la base de données Lexique3 (New et al., 2003)) et taille du vocabulaire chez l'enfant. Stokes et al. (2012) mettent en évidence une corrélation négative entre « densité de voisinage » des mots et taille du vocabulaire chez une population de 208 enfants de deux ans : plus les enfants ont un vocabulaire étendu, plus celui-ci se compose de mots phonologiquement éloignés (dont la densité de voisinage est faible), et à l'inverse, plus le vocabulaire des enfants est réduit, plus les mots sont proches phonologiquement. A l'inverse, la fréquence lexicale est positivement corrélée avec la taille du vocabulaire des enfants : plus l'enfant a un vocabulaire étendu, plus il utilisera des mots fréquents dans la langue. En parallèle à l'étude de Stokes et al. (2012) qui analyse les liens entre taille du vocabulaire, « densité de voisinage » des mots et « fréquence d'occurrence » chez deux

groupes d'enfants de deux ans (dont la taille du vocabulaire est réduite ou étendue), l'étude de Kern & Dos Santos (2012) s'intéresse aux différences entre ces mêmes liens chez deux groupes d'enfants d'âge différent (551 enfants de 8 à 16 mois et 208 enfants de 24 à 30). Les résultats de cette deuxième étude indiquent que chez les enfants les plus jeunes (de moins de 16 mois), il n'existe aucune relation entre taille du vocabulaire réceptif et productif d'une part et « densité de voisinage » et « fréquence d'occurrence » des mots, d'autre part. En revanche, chez les enfants plus âgés (24 à 30 mois), cette relation existe (il s'agit des résultats de l'étude de Stokes et al., 2012).

Ces différents modèles « bottom-up » du développement phonologique permettent donc de mettre en évidence la mise en œuvre conjointe de plusieurs mécanismes et contraintes, notamment articulatoires et perceptives, qui participent à la complexification des premières productions de parole par l'enfant, qui vont peu à peu se rapprocher de celles de l'adulte. La qualité et la quantité de l'input que l'enfant reçoit au cours de ses premières années joue également un rôle important dans l'accès au langage. Les liens entre fréquence lexicale, et diversification du vocabulaire et entre densité de voisinage phonologique et diversification du vocabulaire semblent établis, mais qu'en est-il du processus inverse : le développement de catégories phonologiques de plus en plus clairement distinctes est-il lié à la fréquence et à la complexité des mots dans la langue maternelle et donc dans l'input que l'enfant reçoit ? Parallèlement aux contraintes articulatoires et perceptives, notre étude doit prendre en compte également les caractéristiques de l'input dont dispose l'enfant, pour pouvoir comprendre quels mécanismes favorisent ou entravent le développement phonologique d'enfants ayant eu un accès tardif et partiel à l'oral.

5.3. Modèles de développement plus tardif

Dans ce chapitre, nous avons présenté des modèles du développement phonologique des premiers mois de l'enfant. Il existe très peu de modèles de développement plus tardif chez l'enfant de plus de 18 mois, et même très peu de descriptions des acquisitions phonologiques plus tardives chez l'enfant. Menn & Stoel-Gammon (1995) proposent une analyse du développement phonologique et morpho-phonologique chez des enfants de plus de deux ans : elles distinguent l'acquisition de contrastes et de règles phonologiques saillantes, comme par exemple les règles de voisement, acquises assez tôt, de celle plus tardive des règles morpho-phonologiques, comme par exemple la règle du marquage du pluriel par une sibilante en anglais. Après le stade des premiers mots, les enfants produisent des énoncés de plus en plus complexes, et ont tendance à généraliser des règles phonologiques qui sont en voie d'acquisition. Menn & Stoel-Gammon (1995) notent également que l'acquisition de règles phonologiques et de sandhi segmentaux et tonals reste très peu étudiée et se poursuivrait jusqu'à l'adolescence.

L'analyse de Boysson-Bardies (1996) va dans ce sens d'une acquisition tardive de certaines règles phonologiques. En effet, elle indique que si les règles phonologiques sont acquises par les enfants vers la fin de la deuxième année, avec une régularité des formes sonores des mots de plus en plus grande, accompagnée d'une systématisation de ces règles phonologiques, des difficultés de production subsistent tout de même, en particulier au niveau de la production des clusters consonantiques, de la production des liquides [l] et [r], du dévoisement final inapproprié et du contraste entre fricatives [s] et [ʃ] qui peut être instable jusqu'à l'âge de six ou sept ans.

6. Conclusion

A travers cette présentation d'une sélection de théories d'acquisition du langage et de modèles de développement phonologique différents et divergents, il apparaît clairement nécessaire dans le cadre d'une étude sur la production de la parole de l'enfant de s'interroger sur les facteurs internes et externes qui interviennent dans les premières années de vie de l'enfant et qui influencent sa production de parole. Capacités perceptives et articulatoires de l'enfant, mais également input langagier et interactions avec d'autres locuteurs dès la naissance contribuent à l'émergence de la parole, et nous devrons tenir compte de ces deux aspects, pour pouvoir caractériser les difficultés rencontrées par les enfants porteurs d'implants cochléaires de notre étude et les différentes étapes de leur production de parole.

Les premières années de vie, qui correspondent également aux premières années d'exposition à l'oral, semblent être cruciales dans l'émergence de la parole chez l'enfant typique : l'enfant reçoit un flux d'informations sur sa langue maternelle, lors des interactions avec son entourage, à partir duquel il accède à la phonologie de sa langue maternelle et construit des représentations phonologiques pour chacune des catégories de son système. Au cours de cette période, il va progressivement développer le contrôle de ses articulateurs, et finalement être capable de produire les phonèmes de sa langue maternelle dans tous types de contextes : il se construit donc une représentation motrice associée à chacune des catégories phonologiques. En retour, le développement du contrôle moteur participe à la structuration de l'espace perceptif et peut-être du système phonologique (Dole et al., en préparation).

Cette période dite « critique » correspond chez l'enfant typique aux premières années de vie et aux premières années d'exposition à l'oral. On peut donc se demander lequel de ces deux facteurs est critique : premiers temps de vie ou d'exposition à l'audio. La situation de l'enfant sourd prélingual porteur d'implant cochléaire est particulièrement pertinente par rapport à ce questionnement puisqu'il n'a pas eu accès à l'audio dans ses premiers temps de vie, mais qu'il y est exposé à partir de l'implantation cochléaire.

D'autre part, l'accès à l'oral avec un implant cochléaire reste, par définition, partiel (voir le chapitre 3 ci-après), puisqu'il s'agit d'un dispositif artificiel de remédiation de la surdité qui permet de transmettre certaines caractéristiques acoustiques des sons de parole et des sons de l'environnement mais ne permet pas à l'heure actuelle de recréer une audition parfaite. La question de la construction de représentations phonologiques et de l'émergence de la parole à partir d'une perception auditive partielle est donc particulièrement pertinente et permet de s'interroger précisément sur les rôles conjoints de la perception et de l'articulation dans le développement phonologique.

Toute l'originalité de notre thèse réside donc dans le fait que nous nous intéressons aux effets à long terme de l'utilisation de l'implant chez des enfants sourds pré-linguaux. La question que nous nous posons est la suivante : est-il possible, en ayant passé les premières années de vie sans accès à l'audio, de construire une catégorisation suffisante des sons de l'environnement et notamment des sons de parole ? Par voie de conséquence, la production de parole peut-elle être similaire à celle d'enfants normo-entendants du même âge ?

CHAPITRE 2 : Audition, surdité et implant cochléaire

1. Introduction

Comme nous l'avons vu au chapitre précédent, perception et production de parole sont intimement liées dans le développement du petit enfant et plusieurs auteurs, parmi lesquels Guerrero-Lopez (2010), s'intéressent à la boucle « audio-phonatoire » qui permet à l'enfant de construire sa représentation des sons, puis de les reproduire en utilisant le feedback auditif de sa propre production. L'enjeu de cette thèse est donc de comprendre quels sont les processus d'acquisition de la parole par des enfants sourds, qui accèdent à l'audition à travers l'implant cochléaire, qui est un dispositif qui restaure une perception partielle des sons de parole et des sons de l'environnement. Pour évaluer l'impact de cette perception partielle sur la production de parole et sur l'intelligibilité des enfants qui utilisent ce dispositif, nous nous proposons au préalable de décrire le fonctionnement non pathologique de l'oreille, puis nous expliquerons quels sont les différents types de surdité et de solutions d'appareillage disponibles, ensuite nous détaillerons le fonctionnement d'un implant cochléaire, et enfin nous nous intéresserons aux conséquences de l'utilisation de cet implant sur les capacités de communication et la qualité de vie des enfants qui l'utilisent.

2. Fonctionnement non pathologique de l'oreille

Pour comprendre les problèmes que pose une audition dégradée ou partielle, il nous paraît fondamental de décrire tout d'abord le fonctionnement non pathologique de l'oreille. C'est l'objet de la première partie de ce chapitre, où nous allons expliquer quels sont les mécanismes mis en jeu lorsqu'un son de parole ou un son de l'environnement est capté et traité par le système auditif périphérique, avant d'être traité par le système nerveux central. Pour proposer cette description des mécanismes de l'audition, nous nous appuierons sur les travaux de Johnson (2012), Chouard (2009) et Pujol et al. (2013).

Le système auditif périphérique se décompose en trois niveaux : oreille externe, oreille moyenne et oreille interne. La figure 2.1 ci-dessous propose une représentation schématique de ces trois niveaux constitutifs du système auditif périphérique. La suite de ce chapitre présente en détail chacun de ces niveaux et leur fonctionnement.

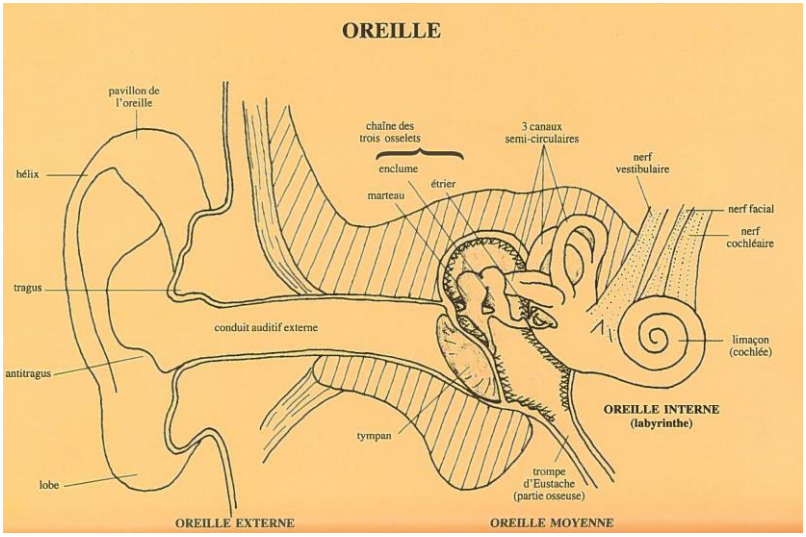


Figure 2.1 : schéma de l'oreille, tiré de Brin et al. (2004)

2.1. Oreille externe

L'oreille externe se compose du pavillon, composé d'un cartilage recouvert de peau, qui permet de capter les sons de l'environnement et du conduit auditif externe qui les transmet jusqu'au tympan, qui est une membrane située à son extrémité. Le tympan est à l'interface de l'oreille externe et de l'oreille moyenne. Dans l'oreille externe, la transmission du son se fait de façon aérienne. Le système composé des deux oreilles externes capte le son en stéréo, ce qui permettra plus tard de localiser également le son dans l'espace.

2.2. Oreille moyenne

L'oreille moyenne se compose du tympan, de trois osselets (marteau, enclume et étrier), de la fenêtre ovale qui est à l'interface entre oreille moyenne et oreille interne et de la trompe d'Eustache. L'onde acoustique arrive par l'oreille externe jusqu'au tympan qui entre en vibration en exerçant à son tour une pression mécanique sur les osselets qui se mettent à bouger l'un après l'autre, en amplifiant les vibrations du son. L'étrier exerce finalement une pression mécanique sur la fenêtre ovale et fait passer le son vers l'oreille interne. La trompe d'Eustache permet de réguler la pression d'air à l'intérieur de l'oreille moyenne.

2.3. Oreille interne

2.3.1. Composition de l'oreille interne

L'oreille interne se compose de la cochlée et du vestibule. A la base de la cochlée se trouvent la fenêtre ovale qui se situe à l'entrée de la rampe vestibulaire, et la fenêtre ronde qui est à l'entrée de la rampe tympanique. La rampe vestibulaire et la rampe tympanique sont les deux circuits qui forment la cochlée. Le vestibule n'intervient pas pour l'audition, mais il participe à la perception des positions et des mouvements du corps dans l'espace, ce qui permet le maintien de l'équilibre général de l'individu. La pression mécanique exercée par l'étrier au niveau de la fenêtre ovale transmet des vibrations qui modifient la pression hydraulique du liquide qui emplit l'intérieur de la cochlée, transformant ainsi cette énergie mécanique en vibrations qui se propagent dans un milieu liquide.

2.3.2. Fonctionnement de la cochlée

La cochlée est un élément de l'oreille interne, reconnaissable à sa forme de coquille d'escargot (du latin *cochlea*, escargot). La figure 2.2 ci-dessous présente une coupe transversale schématique de la cochlée et de ses éléments constitutifs.

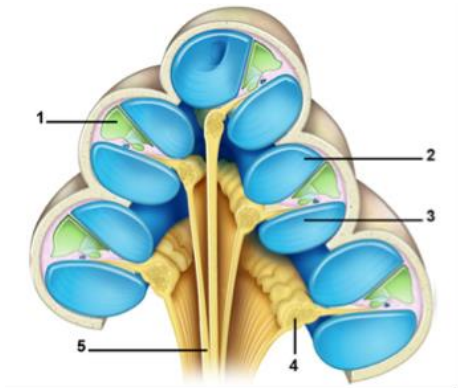


Figure 2.2 : schéma d'une coupe transversale de la cochlée, figure tirée de Pujol et al. (2013), www.cochlea.eu. 1. Canal cochléaire, 2. Rampe vestibulaire, 3. Rampe tympanique, 4. Ganglion cochléaire, 5. Nerf auditif.

La cochlée est un tube d'environ 30-35 mm de long, tapissé d'une membrane (membrane basilaire) et rempli de liquide, et qui s'enroule autour du nerf auditif avec lequel il est relié au niveau de l'organe de Corti. L'organe de Corti comprend notamment des cellules ciliées internes et externes qui transforment les vibrations qui se propagent dans le liquide à l'intérieur de la cochlée en impulsions nerveuses qui vont ensuite stimuler les fibres qui composent le nerf auditif. Il s'agit à ce niveau d'un phénomène de transduction mécano-électrique : l'onde sonore est transformée en influx nerveux.

La cochlée ne transmet pas les fréquences de façon linéaire mais tonotopique, c'est-à-dire que chaque fréquence va être codée à une localisation bien particulière de la membrane basilaire : la distance qui sépare deux zones de la cochlée ne correspond pas à un écart fixe entre deux fréquences traitées par celles-ci. Les cellules ciliées placées le long de la cochlée vont être stimulées différemment selon la fréquence des sons perçus. Les deux schémas présentés sur la figure 2.3 ci-dessous détaillent la répartition des zones de traitement des fréquences dans la cochlée.

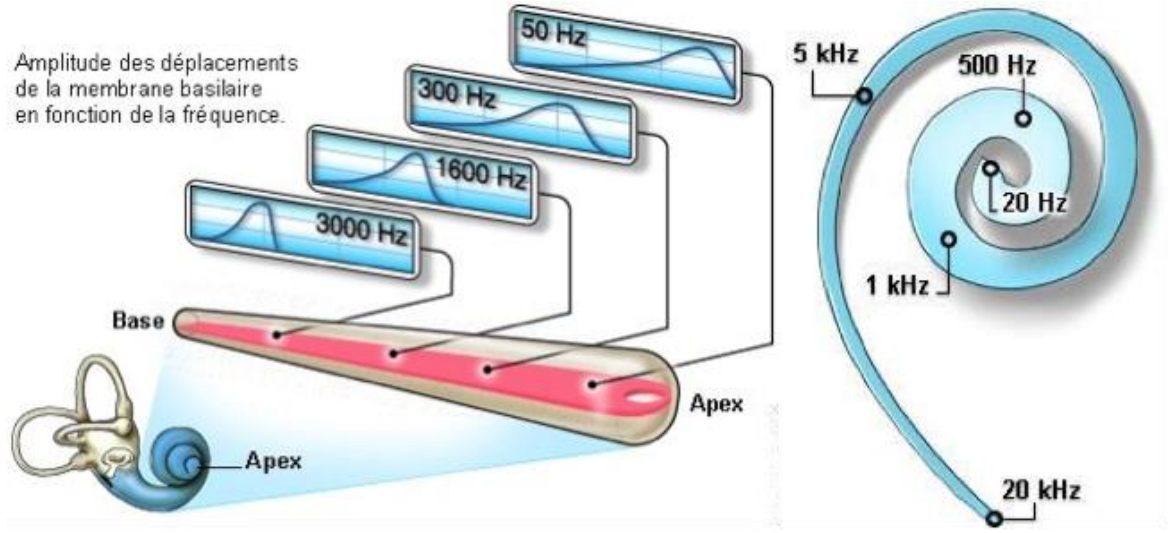


Figure 2.3 : distribution tonotopique des fréquences à l'intérieur de la cochlée, schémas tirés de Pujol et al. (2013), www.cochlea.eu.

La cochlée est plus petite à sa base (et donc entre plus facilement en résonance avec les hautes fréquences), et plus large au niveau de l'apex (et donc entre plus facilement en résonance avec les basses fréquences) : les fréquences les plus hautes sont traitées à la base de la cochlée et les plus basses sont traitées au niveau de l'apex. Les sons sont perçus sur une échelle de perception qui n'est pas linéaire, et plusieurs échelles ont été proposées pour rendre compte de cette répartition tonotopique des fréquences perçues. Les échelles les plus courantes sont ERB (*Equivalent Rectangular Bandwith*, Glasberg & Moore (1990)) et Bark (Traunmüller (1990)).

3. Surdité

3.1. Caractéristiques de la surdité

Il existe plusieurs critères qui permettent de proposer un classement des surdités chez l'enfant. Ces critères sont détaillés dans plusieurs articles, notamment ceux de Mondain et al. (2005), et Lina-Granade & Truy (2005). Nous avons choisi de nous appuyer sur ces deux articles pour proposer un aperçu des caractéristiques des surdités qui peuvent survenir chez l'enfant, et qui peuvent nécessiter le recours à un dispositif auditif. Toutes ces caractéristiques permettent d'établir une classification des surdités en fonction de trois critères principaux : le côté de la surdité, le degré de la surdité et son étiologie. Ces caractéristiques sont représentées dans la figure 2.4 ci-dessous, tirée de Mondain et al. (2005). Ces caractéristiques vont ensuite permettre à l'équipe médicale de proposer un dispositif et une rééducation orthophonique les plus adaptés à l'enfant.

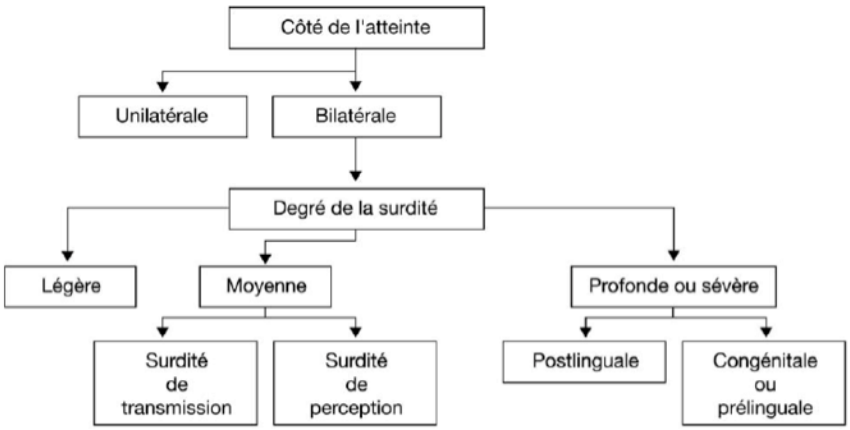


Figure 2.4 : Classification des types de surdités chez l'enfant, tirée de Mondain et al. (2005)

3.1.1. Surdités unilatérales et bilatérales, âge de survenue de la surdité, étiologies

Chez l'adulte, comme chez l'enfant, la surdité peut être unilatérale ou bilatérale, c'est-à-dire qu'elle affecte une seule oreille ou les deux oreilles, ce qui va avoir des répercussions différentes sur l'audition et la qualité de vie de la personne concernée. Chez les enfants, l'âge de survenue de la surdité est également un critère déterminant sur l'évolution des capacités de communication. Mondain et al. (2005) et Lina-Granade & Truy (2005) présentent les différents types de surdité en fonction de l'âge de survenue de celle-ci : si la surdité est présente dès la naissance, il s'agit d'une surdité congénitale ; si elle survient avant l'âge de deux ans, c'est une surdité prélinguale ; entre deux et quatre ans, une surdité périlinguale, et après quatre ans, une surdité postlinguale. Les effets de l'âge de survenue vont être déterminants dans l'accès des enfants à l'oral. En effet, plus la surdité est précoce, moins les enfants vont avoir eu l'occasion de baigner dans un environnement oral, et les éventuelles remédiations de la surdité vont devoir construire des représentations langagières et phonologiques à partir de très peu d'expériences de l'oral.

La surdité peut avoir plusieurs origines, qui vont également déterminer le site de l'oreille qui va être affecté par la perte auditive : la surdité peut être notamment d'origine génétique, ou acquise (liée à un traumatisme, une pathologie, une infection, etc. voir Mondain et al. (2005) pour une revue détaillée de toutes les étiologies de la surdité chez l'enfant).

3.1.2. Degré de surdité

Les surdités peuvent être classées selon le degré de perte auditive subie par la personne sourde. Ce degré de surdité est évalué par des tests audiologiques qui permettent de situer la perte auditive par rapport à des seuils d'audition, définis par le Bureau International d'Audiophonologie. Ces seuils sont calculés à partir de tests de perception de certaines fréquences (500 Hz, 1000 Hz, 2000 Hz, 4000 Hz). La table 2.1 ci-dessous présente ces différents degrés de surdité et la perte auditive correspondante.

Degré de surdité	Perte auditive moyenne de 0,5 à 4 kHz
Déficience auditive légère	Perte de 21 à 40 dB
Déficience auditive moyenne	1er degré : perte de 41 à 55 dB 2ème degré : perte de 56 à 70 dB
Déficience auditive sévère	1er degré : perte de 71 à 80 dB 2ème degré : perte de 81 à 90 dB
Déficience auditive profonde	1er degré : perte de 91 à 100 dB 2ème degré : perte de 101 à 110 dB 3ème degré : perte de 111 à 120 dB

Table 2.1 : degré de surdité et niveau de perte auditive (d'après Mondain et al. (2005) et Lina-Granade & Truy (2005))

3.1.3. Types de surdité

Le dernier critère qui permet de caractériser la surdité est le mécanisme de l'atteinte. Il existe trois types de mécanismes qui vont provoquer la surdité, et qui affectent des sites différents du système auditif : surdité de transmission, surdité de perception et surdité centrale. Il existe également des surdités mixtes (surdité de transmission et surdité de perception conjointes). Les surdités de transmission et de perception (ainsi que les surdités mixtes) affectent le système auditif périphérique alors que la surdité centrale affecte le système auditif central.

3.1.3.1. Surdité de transmission

La surdité de transmission est une atteinte de l'oreille externe ou moyenne. C'est la surdité la plus fréquente chez l'enfant. D'après Mondain et al. (2005), la surdité de transmission est liée à des otites séromuqueuses dans 95% des cas. Elle correspond à une perte auditive généralement inférieure à 60 dB et ne se manifeste pas par une distorsion acoustique. Les personnes atteintes de surdité de transmission ont un retour auditif de leur propre voix par voie osseuse et peuvent ainsi contrôler l'intensité, la mélodie et l'intensité de leurs productions de parole.

3.1.3.2. Surdité de perception

La surdité de perception est une atteinte de l'oreille interne et/ou des voies auditives. Elle est généralement liée à une dégradation ou une absence de cellules ciliées qui tapissent la cochlée et qui transmettent les sons aux fibres qui composent le nerf auditif (voir ci-dessus, §2.3.2, ce chapitre). D'après Mondain et al. (2005), la surdité de perception est d'origine génétique chez 50 à 60% des enfants. La surdité de perception peut être légère, moyenne, sévère, profonde ou totale. Elle correspond à une distorsion fréquentielle et éventuellement à une distorsion d'amplitude.

3.1.3.3. Surdité centrale

La surdité centrale est liée à une atteinte au niveau du système auditif central, c'est-à-dire à une atteinte du nerf auditif ou du tronc cérébral, qui bloque la transmission du son.

3.2. Types de dispositifs de remédiation auditive disponibles

Il existe plusieurs types de dispositifs qui peuvent permettre aux personnes sourdes d'améliorer leur niveau d'audition : dispositifs à conduction aérienne, dispositifs à conduction osseuse et dispositifs par stimulation électrique (Mondain et al., 2005). Chaque type de dispositif est proposé par une équipe médicale, en fonction du type de surdité.

Les dispositifs à conduction aérienne peuvent être soit des prothèses auditives qui captent, amplifient et transmettent les sons, soit des implants de l'oreille moyenne. Les prothèses auditives sont des dispositifs externes et amovibles, qui se fixent au niveau du pavillon.

Les dispositifs à conduction osseuse sont de deux types : les prothèses BAHA, qui sont des systèmes amovibles placés derrière l'oreille, et qui captent le son et le transmettent par des vibrations au niveau des os du crâne, et les prothèses Tactaid qui captent et transmettent le son par des vibrations au niveau du dos ou du torse.

Les dispositifs par stimulation électronique sont soit des implants cochléaires qui stimulent des fibres du nerf auditif au niveau de la cochlée (nous détaillerons le fonctionnement de l'implant cochléaire dans le paragraphe §4 ci-après), soit des implants du tronc cérébral, qui sont placés au niveau du nerf auditif lorsque celui-ci ne transmet plus le son.

Notre projet étant d'étudier les caractéristiques de la parole d'enfants sourds, porteurs d'implants cochléaires, nous avons fait le choix de présenter ce seul dispositif de façon détaillée dans le paragraphe §4 ci-après, en laissant de côté les autres types d'appareillages disponibles.

4. Implant cochléaire

L'implant cochléaire est un dispositif récent de remédiation de la surdité, dont les premiers prototypes ont été élaborés entre 1957 et 1985, avant d'être commercialisés à partir du milieu des années 1980[4]. En 2010, plus de 200 000 personnes dans le monde utilisent des implants cochléaires, dont plus de 10 000 en France (adultes et enfants). Il existe plusieurs modèles d'implants, proposés par quatre fabricants : Neurelec (fabricant français) produit l'implant « Digisonic », Advanced Bionics Corporation (fabricant suisse) produit l'implant « Clarion », Cochlear Corporation (fabricant australien) produit l'implant « Nucleus » et MedEl Corporation (fabricant autrichien) produit l'implant « Combi40 ». Le coût d'un implant cochléaire en France est estimé à 25 000 € (en 2010, d'après Garabédian et Loundon (2010)), et son coût global incluant bilans pré-opératoires, opération, implant et prise en charge post-opératoire (réglages et rééducation) est estimé à 45 000 € d'après l'association CISIC (CISIC (2014)). En France, les coûts relatifs à l'implant et son utilisation sont pris en charge à 100% par la sécurité sociale dans le cadre d'une ALD (Affection Longue Durée), et les familles peuvent également bénéficier d'une Allocation d'Education Spécialisée de la CAF (Mondain et al. (2005)), pour prendre en charge l'appareillage auditif bilatéral et son renouvellement pendant vingt ans, la rééducation orthophonique, le soutien éducatif et le suivi audiologique. Pour mieux comprendre les enjeux de l'utilisation de l'implant cochléaire, nous allons décrire la composition (§4.1 ci-après) et le fonctionnement (§4.2 ci-après) d'un implant cochléaire, puis nous nous intéresserons aux conséquences de l'implant sur les capacités de perception, de communication et sur la qualité de vie des enfants qui l'utilisent. Les paragraphes §4.1 et §4.2 s'appuient sur les articles de Truy & Lina (2003) et Garabédian & Loundon (2010).

4.1. Composition de l'implant cochléaire

L'implant cochléaire est un dispositif à stimulation électrique, qui se compose d'un élément externe, amovible et d'un élément interne, implanté par chirurgie sous anesthésie générale. La figure 2.5 ci-dessous présente l'élément externe et l'élément interne constitutifs d'un implant cochléaire, du fabricant MedEl.

Figure 2.5 : composant externe (gauche) et composant interne (droite) d'un implant cochléaire, tiré du site du fabricant Med El, consulté le 19 mars 2016 : www.medel.com[5]

[4] CISIC (2014), *L'ambition d'entendre - Guide CISIC 2014*, page 7, version électronique disponible sur www.cisic.fr

[5] http://www.medel.com/fr/image-gallery

Le composant externe se compose d'un microphone, d'un processeur et d'une antenne externe, et le composant interne se compose d'une antenne interne, d'un récepteur-stimulateur, d'un porte-électrodes et d'électrodes.

4.2. Fonctionnement de l'implant cochléaire et stratégies de codage

L'implant cochléaire permet de transmettre le son présent dans l'environnement au nerf auditif, en contournant le système auditif périphérique qui, en raison d'une dégradation de certains de ses composants, ne transmet pas suffisamment d'informations au système central pour que la personne puisse percevoir et traiter ces sons.

4.2.1. Fonctionnement général de l'implant cochléaire

La figure 2.6 ci-dessous présente la localisation des différents éléments de l'implant cochléaire, chez une personne sourde ayant reçu cet implant. Dans ce paragraphe, nous allons tout d'abord proposer une description du fonctionnement de l'implant et du rôle de chaque composant dans la transmission du son, puis nous présenterons les principes généraux des principales stratégies de codage qui permettent à l'implant de traiter et transmettre le son.

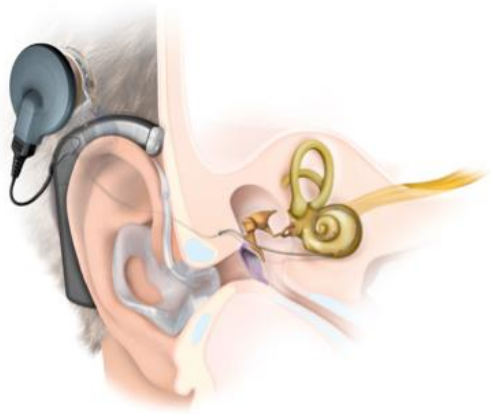


Figure 2.6 : schéma représentant un implant cochléaire en fonctionnement, tiré du site internet du fabricant MedEl, consulté le 19 mars 2016.[6]

Dans un premier temps, un microphone externe capte l'information acoustique disponible dans l'environnement (sons de parole et bruits ambiants), puis convertit cette information acoustique en signal électronique. Ce microphone est relié à un processeur qui code le signal électronique, le filtre et le décompose en portions de signal, en fonction des caractéristiques de fréquence, de durée et d'intensité. L'ensemble microphone-processeur peut avoir la forme d'un boitier ou d'un contour d'oreille, et permet à son utilisateur de contrôler le volume et la sensibilité du dispositif. Les portions de signal traitées par cet ensemble microphone-processeur sont ensuite envoyées vers l'antenne externe, placée derrière l'oreille. Cette antenne externe envoie ensuite le son décomposé vers la partie interne de l'implant auquel elle est reliée par un système magnétique.

Le composant interne capte le son transmis par l'antenne externe au moyen d'une antenne interne, fixée sous la peau derrière l'oreille. Elle reçoit le signal décomposé et le transmet au récepteur-stimulateur qui à son tour traite le signal et stimule les électrodes placées dans la cochlée sur un porte-électrodes. Le porte-électrode est un système plus ou moins rigide, inséré à

[6] http://www.medel.com/fr/image-gallery

l'intérieur de la cochlée, et sur lequel sont fixées les électrodes. Les implants actuels ont entre 15 et 22 électrodes, leur nombre étant variable selon les fabricants et les modèles d'implants. Le récepteur-stimulateur va ainsi pouvoir transmettre les portions de sons au nerf auditif, par des impulsions électriques diffusées au niveau des électrodes. Ces impulsions électriques vont à leur tour stimuler l'intérieur de la cochlée, et ainsi transmettre le son vers les fibres du nerf auditif, remplaçant ainsi les cellules ciliées qui sont déficientes ou absentes.

4.2.2. Stratégies de codage des fréquences des sons de parole par l'implant cochléaire

Il existe plusieurs stratégies de codage des sons de parole par l'implant cochléaire, qui dépendent entre autres des modèles de l'implant, du nombre d'électrodes, de la programmation des processeurs, et des choix des audioprothésistes qui effectuent les réglages. Nous proposons dans cette section un aperçu de ces stratégies de codage, décrites dans Truy & Lina (2003).

Les premières stratégies sont les stratégies d'analyse F0/F1/F2, MPeak (ou multi-peak), Speak (ou spectral peak), qui traitent les sons et s'adaptent au plus près aux caractéristiques de la parole, en stimulant une seule électrode ou un groupe d'électrodes simultanément : des impulsions électriques, qui reproduisent un seul son, stimulent donc plusieurs zones de la cochlée en même temps. Le deuxième type de stratégies se base sur une analyse temporelle du son (alors que les premières utilisent un traitement fréquentiel). Il s'agit des stratégies CA, CIS, multi-pulse stimulation (MPS) et SAS qui fonctionnent par stimulation séquentielle des électrodes (dans un ordre spécifique) pour coder chaque son de parole : le son est reproduit sous la forme d'un enchaînement de stimulations électriques. Le troisième type de stratégies est la stratégie « n of m », où m est un nombre maximal prédéfini d'électrodes à stimuler, et n est le nombre effectif des électrodes utilisées pour la stimulation en fonction du son à coder : le codage d'un son se fait par la stimulation de plusieurs électrodes dont le nombre s'adapte à ce son précis et plutôt qu'à un nombre prédéfini d'électrodes, défini par la personne qui a effectué le réglage de l'implant.

La plupart des stratégies actuelles a une vitesse de 250 impulsions par seconde, mais certaines stratégies utilisent des fréquences de stimulation beaucoup plus élevées (par exemple les stratégies CIS/CA et SAS fonctionnent avec 883 à 1500 impulsions par seconde).

Il est cependant important de noter que toutes ces stratégies ne permettent encore pas, à l'heure actuelle, de reproduire l'intégralité des sons et que l'utilisateur de l'implant n'a accès qu'à une partie de l'information acoustique de ces sons. La figure 2.7 ci-dessous, tirée de Fuller et al. (2014), permet de visualiser la différence entre les sons produits de façon naturelle et avec un implant cochléaire.

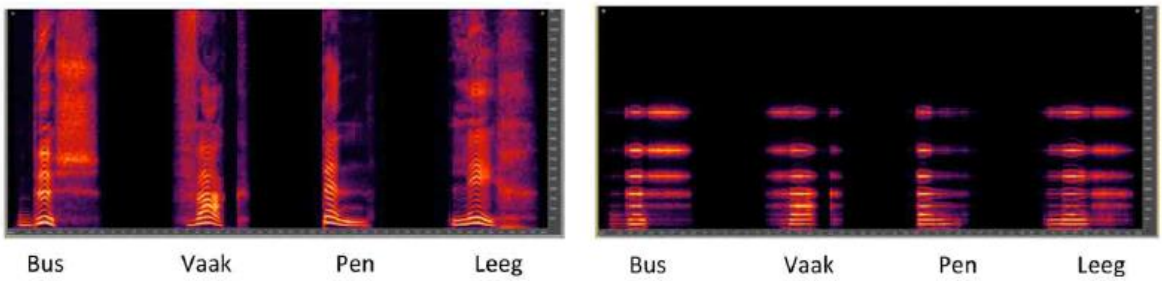


Figure 2.7 : spectrogrammes de quatre mots en néerlandais (de gauche à droite : bus, souvent, stylo, vide), produits de façon naturelle (gauche) et avec un implant cochléaire (droite), tirés de Fuller et al. (2014)

L'implant cochléaire permet donc de transmettre les sons de l'environnement (son de parole et bruits ambiants) à son utilisateur par un système qui contourne le système auditif périphérique et qui stimule directement les fibres du nerf auditif par des impulsions électriques à l'intérieur de la cochlée. Les utilisateurs (adultes et enfants) de l'implant ont donc accès dans une certaine mesure à ces sons, mais ceux-ci ne sont que partiellement perçus, ce qui résulte du traitement et du codage du son par l'implant. L'audition ainsi restaurée n'est pas similaire à celle d'une personne entendante et la perception de la parole et du bruit reste partielle, certaines fréquences et certains sons n'étant pas perçus par les utilisateurs d'implants cochléaires. Ainsi, les enfants sourds prélinguaux qui accèdent à l'audition par ce biais ont des capacités perceptives différentes des enfants normo-entendants, et l'utilisation de l'implant va avoir des conséquences sur les capacités de communication et la qualité de vie de ces enfants. Dans la suite de ce chapitre, nous proposons un bref aperçu des conséquences de l'implant cochléaire sur les capacités perceptives et sur les capacités de communication des enfants sourds porteurs d'implant ainsi que sur leur qualité de vie, à partir d'une sélection d'études disponibles dans la littérature.

5. Conséquence de l'implant sur les capacités de communication et la qualité de vie

5.1. Capacités perceptives

Un certain nombre d'études se sont intéressées aux capacités perceptives pour la parole des enfants porteurs d'implant cochléaire. Parmi celles qui évaluent un ou plusieurs critères de la perception des sons de parole par les enfants, nous avons choisi de présenter quelques études qui nous semblent être représentatives des résultats obtenus dans ce domaine. Cette sélection est non-exhaustive.

Les deux articles de Pisoni (2005) et Bouton et al. (2012) s'intéressent aux capacités des enfants porteurs d'implant cochléaire à percevoir des contrastes de lieux et modes d'articulation, voisement, nasalité, antériorité et hauteur.

Pisoni (2005) présente les résultats d'une étude réalisée auprès d'enfants porteurs d'implant cochléaire et d'enfants normo-entendants du même âge lors de tests de discrimination de traits phonologiques. Plusieurs tests standardisés de reconnaissance de mots de complexité lexicale variable dans des paires minimales, dans des syntagmes, sont utilisés pour évaluer les capacités des enfants à percevoir les sons de parole. Les résultats permettent de distinguer deux groupes d'enfants (« stars » et « low-performing »[7]) qui ont des mécanismes de perception différents : les enfants « stars » arrivent à des résultats approchant ceux des enfants normo-entendants, et peuvent utiliser les informations lexicales pour identifier et discriminer des sons de parole, alors que les enfants « low-performing » ont des difficultés plus marquées pour effectuer les mêmes tâches et s'appuient sur la seule information auditive apportée par l'implant pour effectuer ces tâches : quatre ans après implantation, les enfants « low-performing » arrivent à peine à discriminer des différences de modes d'articulation, mais n'arrivent toujours pas à discriminer le voisement et le lieu d'articulation des consonnes. Pisoni (2005) indique également que les enfants qui n'utilisent qu'une communication « orale » ont des capacités perceptives plus élevées que les enfants qui utilisent une communication « totale » (c'est-à-dire qui utilisent l'oral accompagné de signes ou gestes), ce qu'il explique par une acquisition plus active des représentations lexicales et phonologiques par les enfants qui n'utilisent que l'oral. A la différence de l'étude de Pisoni (2005), qui utilise des tâches d'imitation de mots, Bouton et al. (2012) quant à eux utilisent des tâches de discrimination et d'identification de mots pour évaluer les compétences perceptives des enfants porteurs d'implants et d'enfants normo-entendants, tous francophones. Il ressort de cette étude que les enfants porteurs d'implant cochléaire ont des scores plus bas que les enfants normo-entendants aux deux tâches, mais que ces résultats diffèrent entre les traits phonologiques considérés : pour les consonnes, le mode d'articulation et le voisement sont mieux perçus que la nasalité et le lieu et pour les voyelles, l'antériorité et l'aperture sont mieux perçues que la nasalité, chez les enfants porteurs d'implant cochléaire. Ces résultats moins bons sont d'après Bouton et al. (2012) liés à la période dite « sensible » dans l'acquisition du langage qui correspond à la période pendant laquelle a lieu le rétrécissement perceptif, et qui dans le cas des enfants porteurs d'implant n'est pas restreinte aux deux premières années de vie, comme chez les enfants normo-entendants, mais se poursuit sur une période plus longue à partir du moment où ces enfants commencent à avoir accès à l'audition. La période critique de l'acquisition du langage, et

[7] « très performants » et « peu performants » (notre traduction)

notamment du langage oral, ne serait pas, d'après Bouton et al. (2012) limitée aux deux premières années de vie, mais plutôt aux deux premières années d'exposition à l'oral.

L'étude de Havy et al. (2013) s'intéresse également aux capacités de perception de traits phonologiques dans une tâche d'apprentissage de mots nouveaux par des enfants porteurs d'implant cochléaire : voisement, lieu et mode d'articulation pour les consonnes, et arrondissement, lieu et hauteur pour les voyelles. Les mots nouveaux sont perçus et correctement appris par les enfants porteurs d'implant mais plus lentement que par les enfants normo-entendants. Havy et al. (2013) mettent également en évidence une influence de la durée d'utilisation de l'implant sur cette capacité de perception et d'apprentissage mais ne trouvent pas d'autres facteurs explicatifs comme par exemple l'âge d'implantation. Ce premier ensemble d'études s'intéresse à la perception d'indices segmentaux de la parole. D'autres études, parmi lesquelles celle de Most & Peled (2007) s'intéressent à la perception d'indices suprasegmentaux par des enfants porteurs d'implant cochléaire ou d'aides auditives traditionnelles : intonation (contour intonatif d'énoncé interrogatif ou déclaratif), structure lexicale (nombre de syllabes par mot), accent de mot (localisation sur la première ou la deuxième syllabe de mots bisyllabiques), accent emphatique (marquage ou non de l'emphase). Il ressort de cette étude que les enfants porteurs d'implant cochléaire perçoivent moins bien les indices suprasegmentaux que les enfants utilisateurs d'aides auditives traditionnelles, ce qui peut être lié pour Most & Peled (2007) à un codage moins performant de la fréquence fondamentale par l'implant que par les prothèses auditives traditionnelles, mais également par la rééducation orthophonique qui se concentre sur les caractéristiques segmentales de la parole dans le cas de cette cohorte d'enfants porteurs d'implant cochléaire.

Nous avons choisi de présenter quelques études parmi les nombreuses qui évaluent les capacités perceptives des enfants porteurs d'implant cochléaire. Ce choix très restreint permet de montrer l'effet de l'utilisation de l'implant qui améliore fortement la perception de la parole chez les enfants qui l'utilisent, mais qui ne permet qu'une perception partielle en comparaison d'enfants normo-entendants. Il nous paraît important de noter en outre que les enfants ont des trajectoires de développement de leurs capacités perceptives similaires à celles d'enfants qui n'ont pas été privés de l'audition mais que ces trajectoires restent plus lentes, car elles se font plus tardivement et reposent sur un accès partiel aux sons de parole.

5.2. Capacités de communication et qualité de vie

Il existe peu d'études qui proposent des critères permettant d'évaluer les conséquences de l'implant sur les capacités de communication de l'enfant porteur d'implant cochléaire. Certains critères sont linguistiques et se basent sur des évaluations de la morphologie et de la syntaxe (Le Normand (2004) par exemple), ou sur le développement du lexique (Le Manner-Idrissi et al. (2009)) et d'indices pragmatiques (dans des interactions avec un parent chez Briec (2012) et Briec et al. (2012) par exemple). Ligny et al. (2006) proposent une évaluation de la production de parole chez l'enfant porteur d'implant cochléaire qui prend en compte des critères prédictifs de la réussite de l'implantation. Plecy (2013) présente les enjeux du choix de l'implant d'un point de vue orthophonique, et Virole (2006) propose une double discussion sur les bénéfices, notamment psychologiques, de l'implant sur la vie quotidienne des enfants, et sur les questions éthiques soulevées par le recours à l'implantation, notamment pour les enfants.

Dans cette section, nous allons proposer un aperçu de ces différentes approches d'évaluation de l'effet de l'implant sur les capacités de communication, et sur la qualité de vie des enfants porteurs d'implant cochléaires.

5.2.1.Conséquences de l'implantation sur les capacités langagières de l'enfant

Dans une étude longitudinale réalisée auprès de 50 enfants porteurs d'implant cochléaire et 181 enfants normo-entendants, tous francophones, Le Normand (2004) met en évidence, à partir de mesures de la longueur moyenne d'énoncés (LME, ou « mean length of uterrances », MLU en anglais), un retard des enfants porteurs d'implant dans l'acquisition de la morphologie et des constructions syntaxiques par rapport aux enfants normo-entendants. En effet, il ressort de cette étude que les enfants porteurs d'implant cochléaire ont des difficultés dans le marquage du genre pour les noms, et du temps et de l'accord pour les verbes. Cette étude de Le Normand (2004) indique en outre une différence entre morphologie lexicale qui est acquise plus facilement, et marquage syntaxique dont l'acquisition est plus tardive chez les enfants implantés.

L'étude de Le Normand (2004) permet d'évaluer l'utilisation de deux caractéristiques de la production de parole (morphologie et syntaxe) par des enfants porteurs d'implants. Cette étude ainsi que l'étude longitudinale de Le Maner-Idrissi et al. (2009) qui s'intéresse à la diversification du lexique[8] à partir de l'implantation et au cours des deux années suivantes, montrent que la durée d'utilisation de l'implant a une influence positive sur le développement des capacités langagières des enfants implantés. Le Maner-Idrissi et al. (2009) montrent également l'importance du choix de scolarisation sur ces capacités langagières : cette étude met en évidence les bénéfices liés au choix d'une intégration des enfants dans le système scolaire classique (et non pas spécialisé), avec des enfants entendants.

Briec (2012) et Briec et al. (2012) étudient l'utilisation d'indices pragmatiques pour rendre compte des capacités de communication des enfants porteurs d'implant cochléaire, en les comparant avec celles d'enfants normo-entendants. A partir d'une situation de jeu entre l'enfant et un de ses parents, l'organisation de la prise de parole (prise de tour de parole ou initiation de l'échange) et les actes de langage (assertifs, directifs et expressifs), Briec (2012) et Briec et al. (2012) établissent que les enfants porteurs d'implant ont une utilisation de ces indices pragmatiques relativement proche de celle des enfants normo-entendants, avec une tendance des enfants porteurs d'implant à devenir au cours du temps de plus en plus actifs dans l'échange, ce qui se manifeste notamment par une augmentation du recours aux énoncés assertifs.

5.2.2.Conséquences de l'implantation sur la qualité de vie de l'enfant et sa famille

Plecy (2013) adopte une perspective d'orthophoniste pour présenter les difficultés de l'entourage et les interrogations liées à la surdité et au recours à la remédiation de cette surdité par l'implant cochléaire. Son expérience auprès d'adolescents ayant reçu tardivement un implant met en lumière les difficultés liées d'une part à la représentation de la différence associée à l'utilisation de l'implant, et d'autre part aux difficultés de communication avec un entourage qui n'est pas forcément formé à l'utilisation d'autres modes de communication que l'oral. Plécy (2013) évoque également les interrogations qu'ont les parents entendants sur la capacité de leur enfant à parler, ce qui peut conduire à un refus des parents de parler à l'enfant ou à un risque de l'enfant de repli

[8] La diversification du lexique est évaluée au moyen d'une mesure du nombre de mots différents produits par les enfants dans une situation d'interaction avec un parent.

sur lui-même. Toutes ces difficultés vont jouer un rôle dans le développement des capacités de communication de l'enfant porteur d'implant cochléaire. Plécy (2013) évoque également l'importance d'une communication multimodale, qui participe à l'évolution de la communication avec l'implant : l'utilisation de signes ou de gestes va favoriser l'acquisition de l'oral. Ligny et al. (2006) évoquent également plusieurs étapes dans l'accès à l'audition au moyen de l'implant cochléaire : l'enfant va tout d'abord avoir une période « d'imprégnation sensorielle », qui correspond à un accès indifférencié à tous les sons environnants, puis il va accéder à un travail « d'identification », où il va apprendre à distinguer les sons entre eux, enfin, l'enfant porteur d'implant va pouvoir accéder à la « compréhension » de la parole. Ligny et al. (2006) notent également plusieurs critères qui vont être déterminants pour le développement des capacités de communication de l'enfant porteur d'implant : précocité de l'intervention, niveau d'audition résiduelle, implication familiale, absence de bilinguisme et lien avec les habiletés cognitives de l'enfant.

Virole (2006) indique également dans quelle mesure l'implant peut permettre d'améliorer les relations de l'enfant avec son entourage dans des situations de la vie quotidienne, puisque celui-ci va être en mesure de réagir lorsqu'on l'appelle ou lorsqu'il entend un son inhabituel par exemple. Virole (2006) insiste notamment sur les conséquences de l'accès de l'enfant à l'audition sur le climat familial de l'implant puisque les parents vont pouvoir prévenir d'un danger oralement et avoir un lien avec l'enfant, même si celui-ci ne se trouve pas face à eux.

5.2.3. Modes de communication des personnes sourdes

Adultes et enfants sourds peuvent être amenés à utiliser un ou plusieurs modes de communication, pour pouvoir interagir avec leur entourage et au sein de la société. Ces modes de communication peuvent être utilisés seuls, simultanément ou consécutivement par les personnes sourdes et leurs entourages.

L'oral est le premier mode de communication, utilisé aussi bien par les personnes sourdes utilisant des dispositifs de remédiation (aides auditives ou implants cochléaires par exemple). L'oral peut s'accompagner de la LPC (ou Langue Parlée Complétée) qui est un système de codage qui vient en appui de l'oral, pour apporter une information visuelle supplémentaire permettant de percevoir les phonèmes de la langue maternelle. La LPC diffère selon les langues et transmet des informations qui combinent la forme de la main pour coder les consonnes (huit configurations possibles) et sa position par rapport au visage pour coder les voyelles qui peut être statique ou dynamique mais utilise également la lecture labiale (Leybaert et al., 2015). La figure 2.8 ci-dessous présente la LPC utilisée en anglais américain.

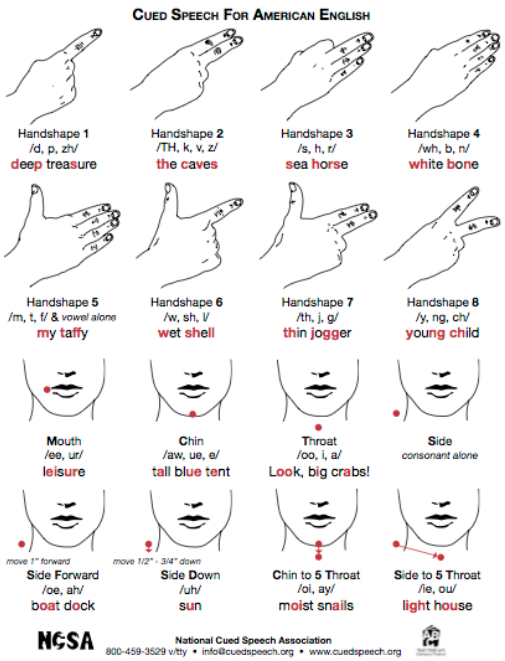


Figure 2.8 : Présentation du code LPC utilisé en anglais américain (Langue Parlée Complétée ou Langage Parlé Complété)

La langue des signes (LSF : Langue des Signes Française) utilise également le signe et diffère selon la langue orale pratiquée au même endroit. Cependant, à la différence de la LPC qui code des sons et syllabes, en accompagnement de l'oral, la LSF est une langue uniquement signée, qui comporte une syntaxe et un lexique propres. La position et le mouvement des deux mains, ainsi que leur emplacement dans l'espace permet de former des mots, et une organisation syntaxique.

Le français signé diffère de la LSF, puisqu'il utilise la syntaxe du français oral, mais les signes correspondant au lexique de la LSF.

Les adultes et enfants sourds, appareillés ou non peuvent donc avoir recours à un ou plusieurs modes de communication, simultanément ou pas, qui peuvent varier selon l'interlocuteur et qui peuvent évoluer au cours de la vie.

5.2.4. Scolarité et intégration sociale de l'enfant porteur d'implant cochléaire

En France, plusieurs dispositifs ont été mis en place, pour accompagner l'enfant sourd avant et pendant sa scolarité. Ces dispositifs sont différents selon l'âge de l'enfant, une description en est proposée dans Mondain et al. (2005) et Lina-Granade et Truy (2005). Les CAMSP (Centre d'Action Médico-Sociale Précoce) et SAFEP (Service d'Accompagnement Familial et d'Education) sont des structures d'accompagnement de l'enfant de moins de six ans. Les SSEFIS (Services de Soutien à l'Education et à l'intégration Scolaire), CLIS (Classes d'Intégration Scolaire), IJS (Institut de Jeunes Sourds) et CROP (Centre de rééducation de l'Ouïe et de la Parole) s'adressent à l'enfant à partir de six ans.

Les SSEFIS proposent un accompagnement et un suivi scolaire, en parallèle de la scolarité l'enfant, alors que les CLIS sont des classes spécifiques en école primaire pour les enfants présentant un handicap (certaines classes, CLIS 2 sont destinées aux enfants présentant une surdité). Les IJS et les CROP accueillent les enfants jusqu'à 20 ans (ils incluent le collège et le lycée).

L'ensemble de ces dispositifs permettent en théorie l'intégration des enfants sourds et leur accès à l'instruction, mais en pratique, de nombreux problèmes subsistent. Un article d'Elsa Maudet, paru le 13 septembre 2016 dans le journal Libération pointe les nombreuses difficultés, voire l'impossibilité de certains enfants de bénéficier d'une instruction équivalente à celle que reçoivent les enfants ne présentant pas de handicap : « sur les 7570 élèves ayant un handicap auditif scolarisés dans les écoles françaises en 2014-2015 (hors établissements spécialisés), moins d'un quart bénéficiait d'un accompagnement humain et autour d'un tiers disposait de matériel adapté (micro sans fil porté par l'enseignant pour les élèves porteurs d'un implant cochléaire, par exemple). »[9] Sont rapportés également des comportements hostiles et de harcèlement, de la part de certaines équipes enseignantes et d'autres élèves, qui témoignent de la difficulté des enfants sourds à suivre leur scolarité, avec pour conséquences un taux d'illettrisme élevé et une insertion professionnelle limitée.

5.2.5. Problèmes éthiques soulevés par l'implant cochléaire

Virole (2006) dresse un aperçu de quelques questions que soulève le recours à un implant cochléaire, du point de vue éthique. En effet, la chirurgie d'implantation n'est pas une chirurgie obligatoire, proposée pour traiter un problème de santé. C'est une chirurgie proposée pour permettre à des enfants d'avoir accès à l'audition, mais il ne s'agit pas d'un traitement médical pour traiter la surdité. Elle ne comporte pas de risques autres que ceux qui sont liés à une chirurgie sous anesthésie générale. Virole (2006) insiste donc sur l'accompagnement nécessaire des enfants et de leur famille lorsqu'une implantation est envisagée, en soulignant les limites de l'implant, qui ne permet pas d'accéder une audition parfaite et instantanée. Dans ce chapitre, Virole (2006) met en garde contre les attentes trop fortes des parents sur la chirurgie, et de l'utilisation de l'implant, qui dans certains cas ne donne pas les résultats escomptés. De même il insiste sur l'accompagnement des familles, qui doivent comprendre les conséquences d'une implantation et l'investissement dans la rééducation orthophonique et le suivi psychologique des enfants. Enfin, une des questions soulevées par l'implant cochléaire est celle de la négation de la surdité, qui est vécue par certaines personnes sourdes comme une identité que l'implant cherche à nier et faire disparaître. Selon Virole (2006), l'implantation est donc à la fois une question médicale et culturelle et ne doit pas être envisagée par les équipes médicales comme la seule solution proposée aux familles d'enfants sourds.

[9] Elsa Maudet, Libération, 13 septembre 2016, http://www.liberation.fr/france/2016/09/13/les-enfants-sourds-n-ont-pas-acces-a-l-instruction_1494486

6. Conclusion

Dans ce chapitre, nous avons présenté les caractéristiques de l'audition dans son fonctionnement normal, pour ensuite expliquer quels sont les différents types et mécanismes de la surdité, et les différentes solutions de remédiation proposées actuellement par les équipes médicales, en nous concentrant sur l'implant cochléaire. Le choix d'un type d'appareillage, notamment celui de l'implant cochléaire, va avoir des répercussions sur la perception des sons de l'environnement et donc sur les capacités langagières et capacités de communication de l'enfant avec son entourage et pose plusieurs questions, notamment psychologiques et éthiques qui doivent être prises en compte lorsque ce choix est envisagé.

Ce chapitre a montré dans quelle mesure l'implant cochléaire permet de restaurer une audition chez des enfants atteints par plusieurs types de surdités, sévères et profondes, liées à une atteinte de l'oreille interne. Cette audition restaurée est cependant incomplète, puisque la qualité du son est dégradée en raison des contraintes technologiques qui ne permettent pas à l'heure actuelle de produire une information auditive synthétique identique à l'information naturelle. Cette qualité du son a des conséquences sur la compréhension de la parole et la capacité des enfants à reproduire des sons qu'ils ne perçoivent pas dans leur intégralité. L'enfant implanté ne devient pas un entendant comme les autres, et n'est plus tout-à-fait un sourd. Dans le chapitre 3, nous allons proposer un aperçu d'études disponibles dans la littérature, qui permettent d'avoir un panorama des difficultés que les enfants porteurs d'implants et d'autres types d'appareillages auditifs rencontrent lorsqu'ils parlent. Ces difficultés sont liées à la production de certains sons de parole et peuvent se traduire par des difficultés à se faire comprendre et à avoir une parole intelligible.

CHAPITRE 3 : production de parole et intelligibilité de la parole des enfants porteurs d'implants cochléaires : état de l'art

Les productions de parole d'enfants porteurs d'implants cochléaires reflètent les difficultés de ces enfants à percevoir certains sons, ce qui a également des conséquences sur leur intelligibilité, avec des répercussions sur leur intégration scolaire et sociale. Dans ce chapitre, nous allons proposer une sélection d'études disponibles dans la littérature et réalisées au cours des dernières décennies. Dans un premier temps, nous présenterons des descriptions de caractéristiques segmentales (voyelles, consonnes et séquences consonnes-voyelles) de la parole d'enfants porteurs d'implants cochléaires, puis dans un deuxième temps, nous proposerons un panorama d'études sur l'intelligibilité de la parole de ces enfants.

1. Caractéristiques segmentales de la parole d'enfants porteurs d'implants cochléaires

1.1. Voyelles

1.1.1. Introduction

Dans la littérature, plusieurs types d'études s'intéressent à la production d'enfants sourds oralistes, porteurs d'implants cochléaires, d'aides auditives ou non. Dans la suite de cette thèse, nous utiliserons les abréviations couramment utilisées dans la littérature en anglais : CI pour « cochlear implanted » (ou porteur d'implant cochléaire), NH pour « normal-hearing » (ou normo-entendant), HA pour « hearing-aid user » (ou utilisateur d'aides auditives). Ces études sont de plusieurs ordres : évaluations perceptives qui proposent des inventaires de voyelles produites par ces populations d'enfants, ou études acoustiques, qui cherchent à définir le plus objectivement possible quels traits des voyelles sont soumis à l'effet d'une perception partielle avec un implant ou une aide auditive, et quels sont les facteurs qui influencent la production des voyelles.

1.1.2. Etudes perceptives / inventaires de voyelles

Plusieurs études utilisant des évaluations perceptives sur la production de voyelles par des enfants sourds porteurs d'aides auditives ou d'implants cochléaires ont permis notamment de décrire l'évolution dans le temps d'inventaires de voyelles chez ces populations d'enfants. Nous appelons « perceptives » les études qui utilisent en majorité des transcriptions ou annotations effectuées à partir de jugements perceptifs. Dans cette section, nous avons choisi de présenter un aperçu de quelques unes de ces études et des résultats obtenus par leurs auteurs.

Dans une étude sur la parole d'enfants sourds oralistes anglophones de six ans non porteurs d'implants cochléaires, Geffner (1980) s'intéresse à l'évaluation des capacités de production de parole de cette population. Il s'agit d'une étude perceptive, où les productions sont évaluées par rapport à une cible à imiter. Cinq critères permettent de caractériser les productions : mode d'articulation, lieu d'articulation, visibilité du phonème, position dans le mot et type d'erreur. En ce qui concerne les voyelles, les meilleurs résultats sont obtenus pour les voyelles postérieures et les voyelles basses ainsi que pour les diphtongues. Geffner (1980) établit une relation entre fréquence fondamentale, formants, intensité, visibilité et production des phonèmes : les voyelles ayant une fréquence fondamentale (F0) basse sont mieux produites que celles qui ont une F0 plus élevée, les voyelles ayant un F2 plus bas sont également mieux produites, les sons les plus intenses (voyelles et diphtongues par exemples) sont mieux produits, et les sons les moins

ambigus visuellement (consonnes labiodentales ou bilabiales) sont produits plus correctement que les sons les plus ambigus.

Une deuxième étude de Tye-Murray et Kirk (1993) propose également une évaluation perceptive des voyelles et diphtongues par des enfants anglophones sourds, cette fois-ci porteurs d'implants cochléaires. Il s'agit d'une étude longitudinale (une session d'enregistrement pré-implantation, puis six, 12, 18, 24 et 36 mois post-implantation), à laquelle participent huit enfants sourds pré-linguaux, implantés entre 3;11 et 7;5 ans. Les résultats indiquent une tendance des voyelles antérieures et postérieures à être produites de façon centralisée avant l'implantation, mais cette tendance diminue au fur et à mesure de l'utilisation de l'implant. Les voyelles centrales sont les mieux produites, et les bons résultats obtenus pour ces voyelles continuent de s'améliorer avec le temps. Les diphtongues sont le plus souvent produites comme des voyelles centrales, mais elles sont également mieux produites lorsque la durée d'utilisation de l'implant augmente.

Dettman et al. (1995) quant à eux s'intéressent à la production des voyelles comme reflet des capacités perceptives des enfants porteurs d'implants. A travers une tâche d'imitation de 20 voyelles et diphtongues en contexte /h_d/, ils mettent en évidence un effet de l'âge chronologique sur la production : les enfants de moins de quatre ans ont des meilleurs résultats que les plus âgés (le nombre de voyelles correctement produites est plus élevé que chez les enfants plus âgés), ce qui correspond à de meilleures capacités perceptives de ce groupe d'enfants, implantés plus jeunes. Dans cette étude, l'âge exact des enfants au moment de l'implantation n'est pas précisé.

Une étude longitudinale de Chin et Pisoni (2000) s'intéresse à l'évolution de la production des segments par un enfant, deux ans après implantation. Ils mettent en évidence que certains sons sont présents dans l'inventaire (comme par exemple les voyelles nasales) alors qu'ils ne font pas partie des sons de la langue maternelle (anglais américain) et que l'inventaire ne possède pas de voyelles postérieures, ce qui est dû à la nature du corpus (des photos de personnes de l'entourage de l'enfant), qui ne contenait pas de voyelles postérieures (sic). Cette étude montre en outre que le système phonologique d'un enfant porteur d'implant a déjà atteint un niveau élevé de complexité, deux ans post-implantation, même si les voyelles produites diffèrent de la cible.

L'étude de Warner-Czyz et Davis (2008) compare les productions segmentales d'enfants CI (âgés de 24.8 (6.8) mois en moyenne, implantés à 16.3 (4.65) mois en moyenne et utilisant l'implant cochléaire depuis 8.5 (2.38) mois en moyenne) et NH (âgés de 13.5 (3.70) mois en moyenne) dans des situations d'interaction parents-enfants, en prenant en compte la fréquence de chaque voyelle dans l'inventaire des enfants et la justesse de la production. Cette étude met en évidence que les voyelles centrales et moyennes sont les voyelles les plus fréquemment produites par les deux groupes, et que les deux groupes produisent plus de voyelles basses que hautes. Les voyelles antérieures sont plus nombreuses que les postérieures chez les enfants CI alors qu'elles sont présentes en quantité comparable chez les enfants NH. Cette étude, qui utilise la transcription pour analyser les erreurs de productions des enfants met en évidence que la dimension d'antériorité est réalisée de façon plus correcte que la hauteur, par les enfants CI.

Cet aperçu d'études perceptives nous permet de dégager quelques facteurs qui jouent un rôle dans la production de voyelles, notamment l'âge lors de l'implantation ou l'âge chronologique. Il apparaît en outre qu'une perception partielle des sons de parole a des répercussions sur la capacité des enfants à reproduire le trait d'antériorité des voyelles.

1.1.3. Etudes acoustiques

Il existe plusieurs types d'études utilisant des analyses acoustiques pour décrire des productions de voyelles chez les enfants sourds porteurs d'implants cochléaires : un premier groupe d'études longitudinales se concentrent sur l'évolution dans le temps de la production des voyelles, un deuxième groupe sur des comparaisons de productions avec ou sans retour auditif (avec implant activé ou désactivé), un troisième groupe d'études proposent des comparaisons transversales de productions d'enfants ayant des capacités perceptives différentes (normo-entendants, porteurs d'implants ou utilisateurs d'aides auditives traditionnelles). Cette section propose un aperçu des caractéristiques de certaines études acoustiques disponibles dans la littérature, dans le but de comprendre quelles sont les caractéristiques acoustiques des voyelles produites par les enfants porteurs d'implants cochléaires, et quelles sont les difficultés rencontrées par ces mêmes enfants.

1.1.3.1. Etudes longitudinales

Les études de Ertmer (2001), Hocevar-Bolthezar et al. (2008), Kunisue et al. (2006) et Liker et al. (2007) s'intéressent à l'évolution dans le temps de caractéristiques acoustiques des voyelles produites par des groupes d'enfants sourds porteurs d'implants cochléaires. Leurs méthodologies diffèrent, que ce soit dans le choix des voyelles étudiées (voyelles cardinales pour Ertmer (2001) et Hocevar-Bolthezar et al. (2008) ou inventaires complets de la langue de l'étude pour Kunisue et al. (2006) et Liker et al. (2007)) ainsi que dans les méthodes d'élicitation des productions de parole (répétition de mots d'après un modèle adulte pour Hocevar-Bolthezar et al. (2008), Kunisue et al. (2006) et Liker et al. (2007) et interactions mère-enfant pour Ertmer (2001)). Ces études longitudinales se concentrent soit sur un nombre restreint d'enfants (un seul enfant CI dans l'étude d'Ertmer (2001), deux enfants CI dans l'étude de Kunisue et al. (2006)) soit sur des cohortes plus importantes (13 enfants CI et 12 adultes CI dans l'étude de Hocevar-Bolthezar et al. (2008) et 18 enfants CI dans l'étude de Liker et al. (2007)). Les études de Hocevar-Bolthezar et al. (2008) et Kunisue et al. (2006) incluent une session pré-implantation, alors que celle de Ertmer et al. (2001) et Liker et al. (2007) ont lieu uniquement post-implantation : l'enfant de l'étude de Ertmer (2001) a été enregistré un mois avant l'implantation (âge à l'implantation : 19 mois), puis tous les mois pendant l'année qui suit l'implantation ; les enfants de l'étude de Hocevar-Bolthezar et al. (2008) avaient un âge moyen à l'implantation de 6;3 ans et ont été enregistrés avant l'implantation, puis six mois après celle-ci ; le premier enfant de l'étude de Kunisue et al. (2006) avait 7;10 ans au moment de l'implantation et le deuxième enfant avait 8;2 ans, ils ont tous les deux été enregistrés un mois avant l'implantation, puis tous les mois pendant l'année qui suit l'implantation ; enfin, l'étude de Liker et al. (2007) a été réalisée auprès d'enfants âgés de 12;1 ans, ayant reçu l'implant à 7;6 ans en moyenne. Ces études longitudinales ne comparent pas ces productions à un groupe contrôle normo-entendant, à l'exception de celle de Liker et al. (2007). Dans ce paragraphe, nous nous intéresserons à la partie longitudinale de cette étude seulement, la comparaison avec le groupe d'enfants normo-entendants étant proposée au paragraphe §1.1.3.3 ci-après.

- **Premier et deuxième formants (F1 et F2)**

Les études de Hocevar-Bolthezar et al. (2008) et Kunisue et al. (2006) indiquent une évolution du F1 consécutive à l'implantation : dans l'étude de Hocevar-Bolthezar et al. (2008) les F1 de /u/ et /i/ diminuent après implantation, et dans l'étude de Kunisue et al. (2006), le F1 de /e/ se rapproche du F1 de /i/ chez le premier enfant de l'étude, et les F1 de /a/, /i/ et /ɯ/, qui étaient

similaires à ceux de sa mère avant implantation, ont tendance à beaucoup varier d'une session d'enregistrement à l'autre, consécutivement à l'implantation, chez le deuxième enfant.

Dans l'étude de Ertmer (2001), qui propose une étude acoustique des productions de voyelles lors de la session d'enregistrement de 10-12 mois post-implantation, les F1 de /ɑ/ se confondent avec celui de /æ/. Les auteurs montrent en outre une distinction nette de hauteur entre /i-u/ et /ɑ-æ/. L'étendue du F1 de /a-æ/ est par ailleurs plus grande que celle de /i-u/.

Comme pour le premier formant, l'étude de Ertmer (2001) met en évidence, à 10 à 12 mois post-implantation, un chevauchement des F2 de /ɑ/ et /æ/. L'étude montre également des étendues de F2 similaires pour /i/, /u/ et /æ/, mais une étendue de F2 plus réduite pour /ɑ/ par rapport aux autres voyelles. L'étude de Liker et al. (2007) montre que le F2 des enfants CI aux trois moments étudiés est significativement différent de celui des enfants NH pour toutes les voyelles, ce qui correspond à des voyelles plus antérieures chez les enfants CI que chez les enfants NH.

- **Espace vocalique**

L'étude de Hocevar-Bolthezar et al. (2008) montre une expansion de la taille de l'espace vocalique suite à l'implantation, ce qui indique une meilleure distinction entre les catégories vocaliques. La figure 3.1 ci-dessous présente cette expansion de la taille de l'espace vocalique pour l'un des enfants de cette étude.

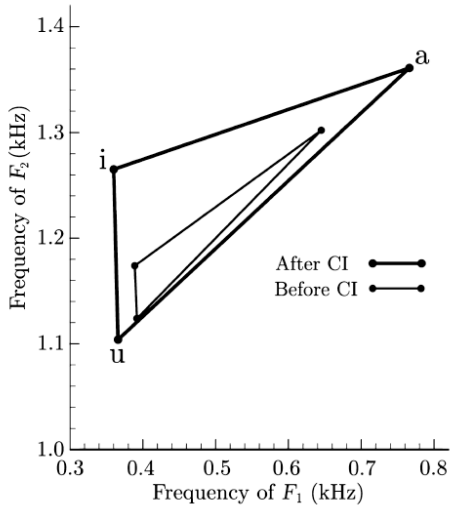


Figure 3.1 : espace vocalique F1-F2 d'un enfant CI dans l'étude de Hocevar-Bolthezar et al. (2008), pré- vs post-implantation

L'étude de Kunisue et al. (2006) indique une évolution de la taille de l'espace vocalique des deux enfants dont les productions sont étudiées, ainsi qu'une évolution de la localisation de la répartition des différentes catégories vocaliques à l'intérieur de cet espace. Les deux enfants ont une évolution différente (pour l'enfant 1, /a/ et /ɯ/ sont stables sur la période étudiée et proches des productions de sa mère alors que /o/ est éloigné du /o/ de la mère, pour l'enfant 2, la distribution des voyelles est relativement similaire à celle de la mère), mais la répartition des voyelles à l'intérieur de l'espace et la taille de l'espace tendent à se rapprocher de plus en plus de celles de leurs mères respectives. La figure 3.2 ci-dessous présente l'évolution de l'espace vocalique de l'enfant 1.

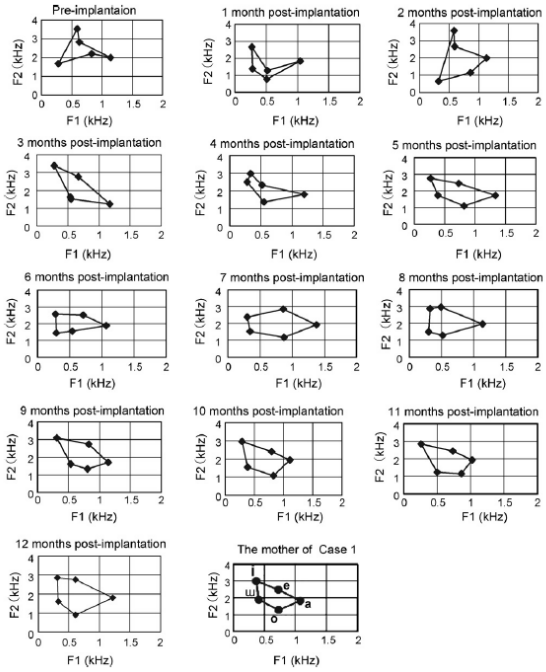


Figure 3.2 : évolution de l'organisation de l'espace vocalique de l'enfant 1 et de celui de sa mère (dernière figure) dans l'étude de Kunisue et al. (2006)

Ces résultats sont cependant en contradiction avec ceux de l'étude de Liker et al. (2007) qui met en évidence un espace vocalique des enfants CI plus grand que celui des enfants NH, mais pas d'évolution dans le temps de la taille de cet espace et de la répartition des catégories vocaliques au sein de cet espace. La figure 3.3 ci-dessous présente les espaces vocaliques moyens des enfants CI aux trois temps de l'étude de Liker et al. (2007), en comparaison avec l'espace vocalique moyen des enfants NH de cette même étude.

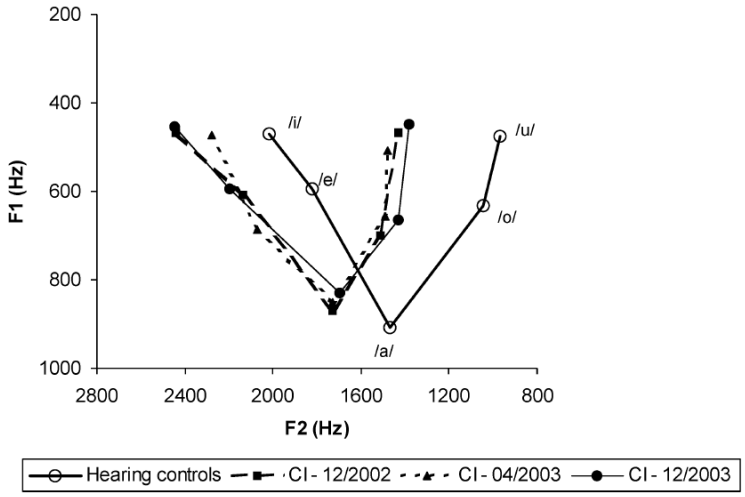


Figure 3.3 : évolution de l'espace vocalique des enfants CI et comparaison avec l'espace vocalique d'enfants NH dans l'étude de Liker et al. (2007)

Les productions de voyelles présentées dans les études de ce paragraphe ont mis en évidence une tendance commune d'instabilité consécutive à l'implantation, que ce soit sur une dimension des voyelles (hauteur ou antériorité), sur la taille des espaces vocaliques, ou sur la répartition des voyelles à l'intérieur de cet espace. Cependant, cette instabilité ne se manifeste pas systématiquement aux mêmes endroits chez tous les enfants de ces études, ce qui pourrait

s'expliquer par la différence des langues maternelles des enfants (japonais, croate, slovène ou anglais), par l'âge auquel ils ont commencé à utiliser leur implant ou éventuellement par des caractéristiques différentes des implants utilisés.

1.1.3.2.Implants ON vs OFF vs enfants NH

L'étude de Bharadwaj et al. (2007) propose une comparaison des caractéristiques acoustiques de voyelles produites dans la L1 (anglais) par quatre enfants CI et de voyelles produites dans une langue inconnue (français) par les mêmes enfants, dans deux conditions différentes (avec l'implant activé ou désactivé). Les auteurs de cette étude utilisent une tâche de répétition de mots /pVt/ et /bVt/ dans des phrases porteuses. Les voyelles étudiées sont les quatre voyelles antérieures de l'anglais (tendues et relâchées) et quatre voyelles du français /y, œ, ɔ̃, a/ qui n'existent pas en anglais. Une deuxième étude, de Poissant et al. (2006) propose une comparaison de la production de voyelles chez six enfants CI avec implant activé ou désactivé. Cette étude ne propose pas directement de comparaison indépendante des valeurs de formants, mais s'intéresse au ratio F1:F2 de cinq voyelles de l'anglais produites lors d'une tâche de production de mots mono- et bisyllabiques en isolation et dans des phrases porteuses, à partir d'images.

- **Implant ON vs OFF**

Bharadwaj et al. (2007) mettent en évidence un effet différent de la condition (ON vs OFF) sur les deux premiers formants : le F1 est différent d'une condition à l'autre pour deux enfants CI sur quatre et le F2 pour tous. En condition OFF, le F1 est plus haut alors que le F2 plus bas qu'en condition ON, ce qui correspond à des voyelles plus basses et plus postérieures lorsque les enfants ne disposent que d'un retour partiel de leur production de voyelles. Enfin, parallèlement à l'étude des premier et deuxième formants, l'étude de Bharadwaj et al. (2007) s'intéresse à des caractéristiques suprasegmentales de la production de parole, et montre un effet de la condition (ON vs OFF) sur la durée des voyelles, qui sont plus longues en condition OFF, ainsi qu'un effet de la condition sur la F0, qui est plus haute en condition OFF.

L'étude de Poissant et al. (2006) montre un effet de la condition (ON vs OFF) sur le ratio F1:F2 chez tous les enfants. Cet effet n'est cependant pas toujours dans la même direction (le ratio F1:F2 est plus élevé en condition OFF pour trois enfants et en condition ON pour les trois autres).

- **Implant ON vs enfants NH**

L'étude de Bharadwaj et al. (2007) montre que les F1 de /ɛ/ et /æ/ de la L1 ont des valeurs plus basses chez les enfants CI en condition ON que chez les enfants NH. Pour la langue inconnue, les enfants CI ont une étendue de F1 plus réduite que les enfants NH.

En condition ON, pour les voyelles /i/ et /ɪ/ de l'anglais (L1), le F2 est plus bas chez les enfants CI que chez les enfants NH, et pour /ɛ/ et /æ/, le F2 est similaire chez les deux groupes. Pour les voyelles du français, le F2 a une étendue comparable chez les deux groupes d'enfants.

1.1.3.3.NH vs CI vs HA

A la différence des études longitudinales et des études avec implants activés ou désactivés, qui permettent de caractériser l'évolution dans le temps de la production des voyelles ou d'évaluer l'importance du retour auditif sur la production de parole, un certain nombre d'études se concentrent sur une comparaison de productions de voyelles par plusieurs populations d'enfants,

normo-entendants, utilisateurs d'aides auditives traditionnelles ou d'implants cochléaires. Dans ces études transversales un seul enregistrement est réalisé, dans le but de mettre en évidence quels sont les facteurs qui affectent la production des voyelles par ces différents groupes d'enfants : âge chronologique, âge d'implantation, durée d'utilisation de l'implant, type d'aides auditives ou d'implants, etc. La table 3.1 ci-dessous présente les différentes cohortes d'enfants enregistrées lors de ces études, et qui seront évoquées dans ce chapitre.

Auteur	Année	Langue	Participants	Âge des participants lors de l'étude (années;mois)	Âge à l'implantation (années;mois)
Baudonck et al.	2011	Néerlandais (Belgique)	40 enfants CI 42 enfants NH 34 enfants HA	8;8 (4;2-15;5) 9;3 (4;3-15;3) 9;16 (4;1-14;2)	2;8 (0;6-10;11) - -
Chuang et al.	2012	Mandarin (Nord de Taïwan)	24 enfants CI 24 enfants NH	7;9 (6-11;1) appariés	3;4 (écart-type 2;3) -
Horga et Liker	2006	Croate	10 enfants CI 10 enfants NH 10 enfants HA	9;7-15;2 9;6-15;1 7;4-15;1	3;11-11;11 - -
Löfqvist et al.	2010	Suédois	12 enfants CI 11 enfants NH	15;4 (12;0-19;1) 8;9 (8;7-9;10)	6;7 (2;7-12;11) -
Neumeyer et al.	2010	Allemand (Sud de l'Allemagne)	5 jeunes CI 5 adultes CI 5 jeunes NH 5 adultes NH	16-27 58-68 15-24 59-68	non précisé non précisé - -
Ryalls et al.	2003	Français (Québec)	10 enfants sourds profonds HA 10 enfants sourds modérés-sévères HA 10 enfants NH	9;4 (7;11-10;9) 8;7 (6;10-9;10) 8;4 (7;6-10;9)	- -
Seifert et al.	2002	Suisse allemand	20 enfants CI valeurs contrôles d'après Huber et al. (2003)	6;5 (3;8-10;2) appariés	4;0 (1;9-5;3) -
Tseng et al.	2011	Mandarin (Taïwan)	15 enfants CI 30 enfants HA	4;1-11;2 3;9-10;5	non précisé -
Uchanski & Geers	2003	Anglais (US)	181 enfants CI 24 enfants NH	8 et 9 ans 8 et 9 ans	non précisé -
Verhoeven et al.	2016	Néerlandais (Belgique)	8 enfants CI 7 enfants HA 90 enfants NH	5-7;3 6;1-7;9 5-7	0;9 (0;5-1;7) - -

Table 3.1 : caractéristiques des cohortes d'enfants, dans la sélection d'études transversales présentées dans ce chapitre

Toutes ces études étudient les productions de voyelles par des enfants porteurs d'implants cochléaires, en les comparant à un groupe contrôle d'enfants normo-entendants (Chuang et al. (2012), Liker et al. (2007), Löfqvist et al. (2007), Neumeyer et al. (2010), Uchanski & Geers (2003) et Verhoeven et al. (2016)) ou porteurs d'aides auditives traditionnelles (Tseng et al. (2011)). Deux études comparent deux groupes d'enfants sourds, porteurs d'aides auditives traditionnelles et d'implants cochléaires à un groupe contrôle d'enfants normo-entendants (Baudonck et al. (2011), et Verhoeven et al. (2016)), une étude compare les productions d'un groupe d'enfants CI à des valeurs de références de la littérature (Seifert et al. (2002)). Nous avons également choisi d'inclure l'étude de Ryalls et al. (2003), même si les groupes dont elle étudie les production

n'incluent pas d'enfants porteurs d'implants cochléaires, mais deux groupes d'enfants sourds porteurs d'aides auditives traditionnelles comparés à un groupe d'enfants normo-entendants. En effet, cette étude est, à notre connaissance, l'une des seules études acoustiques à s'intéresser à des productions de voyelles en français par des enfants sourds (les enfants de l'étude sont tous des enfants québécois francophones).

- **Méthodologies**

Comme pour les études longitudinales, les méthodologies sont variables d'une étude à l'autre, de même que les cohortes d'enfants enregistrées et les analyses acoustiques effectuées. La Table 4.2 ci-dessous propose un aperçu des méthodologies et types d'analyses réalisées dans ces études. Les méthodes d'élicitation les plus courantes sont la production de mots à partir d'images, sans modèle audio (Baudonck et al. (2011), Löfqvist et al. (2010), Seifert et al. (2002)), la répétition de mots ou pseudo-mots (Horga & Liker (2006), Verhoeven et al. (2016)) ou de phrases entières (Chuang et al. (2012), Horga & Liker (2006), Tseng et al. (2011), Uchanski & Geers (2003)). Quelques études utilisent la lecture de phrases (Neumeyer et al. (2010) et Uchanski & Geers (2003)) ou de syllabes (Ryalls et al. (2003)). L'étude de Chuang et al. (2012) enregistre également la production de voyelles tenues, en isolation, pour les comparer aux voyelles enregistrées dans des phrases porteuses répétées.

L'inventaire des voyelles étudiées et les mesures et analyses effectuées sont également variables d'une étude à l'autre et sont présentées dans la table 3.2 ci-dessous : certaines études se basent sur un inventaire complet des voyelles de la langue de l'étude, alors que d'autres se focalisent sur quelques voyelles (par exemple, sur les voyelles cardinales /i, a, u/, sur deux voyelles /i/ et /ɔ/ ou encore sur une seule voyelle /a/).

Dans la plupart de ces études, les deux premiers formants (F1 et F2) sont mesurés, puis analysés soit séparément (Baudonck et al. (2011), Chuang et al. (2012), Horga & Liker (2006), Ryalls et al. (2003), Verhoeven et al. (2016)), soit à travers un ratio F1:F2 (Seifert et al. (2002)), ou au moyen de distances euclidiennes entre les voyelles en deux dimensions F1-F2 ou de tailles d'espaces vocaliques F1-F2 (Baudonck et al. (2011), Chuang et al. (2012), Löfqvist et al. (2010), Neumeyer et al. (2010), Tseng et al. (2011) et Verhoeven et al. (2016)). L'étude de Ryalls et al. (2003) est la seule à proposer une étude des trois premiers formants et l'étude de Uchanski & Geers (2003) ne s'intéresse qu'au deuxième formant.

Auteur	Année	Voyelles étudiées	Formants mesurés	Analyses	Méthode d'élicitation
Baudonck et al.	2011	voyelles cardinales : /i/, /a/, /u/	F1, F2	Distances euclidiennes entre /i/-/a/, /i/-/u/, /a/-/u/, taille espace vocalique	Elicitation de mots, à partir d'images
Chuang et al.	2012	voyelles cardinales : /i/, /a/, /u/	F1, F2	Etendues de F1 de /a/-/i/ et /a/-/u/ Etendue de F2 de /i/ à /u/ Espace vocalique entre /i/, /a/ et /u/	Voyelles tenues, et répétition de phrases
Horga et Liker	2006	voyelles cardinales : /i/, /a/, /u/	F1, F2	Comparaison de F1 et F2 pour chaque voyelle	Répétition de mots et de phrases
Löfqvist et al.	2010	/i/, /e/, /æ/, /y/, /ø/, /u/, /o/, /a/, /ʉ/	F1, F2	Taille de l'esp. vocalique calculé avec : 1) distances euclidiennes des voyelles au centre de l'espace	Elicitation de mots /rVta/, avec accent de mot sur la 1ere voyelle

				F1-F2 2) distances entre les paires de voyelles	
Neumeyer et al.	2010	/iː/, /eː/, /aː/, /oː/, /uː/	F1, F2	Distances euclidiennes de chaque formant (F1 et F2) des voyelles au centre de l'espace F1-F2	Lecture de phrases. Voyelles-cibles dans un mot avec deux conditions de focus (broad vs narrow)
Ryalls et al.	2003	voyelles cardinales : /i/, /a/, /u/	F1, F2, F3	Etude indépendante de chaque formant	Lecture de syllabes (non-mots)
Seifert et al.	2002	/a/	F1, F2, F3	Ratios F1:F2	Elicitation de mots à partir d'images
Tseng et al.	2011	voyelles cardinales : /i/, /a/, /u/	F1, F2	Espace maximal : (F2(i)-F2(u))*(F1(u)-F1(a))	Répétition de phrases avec des images
Uchanski & Geers	2003	/i/, /ɔ/	F2	Différence : F2moy(i)-F2moy(ɔ)	Répétition/lecture de phrases d'après modèle audio adulte
Verhoeven et al.	2016	/i/, /ɪ/, /ɛ/, /u/, /ʏ/, /ɔ/, /ɑ/, /eː/, /yː/, /øː/, /o/, /a/	F1, F2	1) Distances entre /i/-/a/, /i/-/u/, /a/-/u/ 2) Surface délimitée par toutes les occurrences des 12 voyelles	Répétition de monosyllabes /pVt/ et /lVt/

Table 3.2 : caractéristiques des productions enregistrées, dans la sélection d'études transversales présentées dans ce chapitre : voyelles, mesures effectuées, analyses réalisées et méthodologie d'élicitation des voyelles

- **Premier formant (F1)**

Les résultats de l'étude de Seifert et al. (2002) sur la seule voyelle /a/ ne mettent pas en évidence de différences marquées entre le groupe CI et les valeurs de référence de la littérature (Huber et al. (1999)) qui servent de contrôle dans cette étude : 80% des enfants CI produisent un F1 de /a/ à moins d'un écart-type des valeurs moyennes de F1 d'enfants du même âge. Les résultats de cette étude de Seifert et al. (2002) indiquent que ni l'âge chronologique, ni l'âge d'implantation et ni la durée d'utilisation de l'implant n'ont d'effet sur la réalisation du F1. De même, l'étude de Ryalls et al. (2003) chez des enfants québécois francophones montre qu'il n'y a pas de différence significative de réalisation du F1 par les trois groupes d'enfants (enfants sourds profonds, enfants sourds modérés à sévères et enfants NH).

Les résultats sur le F1 des études de Baudonck et al. (2011), Neumeyer et al. (2010) et Verhoeven et al. (2016) diffèrent d'une voyelle à l'autre : dans l'étude de Baudonck et al. (2011), il n'y a pas de différence significative entre les trois groupes CI, HA et NH pour les F1 de /a/ et /i/ mais le F1 de /u/ est significativement plus bas chez les enfants HA par rapport aux enfants NH, et présente une grande variabilité inter-sujets chez les enfants CI. Dans l'étude de Verhoeven et al. (2016), réalisée selon la même méthode que celle de Baudonck et al. (2011) mais en incluant un plus grand nombre de voyelles, les F1 des voyelles /ɛ/, /ɑ/ et /a/ sont plus bas chez les enfants CI que chez les enfants NH et HA, alors que le F1 de /u/ est plus haut chez les enfants CI et HA, et celui de /ɔ/ est plus haut chez les enfants HA. Ces résultats indiquent que les enfants CI produisent des voyelles plus centralisées sur l'axe du F1 que les enfants HA et NH.

Dans l'étude de Neumeyer et al. (2010) réalisée auprès d'adultes et d'adolescents, la distance euclidienne entre le F1 des voyelles et le F1 du centre de l'espace vocalique est plus petite chez les locuteurs CI par rapport aux locuteurs NH pour les trois voyelles cardinales /aː/, /iː/ et /uː/ mais il n'y a pas de différence entre les deux groupes pour /eː/ et /oː/, ce qui indique que les locuteurs CI produisent des voyelles plus centralisées que les locuteurs NH, qui est un résultat proche de ceux de l'étude de Verhoeven et al. (2016).

Enfin, les résultats de l'étude de Chuang et al. (2012) qui compare les F1 des voyelles produites par deux groupes d'enfants NH et CI dans deux conditions, montrent qu'il y a une étendue de F1 entre /a-i/ et /a-u/ plus grande chez les enfants CI et NH pour les voyelles tenues que pour les mêmes voyelles produites dans des phrases porteuses.

Les résultats sur le F1 dans ces études sont très disparates et parfois contradictoires, ce qui peut s'expliquer par des méthodologies de collecte, de traitement et d'analyse des données différentes : les cohortes d'enfants dont les productions sont étudiées sont variables d'une étude à l'autre, le nombre de voyelles étudiées et le mode d'élicitation de la parole (lecture, répétition de mots, parole spontanée, etc.) peut en partie expliquer ces différences de résultats, de même que le nombre de voyelles étudiées et la langue de l'étude.

- **Deuxième formant (F2)**

Comme pour le premier formant, les résultats sur le deuxième formant des études présentées dans ce chapitre sont assez disparates, ce qui s'explique peut-être par le nombre de voyelles qu'elles étudient : l'étude de Seifert et al. (2002) ne se concentre que sur une seule voyelle /a/, celle de Uchanski et Geers (2003), sur deux voyelles /i/ et /ɔ/, les études de Baudonck et al. (2010), Chuang et al. (2012), Ryalls et al. (2003) et Tseng et al. (2011) sur les trois voyelles cardinales /i/, /a/, /u/, enfin, deux études proposent une analyse à partir des voyelles de systèmes vocaliques plus ou moins complets (5 voyelles longues de l'allemand pour Neumeyer et al. (2010), 12 voyelles du néerlandais pour Verhoeven et al. (2016)).

L'étude de Seifert et al. (2002) met en évidence une dispersion plus grande pour le F2 que pour le F1. En effet, alors que les valeurs de F1 sont à moins d'un écart-type des valeurs du F1 moyen de référence pour 80% des enfants, les valeurs de F2 oscillent entre -3,0 et +2,9 écart-types par rapport aux valeurs du F2 moyen de la littérature. Comme pour le F1, il n'y a pas d'effet de l'âge chronologique, de la durée d'utilisation de l'implant ni de l'âge d'implantation sur la réalisation du F2 par les enfants CI de cette étude, qui ne se concentre que sur des productions de /a/. Cette absence d'effets des facteurs étudiés est peut-être liée à la méthodologie de l'étude et au choix de la voyelle /a/ : en effet, cette voyelle n'étant qu'un élément du système vocalique de l'allemand, Seifert et al. (2002) auraient peut-être obtenu des résultats différents sur le F2 d'autres voyelles du système. L'étude de Uchanski et Geers (2003) pose la même question : les résultats sur les deux seules voyelles /i/ et /ɔ/ indiquent une plus grande dispersion du F2 pour les deux voyelles étudiées chez les enfants HA par rapport aux enfants NH, et une dispersion du F2 similaire chez les enfants CI et NH mais une étude de l'ensemble du système vocalique de l'anglais américain chez ces trois populations d'enfants permettrait peut-être de constater des différences ponctuelles sur certaines autres voyelles du système.

Les études de Baudonck et al. (2010), Chuang et al. (2012), Ryalls et al. (2003) et Tseng et al. (2011) sur le F2 des trois voyelles cardinales /i/, /a/ et /u/ ont des résultats également disparates. Pour Tseng et al. (2011), les voyelles étudiées ont des réalisations de F2 variables : la voyelle /a/ est réalisée à l'extérieur de l'espace vocalique maximal pour 21 enfants sur 45 (de façon plus antérieure que /i/ ou plus postérieure que /u/), mais il y a peu de différence entre les F2 de /i/ et /u/ chez huit enfants, et quatre enfants produisent ces deux voyelles en inversant le F2 de celles-ci.

Ryalls et al. (2003) mettent en évidence un F2 plus bas chez les enfants sourds profonds que chez les enfants sourds modérés et sévères et les enfants NH, ce qui indique une tendance des enfants sourds profonds à produire des voyelles plus postérieures que les enfants des deux autres groupes.

Les résultats de l'étude de Baudonck et al. (2011) indiquent un F2 de/a/ et de /u/ plus bas chez les enfants HA que chez les enfants NH, lui-même plus bas que chez les enfants CI, alors que le F2 de /i/ est plus bas chez les enfants HA que chez les enfants CI et NH, de hauteur équivalente : les trois voyelles sont produites de façon plus postérieures chez les enfants HA par rapport aux enfants NH et CI, la seule différence entre les groupes NH et CI se situant au niveau de la voyelle /a/, plus antérieure chez les enfants CI que chez les enfants NH.

L'étude de Chuang et al. (2012), qui compare des productions de voyelles tenues et de voyelles dans des phrases porteuses, s'intéresse à l'étendue de F2 entre /a/ et /u/, plus grande chez les enfants CI et NH pour les voyelles tenues que pour les voyelles dans des phrases porteuses, et à l'étendue de F2 entre /i/ et /u/, plus grande pour les voyelles tenues chez les enfants NH uniquement. Cette étude montre en outre une tendance à la centralisation des voyelles tenues (plus faible étendue de F2) pour les enfants CI, par rapport aux enfants NH.

L'étude de Verhoeven et al. (2016) utilise une méthodologie similaire à celle de Baudonck et al. (2011) avec un inventaire de voyelles plus étendu : 12 voyelles du néerlandais chez Verhoeven (2016) contre 3 voyelles cardinales /i, a, u/ dans l'étude de Baudonck et al. (2011). Verhoeven et al. (2016) mettent en évidence des différences de F2 entre les groupes NH et CI pour 9 voyelles sur 12, des différences entre les groupes NH et HA pour 7 voyelles sur 12, et des différences entre les groupes CI et HA pour 3 voyelles sur 12 : les enfants CI et HA produisent des voyelles plus centralisées sur F2 que les enfants NH. Ce résultat est assez proche de celui de Neumeyer et al. (2010), puisque la distance euclidienne entre le F2 de chaque voyelle et le F2 moyen est plus petite chez les locuteurs CI par rapport aux enfants NH sauf pour le /a:/, et la distance moyenne entre le F2 des voyelles antérieures et postérieures est plus petite chez les enfants CI que chez les locuteurs NH, ce qui correspond à une tendance des locuteurs CI à produire des voyelles plus centralisées que les locuteurs NH.

- **Espace vocalique**

Parallèlement aux comparaisons de la réalisation de F1 et F2 entre différents groupes (CI, HA, NH), certaines études proposent une analyse des relations entre catégories vocaliques, à travers des comparaisons de tailles d'espaces vocaliques par ces mêmes groupes. Plusieurs méthodes de calcul sont utilisées (notamment la surface de l'espace maximal F1-F2 ou la distance moyenne de chaque voyelle par rapport au centre de l'espace).

Dans l'étude de Ryalls et al. (2003), la taille de l'espace vocalique est liée au niveau d'audition avec prothèses auditives traditionnelles : plus le niveau d'audition est bas, plus l'espace vocalique (Figure 3.4 ci-dessous) est petit. Ce résultat est proche de ceux de l'étude de Löfqvist et al. (2010), dans laquelle l'espace vocalique des enfants CI est plus petit que celui des enfants NH, et de ceux de Chuang et al. (2012), dans laquelle les enfants CI ont un espace vocalique plus petit que les enfants NH lorsqu'ils produisent des voyelles tenues en isolation (alors que dans la même étude, la taille de l'espace vocalique, comportant les mêmes voyelles produites dans des phrases porteuses, est similaire chez les deux groupes CI et NH).

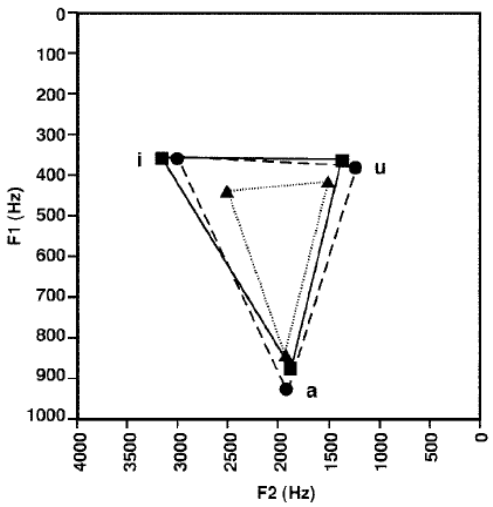


Figure 3.4 : comparaison de l'espace vocalique des enfants sourds profonds HA (triangles et ligne en pointillés), sourds modérés et sévères HA (cercles et ligne en pointillés) et NH (carrés et ligne pleine) dans l'étude de Ryalls et al. (2003)

L'étude de Verhoeven et al. (2016) propose deux méthodes de calcul de la taille de l'espace vocalique : avec un calcul de la surface (en z-scores) délimitée par les douze voyelles (Figure 3.5 ci-dessous), le groupe des enfants CI a l'espace vocalique le plus petit, le groupe des enfants HA a un espace vocalique de taille intermédiaire et le groupe des enfants NH a l'espace vocalique le plus grand. Ces résultats sont en partie similaires avec un calcul de l'espace vocalique (en Hz^2) à partir des trois voyelles cardinales /i/, /a/, /u/ : avec cette méthode, les enfants CI et HA ont des tailles d'espaces vocaliques similaires mais plus petits que les enfants NH.

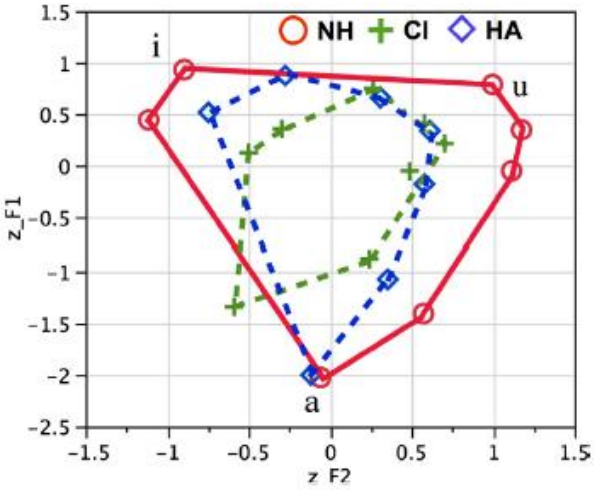


Figure 3.5 : comparaison de l'espace vocalique des enfants CI (signe +), HA (losanges) et NH (cercles) dans l'étude de Verhoeven et al. (2016)

Les résultats de Verhoeven et al. (2016) sont en contradiction avec ceux de Baudonck et al. (2010) : les enfants CI et HA ont des espaces vocaliques (en Hz^2) plus grands que les enfants NH (voir Figure 3.6 ci-dessous). Les résultats contradictoires de ces deux études pourraient s'expliquer par des calculs effectués en Hz, avec des productions d'enfants plus jeunes dans l'étude de Verhoeven et al. (2016), et posent la question du choix d'une méthode de normalisation des données acoustiques qui permet de comparer des productions d'enfants d'âges différents (voir chapitre 4 ci-dessous).

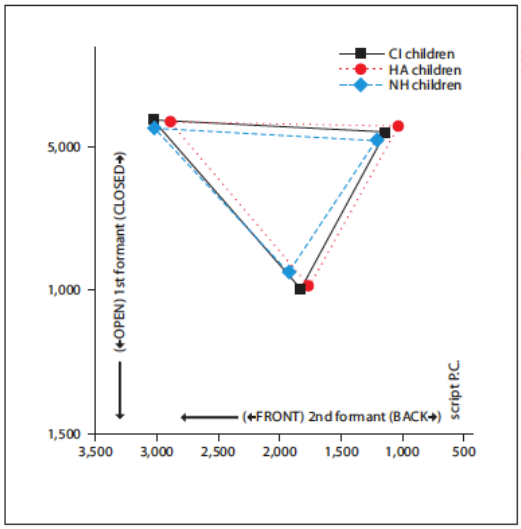


Figure 3.6 : comparaison de l'espace vocalique des enfants CI, HA et NH dans l'étude de Baudonck et al. (2010)

Horga et Liker (2006) trouvent un espace vocalique F1-F2 plus petit chez les enfants HA que chez les enfants CI et NH (Figure 3.7 ci-dessous). L'étude ne précise pas si la différence de taille d'espaces vocaliques entre enfants CI et NH est significative. Ces résultats sont à la fois en contradiction avec ceux de Baudonck et al. (2011) et Verhoeven et al. (2016) : dans les trois cas, les auteurs constatent des différences entre groupes mais ces différences ne vont pas dans le même sens et ne permettent pas de dégager avec certitude des tendances dans la répartition des voyelles à l'intérieur de l'espace vocalique.

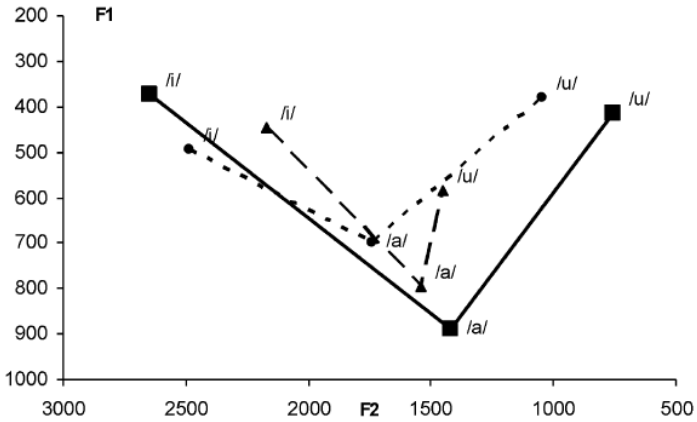


Figure 3.7 : comparaison de l'espace vocalique des enfants CI (cercles et ligne en pointillés), HA (triangles et ligne en pointillés) et NH (carrés et ligne pleine) dans l'étude de Horga et Liker (2006)

Comme pour les analyses isolées des premier et deuxième formants, les analyses des tailles d'espace vocalique et de la répartition des voyelles à l'intérieur de ces espaces permettent de dégager des informations intéressantes : toutes ces études indiquent des différences entre les populations d'enfants étudiées (NH, CI, HA) mais ces différences ne suivent pas les mêmes schémas d'une étude à l'autre. Etant donné que ces études ont été réalisées au cours d'une décennie, les différences observées pourraient être dues à des différences technologiques des implants et prothèses utilisés par les enfants ainsi qu'à des stratégies de codage des appareils ayant évolué au cours du temps, ce qui peut modifier la perception et donc la production de parole chez ces populations d'enfants. Par ailleurs, il est important de noter que ces études prennent en compte des populations d'enfants d'âges hétérogènes sans utilisation systématique d'une méthode

de normalisation des valeurs de formants manipulés. Cette question de la normalisation sera discutée plus en détail dans le chapitre 4 ci-après.

- **Autres résultats**

Parallèlement aux analyses acoustiques des premier et deuxième formants, ainsi qu'à celles de la taille des espaces vocaliques, certaines études proposent des analyses acoustiques supplémentaires : distances entre catégories vocaliques dans l'étude de Baudonck et al. (2010), intersections des catégories vocaliques dans l'étude de Verhoeven et al. (2016) et caractéristiques suprasegmentales dans l'étude de Chuang et al. (2012) et Neumeyer et al. (2010).

On note ainsi que pour Baudonck et al. (2010) les distances entre les centres des catégories vocaliques correspondant aux distances moyennes entre /i/ et /u/, /a/ et /i/, /u/ et /a/, sont plus grandes chez les enfants HA et CI que chez les enfants NH, et que pour Verhoeven et al. (2016), les distinctions entre catégories vocaliques, correspondant au nombre d'intersections entre catégories (analyse effectuée à partir d'ellipses à 95% de confiance) sont plus nettes pour les enfants NH que pour les enfants HA, elles mêmes plus nettes que pour les enfants CI.

Pour les caractéristiques suprasegmentales étudiées, Neumeyer et al. (2010) observent des durées de /e:/ et /i:/ plus grandes chez les locuteurs CI que chez les locuteurs NH, mais pas de différence significative pour les trois autres voyelles de l'étude, et Chuang et al. (2012) notent des débits articulatoires similaires mais des durées de phrases, durées de pauses et proportions de pauses significativement plus élevées chez les enfants CI. Ces études sont en accord avec les résultats de l'étude de Bharadwaj et al. (2007). Il est intéressant de noter que, malgré des différences entre les populations étudiées, ces trois études donnent des informations comparables sur la réalisation de la durée lorsque la perception des sons de parole est partielle : les enfants CI ont des débits d'articulation plus lents lorsque leur implant est désactivé chez Bharadwaj et al. (2007), et des débits de parole et d'articulations plus lents chez les locuteurs CI que chez les locuteurs NH chez Neumeyer et al. (2010) et Chuang et al. (2012).

1.1.4. Conclusion

Cette revue de littérature a tout d'abord permis de mettre en évidence qu'il existe une complémentarité entre études perceptives et acoustiques. En effet, les études perceptives s'appuient sur des jugements et évaluations subjectifs de la production des voyelles, qui permettent de fixer des grandes étapes du développement langagier des enfants sourds, porteurs ou non d'implants cochléaires ou d'aides auditives. On peut considérer que ces jugements impressionnistes sont ensuite complétés par des études acoustiques qui permettent de les valider. Cependant il apparaît également à travers cette revue bibliographique que les résultats obtenus dans le cadre de ces deux types d'études sont très variables et que cette variabilité est liée de près à la méthodologie de collecte et d'analyses des données. Pour pouvoir proposer une description des productions de voyelles par des enfants sourds porteurs d'implants cochléaires, nous devrons donc adopter une méthodologie clairement définie, qui lie clairement les résultats à ce choix méthodologique. Enfin, il n'existe à notre connaissance aucune description acoustique de la production de voyelles par des enfants CI francophones, les deux seules étant l'étude pilote de Scarbel et al. (2012) sur laquelle nous basons notre propre étude et la thèse de Legendre (2014) qui propose une correspondance entre caractéristiques acoustiques des segments (consonnes et voyelles) produits par des enfants CI de 7 à 14 ans, et identification de ces mêmes segments par des locuteurs naïfs. Notre travail de thèse propose donc de pallier ce manque, en proposant une

description des voyelles orales du français, qui est assez complexe par rapport à la plupart des systèmes des langues du monde (voir chapitre 4, ci-après, pour une description du système vocalique du français).

1.2. Consonnes

1.2.1. Introduction

Comme pour les voyelles, l'étude des productions de consonnes par des enfants porteurs d'implants cochléaires se fait par deux biais : jugements perceptifs, et analyses acoustiques. Le premier type d'études se concentre sur des types d'erreurs et des inventaires de consonnes produites par ces populations d'enfants, alors que le deuxième type d'études cherche à déterminer le plus objectivement possible où se situent les difficultés rencontrées par les enfants sourds, porteurs d'implants cochléaires ou d'autres aides auditives, que ne rencontrent pas des enfants normo-entendants du même âge.

1.2.2. Etudes perceptives : inventaires de consonnes produites par les enfants porteurs d'implants cochléaires

Une première étude de Geffner (1980), réalisée auprès de 65 enfants sourds de 6 ans (dans une tâche de dénomination de mots à partir d'images) montre, à partir d'un codage de chaque segment par traits (mode d'articulation, lieu d'articulation, voisement) que les latérales (mode), les consonnes voisées (voisement) et labiales (lieux) sont les mieux produites par les enfants sourds (voir Figure 3.8, ci-dessous). Des résultats plus précis indiquent que [p], [b] et [f] sont les consonnes dont les productions sont les plus correctes. Les consonnes [g], [k] et [ŋ] sont les plus fréquemment omises, alors que [p], [b] et [f] sont les moins fréquemment omises. Cette étude met également en évidence un lien entre la visibilité de la consonne et sa présence dans l'inventaire des enfants sourds.

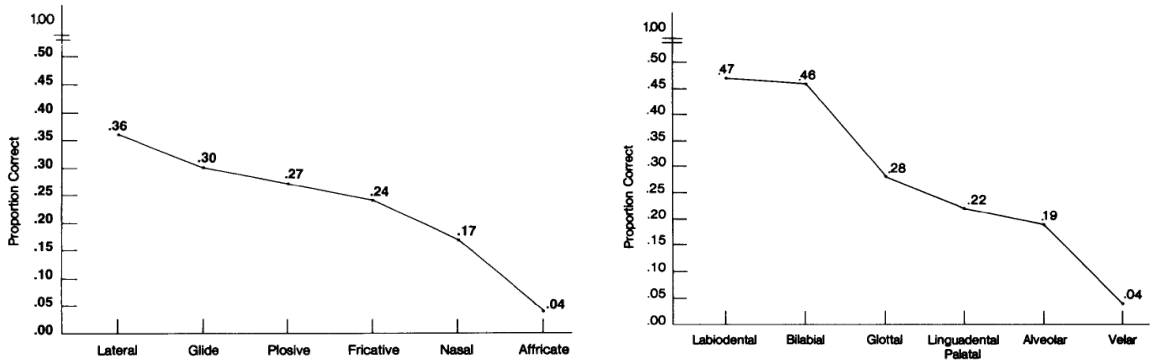
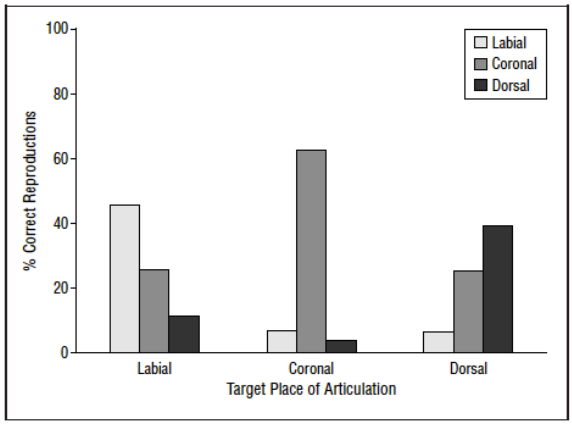


Figure 3.8 : résultats (proportion de consonnes correctement produites) de l'étude de Geffner (1980), pour les modes (gauche) et lieux d'articulation (droite).

L'étude de Dillon et al. (2004) s'intéresse également aux lieux d'articulation des consonnes produites par 88 enfants sourds de huit à dix ans, mais cette fois-ci, porteurs d'implants cochléaires. Les résultats qu'ils obtiennent dans le cadre de cette étude, qui utilise également des jugements perceptifs, divergent de ceux de Geffner (1980), puisqu'ils indiquent une meilleure capacité de cette population d'enfants porteurs d'implants cochléaires à imiter les consonnes coronales que les labiales ou les dorsales. Pour Geffner (1980), les meilleurs résultats des enfants sourds pour les consonnes labiales s'expliquent par une meilleure visibilité des consonnes, ce qui ne semble pas se confirmer dans l'étude de Dillon et al. (2004). La figure 3.9 ci-dessous tirée de Dillon et al. (2004) présente les résultats de cette étude.



Children's place scores as a function of target place of articulation.

Figure 3.9 : résultats (proportion de consonnes correctement produites) de l'étude de Dillon et al. (2004) par lieu d'articulation cible.

L'étude de Warner-Czyz & Davis (2008) propose une description de l'évolution de l'inventaire des consonnes produites et des erreurs de productions de ces consonnes par quatre enfants CI et quatre enfants NH, tous anglophones, à partir de jugements perceptifs. Les enfants CI ont tous reçu l'implant cochléaire avant l'âge de 24 mois, et cette étude longitudinale de Warner-Czyz & Davis (2008) se base sur des productions spontanées enregistrées tous les mois à partir des premiers mots produits et pendant les six mois suivants. Les enfants NH sont appariés en « âge de développement vocal » avec les enfants CI. Dans cette étude, les consonnes sont codées par lieu (labiales, coronales, dorsales) et mode d'articulation (occlusives orales, nasales, fricatives). Les résultats indiquent des différences au niveau des proportions de consonnes produites par lieux et modes d'articulation entre les deux groupes : les enfants CI produisent en majorité des labiales, suivies des coronales, elles-mêmes produites en plus grand nombre que les dorsales, alors que les enfants NH produisent autant de labiales que de coronales, et beaucoup moins de dorsales. La proportion de chaque mode d'articulation des consonnes produites par les enfants NH est stable sur la période de six mois (50% d'occlusives, 30% de nasales et 10% de fricatives), alors qu'elle évolue chez les enfants CI, en particulier pour les nasales qui passent de 45% des consonnes produites à la première session d'enregistrement à 28%, six mois après.

Warner-Czyz & Davis (2008) s'intéressent également aux proportions d'erreurs par lieux et modes d'articulation, calculées à partir d'une classification des consonnes qui sont jugées : 1) correctes, 2) partiellement correctes, 3) incorrectes, 4) omises. Le pourcentage des réalisations incorrectes est comparable chez les deux groupes d'enfants, avec une amélioration de la production au cours des six mois de l'étude. Warner-Czyz & Davis (2008) mettent ainsi en évidence une trajectoire d'acquisition des consonnes similaire chez les enfants CI par rapport aux enfants NH, malgré certaines différences constatées dans les premiers temps d'utilisation de l'implant, mais qui tendent à se combler lorsque la durée d'utilisation de l'implant augmente.

Aucune analyse des types d'erreur n'est proposée dans cette étude, à la différence de l'étude de Kim & Chin (2008) qui s'intéressent aux caractéristiques des erreurs de production des consonnes obstruantes par des enfants CI dans le cadre de la théorie de l'optimalité (OT) pour proposer un modèle de développement phonologique des enfants CI. Dans cette étude, dix enfants CI de 9;8 à 16;1 ans, ayant reçu un implant entre 1;4 et 6;1 ans, sont enregistrés lors d'une tâche de dénomination, à partir d'images, de mots mono- et bisyllabiques contenant toutes les consonnes de l'anglais sauf /ʒ/ dans trois positions (initiales, médianes et finales de mots).

L'évaluation des productions se fait à partir de transcriptions phonétiques. Kim & Chin (2008) s'intéressent aux phénomènes de lenition et fortition des consonnes du corpus, et leurs résultats indiquent une tendance des enfants à produire des consonnes dévoisées (par exemple, des fricatives voisées sont réalisées sans voisement perçu) et à remplacer des fricatives par des plosives, ce qui correspond à deux phénomènes de fortition. Ils notent également des phénomènes de lenition. L'erreur de fortition ou lenition est liée selon Kim & Chin (2008) à la position dans le mot (par exemple, il y a plus de phénomènes de fortition que de lenition en position initiale de mot mais plus de lenition en position médiane). Globalement, les enfants CI font plus d'erreurs de lenition que de fortition, ce que les auteurs expliquent par une minimisation de l'effort articulatoire dans certains contextes phonologiques, et qui correspond, pour Kim & Chin (2008), chez les enfants CI à un pattern d'acquisition similaire à celui d'enfants NH de la littérature (Jakobson, 1941).

L'étude de Baudonck et al. (2010) propose quant à elle une analyse phonétique et phonologique des erreurs de production d'enfants CI, en les comparant à celle d'enfants porteurs d'aides auditives traditionnelles (HA), dans une tâche de dénomination de mots comprenant tous les sons et clusters du néerlandais, à partir d'images. L'analyse phonétique consiste à transcrire phonétiquement les segments produits puis à les comparer avec la cible, pour obtenir une typologie des erreurs (substitution, distorsions, omissions). L'analyse phonologique consiste à classer les types d'erreurs de façon qualitative (dévoisement, réduction de clusters, suppression de consonnes finales, etc.). Les résultats de cette étude indiquent que les enfants HA font plus d'erreurs que les enfants CI et que les erreurs les plus fréquentes chez les deux groupes d'enfants sont les distorsions, puis les substitutions, puis les omissions. Cependant, l'omission est l'erreur la plus fréquente chez les enfants implantés avant l'âge de deux ans et porteurs d'implants récents ainsi que les enfants porteurs d'aides auditives récentes. L'analyse phonologique quant à elle, indique que les erreurs récurrentes chez les deux groupes d'enfants sont liées au lieu d'articulation (substitution de /s/ par /ʃ/ pour les deux groupes et substitution de /n/ par /m/ chez les enfants CI uniquement). La différence entre les deux groupes se situe au niveau du mode d'articulation : les enfants HA remplacent fréquemment les fricatives par l'occlusive de même lieu d'articulation (phénomène de « stopping »), erreur bien moins fréquente chez les enfants CI.

Gaul-Bouchard et al. (2007) étudient l'ordre d'apparition des consonnes chez 22 enfants CI francophones québécois. Cet ordre est établi à partir d'enregistrements de parole spontanée à six, 12 et 18 mois post-implantation. Les auteurs constatent que les productions correctes de consonnes sont présentes dans l'inventaire des enfants dès six mois post implantation. Dans ces enregistrements, les occlusives sont les consonnes les plus fréquentes (55% des consonnes produites), et cette fréquence diminue au cours de l'étude (45% des consonnes produites à 18 mois post-implantation). Les labiales sont les consonnes les plus fréquentes à six mois post-implantation, puis ils observent une augmentation du nombre d'alvéolaires au cours de l'étude. En outre, ils notent que les consonnes produites tout au long de l'étude sont majoritairement des consonnes voisées. Par ailleurs, cette étude met en évidence que la production des consonnes est corrélée au niveau de communication orale avant implantation, plutôt qu'à l'âge d'implantation ou au mode de communication. Cette étude de Gaul-Bouchard et al. (2007) met en évidence une trajectoire de développement phonologique similaire à celle d'enfants normo-entendants, mais avec un délai plus long d'acquisition des consonnes, ce qui est un résultat similaire à ceux de l'étude de Kim & Chin (2008).

L'étude de Chin & Pisoni (2000) propose également un inventaire des productions segmentales (consonnes et voyelles) d'un enfant anglophone CI de 5;8 ans ayant reçu un implant à l'âge de 3;10 ans, à partir d'une tâche d'élicitation de noms de personnes de son entourage. Les productions sont transcrites phonétiquement, à partir de jugements perceptifs. Comme pour les voyelles (voir chapitre 3, §1.), les résultats de l'étude indiquent que l'enfant produit certaines consonnes, non présentes dans l'inventaire de la mère (qui est utilisé comme contrôle dans cette étude), notamment [b̥], [d̪], [ʔ] et [g]. Les consonnes dont les lieux d'articulation sont visibles (labiales et dentales) et moins visibles (alvéopalatales, palatales et vélaires) sont toutes présentes dans l'inventaire de l'enfant. Chin et Pisoni (2000) notent également que certains modes d'articulation (fricatives et affriquées) sont surreprésentés, par rapport aux consonnes-cibles, ils en déduisent que les sons de parole ne sont pas seulement perçus visuellement, mais que l'information acoustique transmise par l'implant participe également à la perception des consonnes moins visibles. Chin & Pisoni (2000) indiquent par ailleurs que l'inventaire des consonnes chez cet enfant CI est relativement complexe et qu'il est similaire à celui des enfants NH, à l'exception d'occlusives non-labiales fréquemment réalisées comme des fricatives alvéopalatales non-voisées, ce qui n'est pas une variation attestée chez des enfants NH. Ces résultats sont globalement similaires à ceux de Gaul-Bouchard et al. (2007), puisqu'ils font état d'un développement phonologique similaire mais plus tardif chez les enfants CI en comparaison avec celui d'enfants NH. Cependant, Gaul-Bouchard et al. (2007) notent que les consonnes les mieux produites sont celles qui sont les plus visibles, ce qui n'est pas le cas dans l'étude de Chin & Pisoni (2000), qui obtiennent des résultats similaires pour les consonnes visibles et non visibles.

L'étude de Peng et al. (2004) s'intéresse à la production des 21 consonnes du mandarin en position initiale de syllabe chez 30 enfants porteurs d'implants cochléaires de six à 12;6 ans (âge moyen : 9;3 ans), ayant reçu l'implant entre 2;3 et 10;3 ans (âge à l'implantation moyen : 5;8 ans) et avec une durée d'utilisation comprise entre 1;7 à 6;5 ans (durée moyenne d'utilisation : 3;7 ans). Cette étude de production se compose d'enregistrements des productions des enfants CI lors d'une tâche d'élicitation de mots à partir d'images, d'un test standardisé de vocabulaire, destiné à évaluer les compétences lexicales des enfants, ainsi que d'un test standardisé ayant pour but d'évaluer le niveau de développement langagier des enfants. Les auteurs notent que pour la plupart des lieux d'articulation, les scores sont compris entre 28% et 74%, ce qui témoigne de difficultés persistantes chez ce groupe d'enfants. Les résultats de l'étude de production des consonnes pour les modes d'articulation sont présentés dans la figure 3.10 ci-dessous.

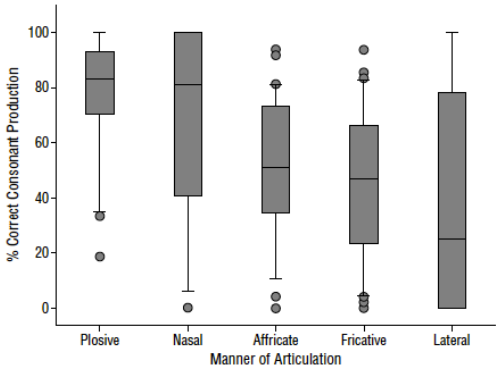


Figure 3.10 : résultats (proportion de consonnes correctement produites) de l'étude de Peng et al. (2004) par mode d'articulation cible.

Les auteurs notent que les différences de productions correctes entre ces modes d'articulation sont significatives : les plosives sont mieux produites que les nasales, que les affriquées, que les fricatives et les latérales, ce qui ne correspond pas aux résultats de l'étude de Geffner (1980) chez les enfants sourds non-implantés. Ces différences peuvent être liées à des capacités perceptives des enfants qui ont reçu un implant cochléaire, mais peuvent également être liées aux caractéristiques de l'inventaire des consonnes (organisation du système consonantique et fréquence des différentes consonnes) de la langue maternelle des enfants (anglais américain dans l'étude de Geffner (1980) et mandarin pour l'étude de Peng et al. (2004)). Par ailleurs, les résultats de Peng et al. (2004) indiquent une corrélation de la production des consonnes avec la compétence lexicale et le développement langagier des enfants.

L'ensemble des études sur la production des consonnes par des enfants sourds porteurs d'implants ou d'aides auditives traditionnelles présentées ci-dessus utilisent des jugements perceptifs et des transcriptions phonétiques des consonnes produites, elles mettent en évidence des difficultés des enfants à produire certains lieux et modes d'articulation, même avec plusieurs années d'utilisation des prothèses auditives ou de l'implant cochléaire. Ces difficultés correspondent pour certains auteurs à une perception partielle des sons de parole qui perturbe l'acquisition de certaines consonnes, mais qui peut correspondre pour d'autres auteurs à un retard de développement langagier des enfants par rapport aux enfants normo-entendants. On peut s'interroger également sur la méthode utilisée, qui repose sur des jugements perceptifs, ce qui introduit un biais subjectif, lié à la personne qui effectue les transcriptions et les jugements de production et à son système phonologique.

La suite de ce chapitre présente une revue de quelques études acoustiques disponibles dans la littérature, et dont le but est de proposer une caractérisation la plus objective possible des productions de consonnes par les enfants sourds, porteurs d'implants cochléaires ou utilisateurs d'aides auditives traditionnelles.

1.2.3. Etudes acoustiques

1.2.3.1. Occlusives

Le premier type d'études acoustiques sur la production des consonnes par les enfants sourds prend en compte des caractéristiques des occlusives, telles que la différence de longueur de VOT entre occlusives voisées et non-voisées. Il existe relativement peu d'études acoustiques sur la production d'occlusives par les enfants sourds, porteurs d'implants cochléaires.

Dans l'étude de Horga & Liker (2006), les VOT des consonnes /t/ et /d/ produites par trois groupes d'enfants (normo-entendants, porteurs d'implants cochléaires et porteurs d'aides auditives traditionnelles) sont mesurées, puis comparés, de façon à pouvoir constater une éventuelle difficulté des enfants sourds à reproduire le trait de voisement. Les résultats de cette étude sont présentés sur la figure 3.11 ci-dessous.

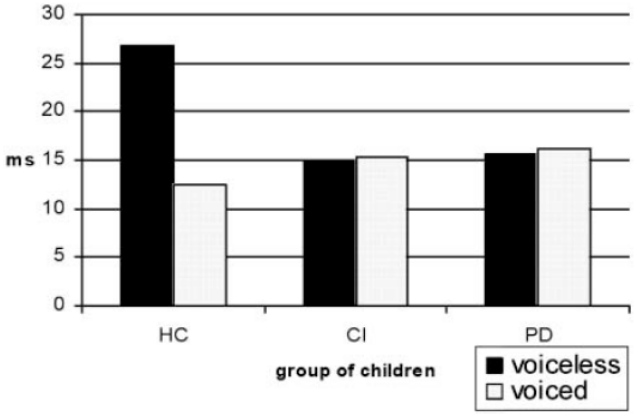


Figure 3.11 : durée de VOT de /t/ (voiceless) et /d/ (voiced) chez les enfants normo-entendants (HC), porteurs d'implants cochléaires (CI) et sourds profonds (PD), dans l'étude de Horga & Liker (2006)

Horga & Liker (2006) mettent en évidence un contraste fort entre occlusives voisées et non-voisées chez les enfants normo-entendants, mais une confusion dans les productions de ce contraste chez les deux groupes d'enfants sourds (CI et HA/PD). Ils constatent également qu'il n'y a pas de gain pour le voisement avec l'implant cochléaire par rapport aux aides auditives traditionnelles.

L'étude de Uchanski & Geers (2003) propose également une description de la réalisation du voisement chez deux populations d'enfants sourds, tous porteurs d'implants cochléaires : le premier groupe d'enfants utilise uniquement l'oral, alors que le deuxième utilise également une langue signée. Les productions d'occlusives de ces deux groupes sont comparées avec un groupe d'enfants normo-entendants. La figure 3.12 ci-dessous présente les résultats de l'étude de Uchanski & Geers (2003) sur le voisement.

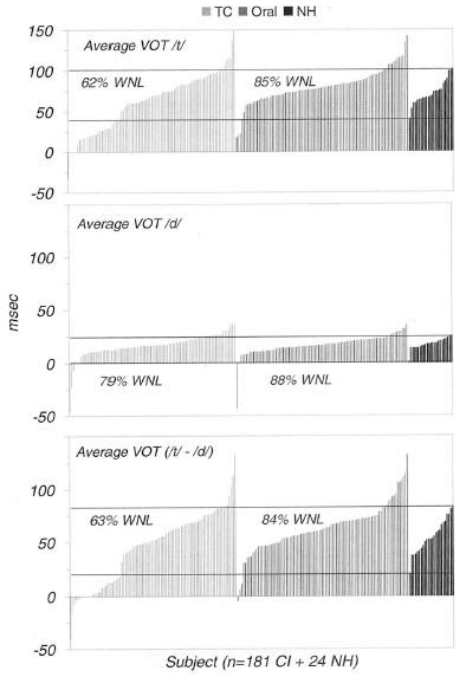


Figure 3.12 : durées des VOT de /t/ et /d/ chez les trois groupes d'enfants (TC : enfants sourds CI, communication signée et orale, Oral : enfants sourds CI, communication orale, NH : groupe contrôle, enfants normo-entendants) dans l'étude de Uchanski & Geers (2003). WNL : « within normal limits »

Uchanski & Geers (2003) indiquent que les enfants CI n'utilisant que l'oral ont des résultats meilleurs que les enfants utilisant deux types de communication (orale et signée). Ces productions de /t/ et /d/ sont en effet plus proches des productions des enfants normo-entendants pour les enfants qui n'utilisent que l'oral par rapport à ceux qui utilisent à la fois communication orale et signée.

1.2.3.2. Fricatives

A la différence des occlusives, il existe un nombre plus conséquent d'études acoustiques sur la production des fricatives.

Bharadwaj et al. (2006) proposent une description de /s/ et /ʃ/ avec implant activé et désactivé par quatre adultes et quatre enfants de 7 à 10;9 ans, locuteurs de l'anglais américain, tous porteurs d'implants cochléaires. Les productions de huit répétitions consécutives du mot « see » et de huit répétitions consécutives du mot « she » ont été enregistrées dans deux conditions : avec implant activé, puis désactivé. Les analyses acoustiques incluent centre de gravité, variance, skewness et kurtosis[10] des fricatives produites par tous les locuteurs dans les deux conditions. Nous nous intéresserons dans ce chapitre aux seuls résultats des productions des enfants. Les résultats indiquent que les centres de gravité ne sont pas significativement différents entre les deux conditions pour les deux fricatives, sauf pour les productions de /ʃ/ d'un seul enfant (C1). La variance n'est pas significativement différente entre les deux conditions, sauf pour les productions de /ʃ/ et /s/ d'un seul enfant (C4 qui n'est pas le même enfant dont les centres de gravité étaient significativement différents d'une condition à l'autre). Les résultats de skewness de /ʃ/ chez l'enfant C1 sont également les seuls à être significativement différents d'une condition à l'autre. La condition n'a pas d'effet significatif sur la kurtosis des productions des deux consonnes par les quatre enfants. Les résultats du groupe des enfants CI indiquent des valeurs similaires aux valeurs de la littérature pour les moments spectraux. Par contre, ils indiquent une variance pour /ʃ/ plus grande chez les enfants que chez les adultes, ce qui montre une variabilité plus grande dans leurs productions de fricatives.

Une étude longitudinale de Mildner & Liker (2008), qui fait suite à l'étude de Liker et al. (2007), présente l'évolution de deux fricatives du croate, /s/ et /ʃ/, enregistrés dans une tâche de répétition de chiffres et de mots contenant ces deux sons, par un groupe d'enfants contrôles NH et un groupe d'enfants porteurs d'implants CI dont on étudie les productions sur une période de 46 mois. La figure 3.13 ci-dessous présente les résultats de cette étude. Mildner & Liker (2008) s'intéressent dans cette étude à la bande de fréquence du bruit de la fricative (centre de gravité et variance). Ils notent une différence significative entre /s/ et /ʃ/ chez les enfants NH, ce qui témoigne d'une séparation claire entre les productions des deux consonnes. Lors de la première session d'enregistrement, les fricatives produites par les enfants CI ont toutes une fréquence similaire à celle de la fricative /s/ du groupe contrôle. Chez les enfants CI, /s/ a une variance également similaire à celle de /s/ du groupe NH, mais /ʃ/ a une variance beaucoup plus grande que celle du /s/ des groupes NH et CI de cette première session. Lors de la première session d'enregistrement, les deux fricatives /s/ et /ʃ/ sont donc produites de façon similaire (centre de

[10] Les termes « skewness » et « kurtosis » sont généralement traduits en français respectivement par « coefficient de dissymétrie » et « coefficient d'aplatissement » mais nous préférons utiliser les termes anglais par souci de concision.

gravité et variance) par les enfants CI, avec des caractéristiques proches du /s/ produit par les enfants NH.

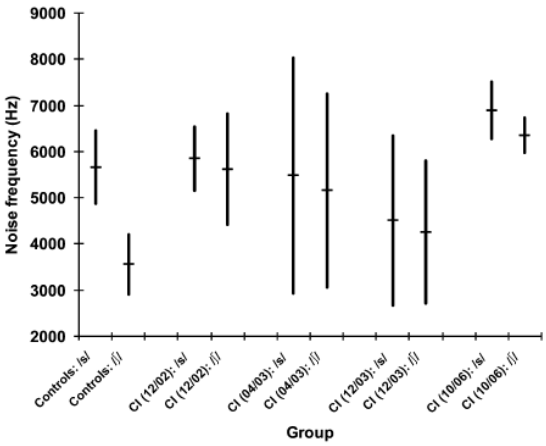


Figure 3.13 : centre de gravité des consonnes /s/ et /ʃ/ du groupe contrôle NH et des quatre sessions d'enregistrement des enfants CI dans l'étude de Mildner & Liker (2008)

Lors des deuxième et troisième sessions d'enregistrement, le centre de gravité des deux consonnes a tendance à diminuer, mais la variance est grande pour les deux consonnes lors des deux sessions d'enregistrement, ce qui traduit une grande variabilité dans la production de ces deux consonnes. Les caractéristiques (centre de gravité et écart-type) des deux fricatives /s/ et /ʃ/ produites par les enfants CI, vont s'approcher progressivement de celles du /ʃ/ produit par les enfants NH lors des deuxième et troisième sessions. Enfin, lors de la quatrième session d'enregistrement, une tendance nouvelle est observée : les enfants CI produisent des fricatives dont le centre de gravité est plus élevé que celui de /s/ produit par les enfants NH. Mildner & Liker (2008) constatent, lors de cette session, une différence significative entre les centres de gravité de /s/ et /ʃ/, ainsi qu'une variance faible pour ces deux consonnes. Ils en déduisent qu'à partir de 34 mois, la différence entre /s/ et /ʃ/ commence à se stabiliser puisque les deux catégories de fricatives sont clairement distinctes (avec des centres de gravité et écarts-types significativement différents) et les deux catégories commencent à émerger, même si elles sont produites dans des zones de fréquence plus hautes que celles des enfants contrôle NH de l'étude.

L'étude de Uchanski & Geers (2003) sur la production de consonnes chez des enfants CI et NH s'intéresse, outre à la réalisation du voisement dont les résultats ont été présentés ci-dessus, aux réalisations des deux fricatives /s/ et /ʃ/. Les auteurs de cette étude s'intéressent à trois moments spectraux (centre de gravité, skewness et kurtosis), convertis en Bark et dont les résultats sont présentés dans la figure 3.14 ci-dessous. Dans cette étude, les enfants répètent des phrases en imitant la production d'un expérimentateur accompagnée d'anglais signé et d'une transcription orthographique.

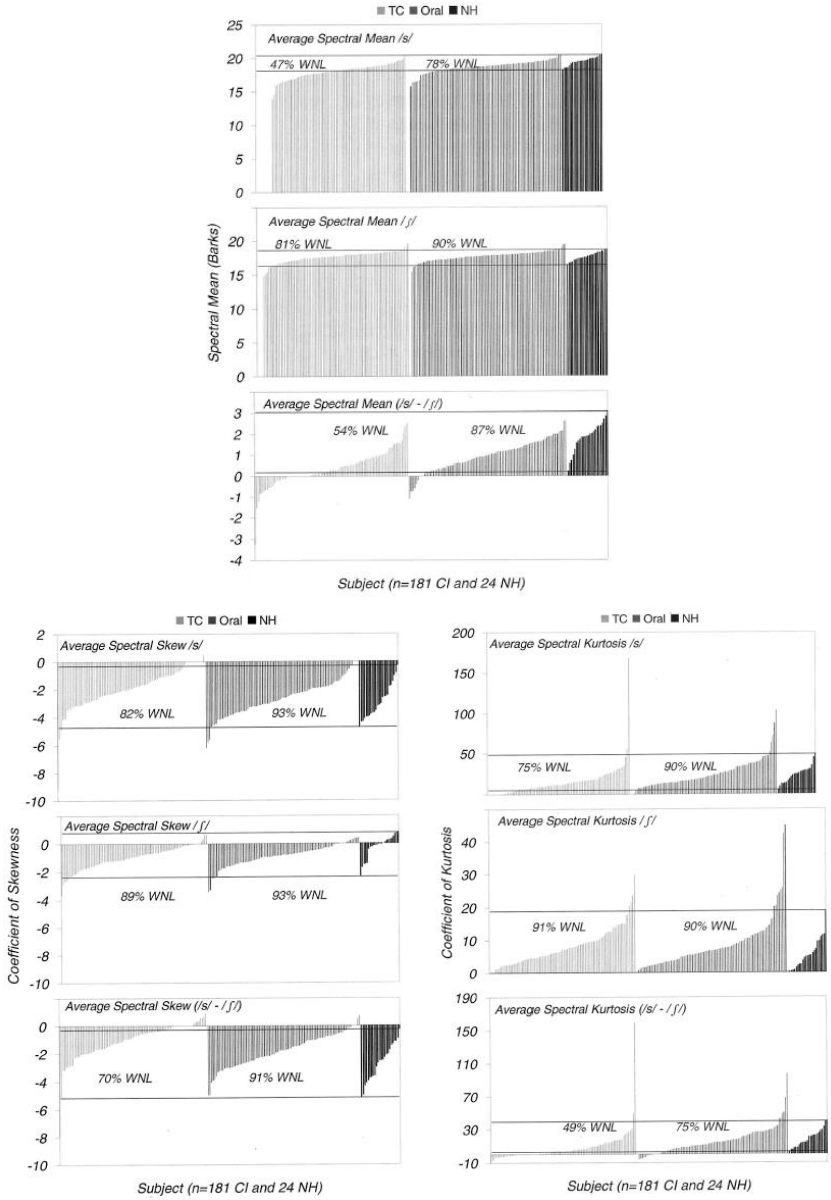


Figure 3.14 : centre de gravité (haut), skewness (bas, gauche) et kurtosis (bas, droite) des consonnes /s/ et /ʃ/ chez les trois groupes d'enfants (TC : enfants sourds CI, communication signée et orale, Oral : enfants sourds CI, communication orale, NH : groupe contrôle, enfants normo-entendants) dans l'étude de Uchanski & Geers (2003)

Les résultats obtenus dans le cadre de cette étude chez les enfants CI sont comparés à ceux d'enfants NH dont les productions ont été enregistrées selon la même procédure. Des jugements de la production des deux fricatives sont également calculés : les enfants CI obtiennent des résultats bien meilleurs (84% de productions correctes) pour /ʃ/, que pour /s/ (36%) et les enfants n'utilisant que l'oral obtiennent des résultats bien meilleurs également (61% de fricatives correctement produites) que les enfants utilisant une communication orale et signée (37%).

Par ailleurs, les analyses acoustiques des trois moments spectraux pour ces deux fricatives indiquent une même tendance pour la majorité des enfants CI : le centre de gravité des deux fricatives a des valeurs similaires à celui des enfants NH (le centre de gravité de /s/ est plus élevé que celui de /ʃ/) chez 71% des enfants CI, le deuxième moment spectral (skewness) a la même orientation que celui des enfants NH chez 81% des enfants CI, et le troisième moment spectral (kurtosis) a la même orientation que celui des enfants NH chez 62% des enfants CI. Uchanski et

Geers (2003) notent en outre que les résultats de ces analyses acoustiques sont meilleurs pour les enfants n'utilisant que l'oral que pour les enfants utilisant une communication orale et signée. Cette étude de Uchanski & Geers (2003) a été réalisée auprès d'enfants CI et NH appariés en âge chronologique, et tous avaient entre huit et neuf ans lors des enregistrements. Une étude de Todd et al. (2011) propose une comparaison des productions des deux fricatives /s/ et /ʃ/ chez des enfants CI, et chez des enfants NH appariés soit en âge chronologique, soit en âge auditif. Tous les enfants CI de l'étude utilisent l'implant depuis au moins quatre ans et l'objectif de Todd et al. (2011) est de comprendre si ces enfants CI ont réussi à combler le retard lié à un accès tardif à l'oral par rapport à des enfants NH du même âge ou si leurs productions de fricatives sont comparables à des enfants NH ayant la même durée d'expérience de l'oral. La figure 3.15 ci-dessous présente les résultats de l'étude de Todd et al. (2011). Dans cette étude, 39 enfants CI de quatre à neuf ans (âge chronologique) et 43 enfants NH appariés en âge chronologique ou en âge auditif.

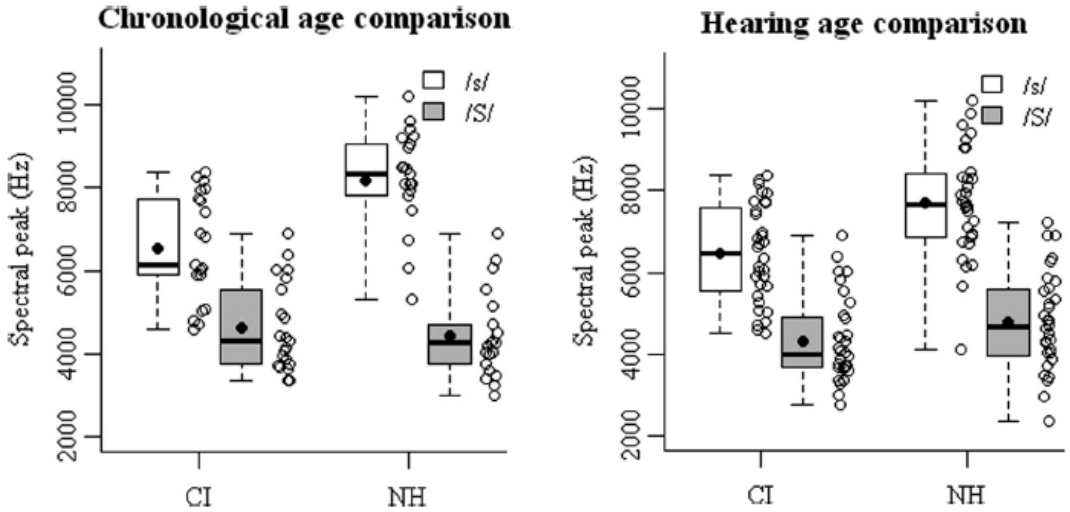


Figure 3.15 : centre de gravité des consonnes /s/ et /ʃ/ chez les deux groupes d'enfants CI et NH en appariés en âge chronologique (gauche) et en âge auditif dans l'étude de Todd et al. (2011)

Les productions de fricatives sont enregistrées lors d'une tâche de répétition de mots d'après un modèle audio adulte et une image. Les fricatives /s/ et /ʃ/ sont toutes en position initiale de mot, et sont suivies de /i, ɑ, u/. Les consonnes sont ensuite transcrites et seuls les sons jugés corrects sont utilisés dans les analyses (les distorsions, substitutions et suppressions de fricatives sont exclues des analyses). Un score correspondant aux productions jugées correctes est calculé pour chaque fricative et chaque groupe : pour tous les enfants, les résultats sont meilleurs pour /ʃ/ que pour /s/, et les auteurs n'observent pas d'effet de groupe (CI, NH appariés en âge chronologique et NH appariés en âge auditif). Pour l'ensemble des enfants, Todd et al. (2011) notent une différence significative entre les pics spectraux des deux fricatives (le pic spectral de /s/ est plus élevé que celui de /ʃ/). Lorsqu'ils comparent les productions des enfants CI avec celles des enfants NH appariés en âge chronologique, ils observent que le pic spectral de /s/ des enfants CI est plus bas que celui des enfants NH, alors que pour /ʃ/, les deux groupes ont des pics spectraux comparables. Lorsqu'ils comparent les enfants CI aux enfants NH appariés en âge auditif, ils constatent que pour les deux fricatives les pics spectraux des enfants CI sont plus bas que ceux des enfants NH. Il ressort de cette étude que les enfants CI produisent des fricatives

dont le pic spectral est plus comparable à celles d'enfants NH appariés en âge auditif qu'à celles d'enfants NH appariés en âge chronologique.

Les résultats de l'étude de Mildner & Liker (2008) semblaient indiquer que dans un premier temps, les centres de gravités des fricatives /s/ et /ʃ/ avaient tendance à être plus bas chez les enfants CI que chez les enfants NH, puis que dans un deuxième temps ils étaient au contraire plus élevés chez les enfants CI, vers la fin de l'étude (46 mois après implantation). Les résultats de Todd et al. (2011) montrent au contraire que les fricatives produites par les enfants CI sont soit similaires, soit plus basses que les enfants NH, qu'ils soient appariés en âge chronologique ou en âge auditif.

1.2.4. Conclusion

Cette revue de littérature sur les productions de consonnes par les enfants sourds, porteurs d'implants cochléaires ou d'aides auditives traditionnelles a permis de mettre en évidence la diversité des populations étudiées ainsi que des méthodologies de recueil et d'analyses de données. Les études perceptives et les études acoustiques sont complémentaires et permettent de mieux comprendre les trajectoires de développement phonologique des enfants sourds appareillés. Les inventaires de consonnes et les types d'erreurs perçues montrent que ces trajectoires ne sont pas linéaires et que certaines caractéristiques des consonnes sont plus difficilement acquises par les enfants. Ces observations subjectives sont la plupart du temps confirmées par les études acoustiques, même si on constate des résultats assez variables d'une étude à l'autre.

Il ressort des études acoustiques sur la production de consonnes, que pour les occlusives, la réalisation du voisement peut être problématique chez les enfants CI et HA alors que pour les fricatives, la distinction entre alvéolaires et post-alvéolaires est difficile chez ces deux groupes d'enfants. Comme pour les études sur les voyelles, il n'existe, à notre connaissance, aucune étude sur la production des consonnes par des enfants CI francophones.

1.3. Coarticulation

1.3.1. Introduction

Les études sur les voyelles et les consonnes présentées dans ce chapitre ont permis de mettre en évidence certaines caractéristiques segmentales de la production de parole par les enfants porteurs d'implants cochléaires. Il apparaît, à travers cette revue de littérature, que les enfants sourds, qui ont une première expérience de la perception orale avec l'implant cochléaire, rencontrent des difficultés de production des sons de parole. Cependant, ces études s'intéressent aux sons de parole en isolation, et ne prennent pas en compte la production de ces sons dans le flux de parole. Quelques rares études pourtant se concentrent sur les interactions des segments et la coarticulation lors de la production de parole par des locuteurs sourds, porteurs ou non d'implants cochléaires ou d'aides auditives traditionnelles. Nous allons présenter ces études dans cette section. Comme pour les études sur les voyelles et les consonnes, l'une des études (Warner-Czyz et al. (2010)) est réalisée sur la base de jugements perceptifs alors que les autres études présentées ici (Monsen (1976), Lane et al. (2001) et McCaffrey-Morrison (2008, 2012)) se basent sur des analyses acoustiques. En raison du nombre très réduit d'études sur la coarticulation chez des locuteurs sourds, nous avons fait le choix de présenter des études sur la coarticulation d'enfants, d'adolescents et d'adultes sourds, qui malgré une grande hétérogénéité des populations

étudiées, permettent de mettre en évidence des difficultés récurrentes dans la production de séquences CV, liées à des difficultés d'audition.

1.3.2.Etude perceptive : types d'erreurs et évolution de la production d'enfants CI

L'étude perceptive de Warner-Czyz et al. (2010) propose d'évaluer l'évolution des productions de séquences consonne-voyelle (CV) chez quatre enfants CI en les comparant avec celles d'enfants NH, tous enregistrés au stade des premiers mots. Tous les enfants sont enregistrés une fois par mois pendant six mois, à partir du moment où ils produisent leurs premiers mots. Les enfants CI, tous sourds prélinguaux, ont un âge chronologique de 24 mois en moyenne, et utilisent un implant depuis au moins six mois, lors de la première session d'enregistrement. Les enfants NH sont un peu plus jeunes, et sont appariés en « âge de développement vocal » (qui permet d'apparier les enfants selon leurs types de productions de parole) avec les enfants CI, pour que les productions des deux groupes soient comparables. Les enregistrements utilisés dans cette étude sont des productions de parole spontanée (interaction de chaque enfant avec l'un de ses parents). Cette étude a pour but de comprendre d'une part dans quelle mesure l'accès tardif à l'audition chez certains enfants affecte la justesse de la production des syllabes CV, et d'autre part quelle est l'évolution des types d'erreurs et des proportions de productions correctes au cours du temps lorsque l'accès à l'oral est plus tardif et perturbé. Les auteurs placent leur étude dans le cadre de la théorie « Frame, then Content », pour expliquer les relations entre contrôle articulatoire et préférence pour un type de séquence d'une part, et émergence de syllabes correctes d'autres part. La théorie prédit notamment une préférence pour des séquences /ba/, /di/, /ku/ (voir chapitre 1, §5.2).

Warner-Czyz et al. (2010) utilisent des transcriptions phonétiques des séquences CV produites, qu'ils comparent avec la cible lexicale attendue. Une syllabe est ensuite jugée correctement produite si à la fois le mode et le lieu d'articulation de la consonne et l'antériorité et la hauteur de la voyelle correspondent à la cible lexicale. Des types d'erreurs et de justesse de la syllabe sont ainsi quantifiés : a) catégorie CV correcte, b) catégorie partiellement correcte, c) catégorie inadaptée, d) omission partielle, e) omission complète. Les résultats indiquent que les séquences consonnes labiales-voyelles centrales sont les plus fréquentes et les mieux produites pour les deux groupes (32% pour les enfants NH et 44% pour les enfants CI), et que les enfants NH ont une préférence pour les séquences CV prévisibles dans le cadre de la théorie « Frame, then Content », alors que les enfants CI n'ont pas cette préférence. Plusieurs hypothèses sont avancées pour expliquer ces différences entre les deux groupes : une production plus limitée de séquences CV chez les enfants CI, l'âge chronologique qui diffère d'un groupe à l'autre (les enfants CI sont plus âgés, puisque l'apparition de leurs premiers mots, utilisée pour apparier les deux groupes d'enfants, est plus tardive), et un effet du choix du lexique sur la production et la fréquence des séquences-cibles. Les auteurs notent par ailleurs des productions plus correctes pour les enfants NH que pour les enfants CI. Les préférences pour certaines séquences CV ne changent pas au cours des six mois de l'expérience pour les enfants CI. Cependant, pour les deux groupes, la justesse des productions CV augmente significativement au cours de l'étude.

Cette étude de Warner-Czyz et al. (2010) propose un premier panorama des premières productions de syllabes chez les enfants porteurs d'implants, qui apparaît légèrement différent de celui des enfants normo-entendants et qui indiquent un effet du feedback auditif, notamment sur le type de consonnes produites par les deux groupes d'enfants.

1.3.3.Etudes acoustiques de la coarticulation

Il existe quelques études acoustiques sur la coarticulation de séquences CV par des personnes sourdes. Ces études s'intéressent en particulier aux transitions formantiques (F2) entre consonnes et voyelles.

La première étude acoustique présentée dans ce chapitre est celle de Lane et al. (2001), qui utilise l'équation de locus pour quantifier les difficultés de production d'adultes sourds porteurs d'implants cochléaires. Il s'agit de sept adultes devenus sourds à l'âge adulte, dont les productions sont comparées à celles de deux adultes normo-entendants. Les productions des adultes porteurs d'implants sont enregistrées lors d'une tâche de lecture des huit voyelles de l'anglais dans des contextes /bVt/ et /dVt/, dans des phrases porteuses. Les enregistrements sont réalisés trois fois avant activation de l'implant, puis cinq fois après activation (0, 4, 8, 12, 24 semaines après implantation). Les auteurs effectuent des mesures de F2 au début du voisement, et au milieu de la voyelle. Ils calculent ensuite un indice de coarticulation (ratio entre F2 au début du voisement et F2 au milieu de la voyelle), et la pente de l'équation de locus à partir des deux mesures de F2. Les résultats indiquent que le changement de capacités perceptives liées à l'implantation n'a que peu d'effet sur les deux indices de coarticulation, et que la coarticulation de ces adultes devenus sourds n'a pas été affectée par la surdité, puis par l'amélioration de l'audition consécutive à l'implantation, puisqu'elle reste dans tous les cas très proche de la coarticulation mesurée chez les deux adultes contrôles, normo-entendants. Les stratégies de coarticulation acquises avant la perte d'audition semblent être fossilisées et ne semblent pas subir d'effets de la perte d'audition puis de l'implantation cochléaire.

L'étude de Monsen (1976), propose une comparaison des transitions formantiques (F2) de séquences CV produites par des adolescents sourds non appareillés (six adolescents de 14 à 18;8 ans) et par des adolescents normo-entendants (six adolescents de 13 à 15 ans). Le but de Monsen (1976) est de comprendre si la coarticulation est liée à l'intelligibilité de la parole des adolescents sourds. Les productions enregistrées sont des mots monosyllabiques CV(C) dans des phrases porteuses. La première consonne est /b/, /d/ ou /f/ et la voyelle est /i/ ou /u/. L'étude s'intéresse exclusivement aux transitions du F2 entre la consonne initiale et la voyelle, qui donne des indications sur l'enchaînement des segments. Le F2 est mesuré en deux points : au début de la transition formantique, et sur une partie stable de la voyelle. Les résultats indiquent à la fois une variabilité inter-sujets pour les deux groupes, une tendance à la réduction vocalique (centralisation de /i/ et /u/) chez les enfants sourds, qui joue un rôle sur la transition formantique. Pour Monsen (1976), l'information consonantique est portée par la transition formantique, et l'impression d'une moins bonne intelligibilité de la parole des enfants sourds serait donc liée à la consonne plutôt qu'à la voyelle. Cette étude est la première qui s'intéresse véritablement à l'interaction des segments dans la parole des locuteurs sourds et donne quelques indications sur les difficultés de ces locuteurs à reproduire un enchainement de segments de manière aussi fluide que des locuteurs normo-entendants.

Deux études de McCaffrey-Morrison (2008, 2012) utilisent également l'équation de locus soit pour quantifier l'évolution de la coarticulation sur une période de six mois chez trois enfants CI ou HA de 11 à 50 mois (McCaffrey-Morrison (2012)), soit pour évaluer les effets sur la coarticulation de l'utilisation d'aides auditives traditionnelles après plusieurs années par quatre adolescents de 13 à 18 ans (McCaffrey-Morrison (2008)). Dans ces deux études, les indices de

coarticulation sont comparés à ceux d'enfants normo-entendants de la littérature (notamment Gibson & Ohde (2007) pour les enfants de 11 à 50 mois, et Sussman et al. (1991) pour les adolescents).

Dans l'étude de McCaffrey-Morrison (2008) chez les adolescents HA, des productions de 30 syllabes /bVt/, /dVt/ et /gVt/, avec neuf voyelles de l'anglais, dans des phrases porteuses ont été enregistrées. Les pentes des équations de locus ont été calculées à partir des mesures de F2 au début de la voyelle et F2 au milieu de la voyelle. Les résultats indiquent que les adolescents HA produisent les syllabes /bVt/ avec moins de coarticulation anticipatoire (pentes des équations de locus plus faibles) que les locuteurs NH de la littérature, les syllabes /dVt/ avec plus de coarticulation anticipatoire que les locuteurs NH de la littérature et que les voyelles antérieures influencent plus les productions de /g/ chez les adolescents HA que les adolescents NH de la littérature. Par ailleurs, McCaffrey-Morrison (2008) note des différences inter-interlocuteurs assez marquées dans la réalisation de la coarticulation anticipatoire. McCaffrey-Morrison (2008) et Monsen (1976) utilisent des mesures de coarticulation différentes : comparaisons de F2 au début de la voyelle et F2 sur une partie stable de la voyelle pour Monsen (1976) et équations de locus pour McCaffrey-Morrison (2008), mais les résultats sont assez similaires, puisque les deux types de mesures permettent d'établir que les adolescents sourds, (dont la surdité est pré- et périlinguistique, pour les deux études) utilisateurs d'aides auditives traditionnelles ne produisent pas des séquences CV strictement similaires à celles d'adolescents normo-entendants du même âge, même après plusieurs années d'utilisation de ces aides auditives.

L'étude de McCaffrey-Morrison (2012) s'intéresse à la réalisation de la coarticulation anticipatoire dans des vocalisations de deux enfants CI et d'un enfant HA de 11 à 50 mois, au stade des premiers mots, en les comparant à celle d'enfants NH de la littérature (Gibson & Ohde (2007)), appariés en âge chronologique et en âge auditif avec les enfants CI. Les vocalisations sont enregistrées lors d'une session un mois avant implantation (pour les deux enfants CI), puis six et douze mois après l'implantation, et six et douze mois après le début de l'utilisation des prothèses auditives pour l'enfant HA. Des séquences /bV/, /dV/ et /gV/ produites lors de ces enregistrements de parole spontanée sont analysées. Les résultats diffèrent d'une séquence CV à l'autre. En effet, pour les séquences /bV/ qui sont déjà produites par les trois enfants, avant implantation, les pentes des équations de locus sont similaires à celles de la littérature pour les enfants NH de même âge chronologique, ce qui indique que la production est facilitée par la perception visuelle de la consonne, et que le développement de la production est liée à la maturation du contrôle moteur, plutôt qu'à l'expérience de l'oral. Pour les séquences /dV/, les résultats de l'enfant HA sont comparables à ceux d'enfants NH de même âge auditif, de même que les enfants CI. Ces deux enfants CI qui ne produisaient pas de séquences /dV/ avant implantation ont des pentes d'équation de locus plus fortes que les enfants NH de même âge chronologique, mais similaires à celles d'enfants NH de même âge auditif, ce qui indique un retard dans l'acquisition de la consonne /d/, liée à l'absence de perception visuelle de cette consonne, qui n'est perçue par les enfants CI que lorsqu'ils accèdent à l'oral avec l'implant cochléaire. Les résultats des séquences /gV/ quant à eux sont similaires aux enfants NH de même âge auditif pour l'enfant HA et aux enfants NH de même âge chronologique pour les enfants CI, 13 mois post implantation.

1.3.4.Conclusion

Les résultats de ces études chez des adultes ou des enfants sourds, semblent montrer que la coarticulation est encodée au niveau moteur après acquisition, et que lorsque la surdité survient tardivement chez des adultes, ceux-ci ne perdent pas ce contrôle. A l'inverse, ces études indiquent un effet de la surdité sur la réalisation des séquences CV et donc sur l'organisation des gestes articulatoires mis en œuvre pour produire ces séquences. Les résultats des jeunes enfants et des adolescents semblent confirmer ces résultats, puisque les enfants ont des productions différentes selon le lieu d'articulation, avec une coarticulation similaire aux enfants NH de même âge chronologique pour /b/ qui est une consonne produite avant même que les enfants aient accès à l'audio. Ces résultats sont également en accord avec ceux de l'étude perceptive de Warner-Czyz et al. (2010) qui montre que les séquences labiales-voyelles centrales sont les premières séquences à apparaître dans les vocalisations des enfants.

Les études présentées dans la première partie de ce chapitre ont permis de mettre en évidence des difficultés des enfants porteurs d'implants cochléaires à produire les voyelles, les consonnes, et à produire des enchaînements de séquences consonnes-voyelles. Ces difficultés peuvent être liées à l'audition partielle de ces enfants, qui ne permet pas systématiquement de distinguer certains sons et à un accès réduit à l'audio dans leurs premiers mois de vie, qui constituent une période critique pour l'acquisition du langage et du développement phonologique. La visibilité du segment (consonne ou voyelle) semble jouer un rôle important dans la capacité à percevoir les sons, puis à les reproduire. Les résultats des études présentées ci-dessous sont également marqués par une très grande hétérogénéité, ce qui peut s'expliquer par des choix méthodologiques disparates et des cohortes d'enfants d'âges très variables et hétérogènes. Les choix méthodologiques de notre étude expérimentale devront prendre en compte ces caractéristiques et nous proposerons une méthode d'analyse des données qui prend en compte cette hétérogénéité des productions de parole.

La suite de ce chapitre est consacrée à l'intelligibilité de la parole d'enfants porteurs d'implants cochléaires qui est le reflet de la production de la parole et qui peut être quantifiée. Comme pour la production de parole, nous avons choisi une sélection d'études d'intelligibilité de la parole d'enfants sourds, porteurs d'implants cochléaires.

2. Intelligibilité de la parole d'enfants porteurs d'implants cochléaires : évaluation et facteurs explicatifs

2.1. Facteurs explicatifs du niveau d'intelligibilité

Le but d'une étude d'intelligibilité est d'évaluer et de quantifier le degré d'intelligibilité par des jugements perceptifs qualitatifs et d'établir des liens entre cette évaluation perceptive et certaines caractéristiques observables et quantifiables dans la production de la parole. Deux types de facteurs peuvent être envisagés pour expliquer un déficit d'intelligibilité de parole : caractéristiques individuelles et environnementales des locuteurs ou caractéristiques de la production de parole de ces locuteurs. Cette problématique est donc complexe, car, qu'il s'agisse de l'un ou l'autre cas, les études d'intelligibilité tentent de définir les liens entre ces corrélats quantifiables objectivement et les jugements d'intelligibilité qui sont des jugements perceptifs, donc subjectifs. Cette question n'est pas seulement abordée dans le cadre d'études sur la parole pathologique et la parole d'enfant mais également sur la parole non pathologique d'adultes. Nous proposerons une réflexion sur les méthodologies utilisées dans le cadre de ces études et qui permettent de quantifier l'intelligibilité de la parole.

2.1.1. Corrélats de l'intelligibilité dans la parole non-pathologique

Dans une étude sur un corpus de locuteurs adultes normo-entendants, Bradlow et al. (1996) s'intéressent aux liens entre taille de l'espace vocalique et intelligibilité. Ils proposent d'étudier la variabilité inter-individuelle de la taille de l'espace vocalique pour comprendre les effets de cette variation sur l'intelligibilité de la parole. Ils cherchent à établir les liens entre caractéristiques des locuteurs, production de parole et intelligibilité. L'étude est à la fois acoustique et perceptive, l'objectif étant de pouvoir dégager les caractéristiques saillantes de production de parole qui influencent l'intelligibilité.

Les auteurs se basent sur trois caractéristiques globales du locuteur (genre, F0, débit de parole) qui sont liées à la totalité des données enregistrées pour chaque locuteur, et deux caractéristiques segmentales (réalisation de la catégorie vocalique et durées relatives des segments).

Les productions analysées dans cette étude sont tirées d'une base de données déjà constituée (Multi-Talker Sentence Database), qui n'est pas construite spécialement pour cette étude : il s'agit de cent phrases produites par vingt locuteurs (dix hommes et dix femmes). Ce sont des phrases ne contenant qu'une proposition, avec cinq mots de contenu par phrase, associés à un nombre variable de mots fonctionnels. Les locuteurs n'ont pas de pathologies du langage et les phrases sélectionnées ne contiennent pas d'hésitations.

Les jugements d'intelligibilité de la parole sont des transcriptions des stimuli par 200 auditeurs. Chaque auditeur transcrit orthographiquement l'intégralité de la production d'un locuteur (100 phrases), et chaque production d'un locuteur est jugée par dix auditeurs différents (200 auditeurs au total).

Une phrase est considérée correctement transcrite si les cinq mots de contenu de cette phrase sont correctement transcrits. Dans le cas contraire, elle est considérée comme incorrecte. Le score moyen d'intelligibilité de chaque locuteur est ensuite calculé à partir des scores donnés par chaque auditeur.

Le groupe des locuteurs se compose de dix hommes et dix femmes, le genre pouvant être considéré comme un facteur explicatif dans la mesure où les groupes diffèrent par la F0 et par des caractéristiques articulatoires. Dans cette étude, les auteurs notent une plus grande intelligibilité, significative statistiquement, des femmes par rapport aux hommes. Pour expliquer cette différence liée au genre, les auteurs proposent l'hypothèse qu'elle découlerait de phénomènes de réduction présents en plus grand nombre dans les productions des hommes.

Les auteurs de cette étude montrent qu'une étendue plus grande de F0 (plutôt que la F0 moyenne), le genre (la parole des femmes est plus intelligible que celle des hommes), un espace vocalique plus grand (ce qui est le cas lorsque les voyelles sont largement dispersées sur la dimension du F1), et la précision articulatoire (l'intelligibilité est plus grande lorsqu'il y a moins de suppressions de segments et d'erreurs de syllabification) sont les caractéristiques qui sont fortement corrélées à l'intelligibilité. L'intelligibilité est donc liée autant à certaines caractéristiques globales des énoncés (comme par exemple l'intonation, le débit de parole, etc.) produits par un locuteur qu'à des informations acoustiques au niveau segmental : l'intelligibilité dépend d'une grande précision articulatoire et de peu de réductions phonétiques. Cette étude n'est pas liée directement à la parole de locuteurs ayant une audition dégradée, mais elle donne des indications pertinentes sur les facteurs qui influencent l'intelligibilité, en particulier la taille de l'espace vocalique et la répartition des voyelles dans cet espace.

2.1.2. Facteurs explicatifs de l'intelligibilité dans la parole d'enfants porteurs d'implants cochléaires

Comme dans l'étude de Bradlow et al. (1996), plusieurs corrélats sont envisagés dans les études sur l'intelligibilité de la parole d'enfants porteurs d'implants cochléaires : des caractéristiques individuelles et environnementales des enfants ou des caractéristiques de la production de parole. Dans le premier cas, les études s'intéressent aux caractéristiques des populations d'enfants étudiés (âge chronologique, âge d'implantation, durée d'utilisation de l'implant cochléaire, entre autres) et n'utilisent les productions enregistrées que pour les jugements d'intelligibilité. Ces enregistrements de productions de parole ne donnent pas lieu à des analyses quantitatives. Dans le deuxième cas, les auteurs des études cherchent à quantifier une variabilité de parole chez les enfants porteurs d'implants, pour expliquer un déficit éventuel d'intelligibilité. Les productions enregistrées sont jugées à la fois perceptivement et acoustiquement.

2.1.2.1. Facteurs individuels ou environnementaux

Les études de Miyamoto et al. (1997), Svirsky et al. (2000), Blamey et al. (2001), Peng et al. (2004), Flipsen et al. (2006), Habib et al. (2010), Khwaileh et Flipsen (2010) et Montag et al. (2014) cherchent à établir l'effet de certaines caractéristiques des enfants telles que l'âge chronologique, l'âge d'implantation, le mode de communication, la stratégie de codage de l'implant, sur l'intelligibilité.

Certains auteurs étudient en outre les liens entre capacités perceptives des enfants et intelligibilité. Calmels et al. (2004) utilisent le « Test for the Evaluation of Voice Perception » (Vieu et al. (1999)).

2.1.2.2. Corrélats phonétiques

Perrin et al. (1999), Horga et Liker (2006), Poissant et al. (2006), Löfqvist et al. (2009) et Chuang et al. (2012) proposent d'étudier l'effet de la variabilité acoustique des productions au niveau

segmental sur l'intelligibilité. Ces études se concentrent sur les caractéristiques des voyelles (formants, taille de l'espace vocalique, fréquence fondamentale des voyelles). L'étude de Horga et Liker (2006) y ajoute une étude acoustique des consonnes (durée d'occlusion, VOT).

Plusieurs études s'intéressent aux effets de la variabilité suprasegmentale sur l'intelligibilité, seule ou en complément de la variabilité segmentale. Ainsi Löhle et al. (1999) étudient le rythme et l'intonation des enfants dans des productions spontanées, Chuang et al. (2012) le débit de parole, Horga et Liker (2006) l'accent de mot et l'accent de phrase, Perrin et al. (1999) le débit d'articulation et Poissant et al. (2006) la durée de mot.

2.2. Résultats des études d'intelligibilité

Les tables 3.3 et 3.4 ci-dessous présentent les corrélats de l'intelligibilité testés par les différents auteurs. Les caractéristiques individuelles et environnementales sont regroupées dans la table 3.3 et les caractéristiques des productions dans la table 3.4. Comme les corrélats testés diffèrent selon les études, nous n'avons indiqué que les résultats des corrélats testés par ces auteurs.

2.2.1. Facteurs individuels et environnementaux de l'intelligibilité

Auteurs	Date	Âge chronologique	Âge à l'implantation	Durée d'utilisation de l'implant	Caractéristiques familiales	Mode de communication/ nombre de L1	Nombre d'implants	Caractéristiques d'utilisation de l'implant (ON-OFF, stratégie de codage, type d'appareil…)	Capacités perceptives	Niveau d'audition
Blamey et al.	2001		non	oui					non	
Calmels et al.	2004		oui						oui	
Flipsen et al.	2006	oui	non	oui						
Habib et al.	2010	oui	oui			oui	non			non
Hassanzadeh	2012				oui					
Khwaileh et Flipsen	2010	non	non	oui						
Löhle et al.	1999		oui						oui	oui
Miyamoto et al.	1997			oui						oui
Montag et al.	2014		oui		oui	oui				
Peng et al.	2004		oui					oui		
Poissant et al.	2006							oui		
Svirsky	2000	oui		oui						

Table 3.3 : effets des caractéristiques individuelles des enfants CI sur l'intelligibilité (classement par ordre alphabétique du premier auteur). Oui : effet sur l'intelligibilité, non : aucun effet sur l'intelligibilité.

2.2.1.1. Durée d'implantation, âge d'implantation et âge chronologique

Parmi les différents facteurs individuels et environnementaux de l'intelligibilité de la production de parole chez les enfants porteurs d'implants, la durée d'utilisation de l'implant est l'un des plus fréquemment étudiés. Les résultats sont identiques pour toutes les études : l'intelligibilité est systématiquement corrélée avec la durée d'utilisation de l'implant dans les études de Blamey et al. (2001), Flipsen et al. (2006), Khwaileh et Flipsen (2010), Miyamoto et al. (1997) et Svirsky (2000). Ces auteurs s'intéressent également aux effets de l'âge chronologique ou de l'âge d'implantation sur l'intelligibilité. A la différence de la durée d'utilisation de l'implant, ces études ne permettent pas à leurs auteurs de mettre en évidence un effet systématique de ces deux corrélats sur l'intelligibilité.

Blamey et al. (2001) montrent une amélioration très forte de l'intelligibilité les trois premières années post-implantation puis une amélioration continue de l'intelligibilité jusqu'à six ans post-implantation, sans effet plateau. Cette amélioration de l'intelligibilité est liée, d'après les auteurs de

cette étude, à la durée d'utilisation de l'implant, mais pas à l'âge d'implantation. Cette étude ne s'intéresse pas à l'effet de l'âge chronologique des enfants.

Miyamoto et al. (1997) et Svirsky (2000) obtiennent des résultats similaires : l'intelligibilité des enfants CI s'améliore progressivement du premier test pré-implantation à la fin de l'étude (sept ans et demi post-implantation), cette amélioration étant liée à la durée d'utilisation de l'implant et de la rééducation orthophonique. Peng et al. (2004) mettent également en évidence un effet de l'âge d'implantation sur l'intelligibilité.

Les résultats de l'étude de Flipsen et al. (2006) montrent simultanément une amélioration de l'intelligibilité pendant les deux années qui suivent l'implantation et une corrélation de l'intelligibilité à l'âge chronologique des enfants. Les auteurs de cette étude indiquent par ailleurs que l'amélioration de l'intelligibilité chez les enfants porteurs d'implants cochléaires est plus lente que celle des enfants normo-entendants, dont l'intelligibilité atteint 100% (selon la méthodologie de l'étude) à l'âge de quatre ans. L'hypothèse selon laquelle, quatre ans post-implantation, les enfants porteurs d'implant auraient atteint le même niveau d'intelligibilité que les enfants normo-entendants de quatre ans n'est pas vérifiée. Khwaileh et Flipsen (2010) obtiennent des résultats légèrement différents : l'intelligibilité est corrélée à la durée d'utilisation de l'implant mais pas à l'âge chronologique ni à l'âge d'implantation des enfants.

Habib et al. (2010) montrent qu'il existe un effet de l'âge chronologique sur l'intelligibilité des enfants porteurs d'implants cochléaires, en particulier chez les enfants qui ont reçu l'implant entre huit et 24 mois. Ils mettent également en évidence des effets de l'âge d'implantation.

Les résultats des tests d'intelligibilité de l'étude de Löhle et al. (1999) montrent que pour les enfants implantés entre 9 et 14 ans, l'intelligibilité est la plus faible, et qu'elle est jugée meilleure avec des mots isolés qu'en parole spontanée. Les enfants implantés entre 5 et 8 ans ont des résultats d'intelligibilité meilleurs que les enfants ayant reçu l'implant plus tardivement. Les enfants implantés entre 2 et 4 sont les plus intelligibles. Les auteurs de l'étude concluent que les capacités de perception de parole sont meilleures si les enfants ont reçu l'implant à un âge précoce, et que l'intelligibilité de ces enfants est également meilleure. L'analyse acoustique indique que les enfants implantés jeunes développent de meilleures capacités à imiter le rythme et l'intonation de façon exacte. Les enfants ayant reçu l'implant le plus tardivement ont des résultats de perception et d'intelligibilité à l'opposé.

2.2.1.2.Caractéristiques de l'implant

Seules deux études s'intéressent aux effets de la nature et de la technologie de l'implant sur l'intelligibilité : Peng et al. (2004) mettent en évidence un effet de la stratégie de codage utilisée sur l'intelligibilité de la parole des enfants porteurs d'implants, l'implant réglé avec la stratégie de codage SPEAK donne de meilleurs résultats que l'implant réglé avec la stratégie de codage MPEAK.

Poissant et al. (2006) s'intéressent à l'effet de l'activation de l'implant sur l'intelligibilité, et mettent en évidence une meilleure intelligibilité lorsque l'implant est activé, malgré des résultats différents selon les enfants.

Miyamoto et al. (1997) et Svirsky (2000) comparent l'intelligibilité des enfants CI à celles d'enfants HA. Leurs études montre un bénéfice de l'implant en terme d'intelligibilité par rapport aux prothèses auditives.

2.2.1.3.Capacités perceptives et niveau d'audition

Calmels et al (2004) montrent un effet des capacités perceptives des enfants sur leur intelligibilité : ils constatent que les tâches de perception (sans lecture labiale) sont difficiles pour les enfants jusqu'à deux ans post-implantation, et qu'il y a une amélioration de l'intelligibilité entre trois ans et cinq ans, alors qu'avant trois ans, l'intelligibilité des enfants CI est jugée assez faible.

A l'inverse, l'étude de Blamey et al. (2001) montrent que l'intelligibilité n'est pas liée aux capacités perceptives des enfants CI. Habib et al. (2010) mettent en évidence que l'intelligibilité n'est pas corrélée au niveau d'audition, contrairement à l'étude de Miyamoto et al. (1997).

2.2.1.4.Caractéristiques familiales, modes de communication et nombre de L1

Habib et al. (2010) mettent en évidence un effet du bilinguisme sur l'intelligibilité : les résultats sont meilleurs chez les enfants les plus âgés et chez les enfants monolingues. Les résultats de l'étude longitudinale de Hassanzadeh et al. (2012) sont similaires chez les deux groupes d'enfants (de parents sourds et de parents normo-entendants) lors du premier réglage, puis s'améliorent lors des sessions de test suivantes, les enfants dont les parents sont sourds non appareillés ayant une amélioration plus forte. Dans le premier cas (Habib et al., 2010), le bilinguisme oral a un effet négatif sur l'intelligibilité des enfants CI, alors que dans le deuxième cas (Hassanzadeh et al., 2012), le bilinguisme oral-gestuel (les enfants dont les parents sont sourds non appareillés utilisent une langue signée pour communiquer avec eux) favorise une meilleure intelligibilité de ce groupe d'enfants.

Montag et al. (2014) mettent en évidence une intelligibilité globale du groupe CI assez élevée, mais plus basse que celle du groupe NH (les résultats d'environ un tiers des enfants CI sont plus bas que les résultats des enfants NH). Par ailleurs, ils notent que les phrases contenant des mots fréquents sont plus intelligibles que les phrases contenant des mots peu fréquents. Pour Montag et al. (2014), ces résultats d'intelligibilité sont liés à des facteurs globaux (âge d'implantation, mode de communication, revenu de la famille).

La comparaison des résultats de ces études indique qu'il est difficile d'établir avec certitude quels sont les corrélats individuels et environnementaux (table 3.3 ci-dessus) de l'intelligibilité de la parole d'enfants porteurs d'implants cochléaires. En effet, si la durée d'utilisation de l'implant, les caractéristiques familiales, le mode de communication et les caractéristiques des implants ont un effet sur l'intelligibilité dans toutes les études où ces caractéristiques ont été testées, les autres caractéristiques testées (âge chronologique, âge d'implantation, capacités perceptives et niveau d'audition) ont un effet variable selon les études.

2.2.2.Corrélats acoustiques de l'intelligibilité

La table 3.4 ci-dessous présente les corrélations testées entre certaines caractéristiques acoustiques et l'intelligibilité de la parole d'enfants porteurs d'implants cochléaires. Les corrélats les plus fréquemment testés sont liées à la production des voyelles : F1, F2, taille de l'espace vocalique. Une étude (Van Lierde et al. (2005)) s'intéresse également aux productions de consonnes. Des caractéristiques suprasegmentales sont également testées : F0, débit et pauses, nasalité, qualité de voix.

Auteurs	Date	Voyelle (F1)	Voyelle (F2)	Taille de l'espace vocalique	Qualité de voix	Nasalité	Débit/ pauses	Fréquence fondamentale (F0)
Chuang et al.	2012		oui				oui	
Horga et Liker	2006	oui	oui	oui				
Löhle et al.	1999							oui
Löfqvist et al.	2009	non	non	non				
Perrin et al.	1999				oui		oui	
Poissant et al.	2006	oui	oui					oui
Van Lierde et al.	2005				?	non		

Table 3.4 : effets des caractéristiques acoustiques des productions de parole sur l'intelligibilité (classement par ordre alphabétique du premier auteur). Oui : effet sur l'intelligibilité, non : aucun effet sur l'intelligibilité.

2.2.2.1.Caractéristiques segmentales

Le second formant (F2) est la mesure la plus fréquemment utilisée pour caractériser la production des voyelles dans les études présentées dans la table 3.4 : les études de Chuang et al. (2012), Horga et Liker (2006) et Poissant et al. (2006), montrent une corrélation entre F2 et intelligibilité de la parole, alors que l'étude de Löfqvist et al. (2009) ne permet pas d'établir cette corrélation. Les mesures réalisées dans le cadre de ces études sont cependant différentes, et les conditions d'élicitation de parole sont également différentes, ce qui peut expliquer les résultats contradictoires obtenus par les différents auteurs.

Poissant et al. (2006) calculent un ratio F1:F2 : celui-ci est plus élevé lorsque l'implant est désactivé pour trois enfants et plus bas pour trois enfants. Les résultats des analyses perceptives par transcription sont meilleurs pour cinq des six enfants lorsque l'implant est activé et les jugements d'intelligibilité sont également meilleurs pour cinq enfants sur six lorsque l'implant est activé. Les résultats de cette étude montrent chez tous les enfants une différence entre les deux conditions d'utilisation de l'implant, mais les résultats des analyses acoustiques et perceptives sont trop hétérogènes pour pouvoir être exploités par les auteurs. Le groupe d'enfants est également relativement petit et hétérogène, et le nombre de productions collectées et analysées est également trop faible pour pouvoir établir clairement des corrélations entre caractéristiques acoustiques de ces productions et intelligibilité des enfants.

Les résultats des analyses acoustiques de l'étude de Horga et Liker (2006) indiquent que les enfants sourds profonds HA ont un espace F1-F2 plus réduit que les enfants CI et NH et l'étude d'intelligibilité des voyelles suit ce même schéma : les enfants CI et NH ont une meilleure intelligibilité que les enfants HA, sauf pour la voyelle /a/, pour laquelle les enfants HA et NH ont une meilleure intelligibilité que les enfants CI.

L'étude de Chuang et al. (2012) confirme les résultats des études de Horga et Liker (2006) et de Poissant et al. (2006), puisqu'elle met en évidence un espace vocalique plus petit chez les enfants CI lorsque les voyelles sont tenues, mais une taille d'espace vocalique similaire à celle des enfants NH lorsque les voyelles sont élicitées dans des phrases porteuses. L'étendue de F2 est plus petite chez les enfants CI, ce qui indique une réduction sur l'axe horizontal. Chuang et al. (2012) montrent une corrélation entre l'étendue de F2 entre /i/ et /u/ (voyelles tenues) et l'intelligibilité. Ils notent également une bonne intelligibilité des enfants CI, même si celle-ci reste moins bonne que celle des enfants NH.

Dans l'étude de Löfqvist et al. (2010), les résultats de l'analyse acoustique indiquent que l'espace vocalique est plus petit chez les adolescents CI que chez les adolescents NH. Les auteurs utilisent

deux méthodes différentes de calcul de la taille de l'espace vocalique. Ils mettent en évidence l'absence de corrélation significative ente la taille de l'espace vocalique et intelligibilité, contrairement aux études de Bradlow et al. (1996) et Horga et Liker (2006). Ils notent en outre une différence significative entre les deux groupes pour l'étendue de F1 mais pas de F2. Par ailleurs, les tests d'identification de voyelles de l'étude de Löfqvist et al. (2009) indiquent qu'il y a une confusion plus grande entre les voyelles testées chez les adolescents CI que chez les NH, avec une différence significative entre les deux groupes : la confusion entre voyelles lors du test d'identification de voyelles est plus grande si la taille de l'espace vocalique est plus petite et si les voyelles sont plus proches les unes des autres. L'une des limites de l'étude de Löfqvist et al. (2010) est qu'elle ne prend en compte que les voyelles dans les analyses acoustiques réalisées, et qu'aucun autre aspect de la production de parole n'est envisagé pour expliquer des problèmes d'intelligibilité.

Dans l'étude de Perrin et al. (1999), l'analyse objective montre que les enfants NH ont des formants plus hauts que les valeurs normales de référence (établies lors d'une étude précédente par les mêmes auteurs auprès de 20 enfants, appariés en âge et en sexe avec les enfants CI et NH de l'étude de Perrin et al., 1999), alors que les enfants CI ont des formants plus bas que ces mêmes valeurs. Les auteurs ont utilisé des valeurs « normales » de référence pour normaliser les valeurs acoustiques mais ces valeurs ne sont peut-être pas adaptées dans ce cas, et ne réduisent peut-être pas toute la variation liée notamment à l'âge des enfants. L'étude est une étude pilote, avec un nombre très réduit d'enfants, dont les âges et la durée d'utilisation de l'implant sont très hétérogènes. Les données de production sont également insuffisantes (une seule phrase lue) pour pouvoir établir avec certitude où se situent les difficultés rencontrées par les enfants CI et les différences avec le groupe d'enfants NH.

Van Lierde et al. (2005) mettent en évidence une intelligibilité plus grande des enfants CI par rapport aux enfants HA, ce qui est partiellement en contradiction avec les résultats de Miyamoto et al. (1997) et Svirsky (2000) : dans ces deux études, l'intelligibilité des enfants CI avant implantation est similaire à celle des enfants HA qui ont le niveau d'audition le plus bas, leur intelligibilité augmentant de façon constante à partir de l'implantation. Cependant, dans l'étude de Miyamoto et al. (1997), même après sept ans et demi d'utilisation de l'implant, l'intelligibilité des enfants CI n'a toujours pas atteint le niveau des enfants HA qui ont le niveau d'audition le plus élevé de l'étude, alors que pour Van Lierde et al. (2005), plusieurs années après implantation (3;4 ans en moyenne), l'intelligibilité des enfants CI atteint un niveau plus élevé que celle des enfants HA. Par ailleurs, les analyses objectives et perceptives de l'étude de Van Lierde et al. (2005) montrent que les difficultés des deux groupes sont localisées sur les productions de consonnes, plutôt que sur les voyelles, même si les enfants CI ont un inventaire de consonnes plus étendu que les enfants HA. Van Lierde et al. (2005) soulèvent également la difficulté de cibler les raisons et les facteurs qui conduisent à une plus grande intelligibilité des enfants CI par rapport aux enfants HA. Les problèmes rencontrés dans la production de certains sons (par exemple les consonnes) ne sont pas les mêmes : substitution d'une autre consonne par les enfants HA, et éloignement de la consonne-cible d'un seul trait (voisement, mode ou lieu d'articulation) pour les enfants CI. Pour Van Lierde et al. (2005), les omissions et substitutions de certains segments (voyelles et consonnes) par les enfants sont la cause la plus importante de perte d'intelligibilité, alors que si la différence par rapport à la cible est localisée sur un seul trait phonétique (ce qui correspond à des distorsions), les enfants ont une intelligibilité plus grande.

Les auteurs en concluent que les problèmes d'intelligibilité sont plus grands chez les enfants HA que chez les enfants CI, ce qui est probablement lié à un nombre plus important de problèmes rencontrés à la fois aux niveaux phonétiques et phonologiques.

Cette étude a par ailleurs quelques limites, notamment en ce qui concerne les groupes d'enfants qui sont assez hétérogènes, de par la taille de chaque groupe et les âges des enfants. On peut noter également que Van Lierde et al. (2005) n'essayent pas d'établir des corrélations entre caractéristiques de la parole et intelligibilité, mais envisagent l'intelligibilité comme l'un des éléments de la parole des enfants porteurs d'implants cochléaires qu'ils étudient. Les auteurs présentent donc des résultats d'intelligibilité mais n'expliquent pas ce qui provoque un déficit d'intelligibilité chez les enfants CI.

Si la grande majorité des études d'intelligibilité chez les enfants CI s'intéressent aux productions de voyelles, l'étude de Horga et Liker (2006) s'intéresse également aux productions de consonnes : dans les analyses de durées d'occlusion et de VOT, les enfants CI obtiennent des résultats similaires aux enfants HA, et moins bons que les enfants NH. Les enfants CI ont des résultats moins bons que les enfants NH en ce qui concerne la différence de voisement entre /t/ et /d/. Cette différence entre enfants CI, HA et NH se traduit également par une intelligibilité moins bonne des enfants CI et HA par rapport aux enfants NH. Horga et Liker (2006) proposent cette analyse à partir d'une comparaison de caractéristiques acoustiques de la production de parole et de l'intelligibilité (à partir d'un test d'identification de voyelles) sans tester la corrélation entre ces deux aspects de la parole des enfants CI.

2.2.2.2.Caractéristiques suprasegmentales et qualité de voix

Perrin et al. (1999) s'intéressent à des caractéristiques suprasegmentales de la production de parole. Ils trouvent que : la durée de la phrase est plus longue chez les enfants CI que chez les enfants NH ; l'intensité perçue est la mesure subjective qui varie le plus, sauf pour l'un des enfants CI ; les pauses sont plus nombreuses et plus longues pour les enfants CI, sauf pour deux d'entre eux, pour lesquels ces résultats sont comparables aux enfants NH ; les voix des enfants CI sont jugées moins agréables sauf la voix d'un des enfants. Les auteurs en concluent que les évaluations subjectives de caractéristiques de voix ne permettent que de dégager des tendances mais pas d'effet fort de l'implant sur l'intelligibilité et les caractéristiques étudiées. L'évaluation objective ne donne pas de différences très marquées, sauf pour la durée de production de la phrase, qui est nettement plus longue pour les enfants CI, ce qui serait une indication que les enfants CI parlent plus lentement que les enfants NH. Chuang et al. (2012) s'intéressent également au débit de parole (qui inclut les pauses), qui est significativement plus lent chez les enfants CI (plus de pauses et pauses plus longues), alors que le débit d'articulation (qui n'inclut pas les pauses) n'est pas significativement différent entre les deux groupes. Chuang et al. (2012) en concluent que les enfants CI ont une stratégie de segmentation des groupes de mots différente de celle des enfants NH, mais cette différence ne semble pas avoir d'effet sur l'intelligibilité.

L'étude de Van Lierde et al. (2005) s'intéresse à la nasalité. Les résultats montrent que la nasalité est moins forte chez les enfants CI que chez les enfants HA, dans les mesures objectives, mais ne se traduit pas par une différence au niveau perceptif.

Löhle et al. (1999) étudient à la fois les effets de caractéristiques individuelles des enfants (trois groupes d'enfants CI selon l'âge d'implantation et quatre groupes d'enfants HA) et les effets du rythme et de l'intonation sur l'intelligibilité. Le groupe des enfants ayant reçu l'implant le plus tôt

a une prosodie jugée naturelle après l'implantation, le groupe des enfants ayant les âges d'implantation intermédiaires est jugé comme ayant une F0 monotone mais une structure rythmique adéquate, alors que les enfants dont l'âge lors de l'implantation était le plus tardif ont une F0 monotone et ne sont pas capables de structurer rythmiquement leur parole. Dans cette étude, Löhle et al. (1999) n'établissent pas clairement le lien entre prosodie et intelligibilité, mais plutôt entre utilisation adéquate de la prosodie et âge d'implantation d'une part et intelligibilité et âge d'implantation d'autre part. Ainsi ils mettent en évidence l'effet de l'âge d'implantation sur l'utilisation de la prosodie et les conséquences de cette utilisation plus ou moins correcte sur l'intelligibilité de la parole des enfants CI.

2.3. Conclusion

L'ensemble des études présentées a montré qu'il existe des corrélats quantifiables de l'intelligibilité : caractéristiques des enfants et de leur utilisation de l'implant d'une part et caractéristiques de la production de parole d'autre part. L'étude de Bradlow et al. (1996) prend en compte ces deux types de corrélats et montre que l'intelligibilité est corrélée à la fois avec les caractéristiques des locuteurs typiques (adultes) et la taille de leur espace vocalique. Les résultats chez les enfants porteurs d'implants cochléaires sont beaucoup moins clairs et on constate une grande variété de résultats contradictoires dans les études présentées : l'intelligibilité est corrélée à la durée d'utilisation de l'implant dans toutes les études mais à certaines caractéristiques des locuteurs comme l'âge chronologique et l'âge d'implantation dans quelques études seulement. De même, l'intelligibilité n'est pas systématiquement corrélée aux caractéristiques acoustiques de la production de la parole, à l'exception du débit de parole et de la fréquence fondamentale qui ne sont analysées que dans deux études, ce qui ne permet pas de généralisation.

Comme pour les études sur la production de parole, cette revue de littérature a montré qu'il existe de grandes disparités entre les études sur l'intelligibilité de la parole. Ces disparités sont liées entre autres aux méthodes de jugements, aux populations d'enfants dont on juge l'intelligibilité. Notre étude expérimentale sur l'intelligibilité de la parole devra donc prendre en compte ces facteurs et nous proposerons une revue détaillée des différentes méthodologies utilisées par ces études pour justifier nos choix expérimentaux. Cette revue méthodologique sera proposée en détail dans le chapitre 6, ci-après.

3. Notre projet

La revue d'état de l'art que nous avons proposée dans ce chapitre a mis en évidence qu'il existe dans la littérature un grand nombre d'études centrées sur la caractérisation du bénéfice immédiat de l'implant sur la production de parole des enfants qui le reçoivent, ces études se concentrant sur les premiers temps d'utilisation de l'implant, ou sur des suivis longitudinaux d'enfants avant et après implantation. Par contre, il existe à l'heure actuelle très peu d'études à long terme de la parole des enfants sourds, utilisateurs d'implant cochléaire, plusieurs années après l'implantation. Les études, disponibles dans la littérature font généralement le choix d'un seul mode d'élicitation de la parole, et leurs résultats sont présentés comme une description de la parole des enfants sourds porteurs d'implant cochléaire à un moment précis ou résultant d'une évolution consécutive à l'implantation.

Comme nous l'avons vu aux paragraphes §1 et §2 de ce chapitre, à l'heure actuelle, mise à part une étude combinant description acoustique de segments produits par des enfants CI de 7 à 14 ans, et identification de ces segments par des locuteurs naïfs (Legendre, 2014), il n'existe à notre connaissance pas d'études des effets à long-terme de l'implantation cochléaire sur la production de parole chez des enfants francophones, prenant en compte les caractéristiques acoustiques de leur production. Notre projet a donc été conçu dans le but de proposer une caractérisation de la production de parole et de l'intelligibilité d'enfants francophones porteurs d'implants cochléaires, de façon à comprendre quelles sont les difficultés persistantes de ces enfants, même après plusieurs années, à produire des contrastes phonologiques du français, en ayant un accès partiel à l'audition.

3.1. Problématique et hypothèses de l'étude de production

3.1.1. Problématique

La première partie de notre étude expérimentale est consacrée à la caractérisation segmentale de la production de parole d'enfants sourds, porteurs d'implant cochléaire depuis plusieurs années, par rapport à celle d'enfants typiques et normo-entendants francophones, du même âge.

Lorsque nous avons choisi de nous intéresser aux caractéristiques segmentales de la production de parole chez des enfants francophones, nous avons été confrontées au manque de descriptions acoustiques de la parole d'enfants francophones et normo-entendants au-delà de l'âge de trois ans, disponibles dans la littérature : l'effet de l'âge chronologique en particulier est très peu pris en compte dans les études développementales chez l'enfant normo-entendant. L'un des enjeux de notre travail est donc de créer un corpus de référence sur le développement tardif de la production de parole des enfants typiques. Ce travail nécessite d'isoler la part de variation due aux effets de la croissance de l'enfant et de son conduit vocal sur la production de parole afin d'identifier la variation liée au développement du contrôle articulatoire. Pour cette raison, nous devrons aborder la question cruciale de la normalisation des données acoustiques.

La constitution d'un corpus de production d'enfants nécessite une méthodologie clairement définie et contrôlée, de façon à pouvoir comparer des productions enregistrées dans les mêmes conditions pour tous les enfants. Nous avons donc conçu une étude de production en trois étapes : 1) une tâche d'élicitation de mots avec d'après un modèle audio adulte pré-enregistré et une image, 2) une tâche de dénomination des mêmes mots d'après la même image sans modèle audio et 3) une tâche de narration, destinée à recueillir des productions de parole spontanée. La

comparaison des deux premières tâches nous permettent de comprendre le rôle de l'input auditif immédiat sur la production de parole.

Notre but étant de caractériser acoustiquement la production de parole de deux groupes d'enfants, nous avons choisi de mesurer les trois premiers formants des voyelles, le VOT des occlusives, quatre moments spectraux des fricatives et la transition du deuxième formant dans des séquences occlusive-voyelle (en utilisant la pente de l'équation de locus) pour étudier la coarticulation anticipatoire.

Nous avons choisi dans cette thèse, de nous intéresser en particulier 1) à l'effet de l'audition sur la production de parole des enfants de 5 à 11 ans, en comparant un groupe d'enfants CI à un groupe d'enfants NH appariés en âge chronologique, 2) à l'effet de l'âge chronologique sur la production de parole par les enfants NH, 3) à l'effet de l'âge à l'implantation et de la durée d'utilisation de l'implant, 4) à l'influence des capacités perceptives sur la production de parole à travers deux situations d'élicitation de parole (avec ou sans modèle audio disponible), et 5) à l'effet de la complexité articulatoire du mot sur la production de parole en utilisant des mots de deux niveaux de complexité articulatoire (élicitation de mots simples ou complexes). Nous cherchons à caractériser la parole des enfants sourds, porteurs d'implant cochléaire en prenant en compte leurs difficultés d'imitation liées à une audition partielle et à la complexité lexicale de l'input proposé, qui peut perturber leur production de parole. Ces deux aspects de notre travail vont donc nous permettre de proposer une description de la réalisation de contrastes phonologiques du français en prenant en compte ces deux aspects de la production de parole, et de comprendre quels sont les obstacles dans le processus de stabilisation de ces contrastes phonologiques rencontrés par des enfants qui ont eu un accès plus tardif à l'oral.

3.1.2. Hypothèses

Notre problématique nous a amenées à formuler plusieurs hypothèses pour pouvoir caractériser les différents facteurs qui affectent les caractéristiques acoustiques de la parole produite par les deux groupes d'enfants. Ces hypothèses sont présentées dans les paragraphes suivants.

3.1.2.1. Effet du groupe sur la réalisation des segments et de la coarticulation par les enfants CI et NH

H0 : nous n'observons pas d'effet du groupe sur les productions de parole des enfants CI et NH, ce qui indique qu'une audition partielle ne joue aucun rôle dans l'acquisition des contrastes phonologiques considérés ou qu'une durée d'utilisation de l'implant suffisante a permis de surmonter ces difficultés.

H1 : nous observons un effet du groupe sur toutes les mesures acoustiques réalisées, ce qui indique que l'audition joue un rôle majeur dans l'acquisition des contrastes phonologiques de la langue maternelle.

H2 : nous observons un effet du groupe sur certaines mesures acoustiques, ce qui indique que l'acquisition des contrastes phonologiques est liée à l'audition et qu'elle se fait selon des étapes successives.

3.1.2.2. Effet de l'âge chronologique sur la réalisation des segments et de la coarticulation par les enfants NH

H0 : nous n'observons pas d'effet de l'âge chronologique sur les productions de parole des enfants NH, ce qui indique une maîtrise des contrastes phonologiques étudiés dans le cadre de notre étude.

H1 : nous observons un effet de l'âge sur toutes les mesures acoustiques réalisées, ce qui indique que l'acquisition des contrastes phonologiques n'est pas terminée chez des enfants de 5 à 11 ans.

H2 : nous observons un effet de l'âge sur certaines mesures acoustiques réalisées, ce qui nous laisserait penser que l'acquisition du vocalisme, du voisement, du lieu d'articulation des fricatives et de la coarticulation se fait à des rythmes différents.

3.1.2.3. Effet de l'âge à l'implantation et de la durée d'utilisation de l'implant sur la réalisation des segments et de la coarticulation par les enfants CI

H0 : nous n'observons aucun effet de l'âge à l'implantation ou de la durée d'utilisation de l'implant, indiquant un bénéfice limité de l'implantation précoce et un processus d'acquisition moins lié à l'audition qu'à des caractéristiques du développement moteur des enfants CI

H1 : nous observons un effet de l'âge à l'implantation ou de la durée d'utilisation de l'implant, argument pour une implantation précoce et pour une amélioration de la production avec l'implant, même à long terme.

3.1.2.4. Effet de l'input auditif immédiat sur la réalisation des segments (consonnes et voyelles) et de la coarticulation CV

H0 : nous ne constatons pas de différence entre les deux tâches (répétition de mots d'après un modèle audio et dénomination des mêmes mots sans modèle audio) sur la réalisation des segments, chez les deux groupes, ce qui indique que l'input auditif immédiat n'a donc pas d'effet sur la réalisation des segments (consonnes et voyelles)

H1 : nous constatons une variation plus importante dans la réalisation des segments lors de la tâche de dénomination, sans modèle audio, ce qui indique que la représentation phonologique ou motrice des catégories vocaliques ou consonantiques n'est pas encore stabilisée chez les enfants de 5 à 11 ans.

H2 : nous constatons une moindre variation entre les deux tâches dans la réalisation des segments et de la coarticulation CV chez les enfants NH que chez les enfants CI, ce qui indique que le processus d'encodage des routines motrices de la langue maternelle est en voie d'achèvement chez les enfants NH, alors que ce processus est toujours en cours chez les enfants CI et que ce processus est influencé par l'information auditive immédiatement disponible.

3.1.2.5. Effet de la complexité du mot sur la réalisation des segments (consonnes et voyelles)

H0 : nous ne constatons pas de différence entre la réalisation des segments (consonnes et voyelles) dans des mots de deux niveaux différents de complexité du mot, ce qui indique que la complexité de la planification articulatoire et de la structure phonologique n'a pas d'effet sur la réalisation des segments (consonnes et voyelles).

H1 : nous constatons une variation plus importante dans la réalisation des segments (consonnes et voyelles) des mots plus complexes, ce qui indique que le contrôle articulatoire ou la représentation phonologique des catégories segmentales sont en cours d'acquisition et que la complexité de la planification motrice influence la production segmentale.

H2 : nous constatons une moindre variation entre les deux niveaux de complexité du mot dans la réalisation des segments chez les enfants NH que chez les enfants CI, ce qui indique que le processus d'acquisition de la phonologie de la langue maternelle est en voie d'achèvement chez les enfants NH, alors que ce processus est toujours en cours chez les enfants CI et que ce processus est influencé par la complexité motrice.

3.2. Problématique et hypothèses de l'étude d'intelligibilité

Les résultats de notre étude de production (chapitre 5) s'intéresseront aux différences entre enfants CI et NH dans la réalisation de certains contrastes phonologiques ainsi qu'aux effets de plusieurs variables sur la production de parole par les enfants des deux groupes. Le chapitre 6 présentera le volet perceptif de notre thèse, consacré à la caractérisation de l'intelligibilité chez les deux groupes d'enfants CI et NH.

La revue bibliographique présentée au paragraphe §2 de ce chapitre a permis de mettre en évidence une grande disparité de résultats dans les études d'intelligibilité réalisées par différents auteurs dans des langues très différentes. Notre étude d'intelligibilité a pour but de confronter les résultats de notre étude de production obtenus lors d'une tâche de dénomination de mots, à l'intelligibilité de la parole continue enregistrée lors d'une tâche de narration.

Nous cherchons à étudier 1) les différences d'intelligibilité des deux groupes d'enfants NH et CI 2) l'effet de l'âge chronologique sur l'intelligibilité de la parole des enfants NH, 3) les effets de l'âge à l'implantation et de la durée d'utilisation de l'implant sur l'intelligibilité de la parole des enfants CI et 4) les liens entre difficultés des enfants à réaliser des contrastes phonologiques et intelligibilité de la parole.

Pour répondre à ces questions, nous faisons les hypothèses suivantes :

1) il y a des différences d'intelligibilité entre les enfants NH et CI,
2) l'âge chronologique n'a pas d'effet sur l'intelligibilité de la parole des enfants NH,
3) l'âge à l'implantation et la durée d'utilisation de l'implant cochléaire ont un effet sur l'intelligibilité des enfants CI,
4) certaines différences inter-groupes en production notamment au niveau des fricatives et de la coarticulation peuvent expliquer les différences inter-groupes au niveau de l'intelligibilité.

DEUXIÈME PARTIE :
Production de parole (1ère étude expérimentale)

CHAPITRE 4 : Étude de production (1ère étude expérimentale) : méthodologie

La revue de littérature proposée dans la première partie de cette thèse (chapitres 1, 2 et 3) a présenté les enjeux de notre étude sur l'acquisition de la parole par les enfants sourds, porteurs d'implant cochléaire. Le chapitre 3 ci-dessus, a présenté en particulier l'évolution des caractéristiques segmentales par des populations d'enfants après implantation et les conséquences de cette implantation sur l'intelligibilité des enfants et le manque de données disponibles sur la production de parole par des enfants porteurs d'implant cochléaires, francophones. La deuxième partie de ce travail (chapitres 4 et 5) présente l'étude expérimentale que nous avons élaborée, pour caractériser acoustiquement la production de parole d'enfants francophones, porteurs d'implant cochléaire, plusieurs années après implantation. Le chapitre 4 présente la méthodologie de notre étude et de nos analyses acoustiques et statistiques. Après avoir présenté notre protocole expérimental, nous proposons une description articulatoire et acoustique de la production de parole en français (voyelles, consonnes et coarticulation), ainsi que l'utilisation qui est faite de ces caractéristiques acoustiques dans la littérature sur le développement phonologique. Le chapitre 5 sera ensuite consacré à la présentation des résultats de notre étude.

1. Recueil des données

1.1. Participants

La population d'enfants que nous avons enregistrés dans le cadre de cette étude se compose d'un groupe contrôle d'enfants normo-entendants et d'un groupe expérimental d'enfants sourds, porteurs d'implants cochléaires uni- ou bilatéraux (figure 4.1 ci-dessous). Les deux groupes sont appariés en âge chronologique (t(28)=-1,031, p=0.3117).

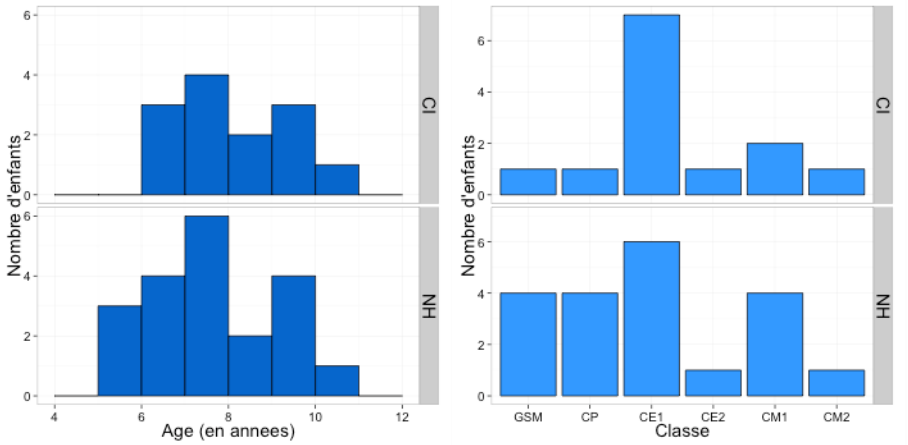


Figure 4.1 : âge (gauche) et classe (droite) des enfants (groupes NH et CI) ayant participé à l'étude de production

Les données et enregistrements ont été anonymisés : à chaque enfant correspond un code NH/CI suivi d'un nombre (NH pour normo-entendant, CI pour porteurs d'implant cochléaire). Les nombres sont attribués selon l'ordre de la réalisation des enregistrements. Cette étude fait suite à une étude pilote (Scarbel (2012), Scarbel et al. (2013)), et l'attribution des chiffres se fait à la suite de ceux des enregistrements de cette étude pilote.

1.1.1. Enfants normo-entendants

1.1.1.1. Recrutement des enfants normo-entendants

Les enfants normo-entendants ont été recrutés par plusieurs biais : bouche à oreille, listes de diffusion du laboratoire Gipsa-Lab et site internet du RISC[11]. Un formulaire d'information et de consentement a été remis au parent présent qui l'a signé. Le formulaire de consentement vierge est présenté en Annexe 1. Il a été également expliqué à chaque enfant le déroulement des enregistrements, en précisant qu'il pouvait demander à y mettre fin à tout moment, s'il le souhaitait.

1.1.1.2. Age et origine géographique des enfants normo-entendants

Les enfants normo-entendants sont tous francophones et vivaient en Isère au moment des enregistrements. Le groupe contrôle se compose de 20 enfants (11 filles et 9 garçons), âgés de 5;7 ans à 10;6 ans : âge chronologique moyen 7;8 ans (Ecart-type : 1;6 ans).

1.1.1.3. Audition des enfants normo-entendants

Afin de détecter d'éventuels problèmes d'audition, tous les enfants normo-entendants ont passé un test audiométrique (audiométrie tonale) avec un Audiomètre Electronica 9910. Tous ont obtenu des résultats témoignant d'une audition normale, aux deux oreilles.

1.1.1.4. Développement langagier des enfants normo-entendants

De la même manière, pour s'assurer qu'ils ont un développement langagier typique et ne présentent pas de pathologie du langage, tout le groupe des enfants normo-entendants a passé deux épreuves du bilan de langage oral ELO (Khomsi (2001)), utilisé fréquemment en orthophonie : l'épreuve 1, « lexique en réception » (LexR) qui teste les connaissances du vocabulaire des enfants et l'épreuve 5, « production d'énoncés » qui teste les compétences morphosyntaxiques des enfants. Il s'agit d'un test calibré par classe, et tous les enfants ont obtenu des résultats au moins égaux au niveau correspondant à leur niveau scolaire.

1.1.2. Enfants sourds, porteurs d'implant cochléaire

1.1.2.1. Recrutement des enfants sourds, porteurs d'implant cochléaire

Les enfants porteurs d'implant cochléaire ont été recrutés par l'intermédiaire du service ORL du Pr. S. Schmerber (avec l'aide de A.-M. Fluttaz et C. Rebière, orthophonistes) au CHU de Grenoble, et du service ORL du Pr. E. Truy de l'Hôpital Edouard Herriot de Lyon (avec l'aide de F. Dambelin, coordinatrice au centre d'implantation). Un formulaire d'information et de consentement, a été remis au parent présent lors de l'étude, qui l'a signé, ainsi qu'un questionnaire destiné à recueillir des informations sur les pratiques langagières des enfants participant à l'étude. Il a été également expliqué à chaque enfant le déroulement des enregistrements, en précisant qu'il pouvait demander à y mettre fin à tout moment, s'il le souhaitait. Deux enfants porteurs d'implants ont ainsi commencé à participer et n'ont pas poursuivi l'étude jusqu'au bout, car ils avaient beaucoup de difficultés à rester concentrés et à répéter les mots de la première tâche. Ces deux enfants ne sont pas inclus dans les analyses de ce chapitre et ne font pas partie des 13 enfants du groupe expérimental. Le formulaire de consentement vierge est présenté en Annexe 2 et un questionnaire vierge est présenté en Annexe 3. Notre étude porte sur les conséquences de l'implantation à long terme dans le cadre d'une surdité prélinguistique. Les critères d'inclusion de notre étude étaient donc une surdité survenue avant l'âge de deux ans, et une durée d'utilisation de l'implant d'au moins deux ans. De façon générale, il s'agit d'une population d'enfants réduite

[11] RISC : Relais d'Information sur les Sciences de la Cognition, http://expesciences.risc.cnrs.fr

et difficile d'accès, nous avons donc été amenés à assouplir ces critères d'inclusion, et la population d'enfants dont nous étudions les productions est plus hétérogène que ce que nous aurions souhaité initialement. Nous devrons donc tenir compte de cette hétérogénéité lors des analyses que nous proposerons.

1.1.2.2.Age et origine géographique des enfants sourds, porteurs d'implant cochléaire

Le groupe expérimental des enfants sourds, porteurs d'implant cochléaire se compose de 13 enfants (6 filles et 7 garçons), âgés de 6;6 ans à 10;7 ans : âge chronologique moyen 8;2 ans (1;3 ans). L'âge moyen à l'implantation est de 3;2 ans (1;9 ans) et la durée moyenne d'utilisation de l'implant est de 5;3 ans (2;3 ans). Tous les enfants habitaient en région Rhône-Alpes au moment des enregistrements (Isère (5), Rhône (5), Saône-et-Loire (2) et Haute-Savoie (1)). Ils sont francophones, se sont déclarés monolingues et utilisent le français dans le cadre familial, à l'école et pour leur rééducation orthophonique. Les âges, domiciles, langues maternelles des enfants porteurs d'implants cochléaires et langues maternelles de leurs parents sont données dans la table 4.1 ci-dessous.

Enfant	Age chronologique (années;mois)	Age à l'implantation (années;mois)	Durée d'utilisation de l'implant (années;mois)	Sexe	Classe	Domicile au moment de l'enregistrement	L1 de l'enfant	Langue parlée à la maison	L1 de la mère	L1 du père
CI6	8;0	5;10	2;2	M	CE1	38	FR	FR	FR	FR
CI7	7;11	1;7	6;4	M	CE1	38	FR	FR	IT	IT
CI10	7;8	4;6	3;2	F	CE1	74	FR	FR	FR	FR
CI11	9;1	2;2	7;8	M	CM1	69	FR	FR+AR	AR	AR+FR
CI12	7;11	4;4	3;7	F	CE1	71	FR	FR	FR	FR
CI13	6;6	2;1	4;5	M	GSM	38	FR	FR	FR	FR
CI14	6;10	1;6	5;4	F	CP	69	FR	FR	AR+FR	AR+FR
CI16	9;4	1;11	7;5	M	CE1	71	FR	FR	FR	FR
CI17	9;2	6;6	2;8	F	CE2	38	FR	FR	FR	FR
CI19	6;8	3;11	2;9	M	CE1	38	FR	FR	FR	FR
CI20	10;7	1;6	9;1	F	CM2	69	FR	FR	FR	FR
CI21	7;1	1;1	6;0	F	CE1	69	FR	FR	FR	ESP
CI22	8;1	3;9	7;7	M	CM1	69	FR	FR	FR	FR
Moyenne (écart-type)	8;2(1;3)	3;2 (1;9)	5;3 (2;3)							

Table 4.1 : âge chronologique, âge à l'implantation, durée d'utilisation de l'implant, sexe, classe, domicile, L1 des enfants, langues parlée à la maison et L1 des parents des enfants

1.1.2.3.Audition des enfants sourds, porteurs d'implant cochléaire

La table 4.2 ci-dessous présente les caractéristiques des implants (nombre et type) utilisés par les enfants, ainsi que les caractéristiques de la surdité avant implantation et audition post-implantation ainsi que les éventuels troubles associés à la surdité des enfants participant à l'étude.

Enfant	Nombre d'implants	Aides auditives avant implantation	Type d'implant	Type de surdité avant implantation	Audition post implantation
CI6	1	oui	Cochlear	Sévère OD Profonde OG	seuils autour de 30 dB
CI7	2	oui x2	Cochlear CP 810	Profonde bilatérale	seuils autour de 20 dB OD 20/25 dB OG
CI10	2	oui x2	Cochlear CP810 CP910	Sévère OD Profonde OG	seuils autour de 20/25 dB
CI11	1	oui CMPS	Cochlear CP900	Profonde	seuils autour de 25 dB
CI12	1	oui prothèse + micro HF	Cochlear CP810 CP900	Sévère	seuils autour de 20 dB OD
CI13	2	oui	Cochlear CP810	Profonde	seuils autour de 20 dB
CI14	2	oui	Cochlear CP 810 CP 200 AG	Profonde	seuils autour de 25 dB OD 30 dB OG
CI16	1	oui	Neurelec Saphyr S-P	Profonde	seuils autour de 20 dB
CI17	1	oui	Cochlear Nucleus	Sévère OD Moyenne OG	seuils autour de 20 dB OD surdité moyenne OG
CI19	1	oui x2	Neurelec Saphyr	Profonde	seuils autour de 20 dB
CI20	2	oui	CP 810 et CP 910	Profonde	seuils autour de 20 OD 25 dB OG
CI21	2	oui x2	Cochlear CP810	Sévère à profonde	seuils autour de 25 OG 20 db OD
CI22	1	oui	Cochlear CP 810	Profonde	seuils autour de 35 dB

Table 4.2 : nombre et types d'implants utilisés par les enfants, degrés de surdité avant implantation, et niveau d'audition après implantation (données récoltées par Corinne Aubert (Aubert, 2016), stagiaire au Gipsa-Lab en 2015-2016 dans le cadre de son stage et mémoire d'orthophonie) – OD : oreille droite, OG : oreille gauche

1.1.2.4. Modes de communication des enfants sourds, porteurs d'implant cochléaire

La Table 4.3 ci-dessous montre l'utilisation de la LSF, du LPC et de l'oral par les enfants de notre étude, avant et après implantation. Il s'agit des réponses au questionnaire que les parents ont rempli.

	Avant implantation			Après implantation		
	LSF	LPC	Oral	LSF	LPC	Oral
CI6	✗	✓	✓	✗	✓	✓
CI7	✗	✓	✓	✗	✓	✓
CI10	✗	✗	✓	✗	✓	✓
CI11	✗	✗	✗	✗	✓	✓
CI12	✓	✗	✓	✓	✓	✓
CI13	✓	✗	✓	✓	✓	✓
CI14	✗	✗	✓	✗	✗	✓
CI16	✓	✗	✓	✓	✓	✓
CI17	✗	✗	✓	✗	✗	✓
CI19	✓	✗	✓	✓	✗	✓
CI20	✗	✓	✓ puis ✗	✗	✗	✓
CI21	✓	✓	✓	✗	✓ (presque ✗)	✓
CI22	✓	✗	✓	✓	✗	✓
Total	6	4	11-12	5	7-8	13

Table 4.3 : Utilisation de LSF (Langue des Signes Française), LPC (Langage Parlé Complété) et Oral par les enfants sourds, porteurs d'implant cochléaire, avant et après implantation (le symbole ✓ signifie oui, et le symbole ✗ signifie non).

La majorité des enfants utilisaient l'oral avant l'implantation, seuls deux ne parlaient pas. Huit enfants sur treize utilisaient un mode de communication par signes/gestes (LSF ou LPC) en plus de l'oral, un seul enfant utilisait à la fois LSF et LPC en plus de l'oral, trois enfants n'utilisaient que l'oral et un enfant n'utilisait ni oral ni gestes/signes. On note que plusieurs années après l'implantation (ce qui correspond au moment où les enfants ont participé à notre étude), tous utilisent l'oral. Tous les enfants qui utilisaient la LSF avant implantation continuent à l'utiliser,

sauf un et il est intéressant de noter que c'est le seul enfant qui utilisait à la fois LSF et LPC avant l'implantation, et qu'il n'utilise désormais presque exclusivement que l'oral. Par ailleurs, trois des enfants qui n'utilisaient pas le LPC avant implantation l'utilisent désormais, et deux enfants n'ont pas changé de mode de communication après l'implantation. Un seul enfant a cessé d'utiliser le LPC après l'implantation. L'utilisation de LSF se fait majoritairement dans le contexte familial (pour six enfants sur sept avant implantation et cinq sur cinq après implantation), mais également à l'école (pour trois enfants sur sept avant implantation et trois enfants sur cinq après implantation). Le LPC est utilisé majoritairement avec la famille (pour cinq enfants sur cinq avant implantation et pour cinq enfants sur huit après implantation), et à l'école (pour un enfant sur cinq avant implantation mais quatre enfants sur huit après implantation). Ces informations nous permettent d'avoir quelques indications sur les modes de communication auxquels les enfants de notre étude ont eu recours avant l'implantation et l'évolution de ces modes de communication après l'implantation. Cependant, comme notre étude porte sur une cohorte d'enfants porteurs d'implant assez restreinte et qu'il n'existe pas de profils-types d'évolution de l'utilisation de chaque mode de communication, nous proposons ces informations à titre indicatif seulement, et nous ne pourrons pas inclure ces facteurs dans nos analyses de la production de parole et de l'intelligibilité des enfants. Nous signalons également que les chiffres obtenus pour ce groupe d'enfants ne peuvent en aucun cas être généralisés à l'ensemble des enfants porteurs d'implant cochléaire.

1.2. Protocole

1.2.1. Tâches

1.2.1.1. Motivation du choix des tâches

Notre problématique a motivé le choix des corpus que nous avons constitués et de la méthode que nous avons utilisée pour les constituer. Dans un chapitre consacré à la constitution de corpus oraux dans le cadre d'études en phonétique et phonologie, Delais-Roussarie & Yoo (2014) proposent une distinction entre « données construites » (« experimental data ») et « données brutes ou produites naturellement » (« raw or naturally produced data ») :

> « The distinction made between experimental data and raw data relies not only on the way the data have been collected - experimental data being produced according to a clear experimental protocol, and raw data in a very natural communicative setting - but also on the basis of the criteria that underlie their selection. In the case of experimental data, the criteria used to design the corpus and gather the data are internal, and almost always linked to the linguistic facts under investigation. By contrast, in larger corpora, the criteria used to gather data are mostly external and independent of any specific research question. »[12]

[12] « La distinction faite entre données construites et données brutes tient non seulement à la façon dont les données ont été collectées – les données construites sont produites selon un protocole expérimental clair et les données brutes dans une situation de communication très naturelle – mais également à la base des critères qui motivent leur sélection. Dans le cas des données construites, les critères utilisés pour élaborer le corpus et collecter les données sont internes, et presque toujours liés aux faits linguistiques étudiés. En revanche, dans le cas de gros corpus, les critères utilisés pour collecter les données sont le plus souvent externes et indépendants de toute question de recherche. » (notre traduction)

Le premier type de données dites « construites » résulte d'un protocole strictement défini qui permet de réaliser des enregistrements comparables pour tous les locuteurs d'une étude, puisque dans ce cadre, chaque locuteur est soumis aux mêmes tâches et aux mêmes conditions d'enregistrement, alors que le deuxième type de données, dites « brutes ou produites naturellement », a l'avantage de s'approcher le plus possible de la production naturelle de la parole par le locuteur. Nous avons fait le choix d'enregistrer des données « construites » : les trois tâches (répétition, dénomination et narration) que nous avons utilisées nous permettent d'une part d'enregistrer des productions comparables chez tous les enfants et d'enregistrer des sons de parole sélectionnés pour rendre compte de plusieurs contrastes phonologiques du français, mais également d'enregistrer des productions assimilables à des productions de parole semi-spontanée, que nous allons ensuite utiliser pour notre étude d'intelligibilité de la parole des enfants sourds, porteurs d'implant cochléaire (voir chapitre 5, ci-après).

Cette étude de production a pour but d'étudier les productions des enfants porteurs d'implant cochléaire plusieurs années après implantation, pour comprendre 1) s'il reste des différences significatives entre les groupes CI et NH, plusieurs années après l'implantation des enfants CI, 2) comment les enfants utilisent le retour auditif pour produire des sons de parole 3) si la complexité du mot (nous considérons que les mots sont complexes s'ils présentent une complexité articulatoire résultant de leur longueur ou de la présence de clusters consonantiques) joue un rôle dans la production de ces mêmes sons de parole et 4) quels facteurs influencent l'intelligibilité de la parole des enfants. Dans cette optique, nous avons choisi d'enregistrer des mots contenant la plupart des sons du français (voyelles et consonnes) dans deux tâches (répétition des mots d'après un modèle audio adulte et dénomination des mêmes mots sans modèle) pour tenter de répondre à la première question, d'enregistrer des mots de deux niveaux de complexité articulatoire dans la seule tâche de répétition d'après un modèle audio adulte pour répondre à la deuxième question et d'enregistrer des productions de parole semi-spontanée, que nous soumettons ensuite au jugement d'auditeurs francophones, pour répondre à la troisième question.

Nous avons fait le choix de ne pas placer les mots dans des phrases porteuses mais de les enregistrer en isolation pour que la tâche soit la moins contraignante et répétitive possible pour des enfants assez jeunes. En effet, utiliser une même phrase porteuse nous semblait compliquer une tâche qui se devait avant tout d'être simple et rapide à effectuer pour éviter une lassitude et un arrêt prématuré de la tâche par l'enfant. Nous pouvons cependant a posteriori regretter justement d'avoir élaboré un protocole dont les tâches sont trop simples et ne permettent pas de mettre en évidence des différences très marquées entre les productions de mots isolés chez les deux groupes d'enfants.

1.2.1.2. Répétition

La première tâche est une tâche de répétition de mots, en isolation. Pour cette tâche de répétition, l'enfant était placé devant un ordinateur portable, sur lequel était présenté oralement un mot pré-enregistré accompagné d'une image qui l'illustrait et qu'il devait répéter. Les images que nous avons utilisées sont des photos ou des dessins qui représentent les objets ou personnes qui illustrent les mots du corpus, trouvés sur internet (à partir d'une recherche sur Google) et libres de droits. Chaque mot est présenté deux fois, en ordre aléatoire, avec une pause après la première série de 65 mots.

La tâche de répétition nous permet entre autres d'étudier l'effet de la complexité du mot sur la réalisation de consonnes et voyelles, nous avons donc deux listes de mots, classés par degré de complexité (mots simples et complexes) qui vont nous permettre de comparer la réalisation de segments dans des mots simples et des mots complexes. En outre, nous avons choisi de ne pas étudier l'effet de la complexité du mot sur la réalisation de la coarticulation, nous n'avons donc pas sélectionné de mots dits « complexes » pour en étudier la coarticulation.

1.2.1.3.Dénomination

Tout comme la tâche de répétition, la tâche de dénomination est une tâche d'élicitation de mots en isolation à partir d'images, mais cette fois, sans modèle audio. Les mots utilisés pour cette tâche ont tous été utilisés lors de la tâche de répétition, avec les mêmes images. La tâche de dénomination est réalisée immédiatement après la tâche de répétition qui lui sert donc d'entrainement. La comparaison des mots produits lors de la tâche de répétition et de la tâche de dénomination nous permet d'étudier l'effet de l'input auditif immédiat sur la réalisation des segments et de la coarticulation. Cette tâche est réalisée à partir des 35 mots de la liste des mots dits « simples » de la tâche de répétition.

Comme pour la tâche de répétition, l'enfant est placé devant un ordinateur portable, sur lequel l'image correspondant au mot-cible lui est présentée, mais cette fois sans modèle audio. Chaque mot est présenté deux fois, en ordre aléatoire, avec une pause après la première série de 35 mots.

Nous souhaitons étudier l'effet de l'input auditif immédiat sur la réalisation segmentale ainsi que sur la réalisation de la coarticulation. Dans cette optique, nous utiliserons la seule liste de mots simples dans la tâche de répétition et la tâche de dénomination, pour étudier la réalisation des segments.

1.2.1.4.Narration

La troisième tâche de notre étude expérimentale de production est une tâche de narration. Notre but étant de recueillir des échantillons de parole représentatifs des productions habituelles des enfants, destinés à être utilisés pour construire les stimuli de notre étude sur l'intelligibilité de la parole, nous avons choisi de leur proposer de visionner un dessin animé (« Hatch Up Your Troubles », Hanna & Barbera (1949)), et de leur demander de raconter à l'expérimentatrice présente tout ce qu'ils avaient vu dans ce dessin animé. Plusieurs raisons ont motivé le choix de cette tâche. En effet, enregistrer la parole d'enfants présente plusieurs difficultés exposées par Rose (2014), ce qui nécessite un protocole expérimental particulier, lié au type d'analyses à réaliser. Notre choix d'une tâche de narration permet de collecter des données de parole semi-spontanée, similaires et comparables chez l'ensemble des enfants puisque la production de parole est guidée. L'intérêt est que le lexique employé va être similaire d'un enfant à l'autre, ce qui permet de recueillir des énoncés représentatifs de la production de parole par chaque enfant, mais cette tâche ne permet pas de rendre compte de toute l'étendue des productions des enfants en situation de production de parole naturelle et habituelle.

Le dessin animé était divisé en huit séquences d'une vingtaine de secondes chacune. Après chaque séquence visionnée par l'enfant, celui-ci devait raconter à l'expérimentatrice ce qu'il venait de voir, et elle lui pouvait également lui poser des questions pour l'encourager à donner le plus de détails possible.

1.2.2. Corpus

1.2.2.1. Critères de choix des mots utilisés dans les tâches de répétition et de dénomination

Les mots ont été choisis sur plusieurs critères : il s'agit d'une liste de mots mono- ou plurisyllabiques qui comportent tous une séquence-cible CV ou CVC initiale. Les sons de parole cibles sont la consonne initiale de mot et la voyelle de cette séquence initiale.

Le choix d'utiliser des mots plutôt que des logatomes ou des pseudo-mots est quant à lui motivé par notre problématique liée à la complexité lexicale. En effet, le choix de logatomes ou de pseudo-mots qui est un choix courant dans la littérature (voir par exemple les études sur les voyelles produites par des enfants CI de Bharadwaj et al. (2007), Löfqvist et al. (2010), Ryalls et al. (2003), Verhoeven et al. (2016)), a l'avantage de contrôler au maximum le contexte d'élicitation des segments (contexte consonantique pour les voyelles et contexte vocalique pour les consonnes), mais ne permet pas de prendre en compte la dimension lexicale de la parole produite. Cependant, nous souhaitions obtenir des productions sans modèle lors de la tâche de dénomination. Pour cette raison, nous avons choisi d'utiliser des mots plutôt que des logatomes ou des syllabes car nous souhaitions que les tâches soient faciles à comprendre par de jeunes enfants, et nous ne souhaitions pas leur imposer la tâche cognitive d'apprendre des pseudo-mots. Nous avons donc choisi des mots du français en contrôlant le contexte vocalique droit immédiat pour l'étude des consonnes et le contexte consonantique gauche immédiat pour l'étude des voyelles.

Notre choix de placer les sons et séquences-cibles dans des syllabes en position initiale de mot est quant à lui motivé par des phénomènes de renforcement des segments et des syllabes en position initiale de groupe accentuel. En effet, ces phénomènes ont des conséquences sur l'articulation des segments et donc sur les caractéristiques acoustiques de ceux-ci (voir notamment Fougeron & Keating (1997) pour une description du renforcement des consonnes en anglais, Keating et al. (2004) pour une étude comparée du renforcement des consonnes en anglais, français, coréen et taïwanais, ou Georgeton & Fougeron (2014) pour une description du renforcement des voyelles en français). Comme l'indiquent Fougeron & Keating (1997), il est nécessaire de contrôler la prosodie des productions enregistrées, pour réduire le plus possible toute variation segmentale liée à la prosodie :

> « Our results underline a point made by a few other researchers: the importance of understanding, controlling, and reporting the prosody of speech materials in articulation experiments. For the individual experimenter, unsought variation in prosody is a potential confound both within and across speakers, as our own experiment shows. It can also make comparisons across studies difficult or impossible, as researchers have always known. »[13]

[13] « Nos résultats soulignent un argument donné par quelques autres chercheurs : l'importance de comprendre, contrôler et décrire la prosodie des productions de parole dans des expériences sur l'articulation. Pour l'expérimentateur, une variation non désirée au niveau de la prosodie est un facteur confondant inter- et intra-locuteurs, comme notre propre expérience le montre. Les comparaisons entre études peuvent se révéler de ce fait difficiles ou impossibles, comme les chercheurs le savent. » (notre traduction)

L'homogénéité de la position des séquences CV cibles en position initiale de mot nous permet donc de contrôler une éventuelle variation articulatoire et donc acoustique liée à la position du son ou de la séquence CV dans le groupe accentuel (qui correspond à un mot sans article dans notre étude). L'étude pilote de Scarbel (2012) avait fait le choix d'analyser les caractéristiques acoustiques des voyelles en position initiale de mot, ce qui s'est avéré problématique, puisqu'un nombre important de voyelles n'étaient pas exploitables car elles étaient élidées, laryngalisées, dévoisées, ou même trop courtes pour pouvoir être analysées. Nous avons donc décidé, pour notre étude, d'étudier les voyelles en contexte /#b_/, de façon à éviter dans la mesure du possible ces phénomènes rendant les voyelles inexploitables : le voisement de la consonne se propage à la voyelle qui la suit et nous évitons ainsi les nombreux dévoisements de voyelles observés dans l'étude pilote de Scarbel (2012). Le choix d'un seul mode d'articulation pour cette consonne initiale nous permet de contrôler les effets éventuels de la coarticulation sur la production de la voyelle, puisque le contexte est identique pour toutes les voyelles. Ce choix est également motivé par le fait que cette consonne est labiale et non protruse, ce qui ne devrait affecter que légèrement la nature articulatoire et acoustique de la voyelle. Pour l'étude de la coarticulation, nous avons choisi cinq contextes vocaliques différents qui constituent une représentation équilibrée bien que réduite de l'espace vocalique du français.

Dans les paragraphes suivants (§1.2.2.2 à §1.2.2.4), nous allons détailler les deux tâches d'élicitation de mots (répétition et dénomination), qui ont été proposées aux deux groupes d'enfants (NH et CI) de façon identique, pour recueillir leurs productions de parole.

1.2.2.2. Choix des mots du corpus pour la tâche de répétition

Cette tâche de répétition comporte 65 mots du français, qui nous permettent d'étudier des productions de voyelles, de consonnes et de séquences syllabiques CV. Les consonnes initiales sont les occlusives /p, t, k, b, d, g/ et les fricatives /f, s, ʃ/, les voyelles sont les voyelles orales /i, e, ɛ, y, ø, œ, a, u, o, ɔ/ et nasales /ɑ̃, ɔ̃, ɛ̃/. Pour l'étude des consonnes, nous avons choisi des contextes vocaliques /#_i/ et /#_u/, pour l'étude des voyelles, nous avons choisi des contextes consonantiques /#b_/ et pour étudier la coarticulation CV, nous avons choisi des mots commençant par des séquences /bV/, /dV/ et /gV/, suivis des voyelles /i, e, u, o, a/.

1.2.2.3. Choix des mots du corpus pour la tâche de dénomination

Les mots présentés sont 35 mots mono- ou bisyllabiques comprenant une séquence CV ou CVC initiale. Les consonnes initiales sont les occlusives /p, t, k, b, d, g/ et les fricatives /f, s, ʃ/, les voyelles sont les voyelles orales /i, e, ɛ, y, ø, œ, a, u, o, ɔ/ et nasales /ɑ̃, ɔ̃, ɛ̃/. Pour l'étude des consonnes, nous avons choisi des contextes vocaliques /#_i/ et /#_u/, pour l'étude des voyelles, nous avons choisi des contextes consonantiques /#b_/ et pour étudier la coarticulation CV, nous avons choisi des mots commençant par des séquences /bV/, /dV/ et /gV/, avec les voyelles /i, e, u, o, a/.

1.2.2.4. Listes des mots du corpus pour la tâche de répétition et la tâche de dénomination

Les mots du corpus sont présentés dans les tables 4.4 (mots utilisés pour l'étude des consonnes), 4.5 (étude des voyelles) et 4.6 (étude de la coarticulation de séquences CV) ci-dessous. Les mots utilisés lors de la tâche de répétition sont les mots dits « simples » et « complexes », dont le choix

devait prendre en compte les contraintes liées à notre problématique ainsi que l'âge des enfants dont nous étudions les productions.

Le degré de complexité des mots du corpus est lié d'une part à la fréquence lexicale de ces mots, et d'autre part à des caractéristiques articulatoires des mots (longueur des mots, présence de clusters consonantiques). Les mots « simples » sont tous mono- ou bi-syllabiques, alors que les mots « complexes » sont tous bi- ou tri-syllabiques. La fréquence lexicale est obtenue à partir la base de données Manulex[14] (Lété et al. (2004)) : nous avons choisi l'indice SFI (Standard Frequency Index)[15] qui permet de rendre compte de la fréquence relative standardisée de mots figurant dans des manuels scolaires utilisés à l'école primaire. Nous avons choisi la fréquence des mots dans la base de données allant du CP au CM2, puisqu'il s'agit des niveaux scolaires des enfants participant à notre étude. Quatre mots n'apparaissent pas dans le corpus de livres scolaires de la base Manulex, nous n'avons donc pas pu en fournir la fréquence lexicale. Les fréquences lexicales des mots de notre corpus sont fournies en annexe 4. Un test t indique que les fréquences lexicales des mots simples et complexes de notre corpus, obtenues à partir de la base de données Manulex, sont significativement différentes : $t(39{,}081)=2{,}6554$, $p<.05$.

Plusieurs mots sont utilisés pour l'étude d'une consonne, d'une voyelle et d'une séquence CV. Ces mots sont présents dans chaque liste (Tables 4.4, 4.5 et 4.6) mais n'ont été présentés qu'une seule fois dans la tâche de répétition. Les mots du corpus utilisés pour la tâche de dénomination sont les listes de mots dits « simples ». Comme pour la tâche de répétition, certains mots sont utilisés pour plusieurs analyses acoustiques différentes mais ils ne sont produits qu'une seule fois par les enfants.

[14] Base de données Manulex, Lété et al. (2004) : www.manulex.org, consultée le 26 mars 2016

[15] Définition de l'indice SFI sur le site http://www.manulex.org/docs/fr/Manulex.pdf : « SFI : Standard Frequency Index : un indice de fréquence courant calculé à partir de U par transformation logarithmique. […] Un SFI de 90 signifie une rencontre (une occurrence) tous les 10 mots lus. Un SFI de 80 traduit une rencontre tous les 100 mots. Un SFI de 70, tous les 1000 mots, etc... Un SFI de 40 signifie une seule rencontre dans un million de mots. Le SFI est calculé à partir de U en utilisant la formule: $SFI = 10*(\log_{10}(U) + 4)$ », avec U : Fréquence estimée d'Usage pour 1 million de mots.

Consonne	Mots simples	Mots complexes
p	pipe	pistache
p	poupée	poussière
t	ticket	T-shirt
t	toupie	tournesol
k	kiwi	kimono
k	couteau	couverture
b	bille	bicyclette
b	bouton	bourdon
d	dix	dinosaure
d	doudou	douze
g	guitare	guirlande
g	goutte	gourmand
s	sirop	cygne
s	soupe	sous-marin
f	filet	ficelle
f	four	fourchette
ʃ	chiffon	chinois
ʃ	chou	choucroute

Table 4.4 : mots utilisés lors de la tâche de répétition pour l'étude des consonnes classés par degré de complexité articulatoire

Voyelle	Mots simples	Mots complexes
a	bateau	balançoire
e	bébé	bétonneuse
ɛ	bec	berceau
i	bille	bicyclette
o	bobo	bocal
ɔ	botte	bol de soupe
u	bouton	bourdon
y	bureau	bûcheron
ø	belette	besace
œ	bœuf	beurre fondu
ã	banc	banquise
ɔ̃	bonbon	bonbonne
ɛ̃	bain	bungalow

Table 4.5 : mots utilisés lors de la tâche de répétition pour l'étude des voyelles classés par degré de complexité articulatoire

/bV/	/dV/	/gV/
bateau	date	gâteau
bébé	dessin	guépard
bouton	doudou	goutte
bille	dix	guitare
bobo	dauphin	gaufre

Table 4.6 : mots utilisés lors de la tâche de répétition pour l'étude de la coarticulation CV classés par lieu d'articulation de la consonne

1.3. Procédure d'enregistrement

1.3.1. Matériel, logiciels et conditions de passation

Tous les enregistrements (stimuli et productions des enfants) ont été réalisés avec un enregistreur numérique de type Marantz PMD 670, avec une fréquence d'échantillonnage de 44 100 Hz, en mono, 16 bits. Nous avons utilisé un micro cardioïde AKG C1000S placé avec un trépied à environ 40 cm de la bouche des enfants, sur la table où était également placé l'ordinateur portable sur lequel les stimuli étaient présentés.

Les stimuli et images ont été présentés au moyen du logiciel Psychopy (Peirce (2007))[16] qui permet de présenter des stimuli associant un son et une image (pour la tâche de répétition) ou une image seule (pour la tâche de production), de façon aléatoire. Pour notre étude, le passage au stimulus suivant se faisait par la touche « espace », les enfants n'avaient pas de contrainte de temps pour répéter ou produire chaque mot.

L'ordre des tâches était le même pour tous les enfants : répétition, puis dénomination, puis narration.

Les enfants recevaient un livre (de type bande dessinée ou album) en remerciement de leur participation.

1.3.2. Elaboration des stimuli de la tâche de répétition

Pour construire les stimuli à utiliser dans la tâche de répétition, nous avons enregistré une locutrice adulte de 39 ans, qui a l'habitude de s'adresser à des enfants, de façon à avoir des stimuli suffisamment articulés pour être compris par les deux groupes d'enfants. La locutrice a produit les mots en accentuant la première syllabe de chaque mot. Les règles d'accentuation du français prévoient que la dernière syllabe du groupe accentuel, qui correspond généralement à un ou plusieurs syntagmes (et qui est composé ici d'un seul mot) porte un accent, c'est-à-dire que cette dernière syllabe est proéminente (pour une description de l'accent primaire du français, voir entre autres Lacheret-Dujour & Beaugendre (1999), DiCristo (1998), DiCristo (2000), Jun & Fougeron (2000) ou Martin (2009)). Cependant, comme nous souhaitons comparer des segments et des syllabes en position initiale de mots, nous avons demandé à la locutrice de produire les mots du corpus en plaçant un accent sur la séquence syllabique cible, qui peut être vu comme un accent emphatique ou un accent de focalisation étroite (pour une description de ces deux types d'accents, voir entre autres DiCristo (2000), Delais-Roussarie et al. (2015)). Le phénomène de convergence va influencer la production des enfants, qui vont à leur tour produire les mots avec

[16] Peirce, J.W. (2007), PsychoPy - Psychophysics software in Python. *Journal of Neuroscience Methods, 162 (1-2),* 8-13, version 1.80.03, téléchargée sur le site www.psychopy.org

un accent initial de mot. Comme la tâche de répétition est envisagée également comme un entraînement à la tâche de production sans modèle audio, la répétition systématique de mots accentués sur la syllabe initiale conduira les enfants à reproduire cette accentuation particulière lors de la tâche de production.

1.3.3. Lieux d'enregistrement

Les enregistrements de la locutrice adulte utilisés pour la construction des stimuli audio, ainsi que ceux de tous les enfants normo-entendants ont tous été réalisés en chambre sourde, au laboratoire Gipsa-Lab, à Grenoble. Les enregistrements des enfants sourds, porteurs d'implant cochléaire ont été réalisés 1) dans une cabine d'audiométrie du service ORL du CHU de Grenoble pour les enfants CI6 et CI7, et 2) dans une pièce calme du service ORL de l'Hôpital Edouard Herriot de Lyon pour les enfants CI10, CI11, CI12, CI13, CI14, CI16, CI17, CI19, CI20, CI21 et CI22.

1.4. Traitement des enregistrements

Les trois tâches ont été enregistrées sur un seul fichier audio au format .wav. Pour pouvoir exploiter les données ainsi recueillies, il a donc été nécessaire de les organiser et de les annoter, de façon à pouvoir extraire ensuite les mesures acoustiques que nous souhaitions utiliser pour nos analyses.

1.4.1. Organisation du corpus

Dans un premier temps, nous avons redécoupé le fichier son correspondant à l'enregistrement de chaque enfant, de façon à isoler chaque mot prononcé. Pour renommer chaque fichier son, nous avons utilisé un codage qui permet d'identifier l'enfant et le groupe auquel il appartient, l'âge de l'enfant (en années), la tâche effectuée (R1, R2, P1, P2 pour 1ère répétition, etc.) et le mot, sous forme d'un code à 4 lettres (exemple de nom de fichier : NH1_8_P1_bato.wav : GroupeEnfant_Âge_Tâche_Mot.wav). Les codes utilisés pour identifier les mots produits sont donnés en Annexe 5.

1.4.2. Annotations

L'ensemble du corpus des deux tâches d'élicitation de mots (répétition d'après modèle audio et dénomination sans modèle) a été annoté avec le logiciel Praat (Boersma, P. & Weenink, D. (2015)) au format .TextGrid. Nous avons annoté tout le corpus de la même façon, puisque nous utilisons certains mots à la fois pour l'étude des consonnes et des voyelles. Nous avons fait le choix d'une annotation à 7 niveaux, sur des tires différents : 1) segments (consonne et voyelle cibles), 2) syllabe, 3) début du voisement (dv) et du burst (b), 4) début de la consonne (C) et de la voyelle (V) cibles et fin de la voyelle (F), 5) erreur sur le mot (1 : aucune erreur, 0 : erreur), 6) erreur sur la syllabe (1 : aucune erreur, 0 : erreur) et 7) description des erreurs éventuelles (dévoisement d'un segment, mot inconnu, etc.). Nous avons décidé de ne pas caractériser perceptivement les types d'erreurs commises par les enfants, mais d'annoter les erreurs qui nous empêchent d'exploiter les segments cibles : cette annotation nous permet donc d'écarter ces productions non-exploitables de nos analyses. Par exemple, il peut arriver qu'un mot n'a pas du tout été prononcé par un enfant ou qu'il est substitué par un autre (par exemple, substitution de /bœf/ par /vaʃ/), il est noté « 0 » et la syllabe-cible est également notée « 0 », le mot sera directement exclu des analyses. Dans d'autres cas, le mot a pu être mal prononcé (par exemple : /biʃiklɛt/ pour /bisiklɛt/, le mot est alors noté « 0 » mais la syllabe-cible (dans cet exemple /bi/)

qui est exploitable, est notée « 1 », les segments de ce mot seront inclus dans les analyses. Il peut également arriver qu'un mot soit noté « 1 » et la syllabe « 0 » (ce qui peut arriver lorsque la voyelle-cible est dévoisée par exemple), dans ce cas, les segments de ce mot seront exclus des analyses.

Les annotations ont été réalisées manuellement, à l'aide d'un script Praat (Boersma, P. & Weenink, D. (2015)) qui crée un fichier TextGrid avec le nombre de tires choisis et qui sauvegarde ce fichier d'annotation automatiquement. La Figure 4.2 ci-dessous présente un exemple d'annotation d'un mot du corpus.

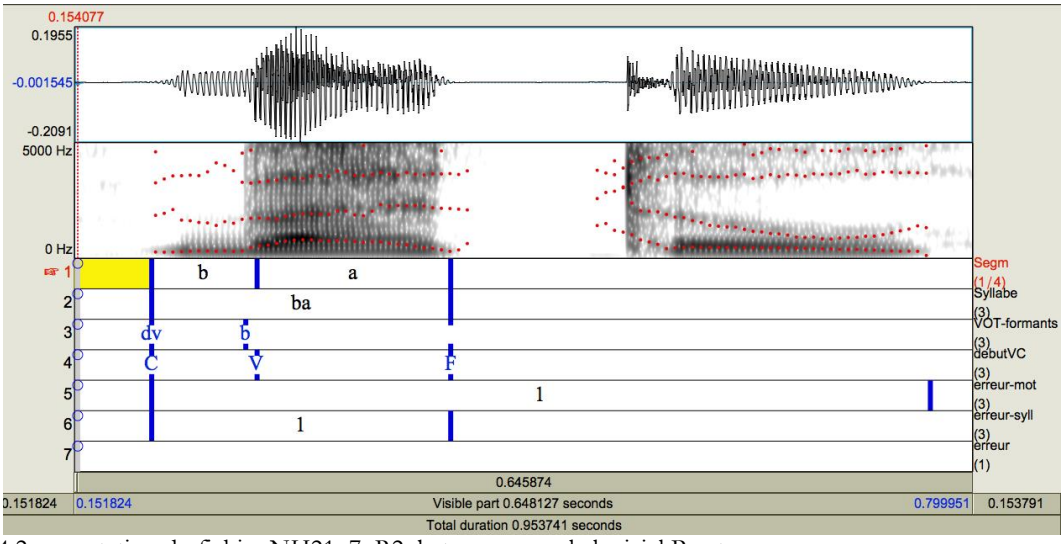


Figure 4.2 : annotation du fichier NH21_7_R2_bato.wav avec le logiciel Praat

Pour les transcriptions des segments et syllabes des tires 1 et 2, nous avons choisi l'alphabet phonétique SAMPA (Wells (1997)), qui utilise des caractères standard, tous disponibles sur un clavier d'ordinateur, ce qui nous permet d'extraire et de traiter les données sur des logiciels qui ne traitent pas les caractères de l'Alphabet Phonétique International (API).

La correspondance entre l'alphabet phonétique international et l'alphabet SAMPA pour les symboles que nous utilisons est donnée dans la table 4.7 ci-dessous.

	consonnes	voyelles
Alphabet Phonétique International (API)	p t k b d g f s ʃ	i e ɛ ɔ u o a y ø œ ɛ̃ ã ɔ̃
Alphabet SAMPA	p t k b d g f s S	i e E O u o a y 2 9 e~ a~ o~

Table 4.7 : correspondance entre l'Alphabet Phonétique International (IPA) et l'alphabet SAMPA

Notre but étant de proposer une caractérisation acoustique de la réalisation des sons de parole par les deux groupes d'enfants, normo-entendants et sourds, porteurs d'implant cochléaire, nous avons fait le choix de transcrire le mot cible adulte.

Le choix de cette transcription nous permet, pour un phonème donné (ou une réalisation allophonique d'un phonème) d'extraire des caractéristiques acoustiques, que nous pouvons ensuite comparer entre elles, et donc de proposer une caractérisation acoustique de ce phonème (ou de cet allophone).

2. Analyses acoustiques

Les mesures acoustiques (calculs de formants, de durée de VOT, de moments spectraux) sont réalisées avec des scripts Praat (Boersma, P. & Weenink, D. (2015)). Dans les paragraphes suivants, nous allons tour à tour présenter la méthodologie des analyses acoustiques réalisées pour étudier les voyelles (§2.1), les consonnes (§2.2) et la coarticulation (§2.3) chez les deux groupes d'enfants. Les méthodologies de traitement et d'analyse acoustique (mesures de formants, VOT, moments spectraux, équations de locus) sont les mêmes pour les deux tâches (répétition et production) et pour les deux degrés de complexité (mots simples et complexes), et nous présenterons les résultats ensuite en deux temps, pour tenter de répondre à nos deux problématiques, présentées au début de ce chapitre (§1.1 et §1.2).

2.1. Voyelles orales : formants (F1, F2, F3), espaces vocaliques F1-F2 et F2-F3

Dans un premier temps, nous allons présenter le système vocalique du français, puis nous détaillerons la méthodologie de mesure des formants que nous avons adoptée. Nous présenterons également une étude préliminaire que nous avons réalisée pour comparer des méthodes de normalisation disponibles dans la littérature, pour faire le choix de la méthode la mieux adaptée à notre problématique et à la population d'enfants dont nous étudions les productions de voyelles. Cette étude a été réalisée à partir des enregistrements de l'étude pilote de Scarbel et al. (2012), dont le corpus est légèrement différent de celui que nous utilisons dans notre propre étude : les détails de l'étude de Scarbel et al. (2012) seront précisés dans le paragraphe §2.1.3.5 ci-après.

2.1.1. Description du système vocalique du français

2.1.1.1. Caractéristiques articulatoires des voyelles du français

Plusieurs descriptions du vocalisme du français sont proposées dans la littérature, et ces descriptions diffèrent quant au nombre de voyelles qu'elles considèrent. Vallée (1994) propose une typologie des systèmes vocaliques dans les langues du monde à partir de la base de données UPSID (UCLA Phonological Segment Inventory Database, 2014) qui reprend, pour le français, la description de Sten (1963). Dans cette description, le français compte 16 voyelles : dix voyelles orales courtes /i/, /e/, /ɛ/, /a/, /o/, /ɔ/, /u/, /y/, /ø/, /œ/, deux voyelles orales longues /ɛ:/ et /o:/, et quatre voyelles nasales /ɒ̃/, /õ/, /œ̃/ et /æ̃/. Vallée (1994) indique que seuls 2,8%[17] des langues répertoriées dans la base de données UPSID ont un système à 16 voyelles (ce qui correspond à neuf langues sur 317), ce qui fait du français un système assez complexe puisque la grande majorité des langues (80%) possèdent entre 3 et 10 voyelles. Par ailleurs, le français fait partie des 0,3%[18] des langues du monde qui possèdent la série des voyelles antérieures arrondies /y/, /ø/, /œ/ et des 22,4%[19] des langues du monde qui possèdent des voyelles nasales.

Cette description du système vocalique du français est partiellement en accord avec celle de Walter (1977), qui compte 12 voyelles orales /i/, /e/, /ɛ/, /ɛ:/, /a/, /ɑ/, /o/, /ɔ/, /u/, /y/, /ø/ et /œ/ et quatre voyelles nasales /ã/, /õ/, /œ̃/ et /ɛ̃/. La description de Léon (2007) fait

[17] Vallée, N. (1994). *Systèmes vocaliques: de la typologie aux prédictions* (Thèse de doctorat, Grenoble 3), p 88.
[18] Ibid., p 118.
[19] Ibid., p 99.

état de 12 voyelles orales /i/, /e/, /ɛ/, /a/, /ɑ/, /o/, /ɔ/, /u/, /y/, /ø/, /œ/ et /ə/ et des mêmes voyelles nasales que Walter (1977). Dans Calliope (1989), le vocalisme du français est réduit à dix voyelles orales /i/, /e/, /ɛ/, /a/, /o/, /ɔ/, /u/, /y/, /ø/ et /œ/ et 3 ou 4 voyelles nasales. En effet, l'évolution de la langue conduit à une indifférenciation de plus en plus fréquente en français standard des voyelles orales [a] et [ɑ] d'une part et des voyelles nasales /œ̃/ et /ɛ̃/ d'autre part.

La table 4.8 ci-dessous présente le système vocalique du français, d'après les descriptions de Sten (1963), Vallée (1994), Calliope (1989), Walter (1977) et Léon (2007).

Antériorité / Hauteur	Antérieures		Centrales	Postérieures	
	Non-arrondies	Arrondies		Non-arrondies	Arrondies
Hautes	i	y			u
Mi-hautes	e	ø			o (o:)
Mi-basses	ɛ (ɛ:) ɛ̃	œ (œ̃)	(ə)		ɔ ɔ̃
Basses	a ã			(ɑ)	

Table 4.8 : système vocalique minimal du français (dix voyelles orales et trois voyelles nasales), d'après Sten (1963), Vallée (1994), Calliope (1989), Walter (1977) et Léon (2007). Les voyelles nasales sont placées en dessous de leur équivalent oral, et les voyelles entre parenthèses sont celles qui ne sont pas communes à toutes les descriptions.

Plusieurs caractéristiques permettent de décrire la production des voyelles d'un point de vue articulatoire. Léon (2007) propose quatre « types » qui résultent de positions différentes des articulateurs mobiles (langue, voile du palais et lèvres) à l'intérieur du conduit vocal :

1) le type ouvert ou fermé qui est lié à l'abaissement plus ou moins grand de la langue, quatre niveaux d'abaissement de la langue ou d'aperture de la voyelle sont attestés en français : les voyelles [i], [y] et [u] sont des voyelles fermées ou hautes, les voyelles [e], [ø] et [o] sont des voyelles mi-fermées ou mi-hautes, les voyelles [ɛ], [œ] et [ɔ] sont des voyelles mi-ouvertes ou mi-basses, et les voyelles [a] et [ɑ] sont des voyelles basses ;

2) le type antérieur ou postérieur qui correspond à la position de la langue sur un axe horizontal : les voyelles [i], [e], [ɛ], [y], [ø], [œ] et [a] sont des voyelles antérieures et les voyelles [u], [o], [ɔ] et [ɑ] sont des voyelles postérieures ;

3) le type oral ou nasal, la voyelle est dite orale lorsque le voile du palais est relevé et le flux d'air passe par le seul conduit buccal, et nasale lorsque le voile du palais est abaissé et le flux d'air passe à la fois par les conduits buccal et nasal : [i], [y], [u], [e], [ø], [o] [ɛ], [œ] et [ɔ] sont des voyelles orales, et [ã], [ɔ̃], [ɛ̃] sont des voyelles nasales.

4) le type labial ou non labial qui correspond à l'arrondissement des lèvres : les voyelles [y], [u], [ø], [œ], [o] et [ɔ] sont dites labiales ou arrondies puisqu'elles sont produites avec un arrondissement des lèvres, et les voyelles : [i], [e], et [ɛ] sont dites non-labiales ou non-arrondies, puisqu'elles sont produites avec un étirement des lèvres.

La figure 4.3 ci-dessous présente des schémas de la production des voyelles du français, classées selon leurs caractéristiques articulatoires (aperture, antériorité, arrondissement, nasalité).

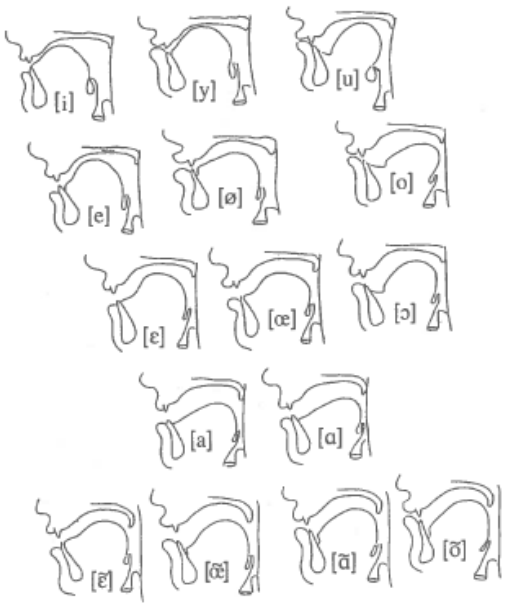


Figure 4.3 : schémas articulatoires des voyelles (orales et nasales) du français, tirés de Léon (2007)

2.1.1.2.Caractéristiques acoustiques des voyelles du français

La théorie source-filtre développée par Fant (1960) a permis de modéliser la production de voyelles : un son (source) est produit par vibration des cordes vocales au niveau de la glotte. Ce son est assimilable à un son périodique, qui se décompose en une fréquence fondamentale (la F0) et des harmoniques. Le son glottique excite le conduit vocal, dont la configuration va stimuler les harmoniques : le conduit vocal peut être modélisé sous forme de tubes qui correspondent chacun à une cavité où certaines harmoniques sont stimulées. Les changements de configuration du conduit vocal qui permettent de produire les voyelles vont modifier les zones d'excitation des harmoniques, ce qui correspond acoustiquement à des fréquences d'excitation qui sont différentes selon les voyelles. Ces zones de renforcement des harmoniques sont appelées les formants. D'après la théorie source-filtre, il existe une corrélation entre caractéristiques articulatoires et acoustiques des voyelles produites : le premier formant est modifié par l'abaissement de la mandibule et/ou de la langue (lorsque la mandibule et/ou la langue s'abaissent, le F1 augmente), le deuxième formant par la position horizontale de la langue à l'intérieur de la cavité buccale (lorsque la langue s'avance, le F2 augmente), le troisième formant par l'arrondissement des lèvres (lorsque les lèvres sont protruses ou arrondies, le F3 diminue). Les voyelles peuvent être représentées graphiquement en deux dimensions sur un plan F1-F2 et sur un plan F2-F3.

Delattre (1948) propose une description des voyelles du français sur un plan F1-F2. Cette représentation se base sur des mesures acoustiques et les valeurs (F1 et F2) utilisées pour cette représentation ainsi que le triangle acoustique correspondant sont proposées dans la figure 4.4, ci-dessous.

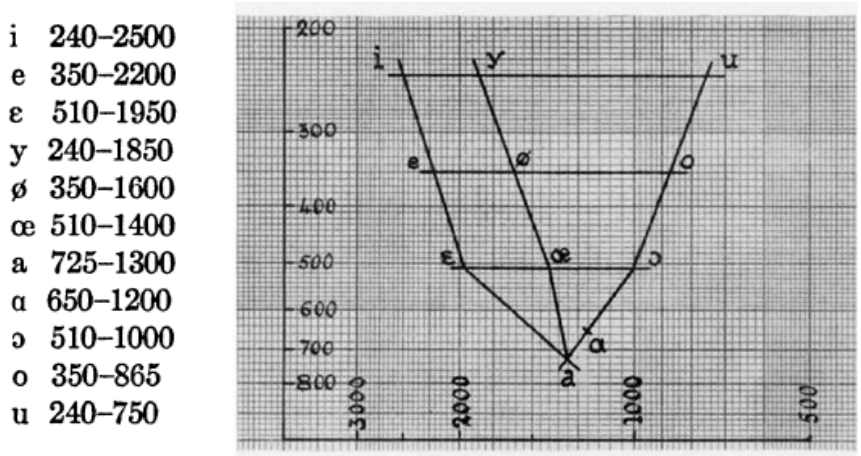


Figure 4.4 : valeurs des 1er et 2ème formants (gauche) et représentation F1-F2 du système vocalique du français (droite), tiré de Delattre (1948)

Delattre (1948) propose cette représentation des voyelles du français à partir de mesures acoustiques des premier et deuxième formants, ce qui lui permet d'établir une correspondance entre caractéristiques articulatoires et acoustiques des voyelles. Sur la dimension verticale du F1, il distingue cinq niveaux de hauteur et sur la dimension horizontale du F2, plusieurs niveaux d'antériorité, qui sont différents selon la hauteur des voyelles, mais qui se caractérisent par une distance constante par couples de voyelles de même hauteur : les distances /i-y/, /e-ø/ et /ɛ-œ/ sont similaires. En outre, selon Delattre (1948), la modification de la taille et de la configuration articulatoire des différentes cavités de résonnance qui constituent le conduit vocal se traduit par une modification des valeurs acoustiques de F1 et F2 : il montre une relation entre hauteur de la voyelle et premier formant, et une relation entre l'antériorité de la position de la langue et l'arrondissement des lèvres et deuxième formant.

L'étude de Chistovich & Lublinskaya (1979) propose une distance maximale entre deux formants consécutifs de 3,0-3,5 Bark, jusqu'à laquelle la voyelle produite est focale. Cette étude a notamment été reprise par Schwartz & Escudier (1989), Schwartz et al. (2004), Menard et al. (2007), qui montrent que la convergence de deux formants a des conséquences perceptives puisque ceux-ci ne sont alors perçus que comme un seul formant.

En effet, les deux études de Ménard et al. (2007) et Schwartz et al. (2004) s'intéressent à la répartition des voyelles dans des espaces à plus de deux dimensions et aux relations entre perception et caractéristiques des voyelles. Ménard et al. (2007) montrent que certaines voyelles sont produites avec des positions extrêmes sur un espace F1-F2-F3, qui résultent de rapprochements de certains formants. La voyelle /i/ est caractérisée par un F3 et un F4 très proches, tandis que les voyelles /u/ et /a/ sont caractérisées par un F1 et un F2 très proches (/u/ ayant des valeurs de F1 et F2 beaucoup plus basses que /a/, ce qui permet de distinguer ces deux voyelles), alors que la voyelle /y/ est caractérisée par un F2 et un F3 très proches. Ménard et al. (2007) montre que ces caractéristiques sont similaires chez trois groupes d'âges différents (enfants de quatre et huit ans, adultes), malgré une différence de maturité et de taille du conduit vocal. La différence constatée étant une moins bonne intelligibilité des enfants de quatre ans, pour la voyelle /y/.

L'étude de Schwartz et al. (2004) se concentre sur la perception de ces voyelles focales. Ils définissent la focalisation comme la « convergence between two consecutive formants in a vowel

spectrum »[20], et montrent que cette focalisation rend les voyelles plus faciles à percevoir. Les voyelles périphériques /u/, /a/, /y/ et /i/ sont vues comme des références pour la perception des autres voyelles du système. Dans cette étude, une expérience de perception où F1 et F2 sont stables, mais où F3 varie (en s'approchant plus ou moins de F2 ou F4) met en évidence que l'identification des voyelles est plus difficile lorsque F3 est équidistant de F2 et F4. Ainsi, la perception des voyelles est facilitée si F3 est proche de F2 ou F4.

Vaissière (2007, 2011) propose également une description des voyelles du français et de la focalisation qui caractérise certaines d'entre elles. Alors que pour Ménard et al. (2007) et Schwartz et al. (2004), seules quatre voyelles étaient dites « focales », Vaissière (2007, 2011) propose six voyelles focales dont les caractéristiques sont présentées dans la table 4.9, ci-dessous.

Palatal	(⇑F3F4)	[i] $^{F3=3200Hz}$	French-type /i/	/i/
	(⇑F2)	[i] $^{F2=2500Hz}$	British-type /i/	
	(F2↓F3)	[y] $^{F2=2100Hz}$	Swedish-type /y/	/y/
	(↓F2F3)	[y] $^{F2=1900Hz}$		
Pharyngeal	⇑(F1F2) or ⇑(F1F2)	[ɑ] $^{F2=1000Hz}$		/ɑ/
Labio-velar	↓(F1F2)	[ɔ] $^{F2=900Hz}$	French-type	/ɔ/
	↓(F1F2)	[o] $^{F2=800Hz}$		/o/
	⇓(F1F2) or ⇓(F1F2)	[u] $^{F2=700Hz}$		/u/

Table 4.9 : description de la convergence des formants des voyelles focales, tirée de Vaissière (2007)

La voyelle /i/ est caractérisée par une convergence de F3 et F4 autour de 3200 Hz (pour un homme adulte) et un F3 haut, /y/ par une convergence de F2 et F3 autour de 1900 Hz, /u/ par une convergence de F1 et F2 avec une concentration d'énergie dans les basses fréquences autour de 700 Hz, /a/ par une convergence de F1 et F2 avec une concentration dans des fréquences les plus hautes possibles vers 1000 Hz (ce qui permet de distinguer /a/ de /u/), et enfin les deux voyelles /ɔ/ et /o/ sont caractérisées par une convergence de F1 et F2 qui augmentent simultanément sur un axe F1-F2 délimité par /u/ et /a/.

Georgeton et al. (2012) proposent également une comparaison des différences F2-F1, F3-F2 et F4-F3 des voyelles du français, qui met en évidence une convergence de certains formants, notamment F1 et F2 pour /a/, /u/, /o/ et /ɔ/, de F2 et F3 pour /y/ et de F3 et F4 pour /i/, avec une limite supérieure de la différence F2-F1, F3-F2 et F4-F3 plus élevée que dans l'étude de Chistovich & Lublinskaya (1979) pour permettre de déterminer une convergence formantique (jusqu'à 6 Bark dans l'étude de Georgeton et al., 2012).

Cette description articulatoire et acoustique des voyelles du français nous a permis de comprendre les difficultés et les enjeux de l'analyse de la production de voyelles par des locuteurs francophones : le français possède un système vocalique complexe (une dizaine de voyelles orales et plusieurs voyelles nasales), qui a des caractéristiques acoustiques particulières (voyelles focales et non-focales) que nous devrons prendre en compte pour caractériser la production de voyelles par les enfants participant à notre étude.

[20] « […] convergence entre deux formants consécutifs sur le spectre d'une voyelle » (notre traduction)

2.1.2. Mesures des formants

Nous avons enregistré des productions des dix voyelles orales et des trois voyelles nasales du français en contexte /#b_/ et /#b_C/, mais notre étude se concentrera sur les seules voyelles orales. En effet, une revue bibliographique a mis en évidence l'impossibilité de caractériser uniquement acoustiquement les voyelles nasales : Amelot (2004) montre par exemple qu'il n'y a pas de corrélation entre l'ouverture vélaire maximale (c'est-à-dire les gestes articulatoires utilisés pour élever et abaisser le velum) et le signal acoustique.

Préalablement à la mesure des formants, nous avons sous-échantillonné les fichiers sons correspondant à chaque mot : en divisant la fréquence d'échantillonnage initiale de 44 100 Hz par quatre, nous obtenons des fichiers avec une nouvelle fréquence d'échantillonnage de 11 025 Hz. Cette manipulation préalable est réalisée automatiquement avec un script Matlab (MATLAB and Statistics Toolbox Release 2014b, version 8.4.0) utilisant les fonctions *downsample* et *wavwrite*. Elle nous permet de concentrer la recherche des 3 premiers formants sur la portion du signal où ils se situent, plutôt que sur la totalité du signal. La figure 4.5 ci-dessous propose une représentation spectrale (fréquence-amplitude) d'une voyelle avec la fréquence d'échantillonnage initiale de 44 100 Hz et avec la fréquence sous-échantillonnée à 11 025 Hz.

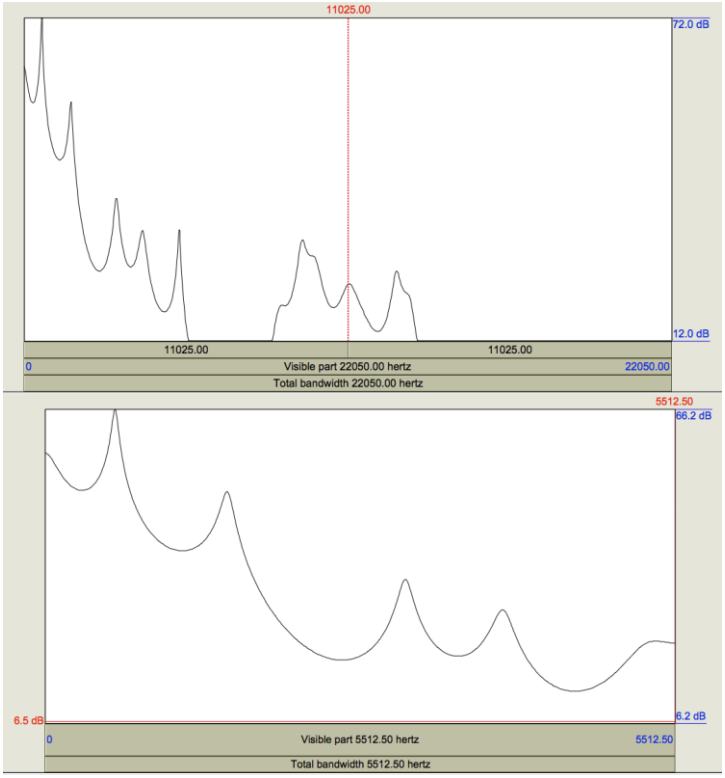


Figure 4.5 : représentation spectrale (fréquence-amplitude) d'une voyelle /a/ produite par l'enfant NH26 lors de la tâche de répétition : en haut, son original (fréquence d'échantillonnage de 44 100 Hz), et en bas, son sous-échantillonné (fréquence d'échantillonnage de 11 025 Hz)

Pour nos analyses acoustiques des voyelles, nous avons choisi de mesurer les formants F1, F2 et F3 au centre de chaque voyelle. Ces mesures ont été effectuées automatiquement (Méthode Burg, 5 formants, 5000 Hz) avec un script Praat (Boersma, P. & Weenink, D. (2015)), puis vérifiées manuellement, pour s'assurer de leur exactitude. Nous avons ensuite normalisé les mesures de formants obtenues, en utilisant la méthode de normalisation des scores-z de Lobanov (Lobanov (1971), voir paragraphe §2.1.3 ci-dessous). Pour nos analyses, nous utiliserons à la fois les

mesures normalisées de chaque formant (F1, F2 et F3) séparément, mais également une mesure de dispersion de la catégorie dans un plan F1-F2 et dans un plan F2-F3. Pour calculer cette dispersion, nous utilisons les formules ci-dessous, qui correspondent à la distance euclidienne de chaque occurrence de voyelle par rapport au centre de la catégorie vocalique correspondante (avec x : occurrence de la voyelle, gr : groupe, cond : condition et i : catégorie vocalique) :

Distance euclidienne F1-F2 :

$$dist(x)_i = \sqrt{(F1(x)_i - F1(moy)_{gr,cond,i})^2 + (F2(x)_i - F2(moy)_{gr,cond,i})^2}$$

Distance euclidienne F2-F3 :

$$dist(x)_i = \sqrt{(F2(x)_i - F2(moy)_{gr,cond,i})^2 + (F3(x)_i - F3(moy)_{gr,cond,i})^2}$$

La condition correspond à la tâche (répétition ou dénomination) ou à la complexité lexicale (simple ou complexe), selon les hypothèses testées et les analyses réalisées.

Nous avons également choisi d'étudier des phénomènes de focalisation ou convergence formantique, de façon à comprendre dans quelle mesure les enfants CI arrivent à produire des voyelles similaires à celles des enfants NH. Nous avons donc calculé une différence F2-F1 et une différence F3-F2, en Bark (Traunmüller, 1990, formule donnée au paragraphe §2.1.3.5 ci-après), une différence F2-F1 ou F3-F2 inférieure à 3,5 Bark (Chistovich & Lubinskaya, 1979) indiquant une convergence formantique.

2.1.3. Normalisation des valeurs de formants : étude préliminaire

2.1.3.1. Introduction

Qu'il s'agisse d'adultes ou d'enfants, la production de voyelles est soumise à deux variabilités qui coexistent et qui se manifestent au niveau acoustique : le premier type de variabilité est lié à la taille du conduit vocal du locuteur, et le deuxième type de variabilité découle du contrôle articulatoire du locuteur. Les voyelles produites par des adultes et des enfants vont donc posséder des caractéristiques acoustiques différentes, à la fois parce que la taille de leurs conduits vocaux diffère et parce que leurs capacités de contrôle diffèrent. Dans le cadre d'une étude sur la parole pathologique, l'analyse des données doit rendre compte de la variabilité de la parole pathologique par rapport à la parole typique, tout en éliminant toute variabilité liée à la taille du conduit vocal des locuteurs, qui est une conséquence de caractéristiques telles que l'âge et le genre. Ce chapitre propose ainsi une réflexion sur la question de la normalisation des données acoustiques dans l'étude des voyelles chez des enfants dont les âges peuvent être hétérogènes et en particulier chez des enfants porteurs d'implants cochléaires.

Dans un premier temps, nous présenterons plusieurs études qui analysent les corrélats acoustiques de la variabilité de la taille du conduit vocal chez les adultes et les enfants dans la production de voyelles ; dans un deuxième temps, nous nous intéresserons aux caractéristiques méthodologiques de plusieurs études acoustiques de la production de voyelles chez des enfants et jeunes adultes sourds utilisateurs d'aides auditives et d'implants cochléaires ; puis nous présenterons les méthodes de normalisation des valeurs de formants les plus fréquemment utilisées dans la littérature ; enfin, nous comparerons six de ces méthodes de normalisation à partir des données d'enfants normo-entendants d'une étude pilote dont le but était de proposer

description acoustique de la production de voyelles par des enfants normo-entendants et des enfants porteurs d'implants cochléaires.

2.1.3.2. Variation dans la production de voyelles chez des adultes et chez des enfants

Plusieurs auteurs se sont intéressés à la production de voyelles et à la représentation spatiale de l'espace vocalique au cours de la croissance. L'étude de Peterson et Barney (1952) a été l'une des premières à mettre en évidence l'évolution de la taille de l'espace vocalique en fonction de l'âge et du genre. Cette étude s'intéresse à la variabilité de la taille de l'espace vocalique en lien direct avec la taille du conduit vocal, qui évolue lorsque les enfants grandissent et qui n'est pas identique chez les hommes et les femmes. A partir de la comparaison des productions de dix voyelles en anglais américain par trois groupes de locuteurs (33 hommes, 28 femmes et 15 enfants d'âges hétérogènes), Peterson et Barney (1952) montrent que dans une représentation spatiale des valeurs en Hz des F1-F2 des voyelles produites par les trois groupes de locuteurs (Figure 4.6 : hommes, femmes et enfants), il existe de nombreuses zones d'intersection entre voyelles, ainsi qu'une distribution continue des voyelles à l'intérieur de l'espace vocalique avec des frontières catégorielles floues.

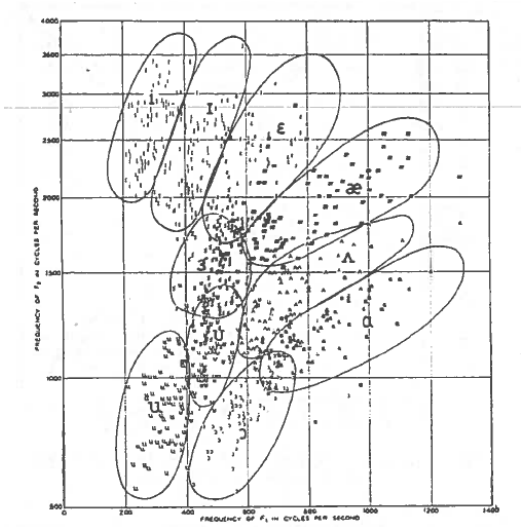


Figure 4.6 : représentation spatiale F1-F2 des voyelles produites par 76 locuteurs (enfants, femmes et hommes) dans l'étude de Peterson et Barney (1952)

Une représentation similaire des voyelles produites par les groupes des locuteurs hommes et enfants (Figure 4.7) indique des zones clairement distinctes pour ces deux groupes, ce qui constitue un argument en faveur d'une production évoluant en fonction de l'âge des locuteurs.

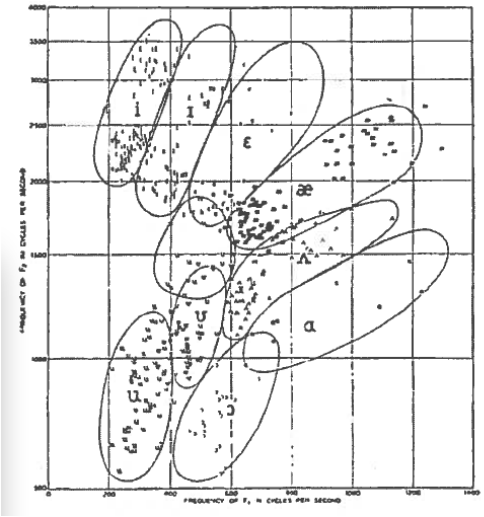


Figure 4.7 : représentation spatiale F1-F2 des mêmes voyelles produites par 48 locuteurs (enfants, hommes) dans l'étude de Peterson et Barney (1952)

Les locuteurs de cette étude produisent des voyelles dont les formants F1, F2 et F3 ont des valeurs moyennes très disparates, ce que Peterson et Barney (1952) expliquent par des différences marquées de tailles de conduits vocaux des locuteurs des trois groupes.

L'étude de Kent et Forner (1979) reprend cette problématique et compare les valeurs des fréquences des formants de deux groupes d'adultes (hommes et femmes), et de plusieurs groupes d'enfants (filles et garçons, âgés de quatre, six et 12 ans). Cette étude a pour but de décrire des variations développementales qui se traduisent par des variations au niveau des valeurs des F1, F2 et F3 produits par les différents groupes, tout en s'interrogeant sur la question de la normalisation des données pour pouvoir comparer des productions d'adultes et d'enfants. Les voyelles enregistrées dans cette étude sont produites par imitation de cinq voyelles de synthèse construites à partir des valeurs de formants des voyelles produites par les locuteurs hommes de l'étude de Peterson et Barney (1952), ainsi que dix autres voyelles qui permettent de représenter l'ensemble de l'espace vocalique de l'anglais. Les analyses montrent des différences significatives entre les valeurs en Hertz des formants F1, F2 et F3 des enfants et des deux groupes d'adultes : les valeurs de formants diminuent lorsque l'âge des enfants augmente et il y a une différence de dispersion des valeurs de formants entre les trois groupes. La transformation des formants en log réduit ces dispersions.

Huber et al. (1999) proposent une étude de l'évolution des valeurs de formants chez des enfants d'âges différents en les comparant avec un groupe d'adultes, tous locuteurs de la même variété d'anglais américain. Les auteurs étudient les variations de la production des voyelles en fonction de l'âge et du genre des enfants. Ils mesurent les valeurs de F1, F2 et F3 en Hertz d'une voyelle /ɑ/ tenue pendant plusieurs secondes chez 20 locuteurs par groupe d'enfants (dix filles et dix garçons) de quatre, six, huit, dix, 12, 14, 16 et 18 ans et d'adultes (dix femmes et dix hommes) de 20 à 30 ans (Figure 4.8 ci-dessous).

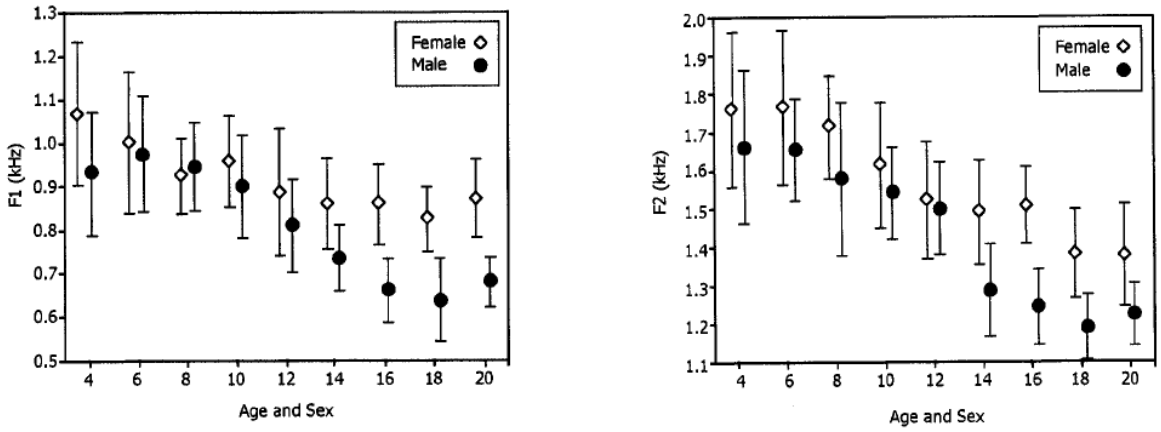


Figure 4.8 : comparaison des valeurs de F1 et F2 (en kHz) de la voyelle /ɑ/, en fonction de l'âge et du genre des participants, dans l'étude de Huber et al. (1999)

Les résultats statistiques de l'étude indiquent un effet de l'âge et du genre sur les valeurs de formants pour les enfants et un effet du genre pour les adultes. Cette étude de Huber et al. (1999) met en évidence une évolution continue des valeurs de formants chez les jeunes enfants jusqu'à 18 ans, signe d'une maturation continue de leur conduit vocal, avec une différence entre filles et garçons : chez les filles, l'évolution est forte jusqu'à 12 ans, puis plus modérée jusqu'à l'âge adulte, alors que chez les garçons, l'évolution est la plus forte entre 12 et 14 ans, puis plus modérée de 16 à 18 ans.

Vorperian et Kent (2007) proposent quant à eux une revue de littérature sur les caractéristiques acoustiques et anatomiques des enfants et dégagent quelques caractéristiques pertinentes pour l'étude des voyelles produites par ces enfants. Leur étude propose d'utiliser le système vocalique entier de la langue des locuteurs dont les productions sont étudiées, plutôt qu'une sélection de voyelles qui ne permettent pas de rendre compte de l'ensemble des caractéristiques acoustiques de l'espace vocalique (taille et répartition des voyelles au sein de cet espace). Ils préconisent également d'accompagner les données acoustiques de données anatomiques pour pouvoir caractériser au mieux l'évolution de la production des voyelles et de prendre en compte simultanément l'âge et le genre des enfants, qui sont deux caractéristiques qui influencent les productions de voyelles.

Ces quatre études réalisées chez des groupes d'enfants et d'adultes au développement typique mettent en évidence la variation liée à l'âge et au genre des locuteurs, qui résulte de différences de taille du conduit vocal, et qui doit être considérée lorsque l'on étudie des caractéristiques acoustiques des voyelles produites par des groupes hétérogènes.

2.1.3.3. Etudes de systèmes vocaliques chez des enfants et jeunes adultes sourds

Comme nous l'avons dit précédemment, l'une des caractéristiques de la parole pathologique, et notamment celle d'enfants porteurs d'implants cochléaires, tient au fait qu'à la variabilité résultant de différences de taille du conduit vocal des sujets dont on étudie la production s'ajoute une variabilité liée à la parole pathologique. Alors que les études de Peterson et Barney (1952), Kent et Forner (1979), Huber et al. (1999) et Vorperian et Kent (2007) ne s'intéressent qu'à la variation liée à l'âge et au genre, les différentes études de la production d'enfants sourds ont pour but de

quantifier la variabilité dans la parole en lien avec des pertes d'audition. Le choix d'une méthode de normalisation des fréquences manipulées est donc nécessaire pour pouvoir neutraliser la variation liée à la taille du conduit vocal tout en conservant la variation liée au déficit d'audition.

Méthode de normalisation	Auteur	Année	Langue	Participants	Âge des participants Moyenne (étendue)	Voyelles étudiées	Formants mesurés	Analyses
Hz	Baudonck et al.	2011	Néerlandais (Belgique)	40 enfants CI 42 enfants NH 34 enfants HA	8;8 (4;2-15;5) 9;3 (4;3-15;3) 9;16 (4;1-14;2)	voyelles cardinales : /i/, /a/, /u/	F1, F2	Distances entre /i/-/a/, /i/-/u/, /a/-/u/ taille de l'espace vocalique
Hz	Ertmer	2001	Anglais (US)	1 enfant CI	De l'âge de 19 à 31 mois (une session par mois)	voyelles cardinales /i/, /æ/, /ɑ/, /u/	F1, F2	Etendues de F1 et de F2
Hz	Hocevar-Bolthezar et al.	2008	Slovène	13 enfants CI 12 adultes CI	impl. à 6;3 (2-13a) impl. à 38;7 (18-62a) Sessions pré-impl., +6m, +12m	voyelles cardinales : /i/, /a/, /u/	F1, F2	Ratios F1:F2 Taille de l'espace vocalique
Hz	Horga et Liker	2006	Croate	10 enfants CI 10 enfants NH 10 enfants HA	9;7-15;2 9;6-15;1 7;4-15;1	voyelles cardinales : /i/, /a/, /u/	F1, F2	Comparaison de F1 et F2 pour chaque voyelle
Hz	Kunisue et al.	2006	Japonais	2 enfants CI 2 adultes NH	7;10 et 8;2 1 session pré-impl. 12 sessions mensuelles post-impl.	voyelles /i/, /e/, /a/, /ɯ/, /u/	F1, F2	Comparaison visuelle
Hz	Liker et al.	2007	Croate	18 enfants CI 18 enfants NH	9;5-12.8 (1ère session) + 3 enregistrements en 20m appariés	/i/, /e/, /a/, /o/, /u/	F1, F2	Comparaison des espaces vocaliques
Hz	Poissant et al.	2006	Anglais (US)	6 enfants CI-OFF 6 enfants CI-ON	7;1 (5;3-10;2) mêmes enfants (+/- 1sem)	/i/, /ɛ/, /a/, /ɔ/, /u/	F1, F2	Ratios F1:F2
Hz	Ryalls et al.	2003	Français (Québec)	10 enfants sourds profonds 10 enfants sourds modérés-sévères 10 enfants NH	9;4 (7;11-10;9) 8;7 (6;10-9;10) 8;4 (7;6-10;9)	voyelles cardinales : /i/, /a/, /u/	F1, F2, F3	Etude indépendante de chaque formant
Hz	Schenk et al.	2003	Allemand	13 adultes sourds (H) 10 adultes sourds (F) 7 adultes NH (H) 11 adultes NH (F)	34-73a 15-62a 22-40a 22-26a	/e/, /a/, /o/	F1, F2	1) Comparaison visuelle des espaces vocaliques 2) Comparaison statistique de chaque formant
Hz	Seifert et al.	2002	Suisse allemand	20 enfants CI valeurs contrôles d'après Huber et al. (2003)	6;5 (3;8-10;2) appariés	/a/	F1, F2, F3	Ratios F1:F2
Hz	Uchanski & Geers	2003	Anglais (US)	181 enfants CI 24 enfants NH	8 et 9a 8 et 9a	/i/, /ɔ/	F2	Différence : F2moy(i)-F2moy(ɔ)
Mel	Chuang et al.	2012	Mandarin (Nord de Taïwan)	24 enfants CI 24 enfants NH	93,8 mois (72-133m) appariés	voyelles cardinales : /i/, /a/, /u/	F1, F2	Etendues de F1 de /a/-/i/ et /a/-/u/ Etendue de F2 de /i/ à /u/ Espace vocalique /i-a-u/
Bark (Traunmüller)	Löfqvist et al.	2010	Suédois	12 enfants CI 11 enfants NH	184m (144-229m) 105m (103-118m)	/i/, /e/, /æ/, /y/, /ø/, /u/, /o/, /a/, /ʉ/	F1, F2	Taille de l'esp. vocalique calculé avec : 1) distances euclidiennes des voyelles au centre de l'espace F1-F2 2) distances entre les paires de voyelles
Bark (Traunmüller)	Neumeyer et al.	2010	Allemand (Sud de l'Allemagne)	5 jeunes CI 5 adultes CI 5 jeunes NH 5 adultes NH	16-27a 58-68a 15-24a 59-68a	/i:/, /e:/, /a:/, /o:/, /u:/	F1, F2	Taille de l'esp. vocalique calculé avec les distances euclidiennes des voyelles au centre de l'espace F1-F2
z-score (Lobanov)	Tseng et al.	2011	Mandarin (Taïwan)	15 enfants CI 30 enfants HA	4;1-11;2 3;9-10;5	voyelles cardinales : /i/, /a/, /u/	F1, F2	Espace maximal : (F2(i)-F2(u))*(F1(u)-F1(a))
HZ et z-score (Lobanov)	Verhoeven et al.	2016	Néerlandais (Belgique)	8 enfants CI 7 enfants HA 90 enfants NH	5-7;3a 6;1-7;9a 5-7a	/i/, /ɪ/, /ɛ/, /u/, /ʏ/, /ɔ/, /ɑ/, /e:/, /y:/, /ø:/, /o/, /a/	F1, F2	1) Distances entre /i/-/a/, /i/-/u/, /a/-/u/ 2) Surface délimitée par toutes les occurrences des 12 voyelles

Table 4.10 : résumé des caractéristiques d'études récentes sur la production de voyelles par des enfants et jeunes adultes porteurs d'implants cochléaires et utilisateurs d'aides auditives traditionnelles (classement par méthode de normalisation puis ordre alphabétique du premier auteur)

Un certain nombre d'études récentes s'intéressent à la production de voyelles par des enfants et jeunes adultes utilisateurs d'aides auditives traditionnelles ou porteurs implants cochléaires. Les caractéristiques méthodologiques de ces études sont présentées dans la Table 4.10 ci-dessus : participants (nombre, âge, type de prothèse auditive), langue de l'étude, voyelles étudiées, mesures acoustiques, analyses réalisées, et méthodes de normalisation utilisées.

La comparaison de ces études permet de mettre en lumière des disparités à l'intérieur des populations étudiées, notamment en ce qui concerne l'âge des participants. Deux types d'études sont présentées : des études transversales et des études longitudinales.

Les études transversales de Baudonck et al. (2010), Horga & Liker (2006), Poissant et al. (2006), Ryalls et al. (2003), Schenk et al. (2003), Seifert et al. (2002) et Uchanski & Geers (2003) s'intéressent à des groupes d'enfants d'âges très disparates, dont l'étendue peut aller, par exemple, de 3;8 à 10;2 ans (Seifert et al. (2002)) ou qui peuvent inclure des adolescents dans un groupe d'adultes, de 15 à 62 ans (Schenk et al. (2003)). Ces études manipulent toutes des valeurs de formants en Hertz, non normalisées, en ne tenant pas compte de l'hétérogénéité anatomique des participants : la variabilité constatée peut donc être physiologique avant d'être liée à un déficit auditif.

Les études longitudinales de Ertmer (2001), Hocevar-Bolthezar et al. (2008), Kunisue et al. (2006), Liker et al. (2007), suivent des enfants et des adultes et comparent les caractéristiques acoustiques des voyelles sans prendre en compte la croissance des enfants sur des durées parfois longues (jusqu'à 20 mois pour l'étude de Liker et al. (2007)), et la modification de la taille de leur conduit vocal pendant cette période. En effet, dans ces études, les valeurs de formants sont mesurées et manipulées en Hz, sans normalisation. Tout comme pour les études transversales, l'absence de normalisation dans les études longitudinales ne permet pas d'établir avec certitude le rôle des deux types de variabilité (de taille de conduit vocal et de contrôle) dans les productions analysées, ce qui peut être problématique.

Les études transversales de Chuang et al. (2012) manipulent des valeurs de formants normalisées en Mel, celles de Löfqvist et al. (2010) et Neumeyer et al. (2010) des valeurs de formants normalisées en Bark (Traunmüller (1990)), et celles de Tseng et al. (2011) et Verhoeven et al. (2016) des valeurs de formants normalisées en scores-z (Lobanov (1971)).

Cette revue d'études sur la parole d'enfants et de jeunes adultes porteurs d'implants cochléaires a permis de montrer l'hétérogénéité des méthodes de normalisation utilisées dans ces études, avec une prédominance des analyses en Hertz.

On peut s'interroger sur la validité des études qui manipulent des valeurs de formants non normalisées et qui ne permettent pas de distinguer si la variation constatée est imputable à une audition perturbée ou à l'âge et donc à la taille du conduit vocal, souvent hétérogène, des enfants des groupes dont les productions sont étudiées. Nous avons donc cherché à utiliser la méthode de normalisation la plus adaptée à l'analyse des voyelles de notre propre étude. Pour ce faire, nous avons donc comparé plusieurs méthodes de normalisation fréquemment utilisées dans les études sur le vocalisme, notamment en sociophonétique.

2.1.3.4. Comparaisons des méthodes de normalisation les plus fréquemment utilisées dans la littérature

Les études de Peterson et Barney (1952), Kent et Forner (1979), Huber et al. (1999) et Vorperian et Kent (2007) sur la parole des enfants, présentées précédemment, ont montré l'importance de la prise en compte de la variabilité liée à l'âge des enfants dans l'analyse acoustiques des voyelles et du choix d'une méthode de normalisation adéquate pour notre étude. Dans la littérature, la question de la normalisation des mesures de formants est notamment envisagée dans des articles de géolinguistique ou de sociophonétique. Les questions posées par l'étude de la parole pathologique sont similaires à l'étude de la variation dialectale ou sociolinguistique : comment neutraliser la variation liée à l'âge et au genre des enfants dont on étudie la production tout en conservant les caractéristiques de la parole pathologique ?

Une étude de Jacewicz et al. (2007) s'intéresse à la taille de l'espace vocalique de locuteurs adultes (femmes et hommes) de trois dialectes de l'anglais parlé aux USA : les auteurs de cette étude cherchent à utiliser la méthode de normalisation la plus adaptée pour éliminer la variation liée à l'âge et au genre des locuteurs mais à conserver la variation dialectale. Ils choisissent de comparer parallèlement les données en Hertz et en scores-z (Lobanov (1971)). Dans cette étude, ils calculent les tailles de l'espace vocalique des locuteurs (9 femmes et 9 hommes, pour chaque dialecte) en utilisant quatre ou cinq voyelles avec ces deux méthodes de normalisation. Les analyses statistiques montrent un effet significatif du genre sur la taille de l'espace vocalique lorsque les analyses sont réalisées en Hertz (espace vocalique significativement plus grand pour les femmes), mais pas d'effet lorsque les mêmes analyses sont réalisées avec les données normalisées en scores-z (Lobanov (1971)). Les résultats et la neutralisation de la variation liée à l'âge sont les mêmes lorsque les calculs sont réalisés avec quatre ou cinq voyelles. Cette étude est un premier argument en faveur de l'utilisation d'une méthode de normalisation plutôt que les données brutes en Hertz pour étudier une variation dialectale ou pathologique.

Deux études comparent les principales méthodes de normalisation utilisées pour comparer des groupes hétérogènes de locuteurs : Disner (1980) et Adank et al. (2004).

L'étude de Disner (1980) compare plusieurs méthodes de normalisation (notamment celles de Gerstman (1968), Harshman (1970), Lobanov (1971) et Nearey (1978)) sur des données de cinq langues germaniques. Le critère utilisé pour cette comparaison est la réduction à l'intérieur de l'espace F1-F2 de la dispersion de chaque catégorie vocalique par rapport aux données brutes en Hertz et la conservation des spécificités de chacune des langues étudiées. La méthode qui donne les résultats les plus satisfaisants selon Disner (1980) est celle de Harshman (1970) qui normalise les données en appliquant une transformation qui les place sur un espace en deux dimensions, avec des axes non orthogonaux, selon un angle différent pour chaque locuteur et chaque catégorie vocalique. Cette méthode est la plus adaptée selon Disner (1980) puisqu'elle permet de réduire la variabilité inter-locuteurs tout en conservant les spécificités des cinq langues étudiées.

L'étude de Adank et al. (2004) propose une comparaison de 12 méthodes de normalisation. Elle distingue deux types de normalisation : les normalisations qui utilisent des caractéristiques intrinsèques des voyelles (« vowel-intrinsic procedures »), c'est-à-dire que les occurrences de chaque voyelle sont normalisées sans tenir compte des autres voyelles du système vocalique et celles qui utilisent des caractéristiques extrinsèques des voyelles (« vowel-extrinsic procedures »),

où chaque réalisation est normalisée à l'intérieur du système vocalique, en lien avec toutes les réalisations de chaque voyelle du système.

La table 4.11 ci-dessous, tirée de Adank et al. (2004) propose un résumé des différentes méthodes utilisées.

	Vowel-intrinsic procedures
HZ	baseline condition, formant frequencies in Hz
LOG	log-transformation of the frequency scale
BARK	bark-transformation of the frequency scale
MEL	mel-transformation of the frequency scale
ERB	ERB-transformation of the frequency scale
S & G	Syrdal and Gopal's (1986) bark-distance model
	Vowel-extrinsic procedures
LOBANOV	Lobanov's (1971) z-score transformation
NEAREY1	Nearey's (1978) single logmean procedure
NEAREY2	Nearey's (1978) shared logmean procedure
GERSTMAN	Gerstman's (1968) range normalization
NORDSTRÖM	Nordström's (1976) vocal-tract scaling
MILLER	Miller's (1989) formant-ratio model

Table 4.11 : d'après Adank et al. (2004), classement des méthodes de normalisation

Les auteurs comparent les procédures suivantes : (1) HZ est la procédure qui conserve les valeurs de F0 et des trois premiers formants en Hertz (qui traite donc les données brutes, sans les normaliser), (2) LOG transforme les valeurs en Hertz sur une échelle logarithmique, (3) BARK transforme les valeurs en Hertz selon une formule proposée par Traunmüller (1990) sur une échelle perceptive (F_i=26,81 x F_i/(1960+F_i) - 0,53), (4) MEL est une transformation similaire à celle de BARK mais utilise une autre formule (F_i= 2595 x ln(1+F_i/700)), (5) ERB est également une méthode similaire à celle de BARK et MEL, avec une formule différente (F_i= 12,4 x ln(0,00437 x F_i +1), (6) S & G est une procédure qui transforme les valeurs transformées en BARK de F1 par rapport à F0 et de F3 par rapport à F2, (7) la transformation en scores-z de Lobanov permet de calculer pour F0 et pour chaque formant une distance par rapport à la moyenne de F0 ou de ce formant pour toutes les voyelles de l'espace vocalique, qui est ensuite divisée par l'écart-type de F0 ou de ce formant, ce qui permet de placer toutes les occurrences de voyelles produites par chaque locuteur sur le même espace vocalique maximal, (8) NEAREY1 transforme les valeurs de formants en log, puis calcule pour chaque formant une distance par rapport à la valeur moyenne de ce formant pour toutes les voyelles du locuteur, (9) NEAREY2 transforme les valeurs de F0 et de formants en log, puis calcule une distance par rapport aux moyennes de F0 et de chaque formant de toutes les voyelles du locuteur, (10) GERSTMAN calcule une distance du formant par rapport à la valeur minimum du formant, puis la divise par l'étendue maximale de ce formant pour toutes les voyelles du locuteur, (11) NORDSTRÖM propose une normalisation qui prend en compte la taille moyenne du conduit vocal de locuteurs hommes et femmes en intégrant cette taille dans un coefficient qui permet de transformer les valeurs brutes en Hertz, (12) MILLER est une méthode similaire à NORDSTRÖM, puisqu'elle prend en compte les caractéristiques de la taille du conduit vocal de locuteurs hommes et femmes.

A partir d'un corpus de voyelles du néerlandais produites par 160 locuteurs (80 femmes et 80 hommes), Adank et al. (2004) effectuent une comparaison statistique de ces méthodes de normalisation pour déterminer quelles méthodes préservent au mieux les catégories vocaliques,

tout préservant les variations dialectales des locuteurs et en minimisant les variations liées aux caractéristiques physiologiques des locuteurs (liées en particulier à des tailles variables du conduit vocal). Ils en concluent que les méthodes qui permettent de répondre au mieux à ces questions sont les méthodes qui utilisent les caractéristiques extrinsèques des voyelles et qui conservent les relations entre ces voyelles à l'intérieur du système vocalique.

2.1.3.5.Notre étude : comparaison de six méthodes de normalisation et application à notre problématique

Afin de déterminer quelle méthode de normalisation est la plus pertinente pour notre étude, nous avons décidé de comparer certaines des méthodes de normalisation présentées dans la table 4.14, en les appliquant aux productions de parole enregistrées dans le cadre de l'étude pilote de Scarbel et al. (2012), que nous avons complétées par de nouveaux enregistrements : dans le cadre de cette étude, des productions de voyelles par des enfants avaient été enregistrées, nous avons décidé d'enregistrer des locuteurs supplémentaires adultes (deux hommes et deux femmes) en suivant la même procédure, pour déterminer quelle méthode de normalisation permet de neutraliser au mieux la variation liée à la taille du conduit vocal. Nous proposons une comparaison de plusieurs méthodes de normalisation utilisées dans les études de production de voyelles présentées dans le paragraphe §2.1.3.4 : données brutes en Hertz, données normalisées en Bark (Traunmüller (1990)), en scores-z (Lobanov (1971)), avec la méthode Nearey 1 (Nearey (1978)), Gerstman (1968), ainsi que notre proposition de calcul de scores-z adaptée de Lobanov (1971).

- **Corpus**

Les participants de cette étude sont les enfants normo-entendants de l'étude pilote de Scarbel et al. (2012) : sept enfants âgés de 6;11 à 8;1 ans (âge moyen : 6;9 ans), et quatre adultes normo-entendants (deux hommes de 40 et 45 ans, deux femmes de 39 et 20 ans), enregistrés avec la même procédure. La table 4.12 ci-dessous présente les caractéristiques (groupe, âge, sexe) des participants dont nous avons utilisé les productions pour notre étude comparative de méthodes de normalisation.

Sujet	Groupe	Âge	Sexe
NH1	Enfant	8,1	M
NH3	Enfant	9	F
NH4	Enfant	7,5	M
NH5	Enfant	7,4	F
NH6	Enfant	6	M
NH7	Enfant	11	F
NH8	Enfant	6,11	F
NH9	Adulte	40	M
NH10	Adulte	39	F
NH11	Adulte	20	F
NH12	Adulte	45	M

Table 4.12 : participants de notre étude de normalisation : enregistrements des enfants NH1 à NH réalisés dans le cadre de l'étude de Scarbel et al. (2012), enregistrements des adultes NH9 à NH12 réalisés par nous-mêmes.

La procédure d'élicitation se composait de deux tâches : une tâche de répétition de mots avec un modèle audio adulte suivie d'une tâche de production des mêmes mots sans modèle audio. Les voyelles-cibles étaient les dix voyelles orales du français /i, e, ɛ, y, ø, oe, a, u, o, ɔ/ en position initiale de mots, avec deux degrés de complexité phonétique ou lexicale de mots (dix mots

simples et dix mots plus complexes). Seules les productions jugées correctes perceptivement ont été analysées acoustiquement.

- **Méthodes de normalisation testées**

La table 4.13 ci-dessous présente les formules de toutes les méthodes de normalisation comparées dans cette section. Dans notre adaptation du score-z de Lobanov (1971), nous utilisons une moyenne des moyennes des valeurs des formants de chaque voyelle, pour éviter une pondération de la moyenne dans le cas où le nombre de formants moyennés diffère d'une personne à une autre : la formule de Lobanov (1971) suppose un nombre constant de données d'une voyelle à l'autre et d'un locuteur à l'autre, ce qui peut être problématique si le nombre d'occurrences de chaque voyelle n'est homogène. En effet, cette méthode (Lobanov, 1971) normalise les données en les plaçant dans un même espace maximal, mais un déséquilibre de la répartition des voyelles produites au sein de l'espace vocalique ne va pas permettre de superposer exactement les espaces vocaliques maximaux de chaque locuteur.

Méthode	**Formule**
Hertz	Données brutes
Bark (Traunmüller (1990)	$F_{ti}^{bark} = 26{,}81x(Fi/(1960+Fi))-0{,}53$
Score-z (Lobanov (1971))	$F_{ti}^{lobanov1} = (F_{ti} - \mu_{ti})/\delta_{ti}$ avec μ_{ti} moyenne des valeurs du formant F_{ti}
Nearey 1	$F_{ti}^{nearey1} = (F_{ti} - \mu_{Dti})$ avec valeurs de F_{ti} en log et μ_{dti} moyenne des log de F_{ti}
Gerstman	$999x(F_{ti} – F_{ti}^{min})/(F_{ti}^{max} – F_{ti}^{min})$
Notre adaptation du score-z de Lobanov	$F_{ti}^{lobanov2} = (F_{ti} - \mu_{ti})/\delta_{ti}$ avec μ_{ti} moyenne des moyennes des valeurs du formant F_{ti}

Table 4.13 : méthodes de normalisation testées (où F_t correspond au formant après transformation et i au numéro du formant)

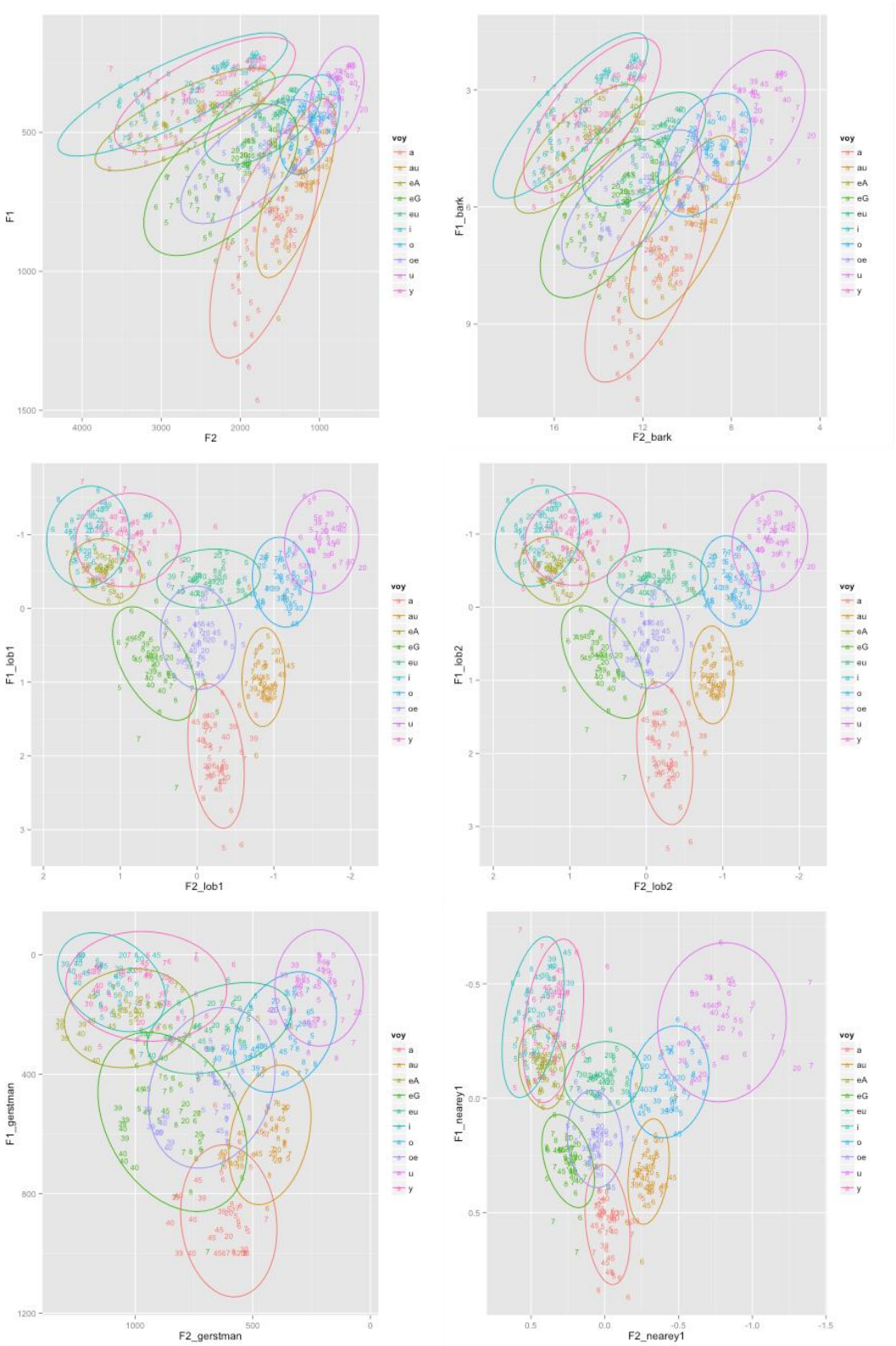


Figure 4.9 : Représentation spatiale F1-F2 des données de l'étude comparative des six méthodes de normalisation (de haut en bas et de gauche à droite : données en Hertz, en Bark, données normalisées en scores-z de Lobanov, avec notre adaptation du score-z de Lobanov, avec la méthode de Nearey 1 et de Gerstman), ellipses de confiance à 95%, les nombres correspondent aux âges des participants.

La figure 4.9 ci-dessus présente les données de notre corpus après application des méthodes de normalisation testées. Toutes les données acoustiques sont analysées avec le logiciel Praat

(Boersma et Weenink (2015)), et tous les calculs et les graphiques sont réalisés avec le logiciel R (R Development Core Team (2012)).

Une inspection visuelle montre clairement un certain nombre de différences entre les différentes méthodes de normalisation. Conformément aux études de Peterson et Barney (1952), Kent et Forner (1979) et Huber et al. (1999), la représentation des données en Hertz montre clairement une grande dispersion des données : les enfants ont les espaces vocaliques les plus grands, et les hommes ont les espaces vocaliques les plus petits. Nous observons également un chevauchement des catégories vocaliques. La représentation en Bark (Traunmüller (1990)) est assez similaire à la représentation en Hertz, le gain de cette méthode par rapport à l'utilisation des données brutes en Hertz est donc peu important. Les représentations des deux méthodes de scores-z (Lobanov (1971) et notre adaptation du score-z) montrent une faible dispersion intra-catégorie, puisque les tailles des ellipses sont les plus petites, en comparaison avec les autres méthodes. La distinction inter-catégorie est également meilleure pour ces deux méthodes, puisque les intersections entre les ellipses sont les moins nombreuses. La méthode Gerstman a une meilleure répartition des voyelles sur l'espace vocalique maximal que les méthodes Hertz et Bark, mais les ellipses sont de très grande taille et présentent de nombreuses intersections. La méthode Nearey1 semble réduire l'étendue de chaque ellipse sur F2 mais pas sur F1, ce qui la distingue des deux méthodes de scores-z, qui agissent sur les deux dimensions : il semble que la méthode Nearey1 produit une plus grande dispersion intra-catégorie que les méthodes de scores-z. Avec cette méthode, les intersections entre ellipses sont en outre moins importantes qu'avec la méthode Gerstman. Nous avons ensuite quantifié les performances de ces différentes méthodes.

- **Comparaison : analyses et résultats**

Notre objectif est de trouver la méthode de normalisation qui élimine au maximum la variation liée à la taille du conduit vocal et donc l'effet de l'âge et du genre sur la taille de l'espace vocalique (F1-F2) et la dispersion des voyelles sur F1 et F2, pour pouvoir ensuite comparer la production des enfants NH et CI d'âges différents.

Pour comparer ces six méthodes, nous avons calculé deux indices : l'indice 1 qui permet de quantifier la distinction inter-catégorie, qui correspond à la répartition des catégories vocaliques à l'intérieur de l'espace vocalique (indice 1 = surfaces des intersections des ellipses/ surface totale des ellipses) et l'indice 2 qui permet de quantifier la dispersion intra-catégories, ce qui correspond à la surface totale des ellipses par rapport à l'espace vocalique maximal (Indice 2 = surface totale des ellipses / ((F1max-F1min) x (F2max-F2min)).

La méthode de normalisation la plus adaptée à notre problématique est celle dont les indices sont les plus bas, puisqu'elle aura ainsi permis d'avoir une distinction inter-catégories la plus grande possible et une dispersion intra-catégories la plus faible possible.

Les résultats des comparaisons des deux indices de notre étude sont présentés dans la table 4.14 ci-dessous.

Méthode de normalisation	Indice 1 : Distinction inter-catégories	Indice 2 : Dispersion intra-catégories
Hz	0,917	1,408
Bark	0,829	1,465
Lobanov 1	0,331	0,732
Lobanov 2	**0,327**	**0,722**
Nearey 1	0,338	0,798
Gerstman	0,751	1,763

Table 4.14 : indices de distinction inter-catégories et de dispersion intra-catégories

Les résultats des calculs d'indices confirment la représentation graphique de la figure 4 : les données en Bark ne permettent pas de minimiser la distinction inter-catégories puisque les deux indices sont à peu près identiques en Bark et en Hertz (les intersections entre ellipses à l'intérieur de l'espace vocalique sont les plus grandes en Hertz) et la dispersion intra-catégories (les ellipses sont plus grandes avec la méthode de Gerstman, en Hertz et en Bark). Les méthodes qui permettent d'éliminer la variance liée à la taille du conduit vocal sont donc la méthode des scores-z de Lobanov (Lobanov (1971)) et notre adaptation de cette méthode, qui gomme la pondération éventuelle des moyennes utilisées pour le calcul des scores-z. C'est donc cette dernière méthode qui est la plus adaptée à l'étude comparée de la production de voyelles chez les deux groupes d'enfants NH et CI. En effet, cette méthode nous permet d'analyser la variation liée à des différences de contrôle articulatoire, tout en éliminant la variation de ces productions dues à l'âge et au sexe des enfants. Cette méthode est en particulier adaptable au fait que nos corpus ne sont pas tout à fait identiques pour tous nos sujets.

2.1.3.6. Conclusion et limites

Nous avons exposé dans ce chapitre les questions que posent l'étude de la production de voyelles chez des enfants dont les caractéristiques physiologiques diffèrent, ainsi que l'absolue nécessité de normaliser les données manipulées dans le cadre d'une étude sur la parole d'enfants.

Notre comparaison quantitative de six méthodes de normalisation a permis de confirmer pour l'étude du développement les résultats de la littérature sociolinguistique : la méthode la plus adaptée pour étudier les voyelles produites par des enfants porteurs d'implants cochléaires est celle des scores-z (Lobanov (1971)), puisqu'elle superpose les centres des espaces vocaliques de chaque locuteur et qu'elle replace chaque voyelle par rapport au centre de cet espace. Il est cependant envisageable que notre choix d'une méthode de normalisation qui élimine totalement l'effet de l'âge chronologique de chaque enfant sur la production des voyelles, masque également certains effets liés à l'évolution du contrôle articulatoire chez ces mêmes enfants, et en particulier certains effets liés à une évolution de la perception partielle des sons par l'enfant porteur d'implant cochléaire.

2.2. Consonnes : VOT, moments spectraux

Les consonnes peuvent être classées selon plusieurs modes articulatoires : mode de fonctionnement laryngien (voisé ou non-voisé, si les cordes vocales vibrent ou ne vibrent pas lors de la production de la consonne), mode de fonctionnement vélaire (oral ou nasal, si le voile du palais est relevé ou abaissé et laisse passer l'air par le nez), mode de fonctionnement articulatoire (occlusive ou constrictive, si le passage de l'air est totalement ou partiellement obstrué lors de la

production de la consonne). Ces modes articulatoires vont donc permettre de décrire la production des consonnes et de représenter le système consonantique correspondant.

Ainsi, plusieurs auteurs ont proposé des descriptions du système consonantique du français dans la littérature sous la forme de classements des catégories consonantiques par mode et par lieu d'articulation : Sten (1963), qui propose une description des consonnes reprise dans la base de données UPSID (UCLA Phonological Segment Inventory Database, 2014), Walter (1977) et Léon (2007). La Table 4.15 ci-dessous présente le système consonantique du français, établi d'après les descriptions de Sten (1963), Walter (1977) et Léon (2007).

	bilabiales	labio-dentales	(apico-) dentales	alvéolaires	post-alvéolaires	palatales	vélaires	uvulaires
occlusives	p b		t d				k g	
fricatives		f v		s z	ʃ ʒ		ɣ	ʁ
nasales	m		n			ɲ	ŋ	
latérales approximantes				l				
trills			r					R

Table 4.15 : système consonantique du français d'après Sten (1963), Walter (1977) et Léon (2007). Les cellules grisées correspondent aux consonnes qui ne sont pas présentes dans toutes les descriptions : /R/ est présente uniquement dans Sten (1963), /r/ et /ɣ/ dans Léon (2007) uniquement et /ʁ/ dans Léon (2007) et Walter (1977).

Les descriptions du système consonantique du français présentées sont globalement similaires et font état de 18 consonnes dans Sten (1963) et Walter (1977), et 20 consonnes pour Léon (2007) : six occlusives, six fricatives (sept dans Léon (2007)), quatre nasales, une latérale approximante et un trill (deux dans Léon (2007)). Les classifications de Sten (1963) et Walter (1977) s'intéressent aux phonèmes du français, alors que la description phonétique de Léon (2007) y ajoute deux allophones du /R/. On peut noter par ailleurs que le trill uvulaire /R/ proposé par Sten (1963) n'est pas présent chez Léon (2007) ni dans la description proposée dans le cadre de PFC (Projet « Phonologie du Français Contemporain »)[21], qui lui préfèrent la fricative uvulaire /ʁ/. Ces descriptions incluent également trois semi-consonnes /j/ /ɥ/ et /w/, que nous n'avons pas incluses dans notre étude expérimentale.

Comme notre étude ne s'intéresse qu'à la réalisation de certains contrastes phonologiques par des enfants sourds, porteurs d'implant cochléaire, nous avons choisi de restreindre notre étude à deux modes d'articulation : occlusives et fricatives, que nous allons décrire plus précisément dans les paragraphes §2.2.1 (occlusives) et §2.2.2 (fricatives).

2.2.1. Description du système consonantique du français : occlusives

Une consonne occlusive peut être définie comme un son produit avec obstruction totale du passage de l'air. Une occlusive est produite en trois temps : 1) une phase d'occlusion qui correspond à la mise en place des articulateurs sollicités pour la production de l'occlusive, 2) une phase de tenue des articulateurs, qui bloquent le passage de l'air, puis 3) une explosion qui correspond au relâchement des articulateurs qui laissent passer le flux d'air.

Le système consonantique du français possède six occlusives orales, classées en trois paires selon leurs lieux d'articulation : /b/ et /p/ sont des occlusives bilabiales, produites par contact des

[21] Projet « Phonologie du Français Contemporain », site web http://www.projet-pfc.net, consulté le 18 avril 2016

deux lèvres ; /d/ et /t/ sont des apico-dentales, produites par contact de la pointe de la langue et des dents supérieures ; /g/ et /k/ sont des dorso-palatales ou dorso-vélaires, produites par contact du dos de la langue et du palais (palais dur lorsqu'elles sont suivies d'une voyelle antérieure ou voile du palais lorsqu'elles sont suivies d'une voyelle postérieure). La figure 4.10 ci-dessous présente les lieux d'articulation des consonnes occlusives du français.

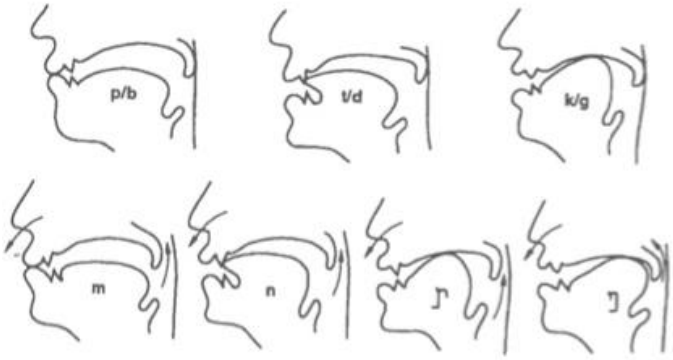


Figure 4.10 : schémas articulatoires des consonnes occlusives (orales et nasales) du français, tirés de Léon (2007)

Chaque paire d'occlusives comprend une occlusive voisée et une occlusive non-voisée. Le voisement correspond à la vibration des cordes vocales lors de la production d'un son, c'est un trait distinctif des consonnes du français : une consonne est dite « voisée » si elle est produite avec vibration des cordes vocales, et « non-voisée » si elle est produite sans vibration des cordes vocales. Les occlusives /b/, /d/, et /g/ sont voisées en français, alors que /p/, /t/ et /k/ sont non-voisées. Le début du voisement des occlusives en français, qui correspond à la mise en vibration des cordes vocales se produit avant l'explosion de la consonne, à la différence d'autres langues, comme par exemple l'anglais. Nous proposerons une description acoustique du voisement des consonnes occlusives dans le paragraphe §2.2.3.1 ci-après.

2.2.2. Description du système consonantique du français : fricatives

A la différence des occlusives, les fricatives sont produites avec une obstruction partielle du passage de l'air. Le bruit de friction perçu résulte d'un passage de l'air par un conduit vocal rétréci.

Comme pour les occlusives, le système consonantique du français possède six fricatives, qui peuvent être classées par paires, selon trois lieux d'articulation : /f/ et /v/ sont des labio-dentales, produites avec un rapprochement de la lèvre inférieure vers les dents supérieures ; /s/ et /z/ sont des (pré-dorso-)alvéolaires, produites avec un rapprochement du dos de la langue vers les alvéoles ; /ʃ/ et /ʒ/ sont des post-alvéolaires ou pré-dorso-palatales, produites avec un rapprochement du dos de la langue vers les alvéoles ou le palais dur. A ces trois paires de phonèmes s'ajoutent les consonnes vélaires [x] et [ɣ] et uvulaires [χ] et [ʁ] qui sont trois variantes contextuelles du /R/ présentes en français et que nous n'inclurons pas dans notre étude de production. La figure 4.11 ci-dessous présente les lieux d'articulation des consonnes fricatives du français.

Figure 4.11 : schémas articulatoires des consonnes fricatives du français, tirés de Léon (2007)

Les fricatives sont produites avec ou sans vibration des cordes vocales, et sont donc soit voisées dans le cas de /v/, /z/ et /ʒ/, soit non-voisées dans le cas de /f/, /s/ et /ʃ/. Nous proposerons une description acoustique des fricatives dans le paragraphe §2.2.4 ci-après.

2.2.3. Caractéristiques acoustiques des consonnes occlusives

Nous avons décrit le système consonantique du français dans les paragraphes §2.2.1 et §.2.2.2 de ce chapitre, à travers une présentation de quelques éléments de phonologie du français et de quelques caractéristiques articulatoires des consonnes (occlusives et fricatives). Le système consonantique du français se caractérise par une opposition de voisement, de lieu et de mode d'articulation. Dans cette partie, nous allons présenter les caractéristiques acoustiques des consonnes occlusives du français, et en particulier l'importance du « Voice Onset Time » (ci-après VOT ou « délai d'établissement du voisement ») dans la description acoustique des occlusives du français. Calliope (1989) propose une description acoustique des occlusives à travers un enchaînement d'évènements acoustiques, visibles sur un spectrogramme, qui correspondent aux trois phases de la production des occlusives proposées au paragraphe §2.2.1) un silence « qui correspond à la phase de tenue articulatoire de l'occlusion complète du conduit vocal »[22], suivi 2) d'une barre d'explosion (ou burst) qui est une « perturbation acoustique (onde de pression) de courte durée (5 à 35 ms), qui peut être intense »[23], 3) puis d'un bruit de friction dont « la durée dépend de la vitesse à laquelle les articulateurs s'écartent »[24], et enfin de transitions formantiques, définies comme les « déflexions fréquentielles rapides des formants que l'on observe au passage d'une consonne à une voyelle et réciproquement »[25].

2.2.3.1.Définition du VOT dans la littérature

Le Voice Onset Time (VOT) est une caractéristique articulatoire des consonnes occlusives, qui peut être mesurée acoustiquement. Il est défini par Lisker & Abramson (1964) comme « the interval between the release of the stop and the onset of glottal vibration, that is voicing », c'est-à-dire l'intervalle qui sépare le relâchement des articulateurs et la mise en vibration des cordes vocales, qui correspond donc à la phase d'explosion et la phase de frication qui lui est consécutive. Dans cette étude comparative de la réalisation du voisement dans 11 langues, Lisker & Abramson (1964) s'intéressent aux frontières catégorielles, lorsque pour le même lieu d'articulation, deux, trois ou quatre catégories de consonnes occlusives sont présentes dans le système consonantique de la langue. En effet, la longueur du VOT peut permettre de percevoir des contrastes de voisement (voisées ou non-voisées), d'aspiration (aspirées ou non-aspirées) ou de force articulatoire (lenis ou fortis) dans les occlusives. Cette étude met en évidence des différences dans la réalisation du voisement selon les langues (par exemple, le /b/ néerlandais a

[22] Calliope (1989), *La Parole et son traitement automatique*. Paris, Masson, page 91
[23] Ibid., page 92
[24] Ibid., page 92
[25] Ibid., page 92

un VOT négatif, ce qui indique que les cordes vocales se mettent à vibrer avant l'explosion consonantique, alors qu'en anglais celui-ci est positif, c'est à dire que les cordes vocales entrent en vibration après l'explosion consonantique), alors que ces consonnes sont perçues comme des consonnes voisées dans les deux cas. Cette première étude montre également que le VOT varie en fonction du lieu d'articulation dans les différentes langues. L'étude de Cho & Ladefoged (1999) qui compare la réalisation du VOT dans 18 langues du monde confirme ces résultats : cette étude, qui vise à distinguer caractéristiques du VOT universelles ou propres à chaque langue, met en évidence que la variation de la durée du VOT est une conséquence physiologique des différents lieux d'articulation. Elle est liée à la vitesse des mouvements des articulateurs et non à l'étendue de la surface de contact entre ces articulateurs. Calliope (1989) explique ce phénomène par une variation de la durée de la partie fricative du VOT selon le lieu d'articulation :

> « La partie fricative est celle où la source de turbulence est le lieu d'articulation. Sa durée dépend de la vitesse à laquelle les articulateurs s'écartent. Cette durée est brève pour les labiales, où les articulateurs sont très mobiles. Elle est longue pour les vélaires en raison de l'inertie de l'articulateur mobile qui est le dos de la langue. La durée pour les dentales est intermédiaire, car la pointe de la langue est un articulateur assez véloce. »[26]

Par ailleurs, Cho & Ladefoged (1999) indiquent que les contrastes d'aspiration ne sont pas directement liés à une durée pré-définie de VOT mais que les consonnes sont distribuées sur un continuum, et que l'augmentation du nombre de catégories d'occlusives dans la langue conduit à une plus grande étendue du VOT sur le continuum.

2.2.3.2. Caractéristiques du VOT en français

Il existe relativement peu d'études acoustiques sur la production d'occlusives en français. Dans ce paragraphe, nous présenterons les résultats de l'étude de Caramazza & Yeni-Komshian (1974), de celle d'Abdelli-Beruh (2009) et de celle de Nearey & Rochet (1994).

L'étude de Caramazza & Yeni-Komshian (1974) est la première à s'intéresser au voisement en français, à travers une comparaison de la durée du VOT dans deux dialectes du français, parlés par des locuteurs monolingues : français standard de France et français canadien. Cette étude s'intéresse aux influences de l'anglais canadien, avec lequel le français canadien est en contact. Les résultats de cette étude montrent que les occlusives voisées du français de France sont produites avec un VOT quasiment toujours négatif aux trois lieux d'articulation testés (consonnes bilabiales, alvéolaires et vélaires) et que la frontière catégorielle du voisement se situe systématiquement à 0 ms pour les trois paires de consonnes. En revanche, les occlusives voisées du français canadien sont produites avec un VOT positif aux trois lieux d'articulation, et la frontière catégorielle entre consonnes voisées et non-voisées ne se situe pas à 0 ms comme chez les locuteurs français, mais elle varie d'un lieu à l'autre et d'un locuteur à l'autre. Cette différence de VOT entre les deux groupes de locuteurs de France et du Canada est statistiquement significative. Caramazza & Yeni-Komshian (1974) constatent que la réalisation du voisement par les locuteurs francophones canadiens est très proche de celle de locuteurs monolingues anglophones canadiens en anglais, ce qu'ils expliquent par la proximité géographique de ces deux

[26] Ibid., page 92

groupes de locuteurs. La durée négative ou positive du VOT est donc une caractéristique distinctive du voisement pour les locuteurs du français de France alors que ce n'est pas le cas pour les locuteurs du français du Canada.

L'étude d'Abdelli-Beruh (2009) s'intéresse à l'effet du lieu d'articulation sur la durée de l'occlusion, du VOT et de l'intervalle non-voisé (qui correspond à la durée totale de l'occlusion et du VOT) dans des productions de mots monosyllabiques CVC par des locuteurs du français standard parisien. Dans cette étude, le VOT (« short-lag VOT ») est mesuré entre l'explosion consonantique (burst) et le début de la transition formantique du F2 de la voyelle qui suit la consonne étudiée : cet intervalle correspond à la durée du burst et de la portion fricative qui lui est consécutive. Avec ces mesures, le VOT des consonnes voisées est donc toujours positif, ce qui diffère d'autres études (notamment Lisker & Abramson (1964) ou Caramazza & Yeni-Komshian (1974)) qui considèrent le VOT comme l'intervalle entre le moment où les cordes vocales entrent en vibration et l'explosion consonantique. Les résultats de cette étude indiquent un effet significatif du lieu d'articulation sur la durée d'occlusion pour les consonnes non-voisées /p/, /t/ et /k/ et voisées /b/, /d/ et /g/, de même que sur la durée du VOT. Abdelli-Beruh (2009) note par ailleurs que la durée du VOT est inversement proportionnelle à celle de l'occlusion : /p/ et /b/ ont le VOT le plus court, mais l'occlusion la plus longue. Ces résultats semblent indiquer que la durée de la consonne est globalement constante, entre les trois lieux d'articulation.

Une dernière étude (Nearey & Rochet (1994)) montre également un effet du lieu d'articulation sur la durée du VOT en anglais et en français, à la fois pour les consonnes non-voisées /p/, /t/ et /k/ et voisées /b/, /d/ et /g/. Par ailleurs, cette étude s'intéresse également à l'effet du contexte vocalique sur la durée du VOT pour les consonnes non-voisées : des logatomes /CVk/ où V est l'une des neuf voyelles orales /i, ɛ, y, ø, œ, a, u, o, ɔ/ sont produits dans des phrases porteuses par des locuteurs du français standard parisien puis analysés. Les résultats indiquent que le VOT des alvéolaires et vélaires est plus long lorsque la consonne est suivie d'une voyelle haute /i/, /y/ et /u/, alors que le VOT des bilabiales est plus long lorsque la consonne est suivie de /y/, /ø/ et /u/, et de longueur intermédiaire lorsqu'elle est suivie de /i/. Dans cette étude, qui propose également les mêmes analyses pour les mêmes consonnes en anglais canadien, Nearey & Rochet (1994) mettent en évidence un effet du lieu d'articulation et du contexte vocalique sur la durée du VOT en anglais, mais ils notent que l'effet du contexte vocalique est moins marqué que pour le français.

2.2.3.3.VOT dans les études développementales

Les études présentées dans le paragraphe précédent ont permis de définir le VOT et de préciser dans quelle mesure son utilisation dans des études de production de parole peut permettre de caractériser acoustiquement des productions de consonnes chez des adultes, dans des langues différentes. Trois études chez l'enfant proposent des trajectoires d'acquisition du contrôle de la production de consonnes à partir de mesures de VOT : Whalen et al. (2007) étudient des productions de babillage chez six enfants francophones et six enfants anglophones de six à 12 mois, Kent & Forner (1980) étudient des productions de deux occlusives non-voisées chez des enfants anglophones de quatre, six et 12 ans, et Whiteside et al. (2003) s'intéressent aux productions d'occlusives chez 46 enfants anglophones de cinq, sept, neuf et 11 ans.

Whalen et al. (2007) proposent une étude longitudinale chez six enfants francophones et six enfants anglophones, en enregistrant leurs productions de babillage à six, neuf et 12 mois. Leurs résultats mettent en évidence un effet de l'âge sur la durée du VOT (le VOT augmente avec l'âge : les enfants produisent des occlusives avec un VOT de plus en plus long au cours de l'étude), un effet du lieu d'articulation (le VOT des vélaires est le plus long, suivi des alvéolaires, puis des bilabiales), ainsi qu'un effet de la langue sur la réalisation du voisement (les enfants francophones ont tendance à produire des occlusives voisées avec un VOT négatif, alors que les anglophones les produisent avec un VOT positif court, ce qui est le reflet de la L1 de chaque groupe). Whalen et al. (2007) déduisent en outre des mesures de VOT de leur étude que l'aspiration qui caractérise la production des consonnes non-voisées de l'anglais et qui correspond à un VOT long n'est pas encore acquise à l'âge de 12 mois par les enfants anglophones, mais que le prévoisement qui caractérise la production des consonnes voisées du français et qui correspond à un VOT négatif est en voie d'acquisition à l'âge de 12 mois par les enfants francophones. Ces résultats montrent que les enfants sont sensibles à la langue de leur environnement mais qu'ils ont des difficultés à maîtriser le contrôle articulatoire nécessaire à la production de ces contrastes de la L1.

L'étude de Kent & Forner (1980) propose de décrire l'évolution de la production de deux consonnes occlusives /k/ et /t/ par des enfants de quatre, six et 12 ans, en les comparant à celle d'adultes. A travers une tâche de répétition d'après un modèle adulte de 12 phrases contenant les deux segments cibles, ils mesurent les VOT de ces segments. Leurs résultats indiquent que le VOT de /k/ est très variable et diminue avec l'âge des enfants, alors que le VOT de /t/ ne subit pas d'effet linéaire de l'âge mais il est très variable chez les enfants les plus jeunes. Ainsi, cette étude semble indiquer que la production de ces consonnes de deux lieux d'articulation différents n'est pas acquise au même rythme par les enfants au cours du développement de leur système phonologique.

Whiteside et al. (2003) quant à eux étudient les productions de consonnes voisées /b/, /d/ et /g/, et non-voisées /p/, /t/ et /k/ par 46 enfants anglophones répartis en cinq groupes d'âges (5, 7, 9, 11 et 13 ans). Ils calculent un coefficient de variation, qui prend en compte la variation de la production par rapport à la moyenne de VOT de chaque enfant. Les VOT des trois lieux d'articulation ne sont pas traités indépendamment, les auteurs ne distinguant les consonnes que par le voisement. Les résultats indiquent un effet de l'âge sur la réalisation du voisement : les enfants de 5, 7 et 9 ans ont une variation de VOT plus grande et les enfants de 11 et 13 ans ont une variation de VOT plus faible. Cette étude ne permet pas de caractériser précisément le contraste de voisement par lieu d'articulation, mais indique qu'il se produit un changement dans la réalisation du voisement des occlusives entre 9 et 11 ans.

Ces quelques études développementales de la production de consonnes occlusives ont permis de mettre en évidence que la production du voisement pour différents lieux d'articulation est un processus relativement lent à mettre en place par les enfants et que la stabilisation interviendrait entre 9 et 11 ans. Ces études n'utilisent malheureusement pas la même méthodologie d'analyse, ce qui ne permet pas véritablement de les comparer entre elles, pour proposer une trajectoire de l'acquisition du voisement par les enfants. Par ailleurs, comme nous l'avions présenté au paragraphe précédent, la réalisation du voisement en français est particulière puisqu'il se fait par un contraste entre VOT positif et négatif, mais il n'existe à notre connaissance aucune étude

développementale du VOT chez des enfants francophones de plus d'un an, et nous manquons donc de repères quant à la réalisation du voisement par des enfants d'âge comparable à ceux de notre étude de production.

2.2.3.4.Mesure du VOT

Cette revue bibliographique sur le VOT a permis de montrer le manque d'études sur la réalisation du VOT par des enfants francophones mais également la pertinence de l'utilisation de cette mesure acoustique pour caractériser les productions de consonnes occlusives : le VOT permet d'étudier la réalisation du voisement pour chaque lieu d'articulation par les locuteurs francophones et la catégorisation du lieu d'articulation pour chaque groupe de consonnes voisées ou non-voisées.

Pour mesurer le VOT des consonnes /p/, /t/, /k/, /b/, /d/ et /g/ en contexte /#_i/ et /#_u/, nous utilisons un script Praat (Boersma, P. & Weenink, D. (2015)) qui, pour chaque son, à partir du fichier d'annotation TextGrid, extrait les temps où nous avons placé les points « dv » (pour début du voisement) et « b » (pour burst). Notre méthode d'annotation est décrite en détails au paragraphe §1.4.2 ci-dessus. Pour chaque mot du corpus utilisé pour l'étude des consonnes, nous avons placé le point « b » sur le pic du burst, ce qui correspond à la perturbation du signal acoustique lors de l'explosion consonantique, et nous avons placé le point « dv » lorsque le signal devient périodique, ce qui correspond à l'apparition d'une structure formantique avant le burst dans le cas d'une occlusive voisée et au début de la transition CV dans le cas d'une occlusive non-voisée. A partir de ces relevés, nous calculons la durée du VOT avec la formule : $VOT=T_{dv}-T_b$. La figure 4.12 ci-dessous présente ces mesures du VOT pour une bilabiale voisée et une bilabiale non-voisée.

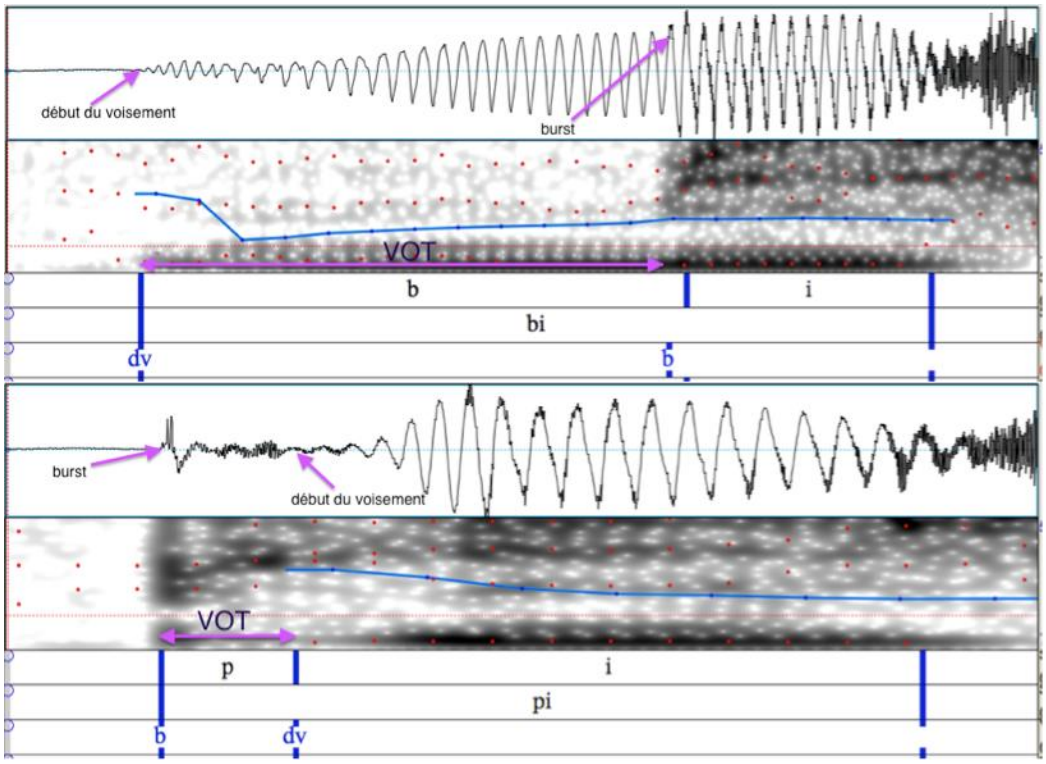


Figure 4.12 : mesures du VOT d'une bilabiale voisée (haut) dans le mot « bicyclette », et d'une bilabiale non-voisée (bas) dans le mot « pistache » produites en répétition par l'enfant NH29 (les échelles temporelles sont différentes pour les deux spectrogrammes)

2.2.4. Moments spectraux

2.2.4.1. Caractérisation acoustique des fricatives

Calliope (1989) définit la fricative comme un « bruit [qui résulte] d'une turbulence aérodynamique qui prend naissance en un ou plusieurs points du conduit vocal en raison de la présence d'un fort resserrement (ou constriction) ou d'un obstacle placé dans le flot d'air expiratoire »[27]. L'intensité du bruit de friction provoque une perturbation acoustique, visible sur le spectrogramme sous la forme de « stries verticales plus ou moins longues, d'intensité variable » (Calliope, 1989), qui peuvent être accompagnées d'une barre de voisement s'il s'agit de fricatives voisées (par exemple /v/, /z/ ou /ʒ/). Plusieurs caractéristiques acoustiques permettent d'analyser une fricative : le centre de gravité qui correspond à la zone de fréquences excitée lors de la production de la fricative, la dispersion du bruit, la (dis-)symétrie du bruit sur la zone de fréquence et l'amplitude du bruit. Ces quatre caractéristiques sont appelées les moments spectraux.

Calliope (1989) propose une caractérisation des fricatives selon le lieu d'articulation : /s/ et /z/ sont produites avec un bruit localisé dans les fréquences comprises entre 4000 et 8000 Hz, /ʃ/ et /ʒ/ avec un bruit de forte intensité et localisé entre 1600 et 2500 Hz, enfin /f/ et /v/ avec un bruit assez peu intense et localisé en deux points à 3500 Hz et 8000-10000 Hz. Ces valeurs sont celles de locuteurs adultes, mais des études de production de fricatives (Nissen & Fox, 2005, Körkkö, 2015, Nittrouer, 1995) indiquent que les fricatives sont réalisées différemment par les adultes et les enfants, avec des réalisations également différentes par les enfants selon le lieu d'articulation.

De façon générale, il existe dans la littérature assez peu d'études sur la production des fricatives et la plupart sont des études sur des populations d'enfants et d'adultes présentant des difficultés de production liées à l'audition (présentées dans le chapitre 3 ci-dessus) ou à une pathologie. Il existe également quelques études développementales utilisant des moments spectraux pour caractériser acoustiquement les productions de fricatives par les enfants. Nous allons proposer un aperçu de ces études développementales dans le paragraphe suivant.

2.2.4.2. Moments spectraux dans les études développementales

Plusieurs études développementales utilisent des mesures de moments spectraux pour caractériser l'acquisition des fricatives chez les enfants. Nous allons présenter les études de Nissen & Fox (2005), Nittrouer (1995), Li et al. (2009) et Körkkö (2015) qui caractérisent des productions de fricatives et leur évolution lors de l'acquisition de la langue maternelle par les enfants.

Nissen & Fox (2005) étudient des productions des fricatives /f, θ, s, ʃ/ par 30 enfants de trois à six ans de mots monosyllabiques CV(C)(C) dans des phrases porteuses, enregistrés lors d'une tâche de répétition. Cette étude chez des enfants anglophones a pour but de proposer une description acoustique de la production typique de fricatives non voisées chez des enfants en la comparant à une production d'adultes, d'étudier les effets de caractéristiques individuelles telles que l'âge, le genre, le lieu d'articulation, le contexte vocalique sur la réalisation des fricatives, puis d'étudier les combinaisons de paramètres acoustiques qui permettent de caractériser les fricatives. Nissen & Fox (2005) prennent en compte plusieurs mesures, dont quatre moments spectraux : 1) la moyenne spectrale (1er moment), qui correspond au centre de gravité dans d'autres études, et

[27] Ibid., page 101

qui est mesurée en Hz, puis transformée en ERB, 2) la variance spectrale ou écart-type (2ème moment), 3) le coefficient de dissymétrie ou « skewness » (3ème moment), et 4) le coefficient d'aplatissement ou « kurtosis » (4ème moment). La figure 4.13 ci-dessous présente les résultats de l'étude de Nissen & Fox (2005).

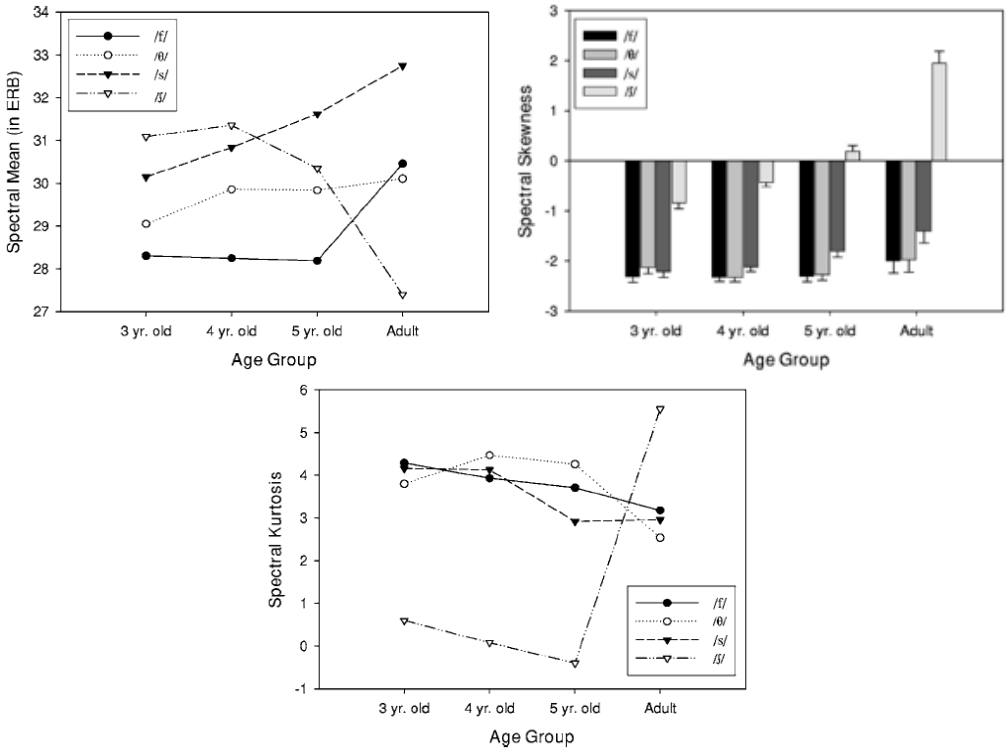


Figure 4.13 : représentation des 1er (haut, gauche), 3ème (haut, droite) et 4ème (bas) moments spectraux, dans l'étude de Nissen & Fox (2005)

Nissen & Fox (2005) montrent notamment un effet du lieu d'articulation sur le centre de gravité des fricatives chez adultes et enfants, ainsi qu'un effet de l'interaction lieu-âge sur ce centre de gravité. Il est important de noter que les trajectoires d'évolution vers la production adulte ne sont pas les mêmes selon les fricatives : le centre de gravité de /f/ est stable chez les trois groupes d'enfants, mais inférieur à celui des adultes, celui de /θ/ augmente légèrement pour atteindre un niveau similaire à celui des adultes, dès quatre ans, celui de/s/ augmente fortement et de façon constante vers la cible adulte, enfin, celui de /ʃ/ a la trajectoire inverse aux autres fricatives, puisqu'il est très élevé chez les enfants les plus jeunes, puis il décroit vers la cible adulte. Cette étude met également en évidence un effet du lieu d'articulation sur l'écart-type qui distingue fricatives sibilantes /s/ et / ʃ/ des non sibilantes /f/ et /θ/ : les sibilantes ont un écart-type plus petit que les non-sibilantes, ce qui correspond à une concentration d'énergie autour du centre de gravité pour les sibilantes. Nissen & Fox (2005) mettent également en évidence un effet du lieu d'articulation sur le coefficient de dissymétrie qui est plus élevé pour les /ʃ/ chez les adultes et tous les enfants, ainsi qu'un effet de l'âge (entre adultes et enfants, mais pas entre les groupes d'enfants d'âges différents) et de l'interaction âge-lieu d'articulation sur le coefficient de dissymétrie. Enfin, les résultats de Nissen & Fox (2005) indiquent un effet du lieu d'articulation (/ʃ/ vs les autres lieux), de l'âge et de l'interaction lieu-genre sur le coefficient d'aplatissement qui est globalement plus bas chez les adultes, sauf pour /ʃ/. Cette étude montre donc que le développement articulatoire n'est pas stabilisé et n'a pas atteint une maturation comparable à celle

des adultes, pour les enfants de cette étude, et les auteurs concluent que ce développement continue même après l'âge de sept ans. Cette étude montre également qu'il existe une différence marquée entre consonnes sibilantes et non-sibilantes dans le développement du contrôle articulatoire par les enfants.

L'étude de Nittrouer (1995) s'intéresse aux productions de deux fricatives /s/ et /ʃ/ chez des enfants de trois, cinq et sept ans, en les comparant à celles d'adultes. Cette étude s'intéresse notamment à la réalisation des fricatives selon le lieu d'articulation et le contexte vocalique. La figure 4.14 ci-dessous présente les résultats de cette étude pour le premier moment spectral (centre de gravité).

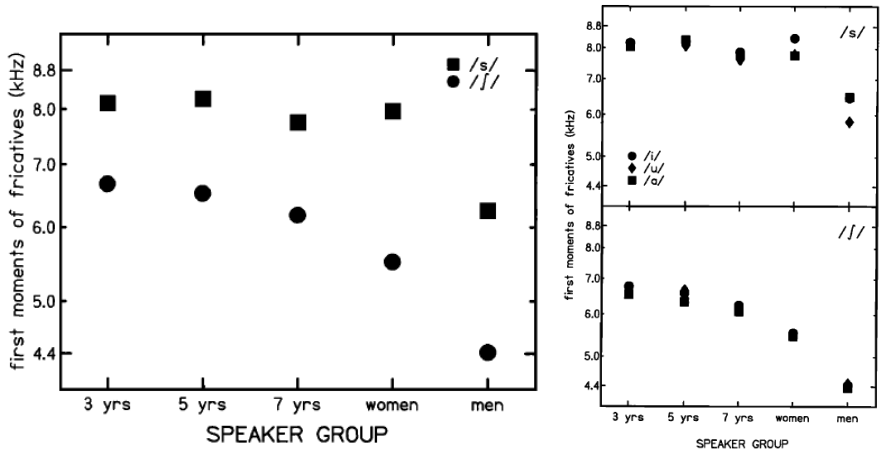


Figure 4.14 : représentation du 1er moment spectral par lieu d'articulation (gauche) et par lieu et contexte vocalique (droite), dans l'étude de Nittrouer (1995)

Les résultats de cette étude montrent un effet de l'âge (adultes vs enfants), du lieu d'articulation et du contexte vocalique sur le centre de gravité : le centre de gravité est plus élevé chez les enfants que chez les adultes, plus élevé pour /s/ que pour /ʃ/ chez les adultes et les enfants, et plus élevé en contexte /#_i/ puis /#_u/ puis /#_a/. Par ailleurs, la différence entre centres de gravité de /s/ et /ʃ/ est plus grande chez les adultes que chez les enfants. Nittrouer (1995) indiquent un effet de l'âge et du lieu d'articulation sur le coefficient de dissymétrie (skewness) : les fricatives produites par les adultes ont un coefficient positif plus élevé que celles qui sont produites par les enfants, et le coefficient de dissymétrie de /ʃ/ est positif et plus élevé que celui de /s/. Enfin, les résultats de cette étude font état d'un coefficient d'aplatissement plus plat pour /ʃ/ que pour /s/ pour tous les groupes d'enfants et d'adultes, ce qui indique que /ʃ/ est produit avec une concentration d'intensité plus faible que /s/. Pour Nittrouer (1995), la production de fricatives d'enfants de sept ans n'a pas encore atteint une maturité de contrôle comparable à celle des adultes, ce qui serait lié à une capacité de constriction des articulateurs en cours d'acquisition, qui ne permet pas encore aux enfants de cet âge de produire distinctement /s/ et /ʃ/. Cette étude propose également une caractérisation de l'évolution de la production des occlusives /k/ et /t/ par les mêmes enfants : à la différence de ce qui se passe pour les fricatives, les enfants de sept ans ont atteint une production des occlusives similaire à celle des adultes. Nittrouer (1995) en déduit que la coordination des gestes articulatoires nécessaires à la production des occlusives serait acquise avant la coordination des gestes utilisés pour la production des fricatives.

Li et al. (2009) proposent une comparaison de l'acquisition du lieu d'articulation de deux fricatives par des enfants anglophones américains et japonophones de deux et trois ans : la

fricative alvéolaire/s/ et la fricative (alvéo-)palatale /ʃ/ de l'anglais ou /ɕ/ du japonais. Pour cette étude, les enfants sont enregistrés dans une tâche de répétition de mots qui contiennent la fricative cible en position initiale de mot. Leurs productions sont analysées acoustiquement (moments spectraux) et par le biais de transcriptions phonétiques fines. Les productions des deux groupes d'enfants sont également comparées à celles d'adultes enregistrés selon la même procédure. Les résultats de l'étude indiquent qu'il existe une plus grande différence de localisation du centre de gravité entre les lieux d'articulation chez les adultes anglophones que chez les adultes japonophones. Pour les deux groupes en revanche, l'analyse de l'écart-type de /s/ et de /ʃ-ɕ/ indique une plus grande variance pour /s/, qui est plus diffus que /ʃ-ɕ/. Le coefficient d'aplatissement est à l'inverse chez les deux groupes : plus plat pour /ʃ/ que /s/ chez les anglophones alors qu'il est plus plat pour /s/ que pour /ɕ/ chez les japonophones. Les résultats des enfants indiquent une moins bonne distinction entre les deux lieux d'articulation que ce qui était observé chez les adultes, mais cette distinction se traduit cependant par des moments spectraux statistiquement différents entre les deux consonnes pour les enfants chez qui ce contraste est présent perceptivement. Li et al. (2009) montrent des différences entre les réalisations des deux types de consonnes par les deux groupes d'enfants : en combinant analyses perceptives (qui correspondent aux transcriptions phonétiques) et analyses acoustiques (avec les quatre moments spectraux), il ressort de cette étude que les enfants anglophones produisent /s/ plus correctement que /ʃ/ alors que les enfants japonophones produisent /ɕ/ plus correctement que /s/. Cette étude contribue à montrer que l'acquisition des fricatives reflète les caractéristiques de la langue maternelle des enfants, et qu'à l'âge de deux ou trois ans, le contrôle articulatoire nécessaire à une production comparable à celle des adultes n'est pas encore acquis, ce qui confirme les résultats de l'étude de Nittrouer (1995) réalisée auprès d'enfants plus âgés, notamment la distinction du lieu d'articulation qui se traduit par une distance plus grande entre les centres de gravité de /s/ et /ʃ/ chez les adultes, et que les enfants n'ont pas encore acquise.

La dernière étude qui s'intéresse à l'acquisition de fricatives est celle de Körkkö (2015). Cette étude propose une caractérisation acoustique de la variation allophonique de /s/ en contexte /i_i/, /y_y/, /u_u/, /æ_æ/ et /ɑ_ɑ/, et l'évolution de cette variation dans l'acquisition de la langue maternelle. Les enfants de cette étude sont regroupés par groupe d'âge (six, neuf et onze ans) et leurs productions sont comparées à celles d'adultes. Dans cette étude, deux moments spectraux (centre de gravité et écart-type) sont utilisés pour caractériser acoustiquement les productions enregistrées. Les résultats indiquent un effet de l'âge et du sexe sur la réalisation du centre de gravité, mais également un effet du contexte vocalique. En contexte /i_i/, le centre de gravité est plus élevé alors qu'en contexte /u_u/, celui-ci est plus bas. Körkkö (2015) observe en outre une variation moins grande de la hauteur du centre de gravité pour tous les contextes vocaliques chez les enfants de neuf et onze ans que chez les adultes, alors que cette variation est similaire à celle des adultes chez les enfants de six ans, sauf pour les contextes /æ_æ/ et /ɑ_ɑ/. Il ressort de cette étude que les adultes sont plus sensibles aux effets de la coarticulation et du contexte vocalique que les enfants, chez qui ces effets sont moins marqués : la production des fricatives n'aurait pas atteint une production d'adultes, même à l'âge de 11 ans, et ce processus d'acquisition se ferait en dents de scies pendant cette période. D'après Körkkö (2015), ces effets du contexte sur la réalisation de la fricative /s/ pourraient être liés à un effet de variation plus important de la production de voyelles chez les enfants.

Ces quelques études centrées sur l'acquisition de la production de fricatives ont mis en évidence une variation assez importante des enfants par rapport aux adultes et une trajectoire d'acquisition qui ne semble pas linéaire, comme ce qui pouvait être le cas dans les études développementales centrées sur la production d'occlusives. Cette acquisition est influencée par le lieu d'articulation pour Nissen & Fox (2005) et Nittrouer (1995), par la langue maternelle pour Li et al. (2009) ou encore par le contexte vocalique pour Nittrouer (1995) et Körkkö (2015). Notre étude de production se propose d'étudier ces facteurs ainsi que l'effet de l'audition sur l'acquisition du mode d'articulation fricatif chez des enfants sourds, porteurs d'implant cochléaire et normo-entendants. Dans le paragraphe suivant, nous allons détailler notre méthodologie de mesure des moments spectraux.

2.2.4.3.Mesure des moments spectraux

Pour mesurer les moments spectraux des consonnes /f/, /s/ et /ʃ/ en contexte /#_i/ et /#_u/, nous adoptons une méthodologie similaire à celle de Li et al. (2009). Dans un premier temps, nous utilisons un script Praat (Boersma, P. & Weenink, D. (2015)) qui, pour chaque son, à partir du fichier d'annotation TextGrid, extrait les temps où nous avons placé les points « dv » (pour début du voisement) et « fr » (pour début de la fricative). Notre méthode d'annotation est décrite en détails au paragraphe §1.4.2 ci-dessus. Pour chaque mot du corpus utilisé pour l'étude des consonnes, nous avons placé le point « fr » au début de la perturbation du signal acoustique de la fricative, et nous avons placé le point « dv » lorsque le signal devient périodique, ce qui correspond à l'apparition d'une structure formantique, qui survient toujours après le point « fr », puisque toutes les fricatives que nous analysons sont non-voisées. A partir de ces relevés, nous calculons la durée de la fricative avec la formule : durée=T_{dv}-T_{fr}. A partir de cette mesure, nous calculons le point temporel correspondant au centre de la fricative. A partir de ce centre, nous définissons puis extrayons une fenêtre de 40 ms (20 ms avant et 20 ms après le centre de la fricative). A partir de cette fenêtre, nous extrayons les quatre moments spectraux suivants : centre de gravité, écart-type, coefficient de dissymétrie et coefficient d'aplatissement. La figure 4.15 ci-dessous présente ces mesures des moments spectraux pour une fricative labio-dentale.

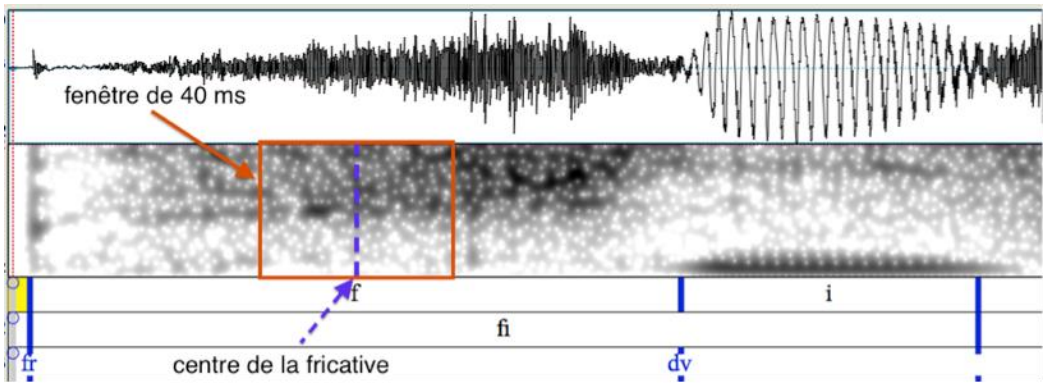


Figure 4.15 : fenêtre utilisée pour mesurer les quatre moments dans le mot « ficelle » produit en répétition par l'enfant NH24

2.3. Coarticulation

La coarticulation est définie par Noiray et al. (2013) comme « the articulatory overlapping of a sound on another one »[28]. La parole est donc constituée d'un enchaînement de sons qui

[28] « […] le chevauchement articulatoire d'un son sur un autre. » (notre traduction)

interagissent les uns avec les autres lors de leur production. Ainsi, parallèlement à l'étude des différents segments (voyelles, consonnes) individuellement, nous proposons une étude de la réalisation de la coarticulation par les deux groupes d'enfants, normo-entendants et porteurs d'implant cochléaire. Nous avons décidé, pour caractériser cette coarticulation, d'utiliser la mesure de l'équation de locus, dont nous allons proposer une description dans la suite de ce paragraphe.

2.3.1. Définition de l'équation de locus dans la littérature

Plusieurs études ont étudié le rôle de la transition du deuxième formant (F2) dans la production des consonnes occlusives dans des séquences CV ou VCV. Delattre et al. (1955) ont proposé une représentation schématique des transitions de F1 et F2 : ils ont défini la notion de « locus », qui leur permet de caractériser la production des consonnes occlusives et de proposer une modélisation de la transition consonne-voyelle pour chaque lieu d'articulation. La figure 4.16 ci-dessous présente les transitions des premier et deuxième formants (F1 et F2) pour trois lieux d'articulation d'occlusives (bilabilales, alvéolaires et vélaires) dans l'étude de Delattre et al. (1955).

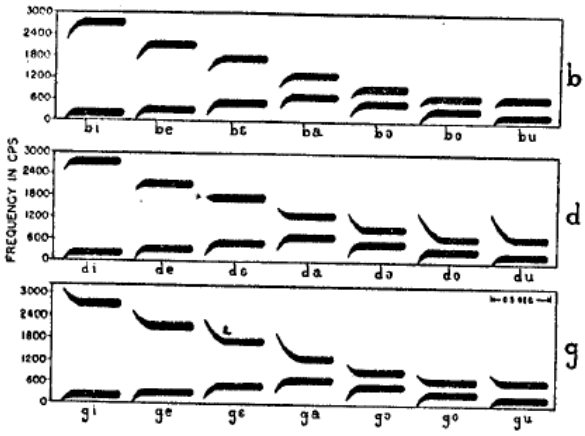


Figure 4.16 : représentation de stimuli synthétiques utilisés pour une expérience de reconnaissance de consonnes (bilabiales, alvéolaires et vélaires) par la transition du F2, tirée de Delattre et al. 1955

Pour les occlusives, Delattre et al. (1955) mettent en évidence que la transition du deuxième formant est un reflet des gestes articulatoires mis en œuvre lors de la production d'une consonne suivie d'une voyelle. Ils montrent par ailleurs que la transition du F1 ne joue aucun rôle dans la reconnaissance de la consonne, à la différence du F2 qui a son origine dans un « locus », défini comme une fréquence extrapolée dans le burst de l'occlusive, caractéristique de chaque lieu d'articulation de cette occlusive mais également d'autres types de consonnes de même lieu d'articulation :

> « The second-formant loci we found here are appropriate not only for b, d and g, but, more broadly for the three places of production (bilabial, alveolar and velar) that these stop consonants represent. »[29]

Ces travaux sur les transitions consonne-voyelle sont poursuivis par Lindblom (1963), à travers une étude sur la réduction vocalique dans des productions de séquences CV. Dans cette étude,

[29] « Les loci du deuxième formant que nous avons trouvés ici correspondent non seulement à b, d et g mais, plus largement, aux trois lieux de production (bilabial, alvéolaire et vélaire) que ces occlusives représentent. » (notre traduction)

Lindblom (1963) s'intéresse à l'effet de la centralisation des voyelles et de la coarticulation CV sur le phénomène de réduction vocalique et aux propriétés dynamiques de l'articulation de voyelles, de durées différentes et de contextes consonantiques différents. Alors que Delattre et al. (1955) utilisaient la transition CV pour caractériser la production de consonnes, Lindblom (1963) considère cette même transition CV pour caractériser la production de la voyelle. Les résultats de cette étude indiquent que les F1 et F2 de la cible de la voyelle (qui correspond au point le plus bas des formants de la voyelle) sont indépendants du contexte consonantique et de la durée de la voyelle.

L'équation de locus permet de caractériser acoustiquement la transition du F2 dans une séquence CV. Les équations de locus peuvent être définies de la façon suivante : « Locus equations are linear regressions of the onset of F2 transitions on their offsets measured in the vowel nucleus »[30], c'est-à-dire une régression linéaire entre le F2 mesuré au début de la transition CV et le F2 mesuré au milieu du noyau vocalique (qui correspond à la cible vocalique définie par Lindblom (1963)).

L'étude de Fowler (1994) met en évidence des pentes différentes selon les lieux d'articulation des occlusives, qui sont représentées sur la figure 4.17 ci-dessous, pour les trois lieux d'articulation (bilabiales, alvéolaires et vélaires).

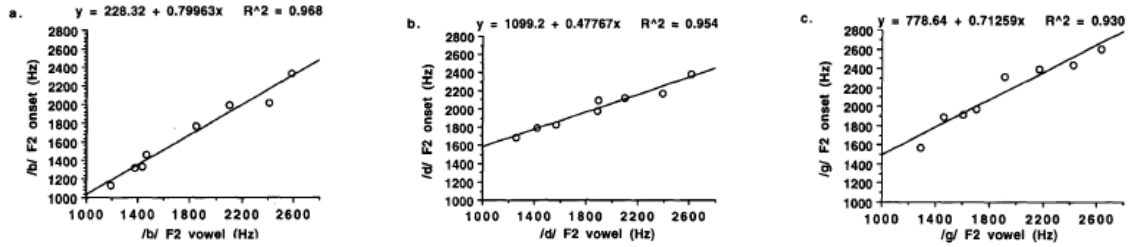


Figure 4.17 : équations de locus chez des adultes pour /b/ (gauche), /d/ (centre) et /g/ (droite), valeurs moyennes du groupe dans des séquences CV, avec huit voyelles différentes dans l'étude de Fowler (1994)

Cependant, l'utilisation de l'équation de locus pour décrire les consonnes est nuancée par Fowler (1994) : en effet, si l'équation de locus permet de distinguer les lieux d'articulation des occlusives entre elles, elle n'est pas suffisante pour identifier la consonne, à la fois par son mode et son lieu d'articulation et l'étude de Fowler (1994) met en évidence que le mode d'articulation a un effet sur l'équation de locus, dans laquelle le lieu est masqué par le mode d'articulation.

Ce résultat diffère de ceux de l'étude de Sussman & Shore (1999), qui proposent une comparaison des pentes d'équation de locus pour un seul lieu d'articulation (alvéolaire), mais cinq modes d'articulation différents (occlusives voisées et non-voisées, fricatives voisées et non-voisées et nasales) : les résultats de cette étude indiquent que les transitions de F2 sont similaires pour les séquences occlusives-voyelles et fricatives-voyelles, pour le même lieu d'articulation.

Par ailleurs, si l'étude de Fowler (1994) indique que les pentes d'équation de locus varient en fonction du lieu d'articulation, elle montre également une tendance différente pour les vélaires selon le contexte vocalique : la voyelle antérieure ou postérieure a un effet sur la pente de l'équation de locus des vélaires, qui est plus basse sous l'influence des voyelles antérieures et plus

[30] « Les équations de locus sont des régressions linéaires du début des transitions de F2 en fonction de leurs cibles, mesurées sur le noyau vocalique. » (notre traduction)

élevée sous l'influence des voyelles postérieures. Ce résultat est présenté dans la figure 4.18, ci-dessous.

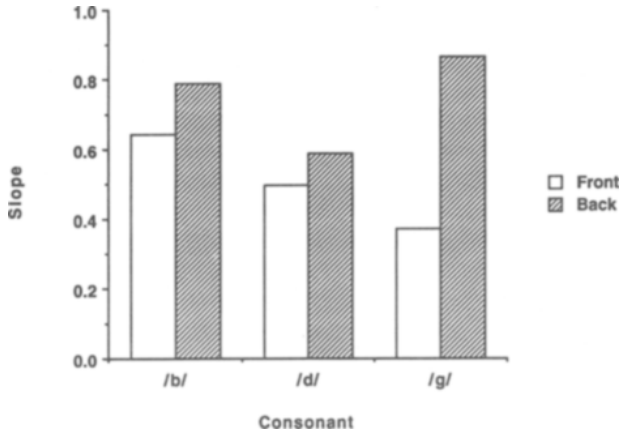


Figure 4.18 : comparaison des pentes (« Slope ») des équations de locus, en fonction de l'antériorité de la voyelle (« front » ou « back »), figure tirée de Fowler (1994)

Enfin, l'équation de locus est également utilisée pour quantifier le degré de coarticulation de séquences CV. La figure 4.19, tirée de Sussman & Shore (1999) présente le principe général de cette caractérisation du degré de coarticulation par l'équation de locus : une pente proche de 0 correspond au degré de coarticulation minimum, alors qu'une pente proche de 1 correspond au degré de coarticulation maximum. En d'autres termes, si la pente est proche de 0, la valeur de F2 mesurée au début de la voyelle reste constante, et ne subit aucune influence de la voyelle alors que si la pente est proche de 1, la valeur de F2 mesurée au début de la voyelle varie en fonction de la voyelle cible.

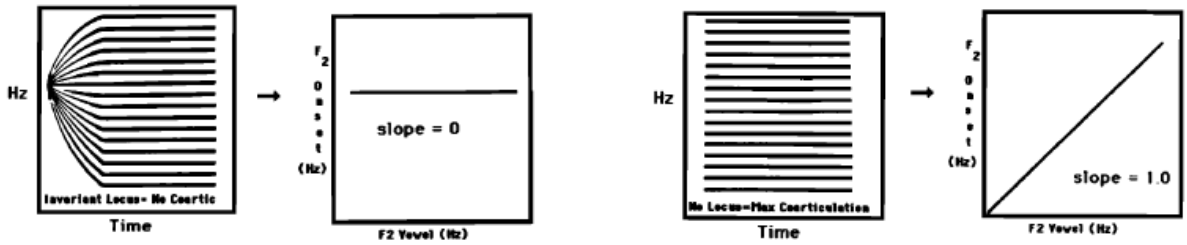


Figure 4.19 : schéma présentant le principe général de l'équation de locus, tiré de Sussman & Shore (1999)

Les résultats de l'étude de Löfqvist (1999) qui s'intéresse aux liens entre caractéristiques articulatoires et acoustiques (équation de locus) de la coarticulation indiquent que la pente de l'équation de locus est un bon prédicteur du lieu d'articulation mais pas du degré de coarticulation. En effet, la pente de l'équation de locus des alvéolaires, par exemple, est très différente de celle des vélaires (la pente de l'équation de locus pour les vélaires est beaucoup plus grande que celle des alvéolaires), alors que les mesures articulatoires ne montrent pas de différence d'amplitude des mouvements de la langue lors de la phase d'occlusion, pour ces deux lieux d'articulation.

Les études présentées dans ce paragraphe ont permis de proposer une définition de l'équation de locus et de son utilisation pour caractériser les influences entre consonnes et voyelles dans des séquences CV. L'équation de locus est utilisée par plusieurs auteurs pour étudier la coarticulation dans différentes conditions de production de parole : Krull (1989) compare la coarticulation dans

des productions de mots tirés d'un corpus de parole spontanée et des mêmes mots produits en isolation et Duez (1992) compare des productions de parole spontanée et de parole lue. Krull (1989) met en évidence un degré de coarticulation plus grand en parole spontanée (où les segments s'enchainent plus rapidement), et une diminution de l'excursion du F2 entre locus et nucleus lorsque le débit de parole augmente (ce qui correspond à une diminution de la durée de la voyelle). Cette étude montre également un effet de l'accent de mot sur le degré de coarticulation. Duez (1992) s'intéresse aux effets du style de parole, de la proéminence syntaxique et de la nouveauté du mot sur la pente de l'équation de locus. Les résultats obtenus dans cette étude indiquent des pentes d'équation de locus plus marquées en parole spontanée qu'en parole lue, ainsi qu'un effet de la proéminence syntaxique dans le discours spontané (les loci et nuclei sont plus proches dans les syllabes non-proéminentes que dans les syllabes proéminentes), ce qui est comparable aux résultats de Lindblom (1963) sur la réduction vocalique. Enfin, les mots connus sont produits avec des pentes plus fortes que les mots nouveaux et les mots lus.

A travers quelques études présentées dans ce paragraphe, nous avons proposé une définition de l'équation de locus et de son utilisation pour caractériser le lieu d'articulation des consonnes occlusives et du degré de coarticulation dans des séquences CV. Le paragraphe suivant est consacré à une présentation de quelques études développementales qui utilisent l'équation de locus pour étudier l'évolution de la production de consonne et de la coarticulation CV par les enfants.

2.3.2. Utilisation de la pente de l'équation de locus dans des études développementales

S'il semble admis par la plupart des auteurs que la pente de l'équation de locus permet de distinguer les occlusives entre elles, les liens entre cette pente et le degré de coarticulation semblent plus controversés et donnent lieu à des analyses différentes selon les auteurs. Cependant, et en gardant à l'esprit cette controverse, il nous semble important d'intégrer une analyse des transitions entre consonne et voyelle à notre étude sur la parole d'enfants sourds, porteurs d'implant cochléaire. Dans ce paragraphe, nous allons proposer un aperçu de quelques études développementales, qui s'intéressent aux transitions de F2 ou qui utilisent l'équation de locus pour décrire l'évolution de séquences CV par de jeunes enfants, pour justifier ce choix.

La première étude développementale qui utilise la transition du F2 pour décrire des trajectoires de coarticulation chez l'enfant est celle de Turnbaugh et al. (1985). Dans cette étude, les productions de consonnes par trois groupes de locuteurs sont comparées dans une tâche de répétition de séquences C_1VC_2 : trois enfants de trois ans, trois enfants de cinq ans et trois adultes. Les séquences C_1VC_2 sont symétriques (C_1 et C_2 sont identiques dans chaque séquence), et les consonnes sont /b/, /d/ ou /g/, avec les voyelles /i/ ou /u/. Les auteurs s'intéressent au F2 mesuré au début de la transition, et une moyenne est calculée pour chaque locuteur et chaque contexte CVC. Un degré de coarticulation en pourcentage est ensuite calculé à partir de ces mesures de F2. Les résultats de cette étude de Turnbaugh et al. (1985) indiquent un degré de coarticulation systématiquement plus faible pour les alvéolaires chez les trois groupes. Chez les enfants de trois ans, le degré de coarticulation des labiales est moins fort que celui des vélaires, alors qu'il sont similaires chez les enfants de cinq ans, et celui des vélaires est moins fort que celui des labiales chez les adultes. Les auteurs de cette étude notent par ailleurs que ces résultats ne

sont pas significativement différents entre les trois groupes. Ils en déduisent que les enfants acquièrent relativement tôt des patrons de coarticulation similaires à ceux des adultes.

Ces résultats contrastent avec ceux de Nittrouer (1993) : dans cette étude de la coarticulation, des productions de séquences schwa-consonne-voyelle dans des phrases porteuses sont enregistrées chez des enfants de trois, cinq et sept ans. Des mesures de durée (du schwa, de l'occlusion de la consonnes, du VOT, de la voyelle) et de fréquences de formants (F1 et F2) à plusieurs instants (30 ms avant l'établissement du voisement, dans les trois premières périodes voisées à partir du début du voisement, puis dans les dix périodes centrales de la voyelle) sont réalisées. Lors de la première mesure (30 ms avant l'établissement du voisement), les résultats sur le F2 indiquent que les enfants ont un F2 plus haut que les adultes, que celui-ci varie selon la qualité de la voyelle, et qu'il y a un effet de la consonne sur la hauteur du F2. A cet instant, Nittrouer (1993) note également que le F2 de /k/ subit une influence plus forte de la voyelle que le F2 de /t/, et que le F2 de /k/ est plus bas que celui de /t/, la différence étant significative pour tous les groupes. Pour la deuxième mesure de F2 (dans des trois premières périodes voisées à partir du début du voisement), le F2 subit des effets de l'âge, de la consonne et de la voyelle. Enfin, pour la troisième mesure (correspondant aux dix périodes centrales de la voyelle), le F2 subit un effet de l'âge, et de la voyelle. De ces mesures de F2, mais également des mesures de durées, Nittrouer (1993) déduit que le développement de la production et de la coordination des gestes articulatoires n'est pas uniforme, et que l'enchainement des gestes linguaux par les enfants subit des contraintes du contexte segmental plus fortes que les adultes.

Sussman et al. (1999) utilisent quant à eux l'équation de locus pour décrire l'évolution de la coarticulation chez un enfant de 7 à 40 mois, dans une étude longitudinale : ils mesurent le F2 au début de la transition CV et au centre de la voyelle, dans des séquences CV du babillage canonique aux premiers énoncés de cet enfant. La figure 4.20 ci-dessous présente l'évolution des pentes des équations de locus ainsi obtenues, sur toute la durée de cette étude. Pour les bilabiales, Sussman et al. (1999) notent une augmentation forte de la pente de l'équation de locus vers la fin de la première année, puis une augmentation plus modérée au cours de la deuxième année, et une stabilité au cours de la troisième année. Pour les alvéolaires, ils notent une forte diminution de la pente de l'équation de locus dans la première année, puis des pentes stabilisées à un niveau inférieur à celui des adultes dans la deuxième et la troisième année. Pour les vélaires, ils notent beaucoup de fluctuations des pentes d'équation de locus dans le babillage canonique, puis une stabilisation à partir de l'âge de 18 mois.

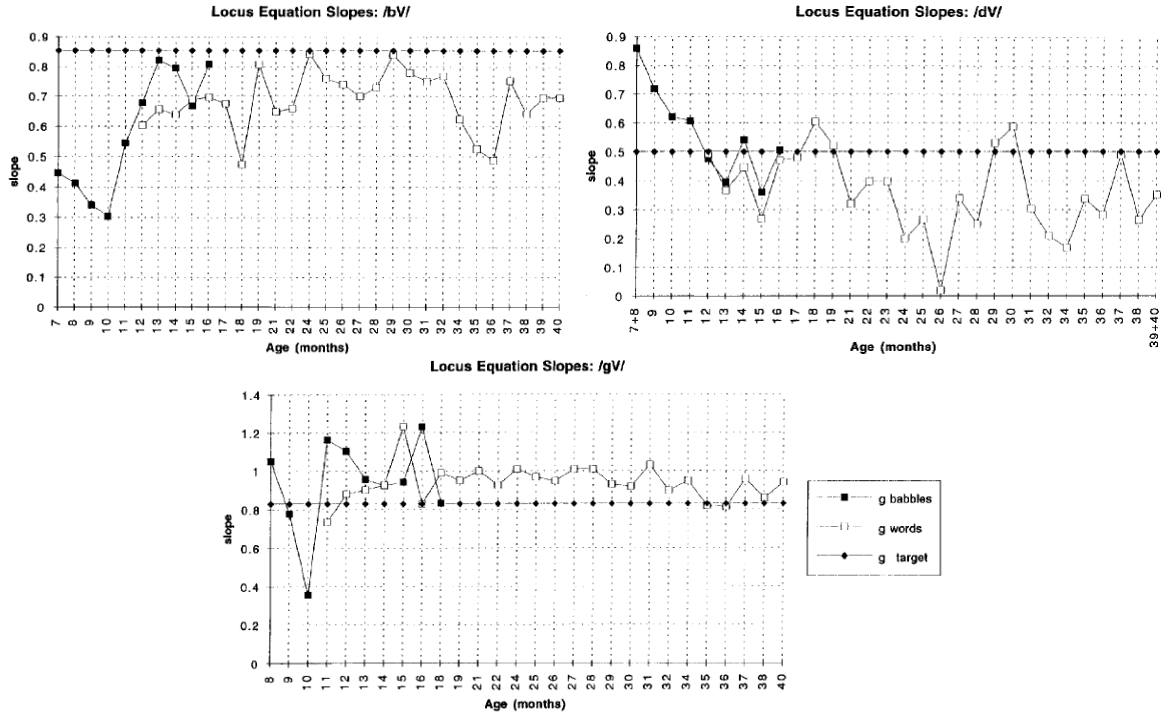


Figure 4.20 : évolution de 7 à 40 mois de la pente des équations de locus des bilabiales (en haut à gauche), des alvéolaires (en haut à droite) et vélaires (bas) dans l'étude de Sussman et al. (1999)

Les résultats de Sussman et al. (1999) indiquent des pentes plus faibles pour les alvéolaires, suivies des vélaires, puis des bilabiales chez l'adulte, et une tendance légèrement différente chez l'enfant vers la fin de l'étude, avec une pente à peine plus forte pour les vélaires que pour les alvéolaires à l'âge de 40 mois. On note que l'évolution des trois pentes n'est pas linéaire mais en dents de scie, ce qui indique un processus très variable dans les premières années de vie.

Gibson & Ohde (2007) proposent également une étude de la coarticulation dans les premières productions de six filles et six garçons, âgés de 17 à 22 mois, sur une période de six semaines. Des productions de séquences CV sont analysées, et des pentes d'équation de locus sont calculés pour les trois lieux d'articulation (bilabiales, alvéolaires et vélaires) à partir de mesures de F2 au début de la transition formantique et au centre de la voyelle. Les résultats de cette étude indiquent une différence significative entre les pentes des équations de locus : les alvéolaires ont les pentes les plus faibles, suivies des bilabiales, puis des vélaires. Ces résultats sont similaires à ceux de l'étude de Sussman et al. (1999), dans laquelle les vélaires avaient des pentes plus faibles que les bilabiales chez l'adulte et mais plus fortes chez l'enfant à la fin de l'étude longitudinale. Ces deux études portaient sur des productions en anglais.

Une étude de Noiray et al. (2013) propose une comparaison articulatoire et acoustique de la coarticulation chez des enfants de quatre et cinq ans et des adultes francophones. L'équation de locus est utilisée pour quantifier le degré de coarticulation selon le lieu d'articulation (bilabial, alvéolaire et vélaire). Les résultats de cette étude indiquent un lien entre données articulatoires et acoustiques : les adultes et les enfants ont des réalisations similaires de la coarticulation (les alvéolaires ont le degré de coarticulation le plus faible, suivies des vélaires, puis des labiales) mais les auteurs notent une grande variabilité inter-sujets à l'intérieur de chaque groupe. Les données articulatoires et acoustiques indiquent que la synergie entre dos de la langue et apex semble avoir été acquise par les enfants dès quatre-cinq ans, mais la grande variabilité des groupes pourrait

indiquer que cette synergie des gestes articulatoires n'est peut-être pas encore stabilisée chez ces enfants. En raison des liens entre données articulatoires et acoustiques, Noiray et al. (2013) indiquent que l'équation de locus peut être utilisée pour évaluer la présence de synergies des gestes moteurs chez l'enfant.

Cet aperçu de quelques études sur la coarticulation nous a permis de comprendre l'importance de la transition du deuxième formant dans l'identification des consonnes dans des séquences CV. Par ailleurs, si la plupart des auteurs s'accordent à dire que la pente de l'équation de locus est un bon indicateur du lieu d'articulation, il semble qu'il n'y ait pas de consensus quant à son utilisation pour quantifier le degré de coarticulation. Cependant, un nombre d'études conséquent indique une corrélation entre caractéristiques articulatoires de la coarticulation occlusive-voyelle et équations de locus. Nous proposons donc d'utiliser l'équation de locus pour comparer la réalisation de cette transition comme indicateur du lieu d'articulation et de la coarticulation chez les deux groupes d'enfants de notre étude de production de parole. Dans le paragraphe suivant, nous allons détailler notre méthode de calcul de la pente de l'équation de locus.

2.3.3. Calcul de la pente de l'équation de locus

Suivant la définition de l'équation de locus, proposée par Sussman & Shore (1999), au paragraphe §2.3.1 ci-dessus, nous avons choisi de mesurer le F2 1) au début du voisement, c'est-à-dire au premier pic de F0 suivant le burst de la consonne, et 2) au milieu de la voyelle. Nous avons choisi d'étudier 15 séquences CV différentes, constituées des consonnes /b/, /d/ et /g/, suivies des voyelles /i, e, u, o, a/ (la liste des mots est fournie dans le paragraphe §1.2.2.4, ce chapitre) dans la tâche de répétition et dans la tâche de dénomination. Les valeurs de F2 sont extraites automatiquement, puis vérifiées manuellement, pour corriger dans la mesure du possible toute erreur de détection.

La figure 4.21 ci-dessous présente un exemple de mesure de F2 réalisée sur une séquence /do/ dans le mot « dauphin », produit par l'enfant NH28.

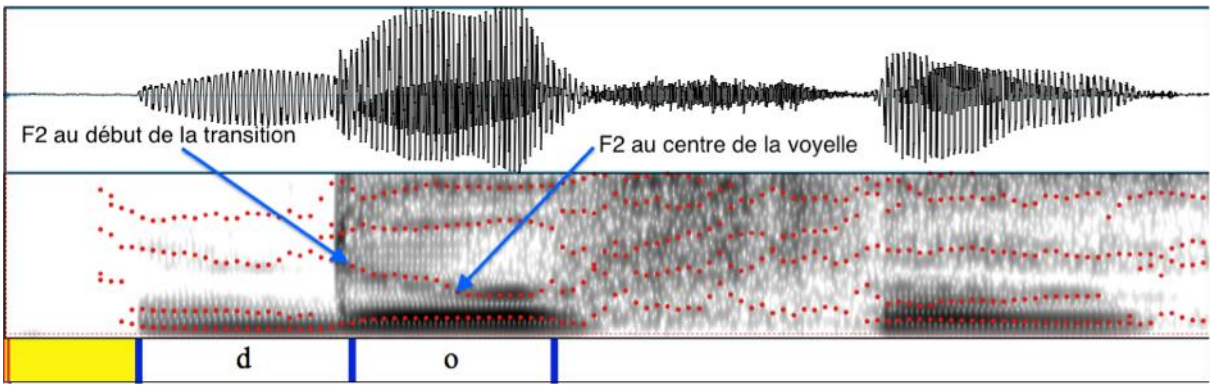


Figure 4.21 : mesure du F2, dans une séquence /do/ du mot « dauphin », produit par l'enfant NH28

Les valeurs de F2 ainsi obtenues ont ensuite été normalisées selon la méthode des scores-z de Lobanov (Lobanov (1971), voir paragraphe §2.1.3), de façon à éliminer la variation liée à l'âge chronologique.

Pour chaque enfant, chaque tâche et chaque lieu d'articulation, nous avons ensuite calculé une régression linéaire, en utilisant la fonction *lm()* avec le logiciel R : nous avons ensuite extrait la pente de chaque régression linéaire, et nous avons regroupé toutes ces pentes dans un tableau,

pour pouvoir les analyser. Les analyses statistiques du chapitre 5, paragraphe §5 seront faites à partir de ces pentes.

3. Analyses statistiques

Toutes les analyses statistiques et représentations graphiques de cette thèse ont été réalisées avec le logiciel R (R Development Core Team (2012)). Nous comparons des données (valeurs de formants, VOT et moments spectraux) qui se caractérisent par une variabilité inter-locuteurs que nous devons prendre en compte : par exemple le VOT, qui est une mesure de durée, sera réalisé différemment chez un enfant dont le débit de parole est lent ou rapide. Nous ne pouvions donc pas comparer des moyennes avec un test ANOVA classique, mais nous devions prendre en compte cet effet du sujet. Nous avons donc choisi d'utiliser des modèles linéaires mixtes, avec un effet aléatoire (ici, effet aléatoire du locuteur). Pour cela, nous utilisons la fonction *lme()* du package « lme4 » dans R.

Notre démarche est toujours la même : nous commençons par établir un modèle contenant toutes les variables (seules ou en interaction) qui nous semblent pertinentes pour expliquer la variation observée. A partir de ce premier modèle, nous effectuons une procédure de sélection de variables avec la fonction *stepAIC()* dans R, qui nous permet d'écarter les variables qui ne sont pas pertinentes dans la modélisation des données. Ensuite, nous utilisons la fonction *anova()* sur ce nouveau modèle, pour étudier l'effet de chaque variable explicative (ou VI, variable indépendante) sur la variable à expliquer (ou VD, variable dépendante). Les variables dépendantes sont les formants pour les voyelles (F1, F2, F3, distances euclidiennes F1-F2 et F2-F3), le VOT pour les occlusives, les moments spectraux pour les fricatives (centre de gravité, écart-type, coefficient de dissymétrie, coefficient d'aplatissement), et la pente de l'équation de locus pour la coarticulation. Les variables indépendantes sont le groupe (CI et NH), l'âge (âge chronologique, âge à l'implantation et âge auditif), la hauteur et l'antériorité (pour les voyelles), le lieu d'articulation et le contexte vocalique (pour les consonnes et la coarticulation), la consonne, la voyelle, la tâche (répétition ou dénomination) et la complexité articulatoire du mot (mots simples ou complexes). La variable « âge chronologique » est commune aux deux groupes, mais nous ne l'intégrons pas au modèle général pour étudier son effet global chez les deux groupes d'enfants : nous l'analysons ensuite séparément pour chaque groupe d'enfants (sous la forme de l'âge chronologique chez les enfants NH et de l'âge auditif et l'âge à l'implantation chez les enfants CI).

Le script suivant, codé sur le logiciel R (R Development Core Team, 2012) est un exemple de script que nous avons utilisé pour modéliser puis analyser les effets du lieu d'articulation et du groupe (désignés ici respectivement « C » et « groupe ») sur le premier moment spectral des fricatives (désigné ici « center ») :

```
modele <- lme(center ~ C*groupe, data = donnees, method="ML", random= ~1|sujet)
anova(modele)
```

Nous nous intéressons à la fois aux effets principaux des variables indépendantes, et à ceux de certaines interactions de ces variables. Lorsque les résultats indiquent un effet d'interaction significatif, nous effectuons un test post-hoc qui permet de tester plus précisément sur quelles associations de variables repose cet effet.

Le script suivant est un exemple de script de comparaisons multiples post-hoc, qui nous permet de tester séparément les effets des variables « lieu d'articulation » et « groupe » dont l'interaction a un effet significatif sur le centre de gravité des fricatives :

```
contrmat <- lsmeans(modele, pairwise~C|groupe, glhargs=list())[[2]]$linfct
comp_mult <- summary(glht(modele, linfct=contrmat))
comp_mult
```

Nous reproduisons cette démarche pour tester les effets des variables indépendantes propres à chaque groupe : âge chronologique pour les enfants NH, et âge à l'implantation et âge auditif pour les enfants CI. De la même façon, nous étudions l'effet de l'input auditif immédiat et l'effet de la complexité articulatoire du mot sur deux jeux de données différents, à partir des listes de mots données aux paragraphes §1.2.2.3 à §1.2.2.4 de ce chapitre, ci-dessus.

Si nos analyses statistiques sont globalement effectuées selon la même procédure, la nature des données nous amène à sélectionner des variables indépendantes légèrement différentes, en fonction de la variable dépendante. Nous présenterons préalablement à chaque analyse les variables dépendantes et indépendantes que nous avons incluses dans cette analyse.

Nous avons choisi, pour cette étude de production, de proposer des analyses comparatives des deux groupes plutôt que des descriptions des productions de chaque enfant, isolément. En effet, notre protocole nous a permis de collecter un nombre de données pour chaque enfant (un maximum de six occurrences de chaque voyelle, six occurrences de chaque occlusive, et six occurrences de chaque fricative pour chaque enfant), qui reste limité mais qui ne pouvait pas être plus élevé, du fait de l'âge et des difficultés éventuelles des enfants à rester concentrer pendant une durée trop longue.

CHAPITRE 5 : Étude de production (1ère étude expérimentale) : résultats

Dans la première partie de ce chapitre, nous allons présenter les résultats de notre analyse de la production de parole chez les enfants CI et NH. Nous nous intéressons principalement à des caractéristiques segmentales (production de voyelles, d'occlusives et de fricatives) mais également à la coarticulation anticipatoire dans des séquences CV. Les résultats des analyses suivront cet ordre. Nous avons choisi de proposer les résultats de chaque effet testé, pour chacune de ces caractéristiques.

1. Voyelles

1.1. Données analysées

Nous avons exclu des analyses sur les voyelles de ce chapitre, les productions de mots non-exploitables car non produits par les enfants ou non identifiables comme le mot cible (comme par exemple, le mot « vache », /vaʃ/ qui remplace le mot « bœuf », /bœf/), et les productions de voyelles non-exploitables, car dévoisées, élidées…

98,7% des mots produits par les enfants NH et 89,2% des mots produits par les enfants CI étaient exploitables et ont été inclus dans les analyses de ce paragraphe. Un test-t (test de Welch) met en évidence une différence significative entre les productions des deux groupes : $t(12,367)=3,26$, $p<.01$. Parmi ces mots exploitables, 80% des voyelles produites par les enfants NH et 74,4% de celles produites par les enfants CI ont été jugées exploitables et ont donc été incluses dans les analyses de ce paragraphe. Un test-t (test de Welch) indique que cette différence entre les deux groupes n'est pas significative ($t(23,845)=0,75$, $p>.05$).

Nos analyses se basent donc sur 1865 productions de voyelles exploitables, dont les caractéristiques par groupe (F1, F2 et F3 moyen, en Hz et en scores-z) sont résumées dans la table 5.1 ci-dessous.

Groupe	Voyelle	Nombre total de voyelles	Nombre de voyelles par enfant (max=6)	F1 moyen (en Hz)	F2 moyen (en Hz)	F3 moyen (en Hz)	F1 moyen (score-z)	F2 moyen (score-z)	F3 moyen (score-z)
CI	i	71	5,46	377 (87)	2847 (455)	3858 (425)	-1,03 (0,59)	1,41 (0,69)	0,95 (0,78)
CI	e	69	5,31	484 (60)	2636 (267)	3679 (292)	-0,27 (0,42)	1,08 (0,36)	0,58 (0,55)
CI	ɛ	69	5,31	622 (109)	2483 (288)	3609 (398)	0,70 (0,61)	0,85 (0,40)	0,39 (0,66)
CI	a	73	5,62	717 (126)	1890 (275)	3469 (335)	1,38 (0,71)	-0,08 (0,38)	0,14 (0,60)
CI	ɔ	74	5,69	620 (97)	1578 (198)	3383 (405)	0,69 (0,66)	-0,57 (0,29)	-0,05 (0,69)
CI	o	75	5,77	483 (77)	1179 (303)	3130 (538)	-0,29 (0,55)	-1,20 (0,45)	-0,54 (0,97)
CI	u	69	5,31	386 (90)	1182 (305)	2680 (54)	-0,96 (0,66)	-1,19(0,44)	-1,46 (1,10)
CI	y	72	5,54	400 (78)	2049 (464)	3398 (483)	-0,84 (0,54)	0,17 (0,72)	-0,01 (0,89)
CI	ø	54	4,15	518 (82)	1814 (457)	3398 (391)	-0,06 (0,49)	-0,21 (0,68)	-0,01 (0,78)
CI	œ	62	4,77	611 (115)	1777 (248)	3439 (280)	0,63 (0,67)	-0,26 (0,35)	0,08 (0,43)
NH	i	120	6,00	393 (70)	2880 (323)	4056 (404)	-1,13 (0,44)	1,38 (0,44)	1,25 (0,72)
NH	e	118	5,90	498 (54)	2702 (303)	3682 (343)	-0,36 (0,33)	1,08 (0,41)	0,53 (0,57)
NH	ɛ	120	6,00	633 (86)	2540 (298)	3641 (315)	0,64 (0,49)	0,83 (0,39)	0,46 (0,52)
NH	a	119	5,95	776 (143)	2009 (204)	3423 (433)	1,64 (0,81)	-0,04 (0,30)	0,03 (0,77)
NH	ɔ	118	5,90	655 (103)	1619 (199)	3443 (318)	0,78 (0,59)	-0,67 (0,31)	0,08 (0,55)
NH	o	120	6,00	514 (67)	1160 (215)	3094 (590)	-0,26 (0,35)	-1,42 (0,30)	-0,56 (1,05)
NH	u	118	5,90	420 (90)	1209 (321)	2700 (668)	-0,95 (0,67)	-1,35 (0,53)	-1,33 (1,19)
NH	y	119	5,95	443 (66)	2186 (340)	3217 (316)	-0,75 (0,42)	0,26 (0,53)	-0,35 (0,60)
NH	ø	111	5,55	515 (55)	2048 (271)	3377 (338)	-0,25 (0,34)	0,03 (0,38)	-0,05 (0,62)
NH	œ	114	5,70	628 (85)	1954 (229)	3371 (298)	0,63 (0,53)	-0,11 (0,33)	-0,06 (0,51)

Table 5.1 : valeurs moyennes des trois premiers formants (avec écart-type) par groupe, par voyelle, en Hz et en scores-z.

Nous ne pouvons pas proposer une comparaison inter-groupes de la taille de l'espace vocalique en Hz qui traduirait des différences liées à la taille du conduit vocal des enfants (et donc à leur âge chronologique), ou en scores-z qui n'est pas pertinente puisque nous avons choisi une méthode de normalisation dont le but est de replacer les voyelles produites sur le même espace maximal F1-F2 ou F2-F3, de façon à éliminer la variation de la taille de l'espace vocalique liée à l'âge (une description détaillée de la méthode de normalisation choisie est proposée au chapitre 4, paragraphe §2.1.3). Nous proposons une comparaison de la catégorisation des voyelles, à travers une étude comparative de la dispersion de chaque catégorie en utilisant une distance euclidienne en deux dimensions (F1-F2 et F2-F3) de chaque occurrence de voyelle par rapport au centre de la catégorie vocalique correspondante. La table 5.2 ci-dessous propose un résumé des distances euclidiennes moyennes ainsi obtenues, pour chaque voyelle et pour chaque groupe.

Groupe	Voyelle	Nombre total de voyelles	Nombre de voyelles par enfant (max=6)	Distance euclidienne moyenne F1-F2 (score-z)	Distance euclidienne moyenne F2-F3 (score-z)
CI	i	71	5,46	1,36 (0,45)	1,89 (0,70)
CI	e	69	5,31	1,00 (0,34)	1,31 (0,49)
CI	ɛ	69	5,31	1,65 (0,48)	1,12 (0,47)
CI	a	73	5,62	2,24 (0,62)	0,88 (0,42)
CI	ɔ	74	5,69	1,80 (0,43)	1,14 (0,43)
CI	o	75	5,77	1,71 (0,35)	1,81 (0,46)
CI	u	69	5,31	1,68 (0,40)	2,06 (0,78)
CI	y	72	5,54	0,79 (0,46)	1,01 (0,68)
CI	ø	54	4,15	1,09 (0,54)	1,00 (0,70)
CI	œ	62	4,77	1,60 (0,55)	0,86 (0,36)
NH	i	120	6,00	1,23 (0,33)	2,04 (0,61)
NH	e	118	5,90	0,97 (0,26)	1,30 (0,45)
NH	ɛ	120	6,00	1,56 (0,43)	1,12 (0,39)
NH	a	119	5,95	2,48 (0,73)	0,91 (0,40)
NH	ɔ	118	5,90	1,90 (0,49)	1,23 (0,32)
NH	o	120	6,00	1,86 (0,28)	2,00 (0,50)
NH	u	118	5,90	1,81 (0,55)	2,21 (0,69)
NH	y	119	5,95	0,59 (0,32)	0,67 (0,44)
NH	ø	111	5,55	0,72 (0,31)	0,74 (0,43)
NH	œ	114	5,70	1,52 (0,45)	0,73 (0,37)

Table 5.2 : valeurs moyennes des distances euclidiennes F1-F2 et F2-F3 par rapport au centre de la catégorie vocalique (avec écart-type) par groupe, par voyelle, en scores-z.

De la même manière, nous nous intéressons à la convergence formantique (ou focalisation) qui caractérise certaines voyelles orales du français. A cet effet, pour chaque voyelle produite, nous avons calculé les différences F2-F1 et F3-F2 en Bark (en utilisant la formule de Traunmüller, 1990, donnée au chapitre 4, paragraphe §2.1.3.4), pour pouvoir étudier les différences de réalisation de focalisation par les deux groupes d'enfants. La table 5.3 propose un résumé des moyennes par groupe et par voyelle de ces deux différences. Les cases colorées correspondent aux valeurs inférieures à 3,5 Bark, qui indiquent que les deux formants convergent pour ne former qu'une seule zone formantique.

Voyelle	Nombre total de voyelles	Nombre de voyelles par enfant (max=6)	Différence F2-F1 moyenne (en Bark)		Différence F3-F2 moyenne (en Bark)	
			Enfants NH	Enfants CI	Enfants NH	Enfants CI
i	71	5,46	11,44 (0,85)	11,45 (1,64)	2,13 (0,75)	1,98 (1,07)
e	69	5,31	10,07 (0,93)	10,04 (0,88)	1,98 (0,79)	2,12 (0,72)
ɛ	69	5,31	8,56 (1,00)	8,51 (1,20)	2,32 (0,68)	2,39 (0,67)
a	73	5,62	5,99 (1,13)	5,96 (1,33)	3,44 (0,95)	4,00 (0,99)
ɔ	74	5,69	5,40 (0,94)	4,65 (1,42)	4,97 (0,85)	6,43 (1,44)
o	75	5,77	4,34 (1,13)	5,50 (1,01)	6,36 (1,48)	5,00 (0,78)
u	69	5,31	5,36 (1,87)	5,55 (1,85)	5,24 (1,85)	5,39 (1,90)
y	72	5,54	9,13 (1,30)	8,99 (1,97)	2,58 (0,99)	3,40 (1,35)
ø	54	4,15	8,07 (0,97)	7,08 (2,08)	3,28 (0,92)	4,30 (1,87)
œ	62	4,77	6,85 (0,78)	6,36 (1,46)	3,58 (0,82)	4,36 (1,05)

Table 5.3 : valeurs moyennes des différences F2-F1 et F3-F2 (avec écart-type) par groupe, par voyelle, Bark (en couleur : cases correspondant à des convergences formantiques, F_{i+1}-F_i < 3.5 Bark)

Nous avons choisi de n'étudier que l'effet du groupe et de la catégorie vocalique sur les différences F2-F1 et F3-F2, car les analyses sont réalisées à partir d'une transformation en Bark qui est une échelle qui permet de rendre compte de la perception de la voyelle, mais qui n'isole pas suffisamment la variation liée à la taille du conduit vocal d'enfants d'âges variables (voir le chapitre 4, §2.1.3). Les deux groupes étant appariés en âge chronologique, nous pouvons en effet étudier l'effet du groupe sur la réalisation de la convergence formantique chez des groupes d'enfants d'âges globalement équivalents mais pas l'effet de l'âge sur cette convergence formantique : la manipulation de données en Bark ne nous permettrait pas d'étudier l'évolution des différences entre formants successifs, caractéristiques de certaines voyelles, qui seraient liées notamment à une perception plus fine des focalisations formantiques (Ménard et al., 2007) par les enfants au cours du temps (qui évoluent donc avec l'âge chronologique, l'âge à l'implantation ou l'âge auditif), tout en faisant abstraction de la variation de la taille du conduit vocal qui est directement liée à l'âge chronologique.

Pour analyser les productions des voyelles, nous avons fait le choix d'utiliser des modèles linéaires mixtes qui prennent en compte à la fois des effets fixes et un effet aléatoire du sujet. Les variables à expliquer (ou variables dépendantes) ainsi que les variables explicatives (ou variables indépendantes) produisant les effets fixes sont détaillées dans la table 5.4 ci-dessous.

Variables dépendantes	Premier formant (F1), mesuré en Hz, puis transformé en scores-z
	Deuxième formant (F2), mesuré en Hz, puis transformé en scores-z
	Troisième formant (F3), mesuré en Hz, puis transformé en scores-z
	Dispersion F1-F2 : mesure de la distance euclidienne entre chaque occurrence de voyelle par rapport au centre de la catégorie vocalique correspondante sur le plan F1-F2, calculée en scores-z
	Dispersion F2-F3 : mesure de la distance euclidienne entre chaque occurrence de voyelle par rapport au centre de la catégorie vocalique correspondante sur le plan F2-F3, calculée en scores-z
	Différence F2-F1, en Bark
	Différence F3-F1, en Bark
Variables indépendantes	Groupe : CI ou NH
	V : Catégorie vocalique
	Hauteur : degré de hauteur (voyelles hautes, mi-hautes, mi-basses et basses)
	Antériorité : degré d'antériorité (voyelles antérieures, antérieures arrondies, postérieures, /a/)
	Age chronologique des enfants (NH)
	Age à l'implantation (enfants CI)
	Age auditif (qui correspond à la durée d'utilisation de l'implant pour les enfants CI)
	Tâche: tâche de répétition (R) ou tâche de dénomination (D)
	Complexité lexicale : mots simples (0) et mots complexes (1)
Variable ayant un effet aléatoire	Sujet

Table 5.4 : variables dépendantes, variables indépendantes et variable ayant un effet aléatoire, pris en compte dans l'analyse des catégories vocaliques

1.2. Effets du groupe et de la catégorie vocalique

La figure 5.1 ci-dessous présente la répartition des catégories vocaliques à l'intérieur des espaces acoustiques F1-F2 et F2-F3 pour chaque groupe (CI et NH). Une première observation de ces espaces semble indiquer que les catégories vocaliques produites par les enfants NH sont plus distinctes les unes des autres, puisqu'elles se caractérisent par des ellipses plus petites et moins de chevauchements inter-catégories que pour les enfants CI.

La comparaison des espaces F1-F2 indique notamment une tendance des enfants CI à produire des voyelles plus centrales sur l'axe du F1, qui correspond à la hauteur de la voyelle, mais également sur l'axe du F2 qui correspond à son antériorité. La représentation F2-F3 semble en revanche indiquer une répartition globalement similaire des voyelles dans l'espace, à l'exception notamment des trois voyelles antérieures arrondies /y/, /ø/ (notée /2/ en SAMPA sur les figures de ce chapitre) et /œ/ (notée /9/ en SAMPA sur les figures de ce chapitre) : ces trois voyelles ont en effet des ellipses notablement plus grandes chez les enfants CI que chez les enfants NH.

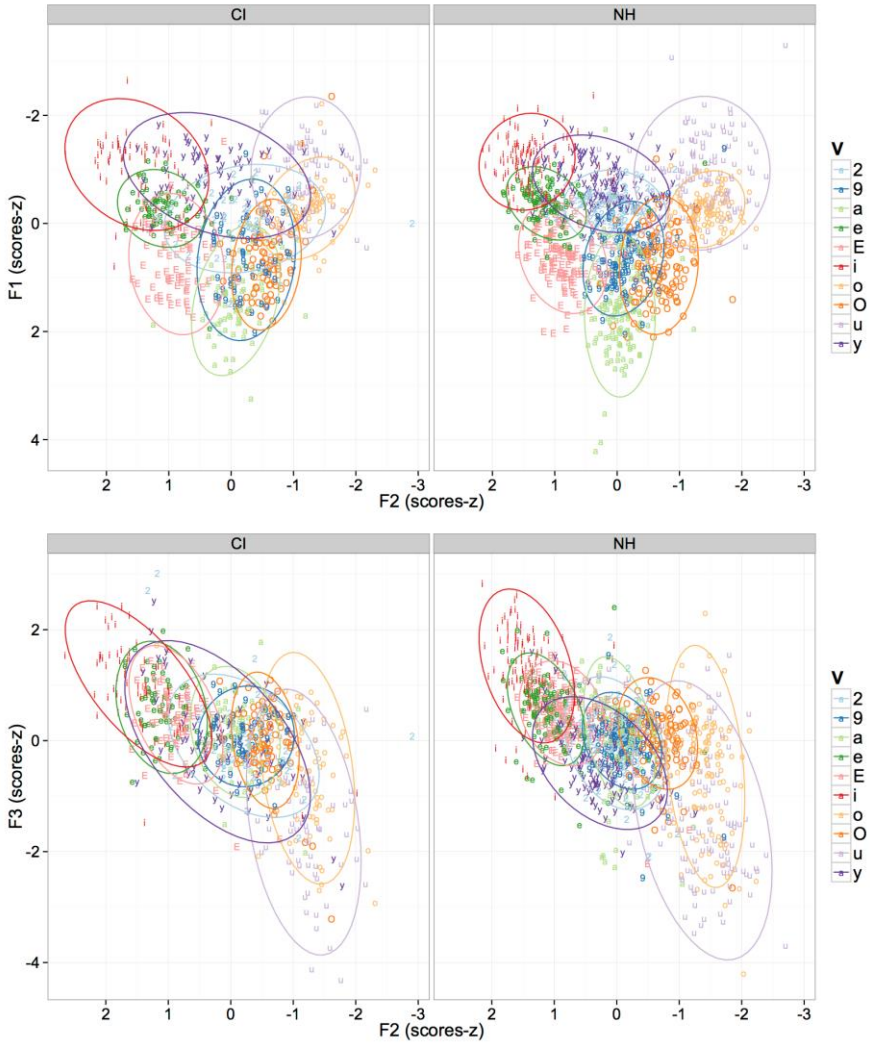


Figure 5.1 : représentation des catégories vocaliques dans l'espace F1-F2 (haut) et F2-F3 (bas) par groupe, en scores-z

A partir de ces premières constatations, nous allons étudier dans un premier temps l'influence de plusieurs facteurs sur la réalisation des voyelles : groupe, catégorie vocalique (ou degré de hauteur sur le F1, degré d'aperture sur le F2), seuls ou en interaction.

1.2.1. F1

Comme le F1 est lié à la dimension de hauteur des voyelles, nous avons choisi, dans un premier temps, de regrouper les voyelles par degré de hauteur : voyelles hautes /i, y, u/(notées /i, y, u/ en SAMPA dans les figures de ce chapitre), mi-hautes /e, ø, o/ (notées /e, 2, o/ en SAMPA dans les figures de ce chapitre), mi-basses /ɛ, œ, ɔ/(notées /E, 9, O/ en SAMPA dans les figures de ce chapitre), et basses /a/ (notée /a/ en SAMPA dans les figures de ce chapitre). La figure 5.2 ci-dessous présente les F1 mesurés en Hz puis transformés en scores-z pour chaque groupe et chaque degré de hauteur. La figure 5.2 ci-dessous semble indiquer que les deux groupes d'enfants ont des réalisations similaires de la hauteur, à l'exception du /a/ qui est réalisé avec un F1 plus haut (et donc un niveau plus bas ou un degré d'aperture plus grand) chez les enfants NH.

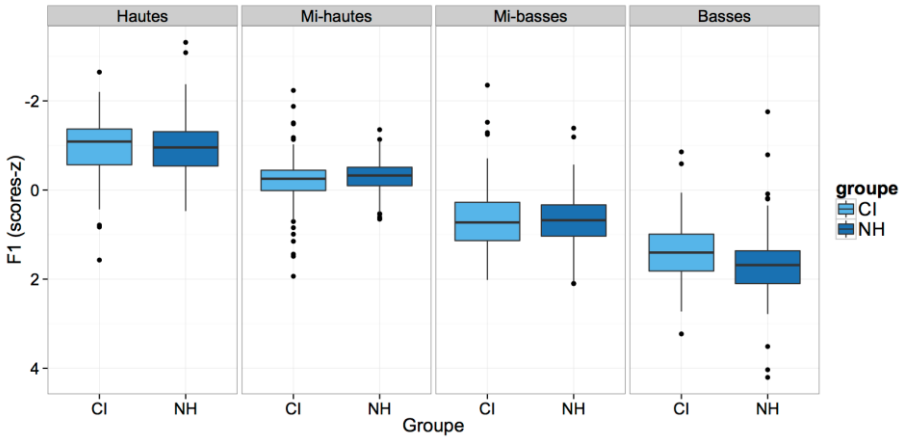


Figure 5.2 : représentation du F1 (en scores-z) par degré de hauteur de la voyelle, pour chaque groupe (enfants CI et NH).

Les résultats du test statistique réalisé selon la méthode d'analyse décrite au chapitre 4, paragraphe §3 sont présentés dans la table 5.5 ci-dessous.

	numDF	denDF	F-value	p-value
(Intercept)	1	1826	0.0537	0.8167
groupe	1	31	0.0170	0.8972
hauteur	3	1826	1356.2797	<.0001 ***
groupe:hauteur	3	1826	4.1174	0.0064 ***

Table 5.5 : tableau de résultats au test statistique sur toutes les données (effets du groupe et du degré de hauteur sur le F1)

Les résultats n'indiquent pas d'effet du groupe seul (F(1,31)=0,02, p>.05) mais un effet du degré de hauteur seul (F(3,1826)=1356,28, p<.001) et en interaction avec le groupe (F(3,1826)=4,12, p<.001). Un test post-hoc de cette interaction entre groupe et degré de hauteur met en évidence que la voyelle basse /a/ est la seule pour laquelle un effet de groupe se révèle significatif. Les résultats du test post-hoc sont résumés dans la table 5.6 ci-dessous.

Degré de hauteur	F1
	CI vs NH
Voyelles hautes	z=0,016, p>.05 (n.s.)
Voyelles mi-hautes	z=1,451, p>.05 (n.s.)
Voyelles mi-basses	z=-0,107, p>.05 (n.s.)
Voyelle basse	z=-3,222, p<.01 **

Table 5.6 : résumé du test post-hoc sur l'effet d'interaction entre degré de hauteur et groupe sur le F1 des voyelles

Un deuxième test post-hoc indique en outre que pour les deux groupes, les F1 des quatre degrés de hauteur sont significativement différents entre eux (Table 5.7).

Groupe	F1		
	Degrés de hauteur testés		
	Voyelles hautes vs mi-hautes	Voyelles mi-hautes vs mi-basses	Voyelles mi-basses vs basses
NH	z=-15,679, p<.001 ***	z=-23,171, p<.001 ***	z=-16,312, p<.001 ***
CI	z=13,234, p<.001 ***	z=-16,176, p<.001 ***	z=-9,245, p<.001 ***

Table 5.7 : résumé du test post-hoc sur l'effet d'interaction sur les F1 des voyelles par paires de degrés de hauteur

Nous nous sommes également intéressées aux réalisations du F1 par voyelle, pour comprendre si le regroupement par degré de hauteur peut masquer certains effets de groupe pour l'une ou

l'autre des voyelles. La figure 5.3 ci-dessous présente la réalisation du F1 (en scores-z) par voyelle pour chaque groupe, qui semble similaire pour les deux groupes, à l'exception du F1 de /ɛ/ et /y/, plus bas chez les enfants NH, et de celui de /a/, plus haut chez les enfants NH.

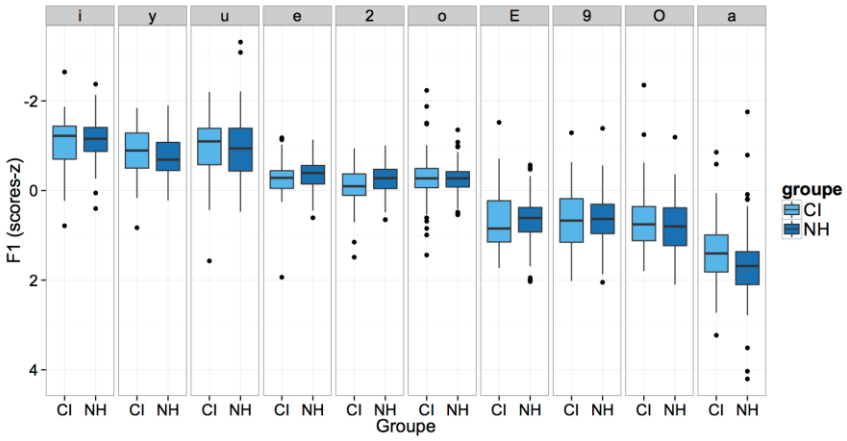


Figure 5.3 : représentation du F1 (en scores-z) par voyelle, pour chaque groupe (enfants CI et NH).

Comme pour le degré de hauteur, nous avons réalisé un test statistique de l'effet de la catégorie vocalique et du groupe sur la réalisation du F1. Les résultats de ce test sont présentés dans la table 5.8 ci-dessous.

	numDF	denDF	F-value	p-value
(Intercept)	1	1814	0.0548	0.8149
groupe	1	31	0.0173	0.8962
V	9	1814	465.5317	<.0001 ***
groupe:V	9	1814	2.3315	0.0131 *

Table 5.8 : tableau de résultats au test statistique sur toutes les données (effets du groupe et de la catégorie vocalique sur le F1)

Comme pour le test par degré de hauteur, nous n'observons pas d'effet significatif du groupe seul sur le F1 de la voyelle ($F(1,31)=0,02$, $p>.05$), mais un effet de la catégorie vocalique seule ($F(9,1814)=465,53$, $p<.001$) et de l'interaction entre groupe et catégorie vocalique ($F(9,1814)=2,33$, $p<.05$). Nous avons réalisé un test post-hoc de cette interaction, qui indique que la voyelle /a/ est la seule catégorie vocalique pour laquelle un effet significatif du groupe ($z=-3,264$, $p<.05$) est observé. Par souci de lisibilité, nous avons fait le choix de ne pas inclure les résultats complets de ce test post-hoc, pour lequel l'effet du groupe sur le F1 des neuf autres voyelles n'atteint pas le niveau de significativité fixé pour notre étude ($p>.05$).

1.2.2.F2

Comme lors des analyses que nous avons effectuées sur le premier formant, nous avons décidé dans un premier temps de regrouper les voyelles pour analyser la réalisation du deuxième formant. Puisque le F2 est lié à la dimension de l'antériorité des voyelles, nous avons regroupé les voyelles en quatre degrés d'antériorité : voyelles antérieures non-arrondies /i, e, ɛ/, voyelles antérieures arrondies /y, ø, œ/, voyelles postérieures arrondies /u, o, ɔ/ et la voyelle /a/. La figure 5.4 présente la réalisation du F2 pour chaque degré d'antériorité. Une première observation de cette figure indique que les voyelles antérieures arrondies des enfants CI semblent plus postérieures que celles des enfants NH (F2 plus bas pour les enfants CI), et que les voyelles

postérieures des enfants CI semblent plus antérieures que celles des enfants NH (F2 plus élevés pour les enfants CI).

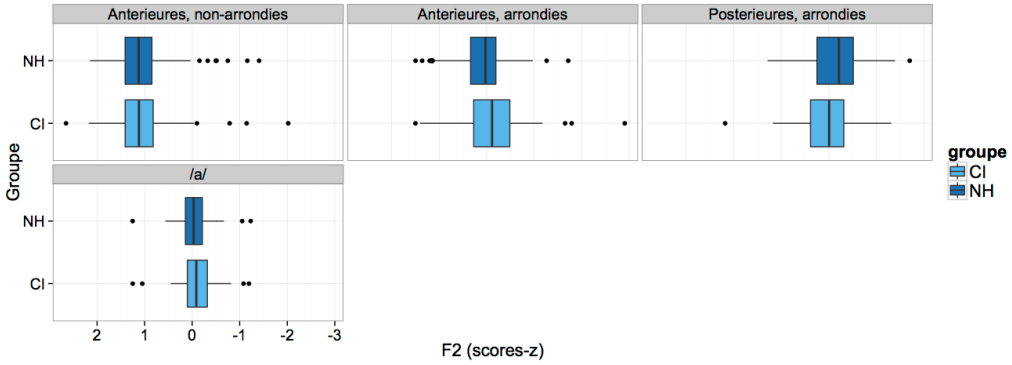


Figure 5.4 : représentation du F2 (en scores-z) par degré d'antériorité, pour chaque groupe (enfants CI et NH).

Nous avons testé l'effet du groupe et du degré d'antériorité de la catégorie vocalique sur la réalisation du F2. Les résultats de ce test statistique sont présentés dans la table 5.9 ci-dessous.

	numDF	denDF	F-value	p-value
(Intercept)	1	1826	0.0001	0.9941
groupe	1	31	0.0271	0.8702
anteriorité	3	1826	1844.6247	<.0001 ***
groupe:anteriorité	3	1826	8.2509	<.0001 ***

Table 5.9 : tableau de résultats au test statistique sur toutes les données (effets du groupe et du degré d'antériorité sur le F2)

Comme pour le F1, nous n'observons pas d'effet significatif du groupe seul (F(1,31)=0,03, p>.05) sur la réalisation du F2 mais un effet du degré d'antériorité seul (F(3,1826)=1844,62, p<.001) et de l'interaction entre groupe et antériorité (F(3,1826)=8,25, p<.001). En effet, un test post-hoc sur cette interaction confirme l'observation visuelle de la différence inter-groupes pour les seules voyelles antérieures arrondies et voyelles postérieures. La table 5.10 ci-dessous résume les résultats de ce test post-hoc.

Degré d'antériorité	F1
	CI vs NH
Voyelles antérieures, non-arrondies	z=0,425, p>.05 (n.s.)
Voyelles antérieures, arrondies	z=-3,159, p<.01 **
Voyelles postérieures	z=-3,384, p<.001 ***
Voyelle /a/	z=-0,534, p>.05 (n.s.)

Table 5.10 : résumé du test post-hoc sur l'effet d'interaction entre degré de hauteur et groupe sur le F1 des voyelles

Nous avons également réalisé un test post-hoc pour évaluer les différences entre voyelles par degré d'antériorité, pour chaque groupe (Table 5.11) : les voyelles se distinguent par leur degré d'antériorité chez les deux groupes, sauf la voyelle /a/ dont le F2 est similaire à celui des voyelles antérieures arrondies.

Groupe	F2		
	Degrés d'antériorité testés		
	Voyelles antérieures vs antérieures-arrondies	Voyelles antérieures-arrondies vs postérieures	Voyelles antérieures-arrondies vs /a/
NH	z=27,649, p<.001 ***	z=32,234, p<.001 ***	z=1,937, p>.05 (n.s.)
CI	z=23,985, p<.001 ***	z=18,295, p<.001 ***	z=-0,008, p>.05 (n.s.)

Table 5.11 : résumé du test post-hoc sur l'effet d'interaction sur les F2 des voyelles par paires de degrés d'antériorité

Pour affiner ces résultats qui mettent en évidence une réalisation plus postérieure des voyelles antérieures arrondies et une réalisation plus antérieure des voyelles postérieures chez les enfants CI, nous avons réalisé des tests par catégorie vocalique. La figure 5.5 présente ces réalisations du F2 par voyelle.

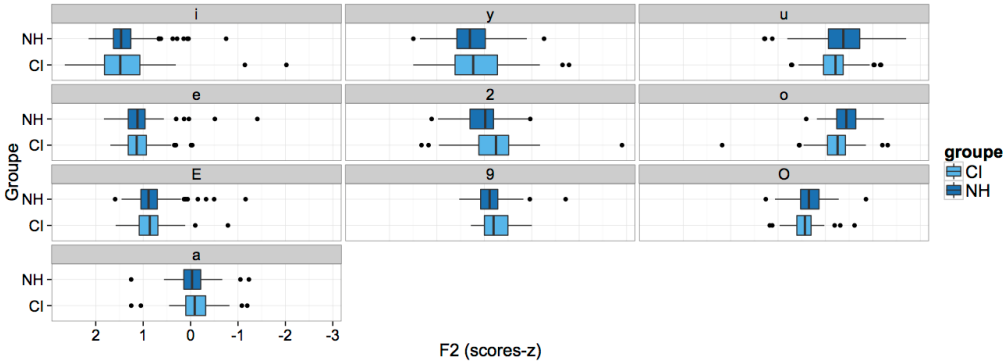


Figure 5.5 : représentation du F2 (en scores-z) par catégorie vocalique, pour chaque groupe (enfants CI et NH).

La table 5.12 présente les résultats au test statistique étudiant les effets du groupe et de la catégorie vocalique seuls et en interaction sur le F2 des voyelles.

	numDF	denDF	F-value	p-value
(Intercept)	1	1814	0.0001	0.9933
groupe	1	31	0.0350	0.8529
V	9	1814	852.4585	<.0001 ***
groupe:V	9	1814	4.1453	<.0001 ***

Table 5.12 : tableau de résultats au test statistique sur toutes les données (effets du groupe et de la catégorie vocalique sur le F2)

Le test statistique n'indique pas d'effet significatif du groupe (F(1,31)=0,04, p>.05) mais un effet significatif de la catégorie vocalique (F(9,1814)=852,46, p<.001) et de l'interaction entre groupe et catégorie vocalique (F(9,1814)=4,15, p<.001). Un test post-hoc sur cette interaction permet de préciser ce résultat : l'effet du groupe est seulement significatif pour les voyelles /ø/ (z=-3,319, p<.001) et /o/ (z=3,487, p<.001) : /ø/ est produit de façon plus postérieure et /o/ de façon plus antérieure par les enfants CI par rapport aux enfants NH. Par souci de lisibilité, nous avons fait le choix de ne pas présenter les résultats du test post-hoc pour les autres voyelles, pour lesquelles aucun effet significatif du groupe n'est observé (p>.05).

1.2.3.F3

La figure 5.6 présente la réalisation du F3 pour chaque catégorie vocalique : seules les voyelles /i/ et /y/ semblent différer d'un groupe à l'autre (le F3 de /i/ étant plus bas chez les enfants CI et celui de /y/ plus haut), à la différence des autres voyelles dont le F3 semble similaire pour les deux groupes.

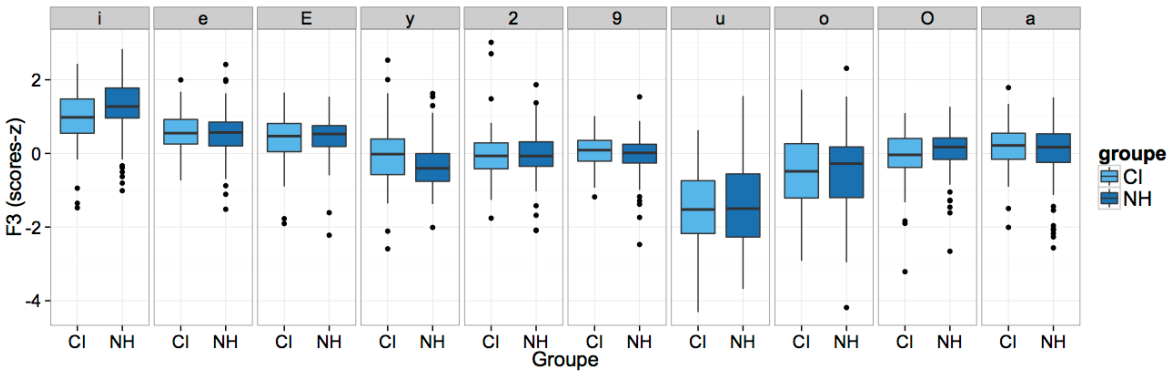


Figure 5.6 : représentation du F3 (en scores-z) par catégorie vocalique, pour chaque groupe (enfants CI et NH).

La table 5.13 présente les résultats du test statistique de l'effet du groupe et de la catégorie vocalique, seuls et en interaction.

	numDF	denDF	F-value	p-value
(Intercept)	1	1814	0.02705	0.8694
groupe	1	31	0.00149	0.9695
V	9	1814	149.03403	<.0001 ***
groupe:V	9	1814	2.47043	0.0084 ***

Table 5.13 : tableau de résultats au test statistique sur toutes les données (effets du groupe et de la catégorie vocalique sur le F3)

Le test statistique n'indique pas d'effet significatif du groupe seul (F(1,31)=0,001, p>.05) mais un effet de la catégorie vocalique seule (F(9,1814)=149,03, p<.001) et en interaction avec le groupe (F(9,1814)=2,47, p<.001). Un test post-hoc de cette interaction indique que seule la voyelle /y/ est sensible à l'effet du groupe (z=3,030, p<.05) et que l'effet du groupe sur la voyelle /i/ est juste au dessus du seuil de significativité (z=-2,725, p=0,0624) fixé à <.05 pour notre étude : le F3 de /y/ est plus haut et celui de /i/ est plus bas chez les enfants CI en comparaison avec les enfants NH.

1.2.4.Dispersion F1-F2

Nous avons choisi d'étudier la dispersion de chaque catégorie vocalique dans le plan F1-F2 à travers une mesure de distance euclidienne en deux dimensions (F1-F2), de chaque occurrence produite par rapport au centre de la catégorie vocalique correspondante, dont la méthode de calcul est détaillée au chapitre 4, paragraphe §2.1.2. La figure 5.7 propose deux représentations de cette dispersion de chaque catégorie vocalique par ellipses à 95% de confiance et par boxplots. Nous observons des différences entre groupes pour les voyelles /i/, /y/, /ø/ et /œ/ : la dispersion de ces voyelles est plus grande pour les enfants CI.

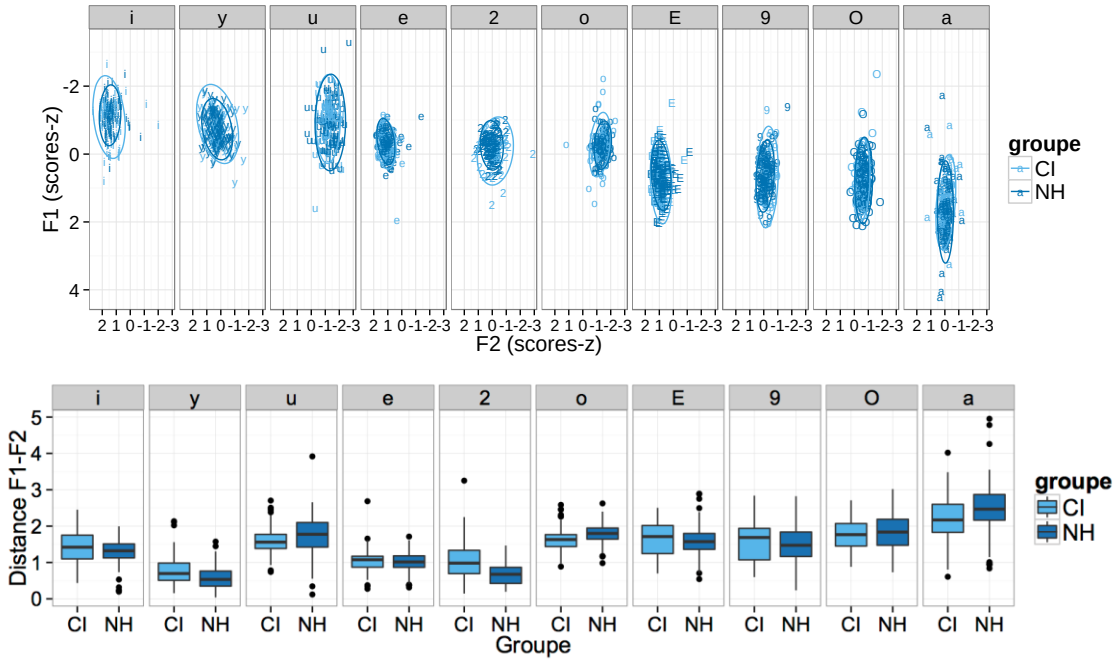


Figure 5.7 : représentations par ellipses (haut) et boxplots (bas) de la dispersion de chaque catégorie vocalique sur F1-F2 (en scores-z) par catégorie vocalique, pour chaque groupe (enfants CI et NH)

Les résultats du test statistique des effets du groupe et de la catégorie vocalique seuls et en interaction sont présentés dans la table 5.14 ci-dessous.

	numDF	denDF	F-value	p-value
(Intercept)	1	279	14682.764	<.0001
groupe	1	31	1.521	0.2268
V	9	279	176.864	<.0001 ***
groupe:V	9	279	5.316	<.0001 ***

Table 5.14 : tableau de résultats au test statistique sur toutes les données (effets du groupe et de la catégorie vocalique sur la dispersion F1-F2)

Le test statistique ne met pas en évidence un effet significatif du groupe (F(1,31)=1,521, p>.05) mais un effet de la catégorie vocalique seule (F(9,279)=176,864, p<.001) et en interaction avec le groupe (F(9,279)=5,316, p<.001). Un test post-hoc de cette interaction nous permet d'affiner cette analyse : l'effet du groupe est seulement significatif pour les voyelles /ø/ (z=4,346, p<.001), /a/ (z=-3,337, p<.01) et /y/ (z=2,874, p<.05) : /ø/ et /y/ sont plus dispersées alors que /a/ est moins dispersée chez les enfants CI.

1.2.5.Dispersion F2-F3

La figure 5.8 présente deux représentations de la dispersion sur F2-F3 (Ellipses à 95% de confiances et boxplots). L'observation de ces deux figures semble indiquer que les deux groupes ont des réalisations différentes pour les voyelles /i/, /y/, /ø/ et /o/, la dispersion de ces catégories étant plus grande pour les enfants CI.

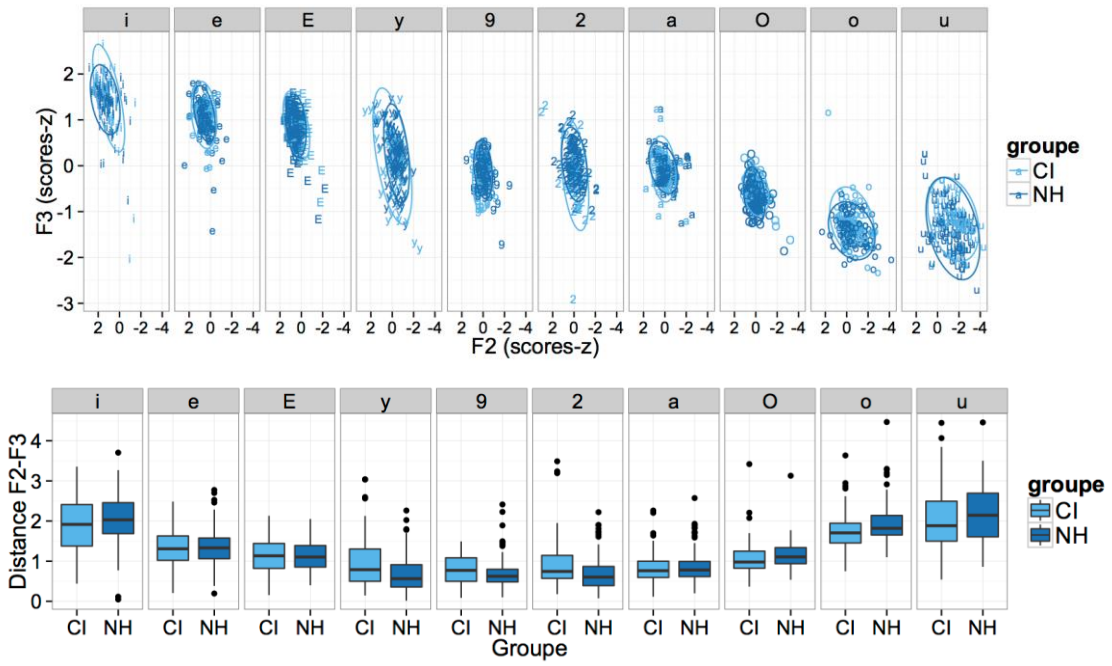


Figure 5.8 : représentations par ellipses (haut) et boxplots (bas) de la distance euclidienne moyenne F2-F3 (en scores-z) par catégorie vocalique, pour chaque groupe (enfants CI et NH)

La table 5.15 ci-dessous présente les résultats du test statistique des effets du groupe et de la catégorie vocalique sur la dispersion F2-F3 par rapport au centre de la catégorie vocalique.

	numDF	denDF	F-value	p-value
(Intercept)	1	279	7990.866	<.0001
groupe	1	31	0.086	0.7718
V	9	279	144.502	<.0001 ***
groupe:V	9	279	2.961	0.0022 **

Table 5.15 : tableau de résultats au test statistique sur toutes les données (effets du groupe et de la catégorie vocalique sur la dispersion F2-F3)

Le test statistique ne met pas en évidence un effet global du groupe (F(1,31)=0,086, p>.05) mais un effet de la catégorie vocalique seule (F(9,279)=144,502, p<.001) et en interaction avec le groupe (F(9,279)=2,961, p<.01). Un test post-hoc de cette interaction met en évidence une différence inter-groupes significative pour la seule voyelle /y/ (z=3,646, p<.01), ce qui ne confirme que partiellement l'observation visuelle des données. En revanche, le test n'indique pas de différence inter-groupes pour les voyelles /i/, /ø/ et /œ/, contrairement à ce que l'observation semblait indiquer.

1.2.6. Focalisation : différences F2-F1 et F3-F2

En plus des analyses des caractéristiques acoustiques précédemment réalisées dans ce chapitre, nous avons décidé de nous intéresser à la focalisation F2-F1 et F3-F2. La figure 5.9 ci-dessous présente les différences F2-F1 et F3-F2 par groupe, pour chaque catégorie vocalique.

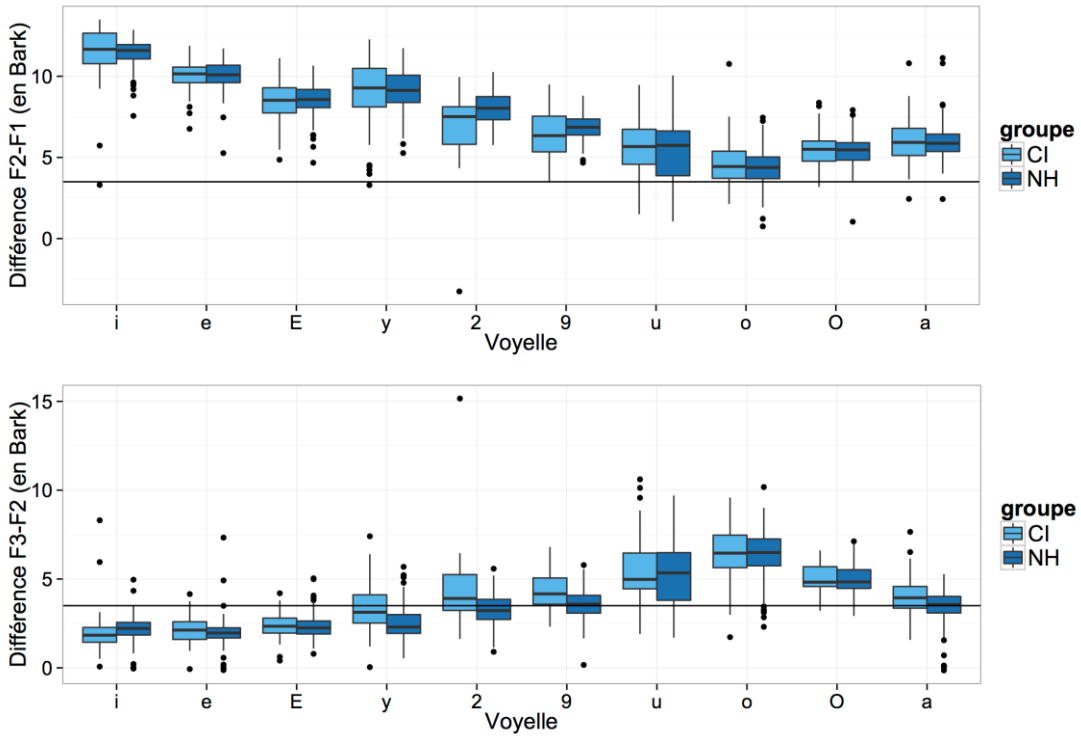


Figure 5.9 : Différences F2-F1 (haut) et F3-F2 (bas) en Bark, par catégorie vocalique, pour chaque groupe (enfants CI et NH), la ligne horizontale étant fixée à 3,5 Bark

Nous avons ajouté à cette représentation la limite de 3,5 Bark, en dessous de laquelle les deux formants successifs convergent en une seule zone formantique, caractéristique d'une voyelle focale (Chistovich & Lublinskaya, 1979, Schwartz & Escudier, 1989, Schwartz et al., 2004 et Ménard et al., 2007).

Georgeton et al. (2012) proposent une description des voyelles focales utilisant une limite supérieure plus élevée, puisque dans cette étude, les voyelles focales en F1-F2 sont /a/ (différence F2-F1 : 6 Bark), /u/ (5 Bark), /o/ (4 Bark) et /ɔ/ (5,5 Bark) et en F2-F3 /y/ (différence F3-F2 : 5 Bark), et en F3-F4 /i/ (6,5 Bark).

Sur la figure 5.9, aucune des différences F2-F1 des catégories vocaliques n'est inférieure à la limite de 3,5 Bark, ce qui semble indiquer que les voyelles produites ne sont pas des voyelles focales. Si cette limite de 3,5 Bark est relevée à 6,5 Bark, nous constatons que les voyelles /a/, /u/, /o/ et /ɔ/ produites par les enfants NH ont des différences F2-F1 similaires à celles de Georgeton et al. (2012), et les enfants CI ont des différences similaires pour /a/ et /u/, mais différentes pour /o/ et /ɔ/, tout en restant inférieures à 5,5 Bark (dans notre étude, les différences F2-F1 des enfants CI de /o/ et /ɔ/ sont similaires à celles de /ɔ/ et /o/ respectivement de l'étude de Georgeton et al., 2012).

La différence F3-F2 est quant à elle inférieure à la limite des 3,5 Bark pour sept voyelles chez les enfants NH (/y/, /i/, /e/, /ɛ/, /ø/, /œ/ et /a/) et pour quatre voyelles chez les enfants CI (/y/, /i/, /e/ et /ɛ/) ce qui correspond dans une certaine mesure aux résultats de Georgeton et al. (2012).

Nous avons réalisé des tests statistiques sur les différences F2-F1 et F3-F2, pour déterminer les effets du groupe et de la catégorie vocalique sur la réalisation de la focalisation, dont les résultats sont présentés dans la table 5.16 (différence F2-F1) et 5.17 (différence F3-F2).

	numDF	denDF	F-value	p-value
(Intercept)	1	1814	14065.365	<.0001
groupe	1	31	1.162	0.2894
V	9	1814	633.560	<.0001 ***
groupe:V	9	1814	3.490	0.0003 ***

Table 5.16 : différences F2-F1 (en Bark), effet du groupe et de la catégorie vocalique

	numDF	denDF	F-value	p-value
(Intercept)	1	1814	4742.906	<.0001
groupe	1	31	10.713	0.0026 **
V	9	1814	338.381	<.0001 ***
groupe:V	9	1814	5.780	<.0001 ***

Table 5.17 : différences F3-F2 (en Bark), effet du groupe et de la catégorie vocalique

Pour les deux mesures, nous observons un effet de la catégorie vocalique seule ($F(9,1814)=633,56$, $p<.001$ pour la différence F2-F1 et $F(9,1814)=338,38$, $p<.001$ pour la différence F3-F2) et en interaction avec le groupe ($F(9,1814)=3,49$, $p<.001$ pour la différence F2-F1 et $F(9,1814)=5,78$, $p<.001$ pour la différence F3-F2). L'effet du groupe n'est visible que sur la différence F3-F2 ($F(1,31)=10,71$, $p<.01$). Nous avons réalisé un test post-hoc de l'interaction entre groupe et catégorie vocalique pour les deux différences F2-F1 et F3-F2. Comme dans ce paragraphe, nous souhaitons caractériser la convergence formantique, nous ne donnerons que les résultats des voyelles focales de la littérature (/a/, /u/, /o/ et /ɔ/ pour la focalisation F1-F2) et des voyelles dont la différence F3-F2 est inférieure à 3,5 Bark dans notre étude. (/y/, /i/, /e/, /ɛ/, /ø/, /œ/ et /a/ pour la focalisation F2-F3). Par souci de lisibilité, nous ne donnerons ici que les résultats pour lesquels une différence inter-groupes significative est constatée.

Le test post-hoc pour les quatre voyelles les plus focales en F1-F2 n'indique pas de différence significative entre les deux groupes. En revanche, le test post-hoc pour les sept voyelles les plus focales en F2-F3 indiquent des différences selon les voyelles considérées. Les voyelles /i/, /e/ et /ɛ/ qui présentent une convergence de F2 et F3 pour les deux groupes ne subissent pas d'effet de groupe, à la différence de la voyelle /y/ ($z=4,37$, $p<.001$). Ce test post-hoc indique en outre un effet significatif de groupe pour les trois voyelles qui sont réalisées avec une convergence formantique chez les enfants NH et sans convergence formantique chez les enfants CI : /ø/ ($z=4,94$, $p<.001$), /œ/ ($z=4,050$, $p<.001$) et /a/ ($z=4,37$, $p<.05$).

1.2.7. Résumé : effets du groupe et de la catégorie vocalique

La table 5.18 présente un résumé de tous les effets testés dans ce paragraphe.

Effets	Caractéristiques acoustiques				
	F1	F2	F3	Dispersion F1-F2	Dispersion F2-F3
Groupe	n.s.	n.s.	n.s.	n.s.	n.s.
Catégorie vocalique	i,y,u<e,ø,o<ɛ,œ,ɔ<a	i,e,ɛ>y,ø,œ,a>u,o,ɔ	i>e>ɛ>y>ø>œ>a>ɔ>o>u	***	***
Groupe*Catégorie vocalique	/a/ : CI < NH **	/ø/ : CI<NH *** /o/ : CI>NH ***	/y/ : CI>NH *	/a/ : CI>NH ** /ø/ : CI>NH *** /y/ CI>NH *	/y/ : CI>NH **

Table 5.18 : résumé des effets des variables groupe, catégorie vocalique et de l'interaction entre groupe et catégorie vocalique sur les formants F1, F2 et F3, ainsi que sur les distances euclidiennes F1-F2 et F2-F4 moyenne au centre de la catégorie

Les tests réalisés dans ce chapitre montrent des différences très localisées sur certaines voyelles, plutôt que des effets globaux de groupe. Les voyelles les plus impactées par l'utilisation de l'implant sont /a/, /y/, /ø/ et /o/. En effet, la seule voyelle pour laquelle une différence entre groupe est observée sur le F1 est la voyelle /a/ qui est plus haute chez les enfants CI. Les autres voyelles (hautes, mi-hautes, mi-basses) sont réalisées de façon similaire par les deux groupes d'enfants. /a/ est également plus dispersée, donc plus variable sur le plan F1-F2, chez les enfants CI. Sur l'axe d'antériorité, les voyelles antérieures-arrondies et en particulier /ø/, mais également les voyelles postérieures, et en particulier /o/, sont les seules à être réalisées différemment par les deux groupes : en effet, /ø/ et /o/ ont tendance à être moins distinctes l'une de l'autre sur le plan F1-F2 chez les enfants CI, et /ø/ est également plus dispersée sur les plans F1-F2 et F2-F3 chez les enfants CI. Enfin, le F3 de /y/ est plus haut, mais également plus dispersé sur les plans F1-F2 et F2-F3 chez les enfants CI, par rapport aux enfants NH. Le F3 étant lié à la protrusion des lèvres (F3 diminue lorsque les voyelles sont protruses), il semble que les enfants CI ont tendance à produire des /y/ moins arrondies, et plus proches de /i/ que les enfants NH. Nous observons également que /i/ est produite avec un F3 plus bas, et donc de façon plus protruse chez les enfants CI (tendance observée, juste au-dessus du seuil de significativité) par rapport aux enfants NH.

1.3. Enfants NH : effet de l'âge chronologique

Comme nous cherchons à isoler les différents effets qui agissent sur la réalisation des voyelles par les deux groupes d'enfants, nous avons choisi d'étudier l'effet de l'âge chronologique sur les caractéristiques acoustiques des voyelles de notre étude. Nous analysons l'âge différemment selon le groupe d'enfants : âge chronologique chez les enfants NH, et âge à l'implantation et âge auditif (ou durée d'utilisation de l'implant cochléaire) chez les enfants CI. Nous n'incluons pas l'âge chronologique dans les analyses que nous réalisons chez le groupe des enfants CI puisque cet effet correspond à un effet à la fois de la taille du conduit vocal et de la maturation du contrôle moteur par les enfants CI, que nous observons également chez les enfants NH. Une étude de l'effet de l'âge chronologique chez les enfants CI n'est donc pas nécessaire, puisque nous étudions déjà cet effet chez les enfants NH. Dans ce paragraphe, nous allons étudier l'effet de l'âge chronologique sur la réalisation des voyelles. Comme nous avons fait le choix de normaliser les données (méthode des scores-z de Lobanov (Lobanov, 1971)) pour éliminer la variance liée à la différence de taille de conduit vocal, nous nous attendons à n'observer aucun effet de l'âge chronologique dû à la croissance du conduit vocal sur les caractéristiques acoustiques des voyelles des enfants NH. En revanche, un effet de l'âge chronologique pourrait indiquer un effet du développement du contrôle moteur, qui traduirait une maturation différente de ce contrôle chez des enfants d'âge différent.

1.3.1. F1

La figure 5.10 présente l'évolution du F1 en fonction de l'âge chronologique pour les enfants NH, par degré de hauteur de la voyelle.

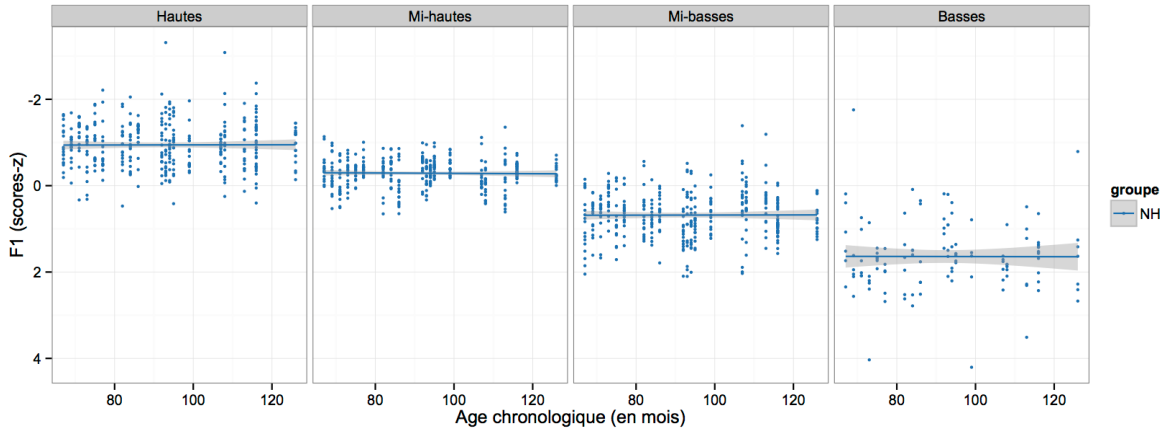


Figure 5.10 : évolution du F1 (scores-z) en fonction de l'âge chronologique (en mois), pour les enfants NH – un point par voyelle produite et par enfant

Le test statistique réalisé par degré de hauteur met en évidence une absence d'effet significatif de l'âge chronologique seul (F(1,18)=0,0037, p>.05) ou en interaction avec le degré de hauteur (F(3,1151)=0,0243, p>.05) sur le F1 des voyelles produites par les enfants NH. Le même test, réalisé en remplaçant le degré de hauteur (quatre degrés) par la catégorie vocalique (dix voyelles) donne le même résultat.

1.3.2.F2

La figure 5.11 présente l'évolution du F1 en fonction de l'âge chronologique pour les enfants NH, par degré d'antériorité de la voyelle.

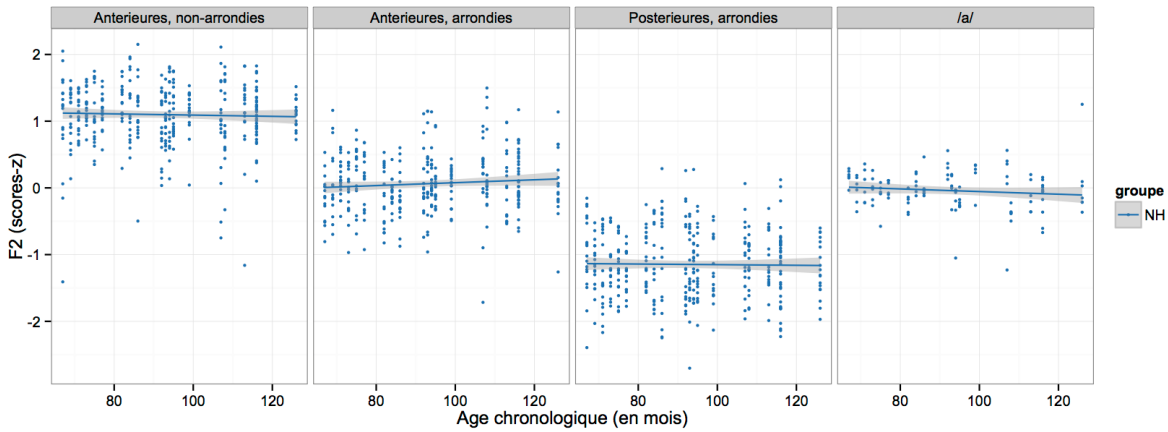


Figure 5.11 : évolution du F2 (scores-z) en fonction de l'âge chronologique (en mois), pour les enfants NH – un point par voyelle produite et par enfant

Le test statistique réalisé sur les données qui prend en compte le degré d'antériorité de la voyelle (quatre degrés) met en évidence que l'âge chronologique seul (F(1,18)=0,0001, p>.05) ou en interaction avec le degré d'antériorité (F(3,1151)=1,1178, p>.05) n'a pas d'effet sur la réalisation du F2 par les enfants NH. Le même test réalisé en remplaçant le degré d'antériorité par la catégorie vocalique donne le même résultat.

1.3.3.F3

La figure 5.12 présente l'évolution du F1 en fonction de l'âge chronologique pour les enfants NH, par catégorie vocalique.

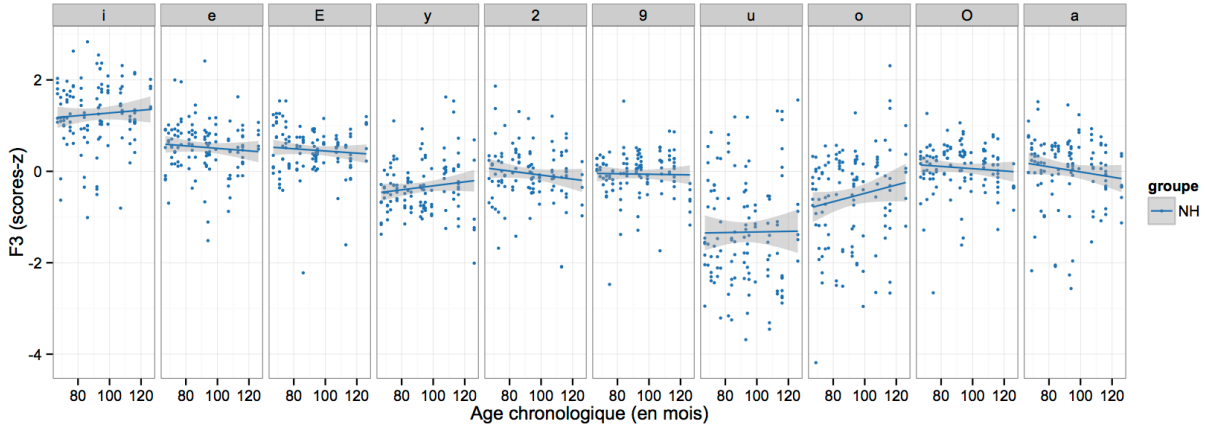


Figure 5.12 : évolution du F3 (scores-z) en fonction de l'âge chronologique (en mois), pour les enfants NH – un point par voyelle produite et par enfant

Le test statistique que nous avons réalisé ne montre pas d'effet de l'âge chronologique sur le F3 des enfants NH, ni seul (F(1,18)=0,002, p>.05), ni en interaction avec la catégorie vocalique (F(9,1139)=1,303, p>.05).

1.3.4.Dispersion F1-F2

La figure 5.13 présente l'évolution de la dispersion F1-F2 moyenne de chaque catégorie vocalique en fonction de l'âge chronologique pour les enfants NH, par catégorie vocalique.

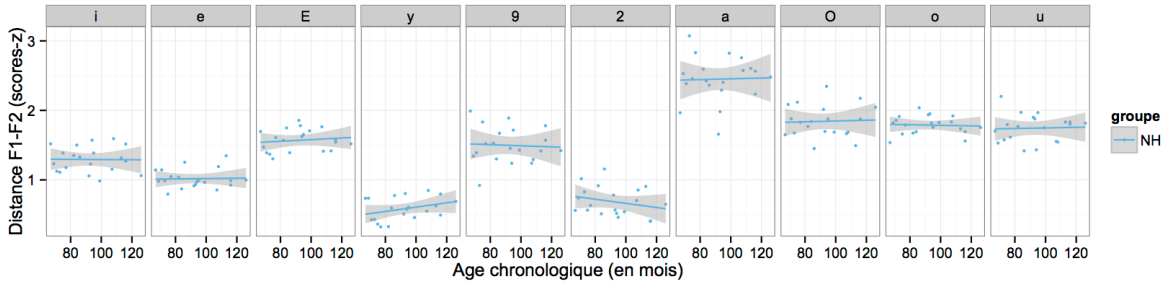


Figure 5.13 : évolution de la dispersion F1-F2 moyenne (scores-z) en fonction de l'âge chronologique (en mois), pour les enfants NH – un point par catégorie vocalique et par enfant

L'analyse statistique met en évidence une absence d'effet significatif sur la dispersion F1-F2 de l'âge chronologique seul (F(1,18)=0,039, p>.05) ou en interaction avec la catégorie vocalique (F(9,162)=0,314, p>.05).

1.3.5.Dispersion F2-F3

La figure 5.14 présente l'évolution de la dispersion F2-F3 moyenne de chaque catégorie vocalique en fonction de l'âge chronologique pour les enfants NH, par catégorie vocalique.

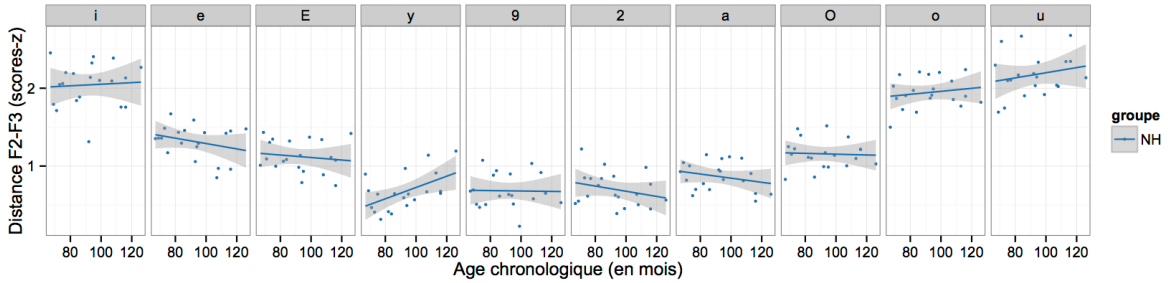


Figure 5.14 : évolution de la dispersion F2-F3 moyenne (scores-z) en fonction de l'âge chronologique (en mois), pour les enfants NH – un point par catégorie vocalique et par enfant

Comme pour les autres mesures de ce paragraphe, nous ne relevons aucun effet significatif de l'âge chronologique sur la dispersion F2-F3 moyenne de chaque catégorie vocalique, ni seul ($F(1,18)=0,033$, $p>.05$), ni en interaction avec la catégorie vocalique ($F(9,162)=1,335$, $p>.05$).

1.3.6. Résumé : effets de l'âge chronologique (enfants NH)

La table 5.19 ci-dessous propose un résumé des résultats obtenus dans ce paragraphe.

Effets	Caractéristiques acoustiques				
	F1	F2	F3	Dispersion F1-F2	Dispersion F2-F3
âge chronologique (NH)	n.s.	n.s.	n.s.	n.s.	n.s.
âge chronologique (NH) * Catégorie vocalique	n.s.	n.s.	n.s.	n.s.	n.s.

Table 5.19 : résumé des effets de l'âge chronologique et de l'interaction entre l'âge chronologique et catégorie vocalique sur les formants F1, F2 et F3, ainsi que sur la dispersion sur F1-F2 et F2-F3

L'âge chronologique n'a aucun effet sur les mesures que nous avons effectuées. Ces résultats confirment que nous avons choisi une méthode de normalisation adaptée à notre problématique. Cette absence d'effet de l'âge chronologique indique également une stabilité du contrôle moteur, puisque la dispersion (sur F1-F2 et F2-F3) des catégories vocaliques chez les enfants NH n'est pas affectée par l'âge chronologique.

1.4. Enfants CI : effet de l'âge à l'implantation et de la durée d'utilisation de l'implant

Nous avons choisi de ne pas étudier l'effet de l'âge chronologique sur la production des voyelles par les enfants CI sous la même forme que chez les enfants NH, mais de considérer l'âge à l'implantation et l'âge auditif (ou durée d'utilisation de l'implant cochléaire) qui sont les deux facteurs qui peuvent influencer la réalisation des voyelles chez cette population d'enfants en particulier.

1.4.1. F1

La figure 5.15 présente l'évolution de la hauteur de la voyelle en fonction de l'âge à l'implantation et de l'âge auditif. Une première observation de cette figure semble indiquer une relative stabilité sur cette dimension de hauteur, avec toutefois une tendance pour la voyelle basse /a/ à s'abaisser (augmentation du F1) avec l'âge à l'implantation et à s'élever (diminution du F1) lorsque la durée d'utilisation de l'implant augmente.

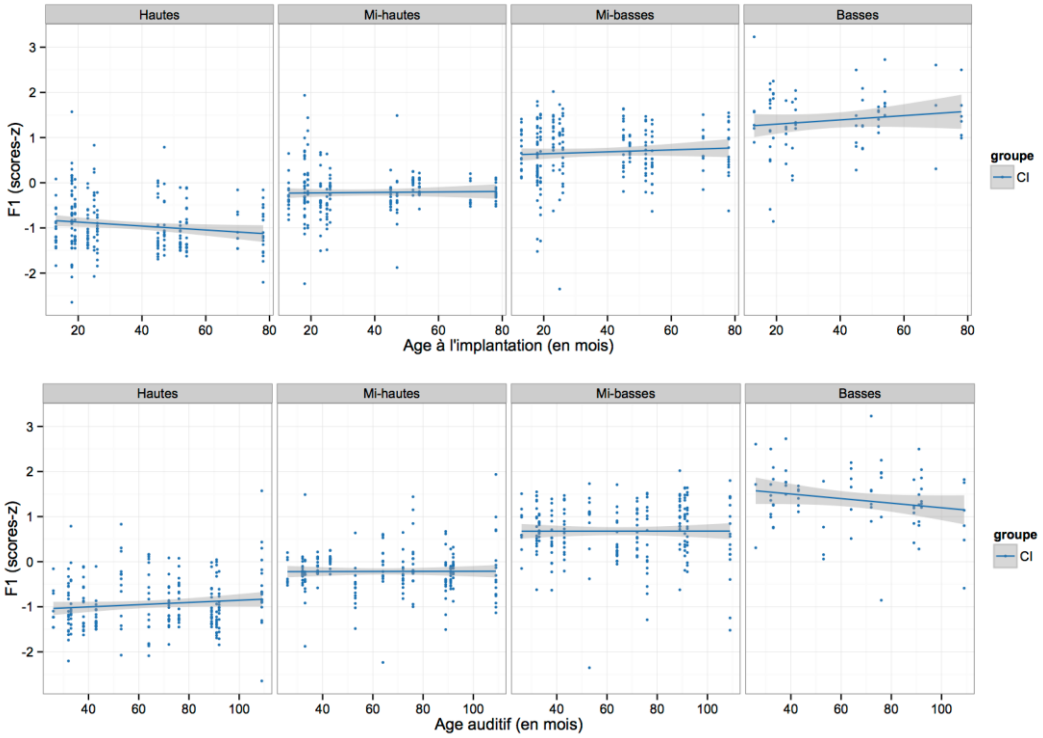


Figure 5.15 : F1 en fonction de l'âge à l'implantation (haut) et de l'âge auditif (bas) en mois, par degré de hauteur des voyelles, pour les enfants CI – un point par voyelle produite et par enfant

Nous avons réalisé deux tests statistiques de l'effet de l'âge à l'implantation et de l'âge auditif sur le F1 de chaque catégorie vocalique ou en regroupant celles-ci par degré de hauteur. Le premier test qui prend en compte chaque catégorie vocalique séparément ne met pas en évidence un effet de l'âge à l'implantation ni de l'âge auditif seuls mais un effet significatif de l'interaction entre catégorie vocalique et âge auditif ($F(9,648)=2,93$, $p<.01$) : les F1 de /i/, /ɛ/ et /a/ diminuent ceux de /y/, /œ/ et /u/ augmentent, et ceux de /e/, /o/, /ɔ/ et /ø/ sont stables lorsque l'âge auditif augmente. En outre, le test regroupant les voyelles par degré de hauteur indique un effet de l'âge à l'implantation juste au dessus du seuil de significativité ($F(3,669)=2,57$, $p=.0531$). Nous avons ensuite effectué un test de l'effet des deux variables d'âge sur le F1 par catégorie vocalique : l'âge auditif a un effet significatif sur la seule voyelle /u/ ($F(1,9)=5,62$, $p<.05$) qui correspond à un F1 plus haut chez les enfants au fur et à mesure de l'utilisation de l'implant et l'âge à l'implantation sur les voyelles /y/ ($F(1,9)=6,27$, $p<.05$) qui est plus élevé chez les enfants implantés plus précocement et /œ/ ($F(1,9)=6,85$, $p<.05$) qui est plus élevé chez les enfants implantés plus tardivement.

1.4.2.F2

La figure 5.16 présente l'évolution du F2 par degré d'antériorité de la voyelle, en fonction de l'âge à l'implantation et de l'âge auditif. Comme pour le F1, nous ne constatons pas d'évolution très marquée du F2 en fonction de ces deux facteurs. Seules les voyelles antérieures arrondies et postérieures semblent évoluer avec l'âge auditif : le F2 des premières semble augmenter (ces voyelles deviennent plus antérieures) et celui des dernières semble diminuer (les voyelles deviennent plus postérieures), lorsque la durée d'utilisation de l'implant cochléaire augmente.

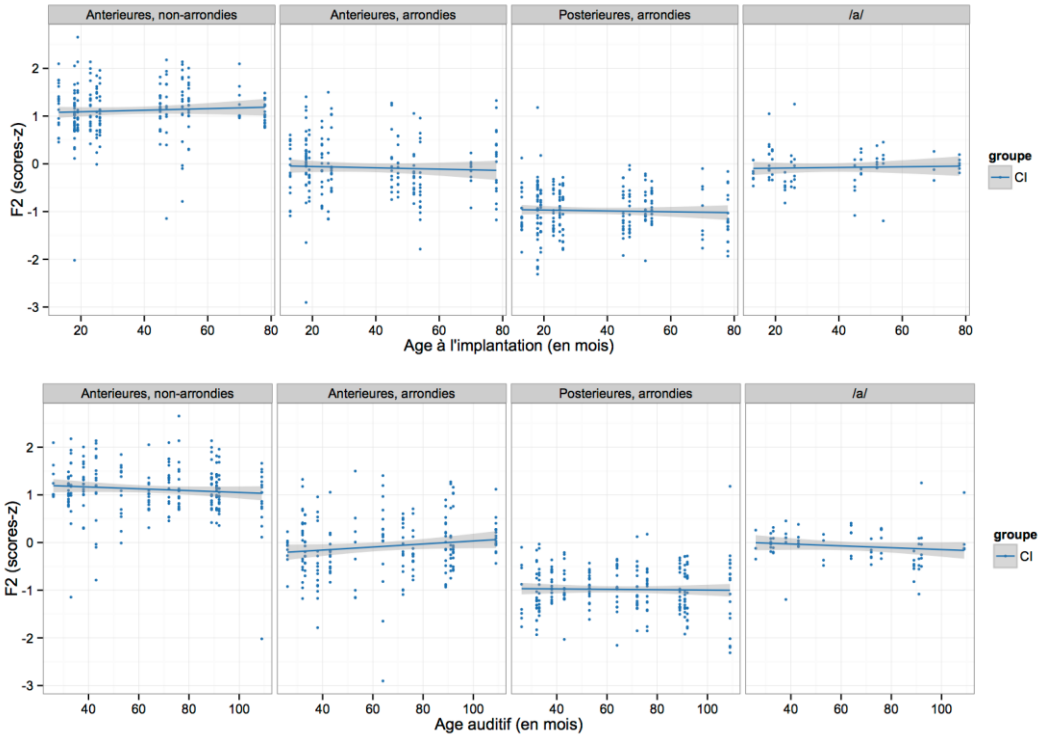


Figure 5.16 : F2 en fonction de l'âge à l'implantation (haut) et de l'âge auditif (bas), en mois, par degré de hauteur des voyelles, pour les enfants CI – un point par voyelle produite et par enfant

Les tests statistiques réalisés sur les voyelles considérées seules ou par degré d'antériorité ne permettent pas de confirmer la tendance observée sur la figure 5.16 : par exemple ni l'âge à l'implantation ($F(1,10)=0,00001$, $p>.05$ pour le test prenant en compte les voyelles individuellement) ni l'âge auditif ($F(1,10)=0,01$, $p>.05$ lors du même test) n'atteignent le seuil de significativité fixé. Seul le test de l'effet de l'âge auditif sur le F2 des voyelles antérieures arrondies approche ce seuil de significativité ($F(1,11)=3,33$, $p=0,0954$) lors de tests réalisés en isolant chaque degré d'antériorité : les voyelles /y/, /ø/ et /œ/ seraient donc les seules à être affectées par l'âge auditif et auraient tendance à devenir de plus en plus antérieures lorsque la durée d'utilisation de l'implant augmente, pour approcher de la production de ces voyelles par les enfants NH.

1.4.3.F3

La figure 5.17 présente l'évolution du F3, pour chaque voyelle, en fonction de l'âge à l'implantation et de l'âge auditif. Nous observons une légère augmentation du F3 pour /ø/, et une diminution pour /ɔ/ avec l'âge à l'implantation, mais également une légère augmentation du F3 pour /y/, /u/, /ɔ/ ainsi qu'une diminution pour /o/ lorsque la durée d'utilisation de l'implant augmente.

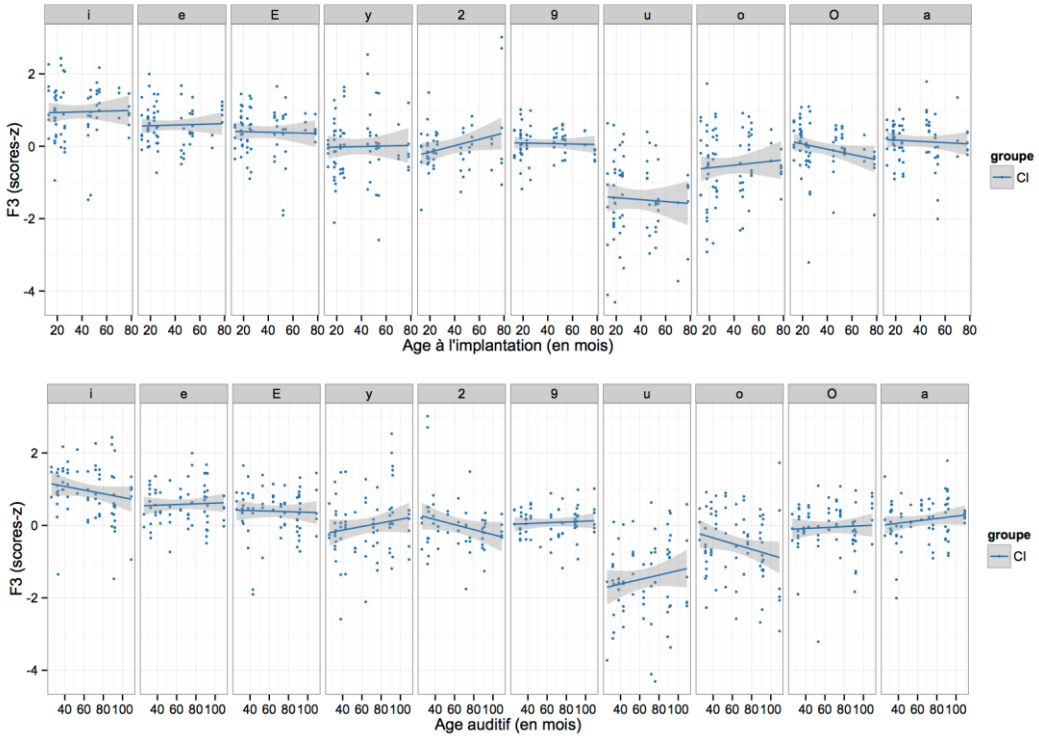
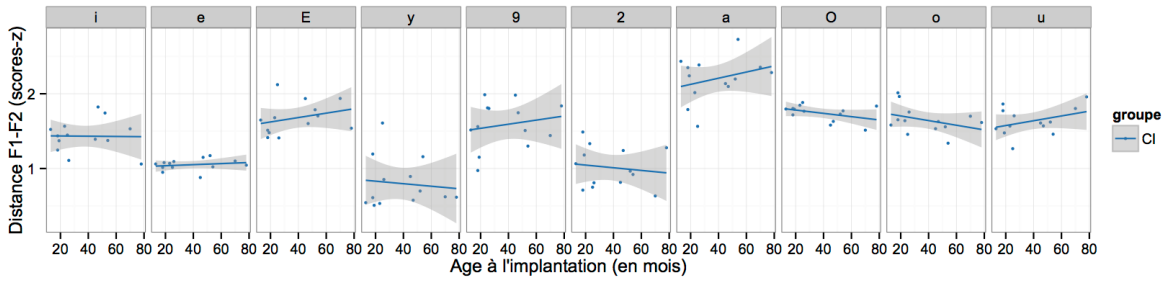


Figure 5.17 : F3 en fonction de l'âge à l'implantation (haut) et de l'âge auditif (bas), en mois, par degré de hauteur des voyelles, pour les enfants CI – un point par voyelle produite et par enfant

Le test statistique n'indique pas d'effet de l'âge à l'implantation ($F(1,10)=0,04$, $p>.05$) et de l'âge auditif ($F(1,10)=0,02$, $p>.05$) seuls, ni de l'interaction entre catégorie vocalique et âge à l'implantation ($F(9,648)=0,081$, $p>.05$) mais un effet de l'interaction entre catégorie vocalique et âge auditif ($F(9,648)=2,93$, $p<.05$). Nous avons ensuite réalisé des tests par voyelle pour lesquels aucun effet de l'âge auditif ne se révèle cependant significatif, ce qui ne nous permet pas d'établir avec certitude à quelles voyelles en particulier est lié cet effet.

1.4.4.Dispersion F1-F2

La figure 5.18 présente l'évolution de la distance euclidienne moyenne F1-F2 par rapport au centre de la catégorie vocalique pour le groupe CI en fonction de l'âge à l'implantation et e l'âge auditif. Nous observons pour certaines voyelles une légère augmentation ou diminution de cette mesure en fonction de l'âge à l'implantation ou de la durée d'utilisation de l'implant.



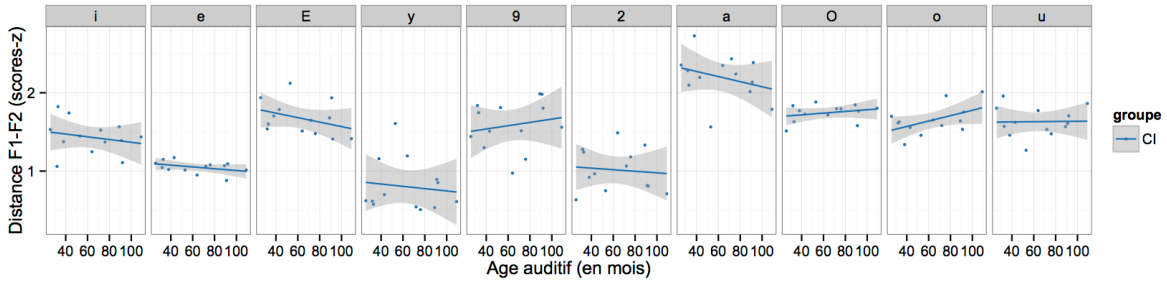


Figure 5.18 : Dispersion F1-F2 (en scores-z), en fonction de l'âge à l'implantation (haut) et de l'âge auditif (bas), en mois, par voyelle, pour les enfants CI – un point par catégorie vocalique et par enfant

Le test statistique que nous avons réalisé pour étudier l'effet de l'âge sur la mesure de dispersion F1-F2 chez les enfants CI indique que ni l'âge à l'implantation ($F(1,9)=0,1935$, $p>.05$), ni l'âge auditif ($F(1,9)=0,0450$, $p>.05$) n'ont d'effet significatif sur la dispersion des catégories vocaliques sur le plan F1-F2. Nous avons cependant réalisé des tests par voyelle pour essayer d'affiner cette analyse, et les résultats indiquent que seule la dispersion sur F1-F2 de /œ/ augmente de façon significative ($F(1,9)=8,94$, $p<.05$) avec l'âge à l'implantation et celle de /o/ avec l'âge auditif ($F(1,9)=6,78$, $p<.05$). Ces résultats sont cependant très localisés et ne nous permettent pas de les généraliser, puisque nous disposons d'une seule mesure de dispersion par enfant et par voyelle. Un nombre plus important d'enfants nous permettrait d'affiner ces résultats.

1.4.5. Dispersion F2-F3

La Figure 5.19 présente l'évolution de la mesure de dispersion de chaque catégorie vocalique sur le plan F2-F3. Nous observons sur cette figure quelques légères augmentations et diminution de la mesure de dispersion avec l'âge à l'implantation et avec l'âge auditif.

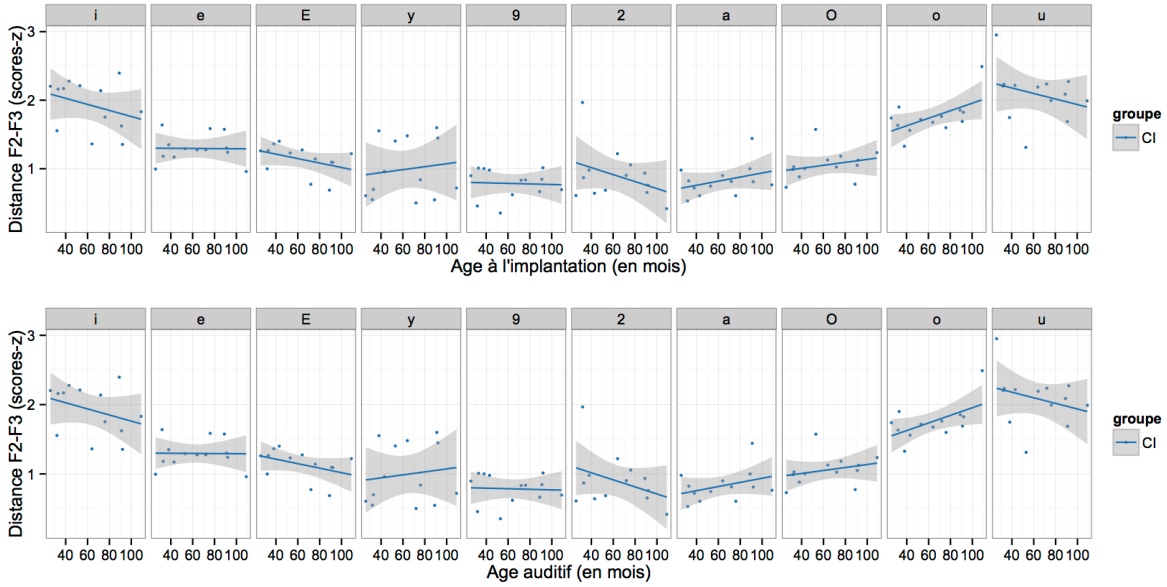


Figure 5.19 : Dispersion F2-F3 par rapport au centre de la catégorie vocalique en fonction de l'âge à l'implantation (haut) et de l'âge auditif (bas), en mois, par voyelle, pour les enfants CI – un point par catégorie vocalique et par enfant

Comme pour la dispersion sur F1-F2, le test statistique réalisé nous indique que ni l'âge à l'implantation ($F(1,9)=0,19$, $p>.05$), ni l'âge auditif ($F(1,9)=0,05$, $p>.05$) n'ont un effet significatif sur l'évolution de la dispersion des catégories vocaliques sur le plan F2-F3. Nous avons

également réalisé des tests par voyelle, et seule l'augmentation avec l'âge auditif de /o/ relevée visuellement est significative (F(1,9)=9,04, p<.05). A nouveau, il nous semble important de nuancer cette absence de résultats qui peut être dû à un nombre de données (dix mesures par enfant) insuffisant.

1.4.6. Résumé : effets de l'âge à l'implantation et de la durée d'utilisation de l'implant (enfants CI)

La table 5.20 résume les effets de l'âge auditif et de l'âge à l'implantation étudiés dans ce paragraphe.

Effets	Caractéristiques acoustiques				
	F1	F2	F3	Dispersion F1-F2	Dispersion F2-F3
âge auditif (CI)	n.s.	n.s.	n.s.	n.s.	n.s.
âge à l'implantation (CI)	n.s.	n.s.	n.s.	n.s.	n.s.
âge auditif (CI) *catégorie vocalique	Augmentation avec /u/*	antériorisation de /y, ø, œ/ (tendance)	n.s.	augmentation pour /o/*	augmentation pour /o/*
âge à l'implantation (CI) *catégorie vocalique	Diminution avec /y/* Augmentation avec/œ/*	n.s.	n.s.	augmentation pour /ø/*	n.s.

Table 5.20 : résumé des effets de l'âge (âge à l'implantation et âge auditif), et de l'interaction entre l'âge et catégorie vocalique sur les formants F1, F2 et F3, ainsi que sur les distances euclidiennes F1-F2 et F2-F4 moyenne au centre de la catégorie

Les tests réalisés dans ce paragraphe ont mis en évidence que ni l'âge à l'implantation, ni l'âge auditif (ou durée d'utilisation de l'implant) n'affectent la réalisation des voyelles. En effet, les effets observés sur des interactions entre âge et catégorie vocalique révèlent des évolutions ponctuelles et très localisées sur certaines voyelles. Les voyelles les plus sensibles à ces effets sont les voyelles antérieures arrondies /y/ dont le F1 évolue avec l'âge à l'implantation (élévation de la voyelle), et le F2 avec l'âge auditif (tendance à l'antériorisation), /œ/ dont le F1 évolue avec l'âge à l'implantation (abaissement de la voyelle) et le F2 avec l'âge auditif (tendance à l'antériorisation) et /ø/ dont le F2 et la dispersion F1-F2 ont tendance à augmenter avec l'âge auditif. Les voyelles postérieures ne sont que très peu influencées par l'âge : la voyelle /u/ s'abaisse avec l'âge auditif et la dispersion de /o/ sur les plans F1-F2 et F2-F3 augmente avec l'âge auditif. Enfin, les voyelles antérieures, non-arrondies /i, e, ɛ/ et /a/ ne sont pas du tout affectées par l'âge.

1.5. Effet de l'input auditif immédiat

Dans ce paragraphe, nous nous intéressons à l'effet de la tâche sur la production de voyelles, pour comprendre si les enfants des deux groupes sont influencés par le modèle audio qu'ils ont lors de la tâche de répétition ou si les voyelles produites lors de cette tâche sont identiques à celles produites lors de la tâche de dénomination, pour laquelle ils n'entendent pas le modèle et doivent baser la réalisation de voyelles sur leur mémoire et sur leur représentation de chaque catégorie vocalique.

1.5.1. F1

La figure 5.20 qui présente le F1 en fonction de la tâche (répétition et dénomination) semble indiquer que les trois degrés de hauteur ne sont pas affectés de la même façon par la tâche : en effet, le F1 des voyelles mi-hautes semble identique dans les deux tâches, alors que les voyelles hautes, mi-basses et /a/ semblent être réalisées différemment selon la tâche proposée. Les F1 des voyelles produites par les deux groupes ne semblent cependant pas être plus proches du modèle dans l'une des deux tâches, ce qui est peut-être dû à la normalisation des données en scores-z.

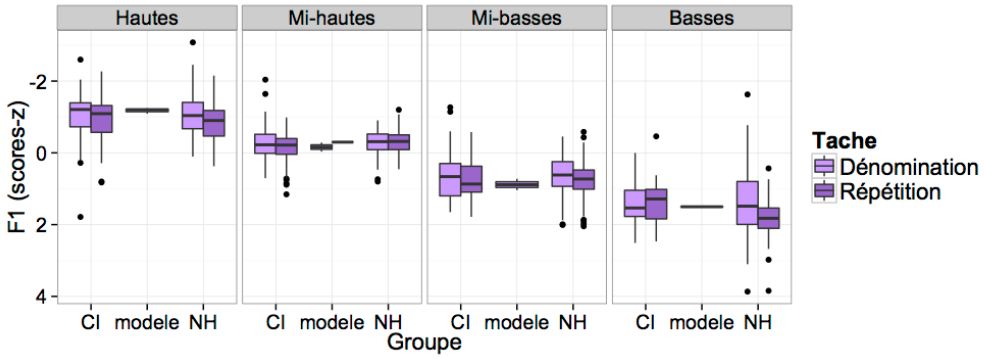


Figure 5.20 : F1 (scores-z) pour chaque degré de hauteur de voyelle, par groupe et tâche (répétition ou dénomination) – « modele » : stimuli (locutrice adulte) utilisés lors de la tâche de répétition

La table 5.21 présente les résultats du test statistique réalisé pour étudier l'effet de la tâche sur la réalisation du F1.

	numDF	denDF	F-value	p-value
(Intercept)	1	1206	0.0957	0.7571
groupe	1	31	0.0518	0.8215
hauteur	3	1206	1027.9885	<.0001 ***
tâche	1	1206	17.1590	<.0001 ***
groupe:hauteur	3	1206	2.6312	0.0488 *
groupe:tâche	1	1206	0.3338	0.5635
hauteur:tâche	3	1206	1.3818	0.2467
groupe:hauteur:tâche	3	1206	3.4137	0.0169 *

Table 5.21 : tableau de résultats au test statistique sur toutes les données (effets de groupe, degré de hauteur des voyelles, tâche)

Les résultats indiquent un effet de la tâche seule ($F(1,1206)=17,59$, $p<.001$) et en interaction avec le groupe et la hauteur de la voyelle ($F(3,1206)=3,41$, $p<.05$). Un test post-hoc de l'interaction entre hauteur et tâche par groupe indique que la tâche n'a pas d'effet sur le F1 des quatre degrés de hauteur chez les enfants CI. En revanche, chez les enfants NH, celle-ci a un effet sur les F1 des voyelles hautes ($z=-2,659$, $p<.05$) et celui des voyelles basses ($z=-4,228$, $p<.001$) mais pas sur celui des voyelles de hauteur intermédiaire. Les résultats sont similaires lorsque le test est réalisé par catégorie vocalique plutôt que par degré de hauteur.

1.5.2.F2

La figure 5.21 présente le F2 par tâche et par degré d'antériorité, pour les deux groupes d'enfants en les comparant au modèle adulte proposé : les deux groupes semblent avoir des réalisations de F2 semblables dans les deux tâches et proches du modèle adulte.

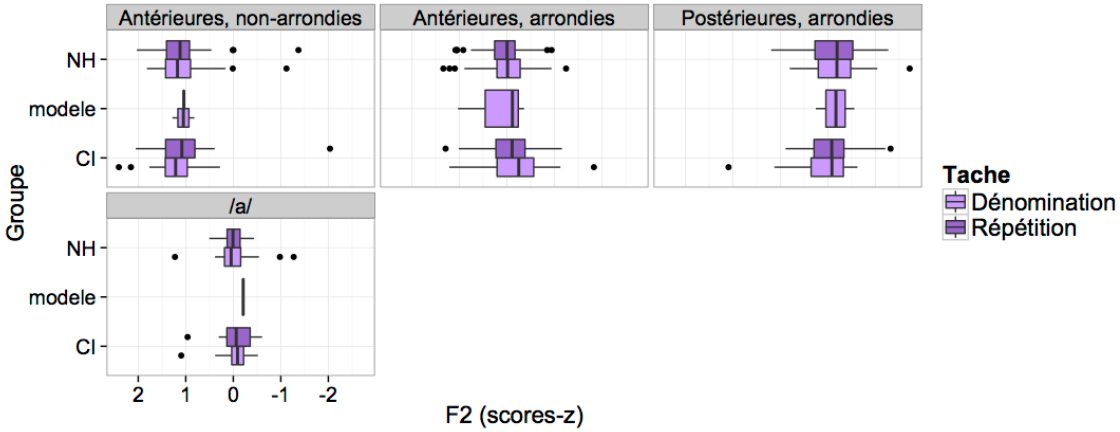


Figure 5.21 : F2 (scores-z) pour chaque degré d'antériorité de voyelle, par groupe et tâches (répétition ou dénomination)

La procédure de sélection de variables que nous avons réalisée nous a conduit à ne pas inclure la tâche dans le modèle utilisé, puisque sa pertinence était jugée trop faible pour modéliser les données. Nous en déduisons donc que la tâche n'a pas d'effet sur la réalisation du F2 par les deux groupes.

1.5.3.F3

La figure 5.22 présente les F3 des voyelles produites par les enfants lors des deux tâches, accompagnés de ceux du modèle adulte. Nous observons quelques différences de réalisation du F3 entre les groupes : le F3 du groupe des enfants CI semble plus sensible à la tâche, notamment pour les voyelles /i/, /ɛ/, /œ/, et /ɔ/ alors que le groupe des enfants NH semble avoir des réalisations similaires du F3, sauf pour les voyelles /u/, /o/ et /a/.

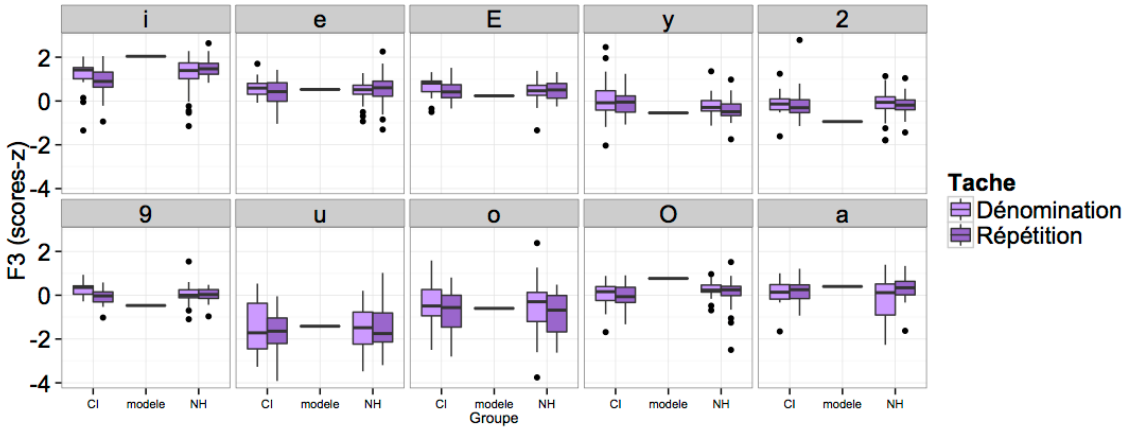


Figure 5.22 : F3 (scores-z) pour chaque degré d'antériorité de voyelle, par groupe et tâche (répétition ou dénomination)

Le test statistique confirme l'observation de la Figure 5.22, puisque si la tâche n'a pas d'effet significatif seule sur le F3 ($F(1,1200)=1,49$, $p>.05$), l'interaction entre groupe et tâche a un effet significatif sur celui-ci ($F(1,1200)=7,56$, $p<.01$). Un test post-hoc indique que le groupe CI est affecté par la tâche ($z=2,946$, $p<.001$) mais pas le groupe NH ($z=-0,724$, $p>.05$). Nous observons en outre que les deux groupes d'enfants ont tendance à produire les voyelles /u/, /e/, /o/ et /a/avec un F3 proche du modèle de la locutrice adulte, mais les voyelles /ɛ/, /y/, /ø/ et /œ/ avec un F3 plus bas et les voyelles /i/ et /ɔ/ avec un F3 plus élevé que ce modèle.

1.5.4.Dispersion F1-F2

Comme les enregistrements dont nous disposons pour le modèle audio adulte, utilisé lors de la tâche de répétition, ne comprennent qu'une seule production de chaque voyelle, il nous est impossible de calculer une dispersion en deux dimensions de chaque catégorie vocalique pour le modèle adulte à partir de cette seule occurrence. Pour les analyses de dispersion F1-F2 et F2-F3, nous allons donc proposer une comparaison de la dispersion des catégories vocaliques des deux groupes dans les deux tâches.

Une première observation de la représentation de la dispersion F1-F2 (figure 5.23), indique que les dispersions sur le plan F1-F2 des productions des deux groupes d'enfants semblent similaires dans les deux tâches, et ce pour toutes les catégories vocaliques, même si les dispersions de /ɛ/, /ɔ/ et /o/ sont légèrement plus faibles dans la tâche de dénomination et celles de /e/, /y/, /œ/, /ø/et /a/, sensiblement plus grandes dans cette même tâche chez les enfants CI, alors que celles de /i/, /ɛ/, /ɔ/et /a/, sont plus faibles dans la tâche de dénomination et celles de /e/ et /o/, sont plus grandes dans cette même tâche chez les enfants NH. Les enfants CI semblent donc produire des voyelles hautes relativement stables dont la dispersion n'est pas affectée par la présentation d'un modèle adulte, et que les enfants NH produisent des voyelles antérieures arrondies ainsi que /u/ et /o/ stables, sans être influencés par la tâche.

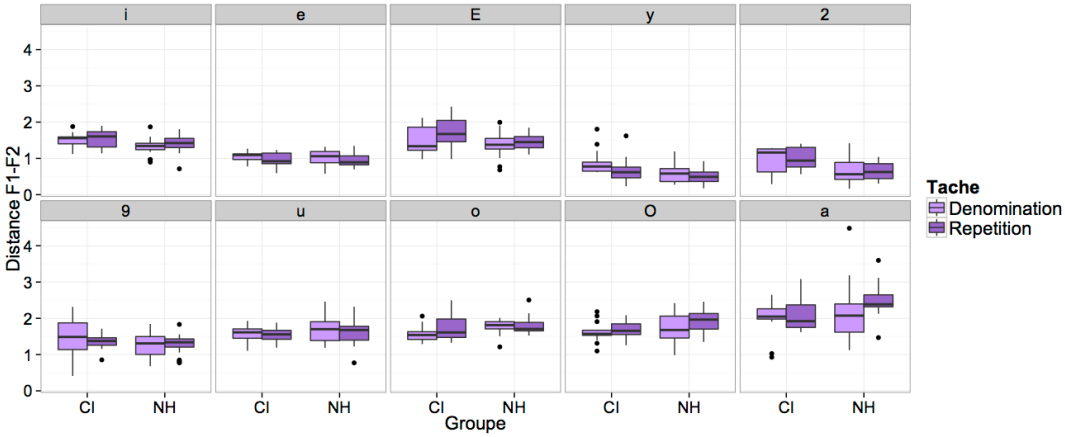


Figure 5.23 : Dispersion F1-F2 (en scores-z) pour chaque voyelle, par groupe et tâche (répétition ou dénomination)

	numDF	denDF	F-value	p-value
(Intercept)	1	574	11205.633	<.0001
groupe	1	31	3.263	0.0806
V	9	574	134.372	<.0001 ***
tâche	1	574	1.223	0.2692
groupe:V	9	574	5.782	<.0001 ***
groupe:tâche	1	574	0.377	0.5395
V:tâche	9	574	1.875	0.0530
groupe:V:tâche	9	574	0.662	0.7434

Table 5.22 : tableau de résultats au test statistique sur toutes les données (effets de groupe, catégorie vocalique, tâche sur la dispersion F1-F2 moyenne de chaque catégorie)

Le test statistique (table 5.22) indique que la tâche n'a pas d'effet sur la dispersion F1-F2 seule (F(1,574)=1,223, p>.05), ni en interaction avec le groupe (F(9,574)=0,38, p>.05) ou la catégorie vocalique (F(9,574)=1,88, p>.05), ce qui ne confirme donc pas l'impression visuelle donnée par la

figure 5.23 : un plus grand nombre de données nous permettrait de confirmer l'impression visuelle initiale.

1.5.5.Dispersion F2-F3

La dispersion des données sur le plan F2-F3 (Figure 5.24) semble indiquer que les enfants CI sont plus sensibles à la tâche proposée et ont tendance à produire des voyelles antérieures arrondies et non-arrondies plus dispersées lors de la tâche de dénomination sans modèle audio. Les voyelles postérieures et le /a/ ne semblent pas subir cet effet de la tâche, à l'exception de /o/. Les enfants NH semblent en revanche produire un /a/ avec une variation plus faible et le /i/ avec une plus grande dispersion lorsque le modèle leur est fourni. Pour les autres voyelles produites par les enfants NH, nous n'observons pas de grandes différences entre les deux tâches.

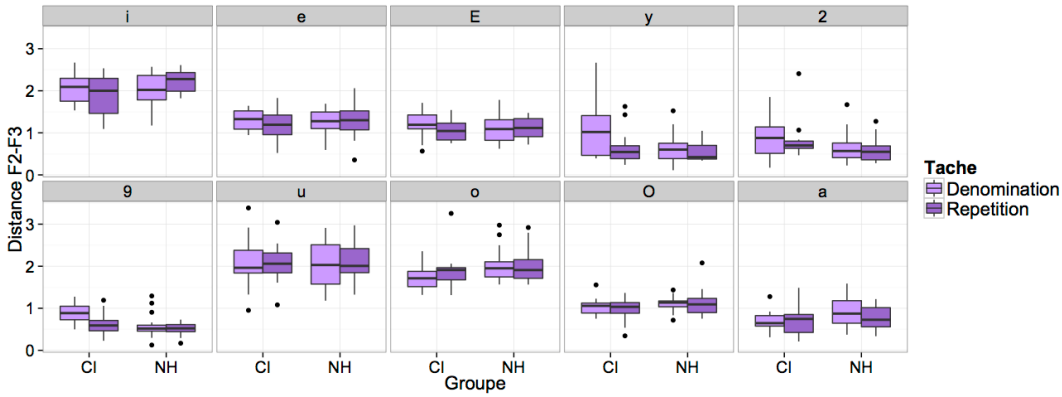


Figure 5.24 : dispersion F2-F3 (scores-z) pour chaque voyelle, par groupe et tâche (répétition ou dénomination)

	numDF	denDF	F-value	p-value
(Intercept)	1	574	8226.077	<.0001
groupe	1	31	0.490	0.4894
V	9	574	153.086	<.0001 ***
tâche	1	574	2.604	0.1072
groupe:V	9	574	3.695	0.0002 ***
groupe:tâche	1	574	2.766	0.0969
V:tâche	9	574	1.165	0.3148
groupe:V:tâche	9	574	0.972	0.4620

Table 5.23 : tableau de résultats au test statistique sur toutes les données (effets de groupe, catégorie vocalique, tâche sur la dispersion F2-F3 moyenne de chaque catégorie)

Le test statistique réalisé (Table 5.23) sur la mesure de dispersion sur le plan F2-F3 indique que celle-ci n'est pas soumise à un effet de la tâche, ni seule ($F(1,574)=2,60$, $p>.05$), ni en interaction avec la catégorie vocalique ($F(9,574)=1,17$, $p>.05$) ni en interaction avec le groupe ($F(1,574)=2,77$, $p>.05$).

1.5.6.Résumé : effet de l'input auditif immédiat

La table 5.24 ci-dessous présente un résumé des effets de la tâche sur les caractéristiques des voyelles de notre étude.

Effets	Caractéristiques acoustiques				
	F1	F2	F3	Dispersion F1-F2	Dispersion F2-F3
tâche	***	n.s.	n.s.	n.s.	n.s.
tâche*groupe	n.s.	n.s.	CI***, NH (n.s.)	n.s.	n.s.
tâche*catégorie vocalique	CI : D=R (n.s.) NH D<R (voyelles hautes*, voyelle basse***)	n.s.	n.s.	n.s.	n.s.

Table 5.24 : résumé des effets de la tâche (tâche de répétition ou dénomination), et de l'interaction entre groupe et tâche, et entre tâche et catégorie vocalique sur les formants F1, F2 et F3, ainsi que sur les dispersions F1-F2 et F2-F3 moyennes de chaque catégorie vocalique

Les résultats indiquent que l'input immédiat proposé lors de la tâche de répétition a des effets sur la production des voyelles par les enfants des deux groupes, principalement localisés sur le F1 chez les enfants NH et sur le F3 chez les enfants CI. En effet, les enfants NH ont tendance à produire les voyelles hautes et les voyelles basses de façon plus haute lors de la tâche de dénomination sans modèle audio, et les enfants CI ont une hauteur de F3 légèrement plus proche du modèle dans la tâche de répétition avec modèle adulte pour la plupart des voyelles (/e, ɛ, ø, œ, a, u, o/). Chez les deux groupes d'enfants, nous n'observons pas de différence de dispersion des catégories vocaliques, ce qui indique que le fait de disposer d'un modèle ne se traduit pas par une production de voyelles plus stables et moins dispersées. Ces différences mineures et localisées sont un argument en faveur d'une production stabilisée de voyelles, qui ne varient que très peu en hauteur, antériorité, protrusion, et dispersion selon la tâche proposée.

1.6. Effet de la complexité du mot

Nous avons choisi d'étudier l'effet de la complexité lexicale sur la production de voyelles dans des mots simples (codés par « 0 » sur nos figures) et complexes (codés par « 1 » sur nos figures) dans la seule tâche de répétition. Nous excluons donc les mots produits lors de la tâche de production sans modèle audio, analysés dans les paragraphes précédents.

1.6.1. F1

L'observation des F1 par degré de hauteur et degré de complexité du mot (Figure 5.25) indique que les deux groupes ont des réalisations similaires de la hauteur, que le mot soit simple ou complexe, à l'exception de la voyelle basse /a/.

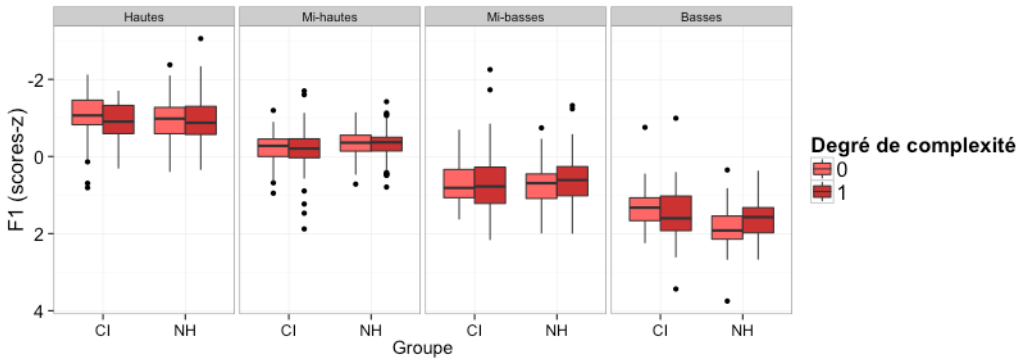


Figure 5.25 : F1 (scores-z) pour chaque degré de hauteur de voyelle, par groupe et degré de complexité du mot (mots simples, codés « 0 » ou mots complexes, codés « 1 »)

La Table 5.25 présente les résultats du test statistique que nous avons réalisé pour étudier l'effet de la complexité du mot sur le F1, seule et en interaction avec le groupe et la hauteur de la voyelle.

	numDF	denDF	F-value	p-value
(Intercept)	1	1214	0.0023	0.9620
groupe	1	31	0.0004	0.9833
hauteur	3	1214	1007.8920	<.0001 ***
complexité_mot	1	1214	0.1285	0.7201
groupe:hauteur	3	1214	6.5836	0.0002 ***
groupe:complexité_mot	1	1214	4.4372	0.0354 *
hauteur:complexité_mot	3	1214	2.7987	0.0389 *
groupe:hauteur:complexité_mot	3	1214	1.4760	0.2194

Table 5.25 : tableau de résultats au test statistique sur toutes les données (effets de groupe, degré de hauteur et degré de complexité du mot sur le F1)

Le test statistique indique une absence d'effet significatif de la complexité du mot seule sur la réalisation du F1 (F(1,1214)=0,13, p>.05, mais en revanche un effet significatif de la complexité du mot, en interaction avec le groupe (F(1,1214)=4,44, p<.05) et avec la hauteur de la voyelle (F(3,1214)=2,90, p<.05). Un test similaire réalisé en prenant en compte la catégorie vocalique plutôt que sa hauteur donne les mêmes résultats. Un test post-hoc par groupe réalisé sur l'interaction entre catégorie vocalique et complexité du mot indique que pour les enfants CI, seul l'effet de la complexité du mot est significatif pour la voyelle /u/ (z=-3,57, p<.01) dont le F1 est plus bas avec les mots simples, alors que pour les enfants NH, cet effet est significatif pour /u/ (z=-6,36, p<.001), dont le F1 est plus bas avec les mots simples, et /a/ (z=2,86, p<.05), /ɔ/ (z=7,055, p<.001) et /y/ (z=3,73, p<.01), dont les F1 sont plus élevés avec les mots simples . Par souci de lisibilité, nous ne donnons que les résultats des tests post-hoc dont les effets sont significatifs.

1.6.2.F2

Le degré de complexité du mot semble agir différemment sur le F2 des voyelles produites par les deux groupes d'enfants, selon le degré d'antériorité de la voyelle (Figure 5.26) : en effet, les voyelles antérieures non arrondies sont plus centralisées lorsque le mot est complexe, que lorsqu'il est simple (le F2 est plus bas lorsque le mot est complexe). Les autres catégories vocaliques ne semblent pas subir d'effet de la complexité du mot.

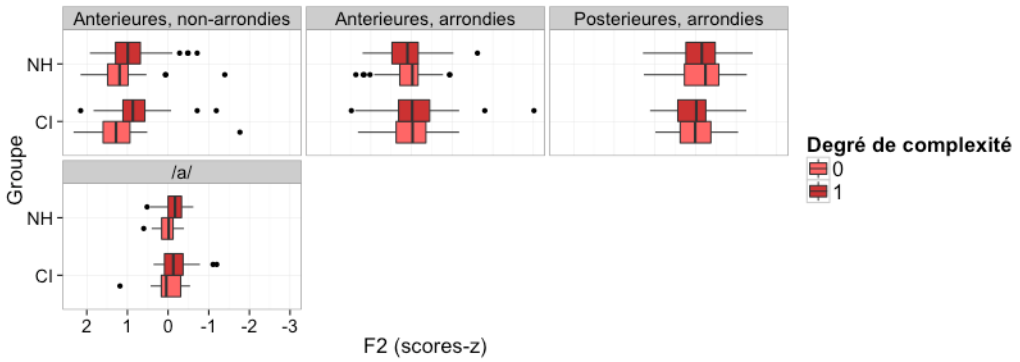


Figure 5.26 : F2 (scores-z) pour chaque degré d'antériorité de voyelle, par groupe et degré de complexité du mot (mots simples, codés « 0 » ou mots complexes, codés « 1 »)

Le test statistique (Table 5.26) indique un effet significatif de la complexité du mot seule (F(1,1214)=9,32, p<.01), mais également en interaction avec l'antériorité de la voyelle (F(3,1214)=11,43, p<.001), ce qui semble confirmer nos premières observations de la figure 5.26.

	numDF	denDF	F-value	p-value
(Intercept)	1	1214	0.0192	0.8898
groupe	1	31	0.0181	0.8937
anteriorité	3	1214	1153.8126	<.0001 ***
complexité_mot	1	1214	9.3232	0.0023 **
groupe:anteriorité	3	1214	4.4328	0.0042 **
groupe:complexité_mot	1	1214	2.0113	0.1564
anteriorité:complexité_mot	3	1214	11.4253	<.0001 ***
groupe:anteriorité:complexité_mot	3	1214	0.5258	0.6646

Table 5.26 : tableau de résultats au test statistique sur toutes les données (effets de groupe, degré d'antériorité et degré de complexité du mot sur le F2)

Comme pour le F1, un test réalisé en prenant en compte la catégorie vocalique plutôt que le degré d'antériorité de la voyelle donne les mêmes résultats pour le F2. Un test post-hoc de l'interaction entre catégorie vocalique et degré de complexité chez les enfants CI indique un effet significatif de la complexité du mot sur le F2 de /i/ (z=3,43, p<.01) et /ɛ/ (z=2,83, p<.05), ce qui est le cas chez les enfants NH pour /i/ (z=4,37, p<.001) et /ɛ/ (z=4,74, p<.001), pour lesquels nous notons également un effet significatif pour /o/ (z=-3,36, p<.01) et /ø/ (z=-4,12, p<.001) : chez les enfants CI et NH, /i/ et /ɛ/ sont plus antérieurs avec les mots simples, et chez les enfants NH seulement, /o/ et /ø/ sont plus postérieurs avec les mots simples.

1.6.3.F3

La représentation du F3 par degré de complexité du mot et par groupe (Figure 5.27) indique trois tendances, similaires pour les deux groupes : le F3 est plus haut avec les mots complexes pour /ø/, /u/ et /o/, plus bas avec les mots complexes pour /i/, /a/ et /e/ et similaire avec les deux degrés de complexité pour /e/, /y/, /œ/ et /ɔ/.

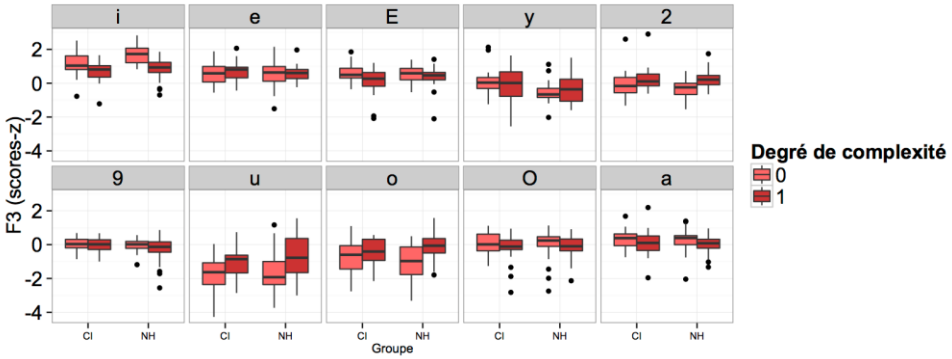


Figure 5.27 : F3 (scores-z) pour chaque catégorie vocalique, par groupe et degré de complexité du mot (mots simples, codés « 0 » ou mots complexes, codés « 1 »)

Le test statistique correspondant (Table 5.27) confirme ces observations graphiques : la complexité du mot seule n'a pas d'effet significatif sur le F3 (F(1,1190)=1,37, p>.05) mais en interaction avec la catégorie vocalique (F(9,1190)=13,47, p<.001).

	numDF	denDF	F-value	p-value
(Intercept)	1	1190	0.19302	0.6605
groupe	1	31	0.07880	0.7808
V	9	1190	106.92469	<.0001 ***
complexité_mot	1	1190	1.36990	0.2421
groupe:V	9	1190	2.99737	0.0015 **
groupe:complexité_mot	1	1190	3.35985	0.0671
V:complexité_mot	9	1190	13.47098	<.0001 ***
groupe:V:complexité_mot	9	1190	1.16305	0.3151

Table 5.27 : tableau de résultats au test statistique sur toutes les données (effets de groupe, catégorie vocalique et degré de complexité du mot sur le F3)

Des tests post-hoc de l'interaction entre degré de complexité du mot et catégorie vocalique, réalisés pour chaque groupe d'enfants indiquent un effet significatif de la complexité du mot pour /u/ (z=-3,14, p<.05) chez les enfants CI, et pour /u/ (z=-6,19, p<.001), /o/(z=-5,75, p<.001), /i/ (z=5,03, p<.001) et /ø/ (z=-3,48, p<.01) chez les enfants NH : le F3 de /u/ chez les enfants CI et NH, ainsi que ceux de /o/ et /ø/ chez les enfants NH sont plus bas avec les mots simples, et le F3 de /i/ chez les enfants NH est plus haut avec les mots simples.

Etant donné que nous n'avons qu'un mot par syllabe CV et par niveau de complexité, il importe de vérifier que l'effet de la tâche sur le F3 n'est pas dû à l'influence du contexte des autres segments adjacents aux voyelles cibles. Par exemple, pour la voyelle /o/, nous avons choisi les mots « bobo » (mot simple) et « bocal » (mot complexe). L'influence du /b/ sur le /o/ dans « bobo » devrait être un abaissement du F3, et le /k/ dans « bocal », devrait également avoir tendance à abaisser le F3 du /o/. L'abaissement de F3 observé dans les mots simples par rapport aux mots complexes ne s'explique donc pas dans ce cas par le contexte phonétique.

1.6.4.Dispersion F1-F2

La dispersion des catégories vocaliques sur le plan F1-F2 semble suivre les mêmes schémas chez les deux groupes d'enfants (Figure 5.28). En effet, seuls deux types d'effets sont observés : la dispersion est plus grande avec les mots simples pour /i/, /e/, /ɛ/, /u/, /o/ et /œ/ mais plus grande avec les mots complexes pour /y/, /ø/, /ɔ/ et /a/.

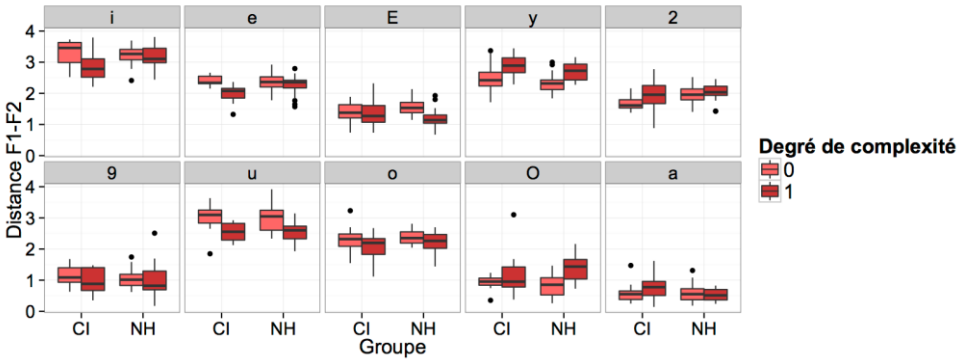


Figure 5.28 : Dispersion F1-F2 par rapport au centre de la voyelle (scores-z) pour chaque degré de hauteur de voyelle, par groupe et degré de complexité du mot (mots simples, codés « 0 » ou mots complexes, codés « 1 »)

Les résultats du test statistique (Table 5.28) semblent confirmer l'observation de la figure 5.28. Le degré de complexité du mot a un effet significatif seul (F(1,580)=10,53, p<.001), mais également en interaction avec la catégorie vocalique (F(9,580)=7,26, p<.001) et avec le groupe et la catégorie vocalique (F(9,580)=2,29, p<.05) sur la dispersion dans le plan F1-F2.

	numDF	denDF	F-value	p-value
(Intercept)	1	580	14822.258	<.0001
groupe	1	31	0.581	0.4518
V	9	580	219.070	<.0001 ***
complexité_mot	1	580	10.532	0.0012 **
groupe:V	9	580	4.627	<.0001 ***
groupe:complexité_mot	1	580	1.844	0.1750
V:complexité_mot	9	580	7.259	<.0001 ***
groupe:V:complexité_mot	9	580	2.285	0.0160 *

Table 5.28 : tableau de résultats au test statistique sur toutes les données (effets de groupe, catégorie vocalique et degré de complexité du mot sur la distance F1-F2 par rapport au centre de la catégorie vocalique)

Des tests post-hoc de comparaisons multiples réalisés par groupe indiquent un effet significatif sur la dispersion F1-F2 chez les enfants CI pour la seule voyelle /i/ (z=3,348, p<.01), mais chez les enfants NH pour les voyelles /ɛ/ (z=3,675, p<.01), /y/ (z=-3,893, p<.001), /u/ (z=4,443, p<.001) et /ɔ/ (z=-5,348, p<.001).

1.6.5.Dispersion F2-F3

La dispersion des catégories vocaliques sur le plan F2-F3, matérialisée au moyen de la mesure de distance euclidienne moyenne entre les voyelles et le centre de la catégorie correspondante semble influencée par le degré de complexité du mot et diffère d'une voyelle à l'autre. En effet, nous observons sur la Figure 5.29 que la dispersion F2-F3 est plus grande avec les mots simples pour les deux groupes avec les voyelles /i/, /e/, /ɛ/, /u/ et /o/, mais plus faible avec les mots simples pour les deux groupes avec /y/ et /œ/, et similaire avec les mots simples et complexes pour /ø/, /ɔ/ et /a/.

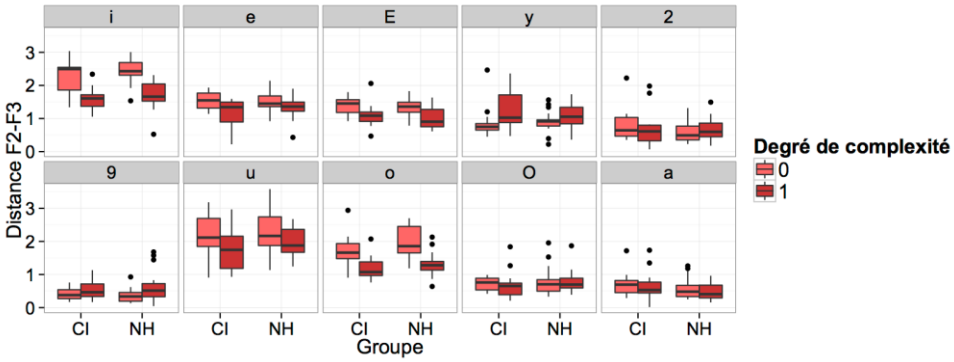


Figure 5.29 : Distance F2-F3 par rapport au centre de la voyelle (scores-z) pour chaque degré de hauteur de voyelle, par groupe et degré de complexité du mot (mots simples, codés « 0 » ou mots complexes, codés « 1 »)

Le test statistique réalisé (Table 5.29) indique un effet de la complexité du mot seule (F(1,58-=128,49, p<.001), mais également en interaction avec la catégorie vocalique (F(9,580)=9,66, p<.001).

	numDF	denDF	F-value	p-value
(Intercept)	1	580	5451.062	<.0001
groupe	1	31	0.294	0.5917
V	9	580	128.487	<.0001 ***
complexite_mot	1	580	33.381	<.0001 ***
groupe:V	9	580	1.247	0.2633
groupe:complexite_mot	1	580	0.038	0.8465
V:complexite_mot	9	580	9.660	<.0001 ***
groupe:V:complexite_mot	9	580	0.379	0.9452

Table 5.29 : tableau de résultats au test statistique sur toutes les données (effets de groupe, catégorie vocalique et degré de complexité du mot sur la distance F2-F3 par rapport au centre de la catégorie vocalique)

Les tests post-hoc réalisés par groupe indiquent que pour les enfants CI, l'effet de la complexité du mot n'est significatif que pour /i/ (z=4,033, p<.001) et /o/ (z=2,945, p<.05) alors que pour les enfants NH, il est significatif pour /i/ (z=6,328, p<.001) et /o/ (z=5,501, p<.001) mais également pour /ɛ/ (z=5,501, p<.001). Dans tous ces cas, la dispersion sur le plan F2-F3 est toujours plus grande lorsque la voyelle est produite dans un mot simple. Cette variation sur F2-F3 en fonction de la complexité du mot peut être interprétée comme une acquisition par ces enfants de stratégies de réduction vocalique semblables à celles des adultes : après une étape du développement au cours de laquelle la production des voyelles est soumise à une forte dispersion due à un manque de contrôle précis des articulateurs, puis une étape de production de voyelles très stéréotypées, une fois la maîtrise des articulateurs acquise, les enfants deviennent capables d'adapter leur production de voyelles à la situation d'énonciation, pour qu'elles puissent être identifiées comme faisant partie de la catégorie cible, malgré une variation importante de la catégorie. Ces stratégies de « formant undershoot » (non atteinte de cibles formantiques) pour les mots simples, sont assimilables à la réduction plus grande des mots de haute fréquence décrite par exemple par Bybee (2002), cit. in Meunier & Espesser (2011). La moindre variation constatée pour les mots complexes correspondrait donc, pour ces voyelles, à une production de voyelles stéréotypées, et moins maitrisées que dans les mots simples.

1.6.6. Résumé : effet de la complexité du mot

La table 5.30 ci-dessous résume les effets de la complexité du mot sur les caractéristiques acoustiques étudiés dans ce paragraphe.

Effets	Caractéristiques acoustiques				
	F1	F2	F3	Dispersion F1-F2	Dispersion F2-F3
complexité lexicale	n.s.	0 < 1 **	n.s.	0 > 1 **	0 > 1 ***
complexité lexicale*groupe	CI : 0<1 *	n.s.	n.s.	n.s.	n.s.
	NH : 0 ≃ 1 (n.s.)				
complexité lexicale* catégorie vocalique	CI : /u/ 0<1	CI : /i,ɛ/ 0>1	CI : /u/ 0<1	CI : /i/ 0>1	CI : /i,o/ 0>1
	NH : /u/ 0<1, /a,u,y/ 0>1	NH : /i,ɛ,ø/ 0>1, /o/ 0<1	NH : /u,o,ø/ 0<1, /i/ 0>1	NH : /y,ɔ/ 0>1, /ɛ,u/ 0<1	NH : /i,o,ɛ/ 0>1

Table 5.30 : résumé des effets de la complexité du mot (« 0 » pour les mots simples, et « 1 » pour les mots complexes), et de l'interaction entre groupe et complexité, et entre complexité et catégorie vocalique sur les formants F1, F2 et F3, ainsi que sur les dispersions F1-F2 et F2-F3 moyennes de chaque catégorie vocalique

Les résultats des analyses de ce paragraphe indiquent que les effets de la complexité du mot diffèrent entre les groupes et les voyelles considérées. En effet, quatre voyelles produites par les deux groupes d'enfants diffèrent lorsque le mot est simple ou complexe : /i/ (pour le F2 ainsi que les dispersions F1-F2 et F2-F3, tous plus élevés lorsque le mot est simple et pour les enfants NH, F3 plus élevé avec les mots simples), /ɛ/ (F2 plus haut avec les mots simples), /u/ (F1 plus

bas et pour les enfants NH, dispersion F1-F2 plus faible avec les mots simples) et /o/ (dispersion F2-F3 plus grande avec les mots simples). En revanche pour les enfants NH, plusieurs voyelles supplémentaires sont réalisées différemment selon la complexité du mot : /a/ (F2 plus haut avec les mots simples), /ɔ,y/ (F1 plus élevé et dispersion F1-F2 plus grande avec les mots simples), et /ɛ/ (dispersions F1-F2 plus faible et F2-F3 plus grande avec les mots simples) et /o,ø/ (F2 et F3 plus bas avec les mots simples). Comme pour les variations constatées entre les deux tâches, les différences entre mots simples et complexes ne sont pas systématiques et affectent les groupes et les catégories vocaliques de façon ponctuelle et localisées. Les voyelles hautes /i,y,u/ ainsi que les voyelles mi-basses /ø,o/ sont les plus affectées par la complexité du mot.

1.7. Discussion

La table 5.31 ci-dessous résume tous les effets sur la réalisation des voyelles, testés dans ce paragraphe.

Effets	Caractéristiques acoustiques				
	F1	F2	F3	Dispersion F1-F2	Dispersion F2-F3
Groupe	n.s.	n.s.	n.s.	n.s.	n.s.
Catégorie vocalique	/i,y,u/<</e,ø,o/<</ɛ,œ,ɔ/<</a/	/i,e,ɛ/<</y,ø,œ,a/<</u,o,ɔ/	i>e>ɛ>y>ø>œ>a>ɔ>o>u	***	***
Groupe*Catégorie vocalique	/a/ : CI < NH **	/ø/ : CI<NH *** /o/ : CI>NH ***	/y/ : CI>NH *	/a/ : CI>NH *** /ø/ : CI>NH *** /y/ CI>NH ***	/ø/ : CI<NH * /y/ : CI>NH ***
âge chronologique (NH)	n.s.	n.s.	n.s.	n.s.	n.s.
âge chronologique (NH) * Catégorie vocalique	n.s.	n.s.	n.s.	n.s.	n.s.
âge auditif (CI)	n.s.	n.s.	n.s.	n.s.	n.s.
âge à l'implantation (CI)	n.s.	n.s.	n.s.	n.s.	n.s.
âge auditif (CI)*catégorie vocalique	Augmentation avec /u/*	/y, ø, œ/ antériorisation (tendance)	n.s.	/o/ : augmentation*	/o/ : augmentation*
âge à l'implantation (CI) *catégorie vocalique	Diminution avec /y/* Augmentation avec/œ/*	n.s.	n.s.	/ø/ : augmentation*	n.s.
tâche	***	n.s.	n.s.	n.s.	n.s.
tâche*groupe	n.s.	n.s.	CI***, NH (n.s.)	n.s.	n.s.
tâche*catégorie vocalique	CI : n.s., NH D<R (voyelles hautes*, voyelle basse***)	n.s.	n.s.	n.s.	n.s.
complexité lexicale	n.s.	0 < 1 **	n.s.	0 > 1 **	0 > 1 ***
complexité lexicale*groupe	CI : 0<1 * NH : 0 ≃ 1 (n.s.)	n.s.	n.s.	n.s.	n.s.
complexité lexicale*catégorie vocalique	CI : /u/ 0<1 NH : /u/ 0<1 /a,u,y/ 0>1	CI : /i,ɛ/ 0>1 NH : /i,ɛ,ø/ 0>1 /o/ 0<1	CI : /u/ 0<1 NH : /u,o,ø/ 0<1 /i/ 0>1	CI : /i/ 0>1 NH : /y,ɔ/ 0>1 /ɛ,u/ 0<1	CI : /i,o/ 0>1 NH : /i,o,ɛ/ 0>1

Table 5.31 : résumé des effets du groupe, de la catégorie vocalique de l'âge chronologique (enfants NH) de l'âge à l'implantation (enfants CI), de l'âge auditif (enfants CI), de la tâche et de la complexité du mot, seuls et en interactions sur la réalisation des caractéristiques acoustiques des voyelles.

1.7.1. Production de voyelles chez les enfants NH

Les résultats de notre étude indiquent que les enfants NH produisent les dix voyelles du français avec quatre niveaux de hauteur (qui correspondent à la mesure du F1) et quatre niveaux d'antériorité (qui correspondent à la mesure du F2) bien distincts et similaires aux descriptions de la production de voyelles d'adultes francophones de la littérature (Delattre, 1948, Sten, 1963, Vallée, 1994, Walter, 1977 et Calliope, 1989 par exemple) : 1) les voyelles /i, y, u/ sont hautes, et se distinguent clairement de /e, ø, o/ mi-hautes et elles-mêmes distinctes des voyelles mi-basses /ɛ, œ, ɔ/, distinctes de la voyelle basse /a/, et 2) les voyelles antérieures non-arrondies /i, e, ɛ/ sont produites avec un F2 distinct de celui des antérieures arrondies /y, ø, œ/ et de /a/, elles-mêmes distinctes des postérieures /u, o, ɔ/.

Il n'y a pas d'effet de l'âge sur les valeurs formantiques pour ce groupe. Ainsi, la catégorisation des degrés de hauteur et d'antériorité semble être acquise chez le groupe des enfants NH de notre étude, qui ont tous entre 5 et 11 ans. Les résultats des voyelles produites par les enfants NH indiquent en outre que la protrusion des trois voyelles antérieures arrondies /y, ø, œ/ est également acquise, puisque ces trois voyelles sont produites avec un F3 significativement plus bas que celui de la série des voyelles antérieures non-arrondies /i, e, ɛ/.

Nous observons en outre que les voyelles produites par les enfants NH de notre étude sont peu affectées par la tâche (répétition avec modèle audio adulte ou dénomination sans modèle adulte) et la complexité du mot puisque seul le F1 des voyelles hautes et de /a/ diffère en fonction de la tâche, et que seules les voyelles /u, a, y/ pour le F1, /i, ɛ, o/ pour le F2 et /u, o, ɔ, i/ pour le F3, sont soumis à un effet de la complexité du mot : quelques différences sont observées pour ces voyelles, ce qui semblerait indiquer que la catégorisation des voyelles est globalement en voie de stabilisation à l'âge de 11 ans.

1.7.2. Production de voyelles chez les enfants CI

1.7.2.1. Effet de l'implant : groupes CI vs NH

Les résultats de notre étude de production de voyelles par les enfants sourds porteurs d'implant cochléaire indiquent qu'à l'image des enfants NH, les enfants CI réalisent des voyelles avec des niveaux de hauteur et d'antériorité clairement distincts.

Le F1 est produit globalement de la même façon chez les enfants CI et NH, sauf pour la voyelle /a/, plus haute chez les enfants CI. Ce résultat sur la voyelle /a/ contraste avec la littérature, notamment l'étude de Seifert et al. (2002) pour laquelle aucune différence entre enfants CI germanophones et enfants NH de la littérature n'était observée, celle de Ryalls et al. (2003) qui n'observaient aucune différence de F1 non plus entre enfants francophones québécois CI et NH, et de façon différente, avec celle de Baudonck et al. (2011) dans laquelle les F1 de /a/ et /i/ étaient similaires chez les enfants néerlandophones CI, HA et NH et celle de Verhoeven et al. (2016) qui notaient une tendance des enfants CI néerlandophones à produire des voyelles plus centralisées verticalement que les enfants HA et NH. Dans l'étude de Verhoeven et al. (2016), en particulier, toutes les voyelles considérées sont produites avec une hauteur ou une aperture moins distinctes chez les enfants CI que chez les enfants HA et NH, alors que dans notre étude, la voyelle /a/ est la seule pour laquelle nous observons une différence de groupe, avec un F1 plus bas chez les enfants CI.

Les F2 des voyelles produites par les enfants CI de notre étude sont globalement similaires à ceux des enfants NH, à l'exception des F2 des voyelles arrondies antérieures et postérieures, en particulier ceux de /ø/ et /o/ qui sont produites de façon plus centralisée chez les enfants CI. Des résultats similaires sur l'antériorité sont observés par Liker et al. (2007) chez des enfants locuteurs du croate, qui produisent des voyelles plus antérieures que les enfants NH, et les études de Baudonck et al. (2011), Chuang et al. (2012), Ryalls et al. (2003) et Tseng et al. (2011) : toutes ces études indiquent une variabilité plus grande du F2 par rapport au F1 entre enfants CI et NH, ce qui est également un résultat observé par Seifert et al. (2002) sur la seule voyelle /a/, dont le F2 est plus dispersé que le F1 chez les enfants CI de cette étude.

Le F3 des enfants CI est similaire à celui des enfants NH, sauf pour la voyelle /y/ dont le F3 est nettement plus haut chez les enfants CI (ce qui correspond à une voyelle moins protruse et moins

arrondie chez les enfants CI), et la voyelle /i/ dont le F3 est très légèrement plus bas chez enfants CI (ce qui correspond à une voyelle plus protruse et plus arrondie chez les enfants CI). Les seules voyelles produites différemment par les enfants CI sont donc plus proches l'une de l'autre que pour les enfants NH.

Les études de Ryalls et al. (2003), Löfqvist et al. (2010), Chuang et al. (2012) et Verhoeven et al. (2016) utilisent une mesure de taille d'espace vocalique pour caractériser la dispersion des voyelles à l'intérieur de l'espace vocalique puisque ces auteurs postulent que plus l'espace vocalique est grand, plus les voyelles sont distinctes à l'intérieur de cet espace. Les résultats qu'ils obtiennent sont identiques : l'espace vocalique des enfants CI est plus petit que ceux des enfants NH auxquels ils sont comparés dans le cadre de ces études, ce qui laisse penser que la distinction inter-catégorie est moins bien acquise par les enfants CI que les enfants NH. Cependant, les études de Baudonck et al. (2011) et Horga & Liker (2006) se distinguent de ces résultats puisque dans le premier cas, les enfants CI ont des espaces vocaliques plus grands que les enfants NH et dans le deuxième cas, aucune différence de taille n'est constatée. A notre connaissance, aucune étude n'effectue de comparaison inter-groupes de mesures de dispersion de catégories vocaliques. Les mesures de dispersions F1-F2 et F2-F3 de notre étude donnent des indications sur la variation intra-catégorie des voyelles produites par les enfants CI : les voyelles /a/, /ø/ et /y/ sont produites avec une dispersion F1-F2 plus grande et /y/ est produite avec une dispersion F2-F3 plus grande chez les enfants CI. Nos résultats indiquent donc que la variation intra-catégorie est réalisée de façon globalement proche chez les enfants CI et NH de notre étude, puisque seules trois voyelles sur dix ont des dispersions F1-F2 plus grandes chez les enfants CI et une seule voyelle sur dix a une dispersion sur F2-F3 plus grande chez les enfants CI.

1.7.2.2.Effet de l'âge à l'implantation et de l'âge auditif

Notre étude met en évidence un effet limité de l'âge auditif sur le F1 de la voyelle /u/ qui est plus basse chez les enfants CI qui utilisent leur implant depuis le plus longtemps, ainsi qu'un effet de l'âge à l'implantation sur les voyelles /y/ plus haute et /œ/ plus basse chez les enfants implantés le plus tôt. Ces résultats sont en partie similaires à ceux de Hocevar-Bolthezar et al. (2008) qui observaient une diminution du F1 de /u/ et /i/ des enfants CI locuteurs du croate avec l'âge auditif, alors que Seifert et al. (2002) n'observent aucun effet sur le F1 de l'âge chronologique, de l'âge à l'implantation ou de l'âge auditif chez des enfants CI germanophones.

Dans notre étude, le F2 de la série des voyelles antérieures arrondies est le seul a être affecté par l'âge auditif : ces voyelles sont plus postérieures chez les enfants CI et ont tendance à s'antérioriser avec une durée d'utilisation de l'implant plus longue.

Enfin, nous notons que l'âge à l'implantation et la durée d'utilisation de l'implant n'impactent pas la réalisation du F3.

Nous observons en outre que la dispersion de /o/ est sensible à l'effet de l'âge à l'implantation et à l'âge auditif.

1.7.2.3.Effet de la tâche et de la complexité du mot

Les F1 des voyelles produites par les enfants CI dans notre étude ne sont pas affectés par la tâche, puisque nous observons des F1 similaires lors de la répétition de mots avec modèles audio adultes et lors de la dénomination des mêmes mots sans modèle, ce qui est un résultat comparable à celui de l'étude de Bharadwaj et al. (2007) dans laquelle des enfants anglophones

ont des réalisations de F1 similaires avec implant activé ou désactivé. Cet absence de l'effet de la tâche, ainsi que l'effet de la complexité du mot limité à la seule voyelle /u/ semble indiquer que les enfants CI de notre étude ont globalement acquis la dimension de la hauteur, puisque celle-ci n'est que marginalement impactée par l'implant.

En outre, comme ce que nous observions sur le F1, la tâche n'affecte pas le F2 des voyelles et la complexité du mot n'affecte que deux voyelles /i/ et /ɛ/. En revanche, les enfants CI produisent des voyelles aux F3 plus proche de celui du modèle adulte lorsque celui-ci est disponible et la complexité du mot n'affecte que le F3 de /u/.

Par ailleurs, les deux mesures de dispersion ne sont pas influencées par la tâche, et seules celles de /i/ et /o/ sont affectées par la complexité du mot, ce qui semble indiquer une acquisition relativement stable de la catégorisation des voyelles par les enfants CI.

1.7.3. Focalisation F1-F2 et F2-F3 chez les deux groupes d'enfants

Si aucune des dix voyelles n'est réalisée avec une différence F2-F1 inférieure à la limite de 3,5 Bark fixée par l'étude de Chistovich & Lublinskaya (1979) puis reprise par Schwartz & Escudier (1989), Schwartz et al. (2004) et Ménard et al. (2007), nous notons que la focalisation F1-F2 est réalisée de la même façon par les deux groupes d'enfants que dans l'étude de Vaissière (2007). En effet, les voyelles /u/, /o/, /ɔ/ et /a/ sont réalisées avec une convergence F1-F2 observée pour ces mêmes voyelles chez Vaissière (2007) et Georgeton et al. (2012), ce qui indique que les enfants NH de notre groupe contrôle réalise la focalisation F1-F2 des voyelles de façon similaire à celle d'adultes francophones et que les différences d'audition n'impactent pas la réalisation de la focalisation F1-F2 chez les enfants CI.

La mesure de focalisation F3-F2 chez les enfants NH indique une convergence de ces deux formants pour les voyelles antérieures non-arrondies /i, e, ɛ/, pour les voyelles antérieures arrondies /y, ø, œ/ et pour /a/, avec une différence F3-F2 inférieure à la mesure de 3,5 Bark. Ce résultat est partiellement en accord avec la littérature sur la focalisation formantique chez les adultes francophones, puisque les études de Vaissière (2007) et Georgeton et al. (2012) ne font état que d'une seule voyelle focale F2-F3 /y/. Nos résultats semblent indiquer que les enfants NH de notre étude n'ont donc pas encore acquis une focalisation F2-F3 similaire à celle d'adultes normo-entendants, ce qui témoignerait d'une perception catégorielle et d'un contrôle moteur dans la production de voyelles encore en voie de stabilistation. A la différence des enfants NH, chez lesquels sept voyelles ont des différences F3-F2 inférieures à la limite de 3,5 Bark, les enfants CI ne réalisent que quatre voyelles /y/, /i/, /e/ et /ɛ/ avec une focalistion F2-F3, ce qui fait que leur production de voyelles focales F2-F3 est plus proche de celles des adultes francophones de la littérature (Vaissière, 2007, Georgeton et al., 2012).

1.8. Conclusion

Nous avons montré que les enfants CI de notre étude ont des difficultés très localisées sur quelques voyelles pour produire des voyelles similaires à celles des enfants NH. Nos analyses acoustiques de la production des voyelles orales chez les enfants CI ont mis en évidence une confusion entre voyelles arrondies antérieures et postérieures sur F2, avec une évolution de la production des enfants CI approchant celle des enfants NH, lorsque la durée d'utilisation de l'implant (ou âge auditif) augmente. La dimension de l'antériorité est la plus difficilement acquise par les enfants CI, ce que nous pouvons expliquer par une perception visuelle ambigue de ces

voyelles et des voyelles postérieures qui sont également arrondies, mais qui ne s'en distinguent que par la position plus ou moins antérieure de la langue. Nous avons également montré que le /y/ est produit avec un F3 plus haut chez les enfants CI. Les analyses de la dispersion de chaque catégorie vocalique ont montré que la dispersion sur F1-F2 de /y/ et /ø/ est plus grande chez les enfants CI, avec un effet de l'âge à l'implantation sur cette dispersion F1-F2 pour /ø/, qui est plus proche de celle des enfants NH pour les enfants implantés le plus précocement. La dispersion sur F2-F3 de /ø/ est plus basse chez les enfants CI, et la dispersion de /o/ sur F1-F2 et F2-F3, qui est plus basse chez les enfants CI, augmente avec la durée d'utilisation de l'implant.

Les informations fréquentielles que les enfants CI perçoivent moins bien que les enfants NH du fait de l'implant n'empêchent donc pas les enfants CI de catégoriser les voyelles, et de produire des voyelles aux caractéristiques globalement similaires à celles des enfants NH. Cependant, il est important de noter que nos résultats concernent des voyelles produites dans des mots isolés, dont nous contrôlons le contexte consonantique gauche immédiat (les voyelles sont toutes en contexte #b_), et le statut accentuel (les voyelles sont toutes placées dans une syllabe initiale de mot, artificiellement accentuée). Cette situation de production de parole est donc peu naturelle et favorise vraisemblablement une hyperarticulation de la part de nos locuteurs. Une analyse complémentaire de la production des voyelles orales en production spontanée et continue pourrait permettre de préciser certaines difficultés observées lors de notre propre étude, en étudiant les effets de la coarticulation consonantique et de l'accentuation sur la production de voyelles.

2. Occlusives

2.1. Données analysées

Notre étude prenant en compte uniquement les occlusives produites par les deux groupes d'enfants, nous avons exclu des analyses les productions non-exploitables : mots non-produits, ou non identifiables comme le mot-cible. Nos analyses sont réalisées à partir de 2183 occlusives exploitables : 99,5% des mots produits par les enfants NH et 93,6% des mots produits par les enfants CI ont donc pu être inclus dans les analyses ; un test-t (test de Welch) indique une différence significative entre les deux groupes : $t(12,262)=2,66$, $p<0.05$. Des productions de mots exploitables, nous excluons ensuite toutes les productions d'occlusives non-exploitables : les occlusives cibles sont élidées, spirantisées, ou produites avec un dévoisement de la voyelle suivante, rendant impossible la détection du début du voisement. Parmi les mots exploitables, 82,6% des consonnes produites par les enfants NH sont des occlusives exploitables et 84,3% des consonnes produites par les enfants CI sont des occlusives exploitables ; un test t (test de Welch) indique que la différence entre les productions des deux groupes d'enfants n'est pas significative : $t(23,85)=-0,41$, $p>0,05$.

Le compte détaillé des productions d'occlusives est donné dans la table 5.32 ci-dessous qui présente les valeurs moyennes du VOT (avec écart-type), par groupe, consonne et contexte vocalique, analysé dans le cadre de notre expérience de production.

Groupe	Mode de fonctionnement laryngien	Contexte vocalique	Bilabiales			Alvéolaires			Vélaires		
			Nombre total	Nombre par enfant (max=6)	VOT moyen (en ms)	Nombre total	Nombre par enfant (max=6)	VOT moyen (en ms)	Nombre total	Nombre par enfant (max=6)	VOT moyen (en ms)
CI	Non-voisées	i	64	4,92	15 (40)	70	5,38	57 (25)	67	5,15	69 (30)
		u	71	5,46	41 (30)	72	5,54	29 (22)	68	5,23	46 (33)
	Voisées	i	69	5,31	-79 (66)	69	5,31	-65 (69)	65	5,00	-39 (68)
		u	74	5,69	-56 (53)	73	5,62	-55 (65)	62	4,77	-51 (85)
NH	Non-voisées	i	116	5,80	30 (23)	101	5,05	71 (22)	117	5,85	87 (39)
		u	109	5,45	39 (18)	113	5,65	38 (16)	110	5,50	62 (24)
	Voisées	i	119	5,95	-78 (52)	116	5,80	-56 (69)	109	5,45	-30 (58)
		u	115	5,75	-55 (58)	116	5,80	-57 (70)	118	5,90	-41 (63)

Table 5.32 : valeurs moyennes du VOT (avec écart-type) par groupe, par consonne et par contexte vocalique

Comme nous cherchons à étudier la réalisation du voisement chez les deux groupes d'enfants de notre étude, nous avons décidé de nous intéresser à la distinction entre paires de consonnes voisées et non-voisées, à travers un calcul de la distance entre les deux types de consonnes pour chaque lieu d'articulation (bilabiales, alvéolaires et vélaires). A cet effet, nous avons calculé une distance moyenne entre les VOT de /p-b/, /t-d/ et /k-g/, par sujet, par contexte vocalique (/i/ ou /u/), par tâche (répétition ou dénomination) et par degré de complexité (mots simples ou mots complexes). Les valeurs moyennes de distance entre paires de consonnes sont données dans la table 5.33 ci-dessous.

Groupe	Contexte vocalique	/p-b/			/t-d/			/k-g/		
		Nombre total de données	Nombre de données par enfant (max=3)	distance moyenne entre VOT (en ms)	Nombre total de données	Nombre de données par enfant (max=6)	distance moyenne entre VOT (en ms)	Nombre total de données	Nombre de données par enfant (max=6)	distance moyenne entre VOT (en ms)
CI	i	33	2,54	92 (60)	37	2,85	121 (69)	33	2,54	105 (64)
	u	38	2,92	96 (51)	39	3,00	85 (46)	35	2,69	99 (72)
NH	i	59	2,95	109 (57)	56	2,80	128 (60)	58	2,90	119 (60)
	u	57	2,85	96 (53)	59	2,95	95 (56)	59	2,95	103 (61)

Table 5.33 : distance moyenne (en sec) entre les VOT de /p-b/, /t-d/ et /k-g/, par groupe et par contexte vocalique

Pour analyser ces occlusives, nous avons fait le choix d'utiliser des modèles linéaires mixtes qui prennent en compte à la fois des effets fixes et un effet aléatoire du sujet. Les variables à expliquer (ou variables dépendantes) ainsi que les variables explicatives (ou variables indépendantes) produisant les effets fixes sont détaillées dans la table 5.34 ci-dessous.

Variables dépendantes	VOT, calculé en secondes
	Distance entre les VOT moyens des paires d'occlusives voisées et non voisées de chaque lieu d'articulation /p-b/, /t-d/ et /k-g/, calculée en secondes
Variables indépendantes	Groupe : CI ou NH
	C : Lieu d'articulation : bilabial /p/ et /b/, alvéolaire /t/ et /d/ et vélaire /k/ et /g/
	V : Contexte vocalique : _i ou _u
	Age chronologique des enfants (NH)
	Age à l'implantation (enfants CI)
	Age auditif (qui correspond à la durée d'utilisation de l'implant pour les enfants CI)
	Tâche : tâche de répétition (R) ou tâche de dénomination (D)
	Complexité lexicale : mots simples (0) et mots complexes (1)
Variable ayant un effet aléatoire	Sujet

Table 5.34 : variables dépendantes, variables indépendantes et variable ayant un effet aléatoire, pris en compte dans l'analyse des occlusives

2.2. Effets du groupe et du lieu d'articulation

2.2.1.VOT

La figure 5.30 ci-dessous présente la distribution des VOT obtenus pour chaque groupe, lieu d'articulation et mode de fonctionnement laryngien (voisé ou non-voisé).

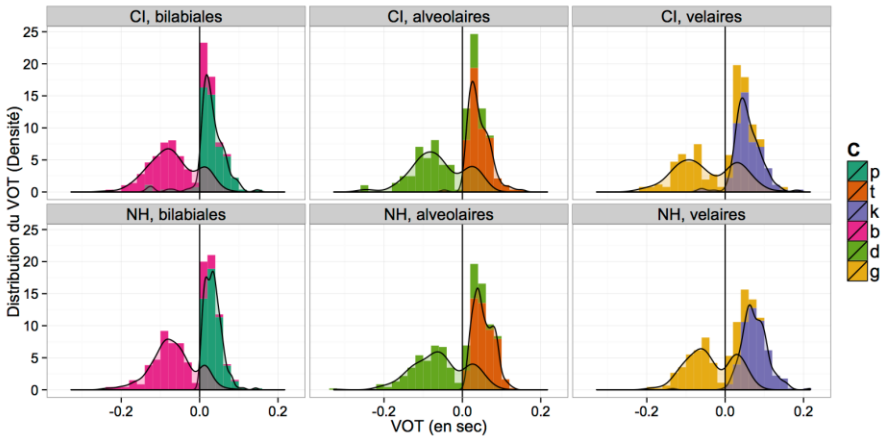


Figure 5.30 : distribution du VOT (en sec) par groupe et lieu d'articulation

Une première observation visuelle de ces distributions indique que la frontière catégorielle ne semble pas aussi nette que ce qui est décrit dans la littérature (Caramazza & Yeni-Komshian, 1974, Abdelli-Beruh, 2009 et Nearey & Rochet, 1994 pour les adultes francophones, et Whalen et al., 2007 pour les enfants francophones), et que les deux groupes d'enfants produisent pour chaque lieu d'articulation des consonnes voisées avec un VOT aussi bien positif que négatif. En revanche, les enfants produisent beaucoup plus rarement des consonnes non-voisées avec un VOT négatif, et les quelques enfants chez qui cette tendance est observée font tous partie du groupe des enfants porteurs d'implant cochléaire (à l'exception d'une seule occurrence de /k/ produite avec un prévoisement par un enfant NH).

La figure 5.31 ci-dessous présente l'effet du lieu d'articulation et du groupe sur le VOT. Nous avons choisi de présenter ces deux effets conjointement, puisqu'ils peuvent donner lieu également à un effet d'interaction entre groupe et lieu d'articulation.

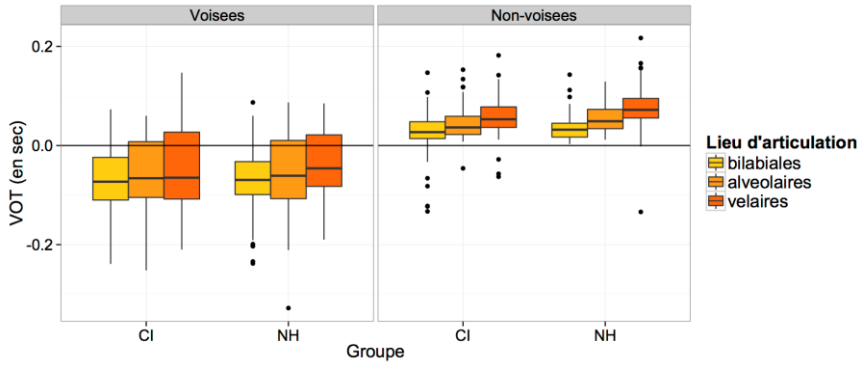


Figure 5.31 : VOT (en sec) par groupe et lieu d'articulation

La table 5.35 ci-dessous présente les résultats au test statistique sur toutes les données : le test ne met pas en évidence d'effet de groupe (F(1,31)=1,40, p>.05) mais un effet du lieu d'articulation (F(5,2140)=559,29, p<.001).

	numDF	denDF	F-value	p-value
(Intercept)	1	2140	1.0149	0.3138
C	5	2140	559.2935	<.0001 ***
groupe	1	31	1.3977	0.2461
C:groupe	5	2140	1.4040	0.2196

Table 5.35 : tableau de résultats au test statistique sur toutes les données (effets de groupe, lieu d'articulation sur le VOT)

D'une part, les consonnes non-voisées /p/, /t/ et /k/ sont produites avec un VOT positif et clairement distinct de 0, alors que les consonnes voisées /b/, /d/ et /g/ sont produites avec un VOT négatif ou légèrement positif. Le contraste de voisement des consonnes est donc bien visible au niveau du VOT qui permet de distinguer consonnes voisées et non-voisées. D'autre part, la figure 5.31 et la table 5.35 mettent en évidence une augmentation du VOT en fonction de la postériorité du lieu d'articulation : les consonnes bilabiales non-voisées sont produites avec un VOT inférieur à celui des alvéolaires non-voisées, lui même inférieur à celui des vélaires non-voisées. Cette augmentation du VOT en fonction de la postériorité du lieu d'articulation, bien que moins visible, est inversée en valeur absolue pour les consonnes voisées : la bilabiale /b/ a le prévoisement le plus long, suivi de l'alvéolaire /d/ puis de la vélaire /g/.

Des tests statistiques réalisés sur les données des consonnes voisées d'une part et non-voisées d'autre part nous permettent d'affiner ces résultats : l'effet du groupe et du lieu d'articulation n'est pas similaire entre les deux modes de fonctionnement laryngien. En effet, si nos résultats indiquent un effet du lieu d'articulation sur le VOT des consonnes voisées ($F(2,1062)=23,00$, $p<.001$) et sur celui des consonnes non-voisées ($F(2,1035)=190,20$, $p<.001$), nous n'observons pas d'effet de l'interaction du groupe et du lieu d'articulation pour les consonnes voisées ($F(2,1062)=0,33$, $p>.01$) alors que nous observons cet effet d'interaction du groupe et du lieu d'articulation pour les consonnes non-voisées ($F(2,1035)=4,46$, $p<.05$).

Un test post-hoc supplémentaire (CI vs NH) présenté dans la table 5.36 ci-dessous nous permet de caractériser plus précisément cet effet d'interaction pour les consonnes non-voisées (nous n'avons pas réalisé de test post-hoc similaire pour les consonnes voisées, puisque nous n'observions pas d'effet d'interaction entre groupe et lieu d'articulation sur ces consonnes) : la vélaire /k/ est la seule consonne pour laquelle un effet significatif du groupe est observé, les enfants NH ayant tendance à produire cette vélaire avec un VOT plus long que les enfants CI.

Consonne	VOT
	CI vs NH
/p/	z=-1,30, p>.05 (n.s.)
/t/	z=-2,13, p>.05 (n.s.)
/k/	z=-3,50, p<.01 **

Table 5.36 : résumé du test post-hoc sur l'effet d'interaction entre groupe et lieu d'articulation sur le VOT des consonnes non-voisées

2.2.2. Distance entre VOT des occlusives voisées et non-voisées

Les études disponibles dans la littérature sur la réalisation du voisement chez les enfants CI font état d'une difficulté persistante de ces enfants à produire des consonnes voisées et non-voisées distinctes (voir entre autres l'étude de Baudonck et al. 2010 en néerlandais, celle de Horga & Liker, 2003 en croate, celle de Uchanski & Geers, 2003, en anglais américain). Une étude sur l'acquisition du voisement chez une population d'enfants NH (Whiteside et al., 2003, en anglais américain), indique que le contraste de voisement serait acquis tardivement, entre les âges de 9 et 11 ans, mais une étude de Whalen et al. (2007) indique que le prévoisement serait acquis dès la fin de la première année de vie chez l'enfant francophone. Notre étude étant réalisée auprès d'enfants francophones, nous nous attendons à constater une réalisation du voisement par ces enfants qui distingue clairement consonnes voisées (caractérisées par un VOT négatif ou prévoisement) et consonnes non-voisées (caractérisées par un VOT positif). Pour cette raison, nous avons décidé d'étudier le contraste de voisement à travers une mesure de la distance entre les deux types de consonnes (voisées et non-voisées).

La figure 5.32 ci-dessous présente les mesures de distances entre moyennes de VOT des consonnes voisées et non-voisées ainsi obtenues, pour chaque lieu d'articulation et pour chaque groupe. Nous constatons que la distance semble similaire chez chaque groupe pour chaque lieu d'articulation et que seules les vélaires semblent présenter une différence entre les groupes.

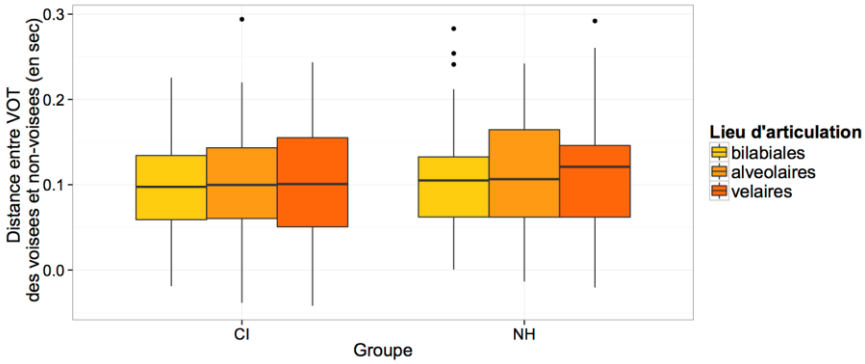


Figure 5.32 : distance entre les VOT des consonnes voisées et non-voisées (en sec) par lieu d'articulation, pour chaque groupe

Nous avons réalisé un test statistique pour étudier l'effet du groupe et du lieu d'articulation sur la distance entre les VOT des consonnes voisées et non-voisées, dont les résultats sont présentés dans la table 5.37 ci-dessous.

	numDF	denDF	F-value	p-value
(Intercept)	1	526	315.90944	<.0001
C	2	526	1.62013	0.1989
groupe	1	31	0.51965	0.4764
C:groupe	2	526	0.08405	0.9194

Table 5.37 : tableau de résultats au test statistique sur toutes les données (effets du groupe sur la distance entre VOT des consonnes voisées et non-voisées)

Le test statistique réalisé ne permet pas de dégager d'effets du lieu d'articulation (F(2,526)=1,62, p>.05) ni du groupe (F(1,31)=0,52, p>.05) seuls ou en interaction (F(2,526)=0,08, p>.05), ce qui semble indiquer que les deux groupes d'enfants ont une réalisation similaire du contraste de voisement et donc de la catégorisation en production des consonnes voisées et non-voisées.

2.2.3. Résumé : effets du groupe et du lieu d'articulation

La table 5.38 ci-dessous présente un résumé des effets du groupe et du lieu d'articulation sur le VOT des occlusives et sur les distances entre VOT des consonnes voisées et non-voisées, testés dans ce paragraphe.

Effets	Mesures		
	VOT		Distance entre VOT des consonnes voisées et non-voisées
	Non-voisées	Voisées	
Groupe	n.s.		n.s.
Lieu d'articulation	/p/ < /t/ < /k/ ***	/g/ < /d/ < /b/ (n.s.)	n.s.
Groupe*Lieu d'articulation	/p/ : n.s.	/b/ : n.s.	n.s.
	/t/ : n.s.	/d/ : n.s.	n.s.
	/k/ : CI < NH **	/g/ : n.s.	n.s.

Table 5.38 : résumé des effets des variables groupe, lieu d'articulation et de l'interaction entre groupe et lieu d'articulation sur le VOT et la distance entre VOT des consonnes voisées et non-voisées

Ces premiers résultats indiquent dans un premier temps que les VOT des consonnes non-voisées varient selon le lieu d'articulation chez les deux groupes d'enfants, ce qui est en conformité avec la littérature sur le VOT en français (Calliope, 1989, Caramazza & Yeni-Komshian, 1974, Abdelli-Beruh, 2009 et Nearey & Rochet, 1994) : le VOT varie en fonction de la postériorité de la consonne. Ce résultat est identique chez les deux groupes d'enfants (CI & NH), ce qui semble

indiquer que les enfants des deux groupes ont acquis une maîtrise similaire de la production des consonnes non-voisées. Les consonnes voisées ne sont quant à elles pas soumises à un effet du lieu d'articulation. La seule étude du prévoisement des occlusives voisées en français (Caramazza & Yeni-Komshian, 1974) ne mettait pas en évidence de distinction entre consonnes voisées selon le lieu d'articulation, et nos résultats vont dans ce sens : les consonnes voisées des deux groupes d'enfants de notre étude se distinguent des consonnes non-voisées, mais nos analyses ne montrent pas d'effet du lieu d'articulation sur le VOT pour les consonnes voisées.

L'effet de l'utilisation de l'implant n'est visible que pour l'occlusive vélaire non-voisée /k/. En effet, si les enfants CI ont tendance à produire des occlusives non-voisées avec un VOT légèrement plus court que celui des enfants NH, /k/ est la seule consonne pour laquelle nous mettons en évidence un effet du groupe.

Dans le paragraphe suivant, nous allons poursuivre ces analyses et affiner ces résultats en prenant en compte l'effet du contexte vocalique sur la réalisation du voisement des consonnes occlusives.

2.3. Effet du contexte vocalique

Pour analyser l'effet du contexte vocalique, nous avons ajouté la variable « contexte vocalique » (désignée « V » dans les tables de ce paragraphe) aux modèles utilisés pour les tests statistiques réalisés aux paragraphes §2.2.1 et §2.2.2 qui ne prenaient en compte que les effets du lieu d'articulation et du groupe sur la réalisation du VOT. Comme pour les analyses précédentes, nous étudions l'effet du contexte vocalique seul mais également en interaction avec le groupe d'une part et le lieu d'articulation d'autre part.

2.3.1.VOT

La figure 5.33 ci-dessous présente une représentation du VOT pour chaque lieu d'articulation, en fonction du groupe et du contexte vocalique. Une première observation graphique semble indiquer que les occlusives voisées et non voisées ne sont pas affectées de la même manière par le contexte vocalique : le contexte vocalique semble avoir un effet moindre sur le VOT des consonnes voisées que sur celui des consonnes non-voisées.

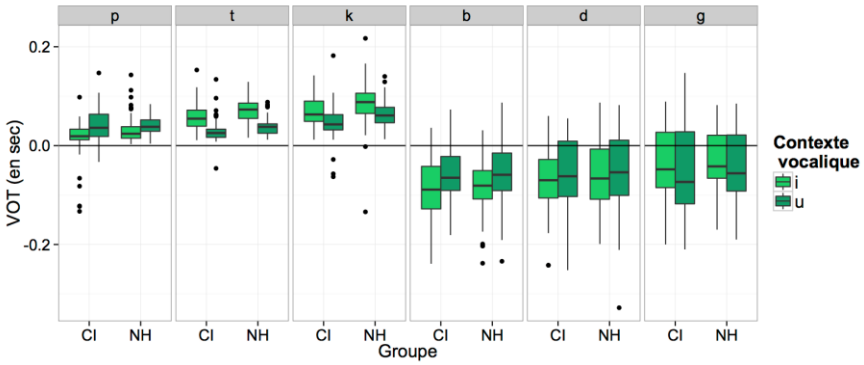


Figure 5.33 : VOT (en sec) par groupe, par lieu d'articulation et par contexte vocalique

La table 5.39 ci-dessous présente les résultats du test statistique de l'effet du contexte vocalique sur la réalisation du VOT, seul et en interaction avec le lieu d'articulation (la procédure de sélection de variables réalisée préalablement nous a conduites à ne pas inclure le groupe dans le modèle utilisé pour le test).

	numDF	denDF	F-value	p-value
(Intercept)	1	2139	0.9380	0.3329
C	5	2139	584.4802	<.0001 ***
V	1	2139	3.4786	0.0623
C:V	5	2139	19.8059	<.0001 ***

Table 5.39 : tableau de résultats au test statistique sur toutes les données (effets de groupe, lieu d'articulation et contexte vocalique sur le VOT)

Les résultats du test statistique indiquent que le contexte vocalique seul n'a pas d'effet sur le VOT ($F(1,2139)=3,48$, $p>.05$) mais que l'interaction entre contexte vocalique et consonne a un effet significatif sur le VOT ($F(5,2139)=19,81$, $p<.001$). L'observation de la figure 5.33 ci-dessus indique que les VOT des deux groupes d'enfants subissent des effets similaires du contexte vocalique : le VOT est plus court lorsque la consonne est suivie de /i/ pour /p/ et /g/, mais plus long lorsqu'elle est suivie de /i/ pour /t/, /k/, /b/ et /d/.

Un test post-hoc de comparaisons multiples indique cependant que cet effet du contexte vocalique n'est significatif que pour /p/, /t/, /k/ et /b/. La table 5.40 ci-dessous détaille ces résultats.

Consonne	VOT
	Contexte vocalique : /i/ vs /u/
/p/	z=-3,09, p<.05 *
/t/	z=6,20, p<.001 ***
/k/	z=4,91, p<.001 ***
/b/	z=-4,93, p<.001 ***
/d/	z=-0,67, p>.01 (n.s.)
/g/	z=2,34, p>.05 (n.s.)

Table 5.40 : résumé du test post-hoc sur l'effet d'interaction entre contexte vocalique et lieu d'articulation sur le VOT

Ces résultats sont en contradiction avec ceux de Nearey & Rochet (1994) qui montraient que le VOT est influencé par le contexte vocalique, mais que cet effet est lié à la hauteur de la voyelle et non à son antériorité : le VOT est plus court lorsque la consonne est suivie d'une voyelle haute ou mi-haute. Les voyelles que nous avons choisies sont toutes les deux des voyelles hautes (/i/ et /u/), et ne devraient pas affecter la réalisation du VOT.

2.3.2. Distance entre VOT des occlusives voisées et non-voisées

La figure 5.34 ci-dessous présente la distance entre VOT des consonnes voisées et non-voisées par lieu d'articulation et contexte vocalique, pour chaque groupe. L'observation de cette figure indique que, comme pour le VOT de chaque consonne, la différence de voisement aux trois lieux d'articulation n'est pas affectée de la même façon par le contexte vocalique : alors que le contraste de voisement des bilabiales et vélaires ne semble pas être affecté par le contexte vocalique, le contraste entre alvéolaires voisées et non-voisées semble fortement affecté par celui-ci puisque la distance entre /t/ et /d/ est visiblement plus marquée lorsque la consonne est suivie de /i/.

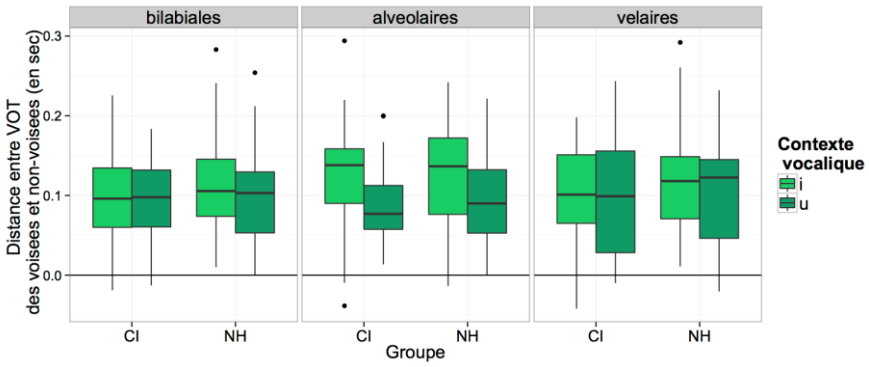


Figure 5.34 : distance entre les VOT des consonnes voisées et non-voisées (en sec) par lieu d'articulation et par contexte vocalique, pour chaque groupe

Comme pour les valeurs de VOT, nous avons ajouté la variable « contexte vocalique » au modèle utilisé au paragraphe §2.2.2. La table 5.41 ci-dessous présente les résultats de ce test statistique étudiant les effets du contexte vocalique sur la distance entre consonnes voisées et non-voisées. Notre procédure de sélection de variables nous a conduites à ne pas inclure la variable « groupe » dans les analyses.

	numDF	denDF	F-value	p-value
(Intercept)	1	525	304.40889	<.0001
C	2	525	1.70485	0.1828
V	1	525	19.70346	<.0001 ***
C:V	2	525	3.84998	0.0219 *

Table 5.41 : tableau de résultats au test statistique sur toutes les données (effets du contexte vocalique sur la distance entre VOT des consonnes voisées et non-voisées)

Les résultats indiquent un effet significatif du contexte vocalique (F(1,525)=19,70, p<.001) seul et en interaction avec le lieu d'articulation (F(2,525)=3,85, p<.05). Un test post-hoc de comparaisons multiples nous permet ensuite d'affiner ce premier résultat (table 5.42).

Consonne	Distance entre VOT
	Contexte vocalique : /i/ vs /u/
/p-b/	z=1,08, p>.05 (n.s.)
/t-d/	z=4,81, p<.001 ***
/k-g/	z=1,77, p>.05 (n.s.)

Table 5.42 : résumé du test post-hoc sur l'effet d'interaction entre contexte vocalique et lieu d'articulation sur la distance entre les VOT des consonnes voisées et non-voisées

Le test statistique confirme l'observation visuelle : les alvéolaires sont les seules consonnes affectées par le contexte vocalique. Le contraste entre /t/ et /d/ est en effet plus marqué lorsque la consonne est suivie de /i/ que de /u/, pour les deux groupes d'enfants.

2.3.3. Résumé : effet du contexte vocalique

La table 5.43 ci-dessous présente un résumé des résultats des analyses réalisées dans ce paragraphe : effets du contexte vocalique et des interactions entre contexte vocalique et groupe et entre contexte vocalique et lieu d'articulation.

Effets	Mesures		
	VOT		Distance entre VOT des consonnes voisées et non-voisées
	Non-voisées	Voisées	
Contexte vocalique	n.s.		n.s.
Contexte vocalique * Groupe	n.s.		n.s.
Contexte vocalique * Lieu d'articulation	/p/ : i < u *	/b/ : i > u ***	/p-b/ : n.s.
	/t/ : i > u ***	/d/ : n.s.	/t-d/ : i > u ***
	/k/ : i > u ***	/g/ : n.s.	/k-g/ : n.s.

Table 5.43 : résumé des effets des variables groupe, contexte vocalique et des interactions entre groupe et contexte vocalique et contexte vocalique et lieu d'articulation sur le VOT et la distance entre VOT des consonnes voisées et non-voisées

Les résultats présentés dans ce chapitre ont pu mettre en évidence un effet de coarticulation anticipatoire de la voyelle sur certaines consonnes : /p/, /t/, /k/ et /b/, ainsi que sur le contraste de voisement des alvéolaires. Ce résultat est surprenant, au regard de la littérature disponible sur la coarticulation anticipatoire de la voyelle sur la consonne (Nearey & Rochet, 1994), qui prédit un effet de coarticulation liée à la hauteur de la voyelle, et nous pouvons émettre l'hypothèse que les consonnes non-voisées d'une part, et les alvéolaires d'autre part qui sont les consonnes les plus affectées par la coarticulation anticipatoire de la voyelle reflètent une acquisition des contrastes de lieu d'articulation et de voisement non stabilisée par les enfants de notre étude.

2.4. Enfants NH : effet de l'âge chronologique

Comme nous cherchons à étudier les facteurs qui influencent la réalisation du contraste de voisement pour pouvoir caractériser son évolution au cours du développement, nous avons choisi d'étudier l'effet de l'âge sur le VOT des occlusives. Nous étudions séparément l'effet de l'âge chronologique chez les enfants NH et celui de l'âge à l'implantation et de la durée d'utilisation de l'implant (ou âge auditif) chez les enfants CI. Dans ce paragraphe, nous allons présenter les résultats chez les enfants NH.

2.4.1.VOT

Dans un premier temps, nous avons choisi d'étudier l'effet de l'âge chronologique chez les enfants NH, en ajoutant ce facteur à notre modèle initial et en prenant en compte le contexte vocalique, puisque celui-ci a un effet significatif sur le VOT. La figure 5.35 ci-dessous présente l'évolution du VOT chez les enfants NH en fonction de l'âge chronologique, par consonne et contexte vocalique.

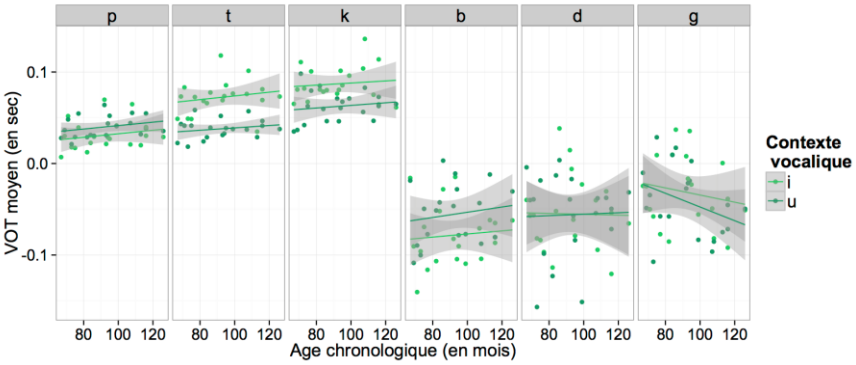


Figure 5.35 : VOT (en sec) en fonction de l'âge chronologique (en mois) par contexte vocalique et lieu d'articulation, pour les enfants NH

Une première observation de la figure 5.35 indique que le VOT semble évoluer de façon différente avec l'âge chronologique, selon la consonne : le VOT des consonnes non-voisées, ainsi que celui de /g/ augmentent légèrement avec l'âge chronologique, alors que celui de /b/ diminue faiblement et celui de /d/ est stable. La table 5.44 ci-dessous présente les résultats du test statistique que nous avons réalisé pour étudier l'effet de l'âge chronologique.

	numDF	denDF	F-value	p-value
(Intercept)	1	1323	0.0345	0.8526
C	5	1323	405.1719	<.0001 ***
V	1	1323	5.5331	0.0188 *
âge_chronologique	1	18	0.0004	0.9834
C:V	5	1323	12.1195	<.0001 ***
C:âge_chronologique	5	1323	3.1112	0.0085 ***

Table 5.44 : tableau de résultats au test statistique sur les données des enfants NH (effets de groupe, lieu d'articulation, contexte vocalique et âge chronologique sur le VOT)

Ces résultats indiquent que l'âge chronologique seul n'a pas d'effet significatif sur la réalisation du VOT ($F(1,18)=0,0004$, $p>.05$) mais qu'il a un effet en interaction avec le lieu d'articulation ($F(5,1323)=3,11$, $p<.001$).

Des tests par lieu d'articulation ont donc été réalisés pour essayer d'affiner ce résultat et pour comprendre si l'âge chronologique affecte différemment le VOT de chaque lieu d'articulation. Les résultats de ces tests par lieu d'articulation se révèlent cependant non significatifs, et ne nous permettent pas de mettre en évidence un effet de l'âge chronologique sur le VOT, peut-être en raison d'un nombre insuffisant de mesures de VOT pour chaque lieu d'articulation.

2.4.2. Distance entre VOT des occlusives voisées et non-voisées

La figure 5.36 ci-dessous présente la distance entre consonnes voisées et non-voisées par lieu d'articulation et contexte vocalique, pour les enfants NH, en fonction de l'âge chronologique. L'observation des distances entre consonnes voisées et non-voisées semble indiquer que la distance entre /k/ et /g/ augmente avec l'âge chronologique, alors que les distances entre /p/ et /b/ d'une part et /t/ et /d/ d'autre part semblent stables.

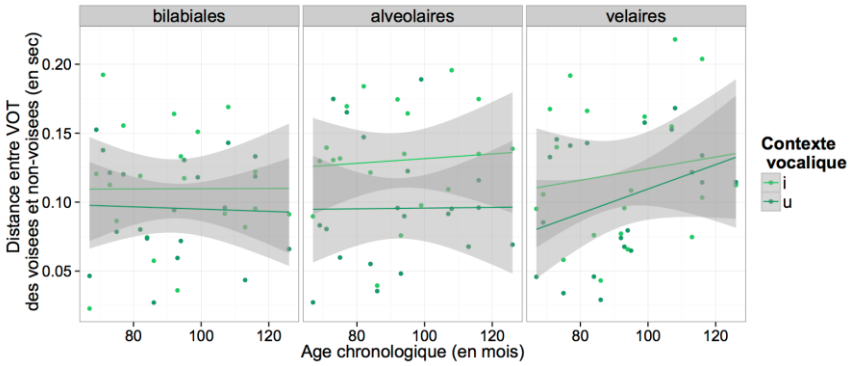


Figure 5.36 : Distance entre VOT des consonnes voisées et non-voisées (en sec) en fonction de l'âge chronologique (en mois) par contexte vocalique et lieu d'articulation, pour les enfants NH

Les résultats du test statistiques de la table 5.45 ci-dessous indiquent que l'augmentation de la distance entre /k/ et /g/ observée visuellement ne se traduit pas par un effet significatif de l'âge chronologique sur cette même distance ($F(1,28)=0,27$, $p>.05$).

	numDF	denDF	F-value	p-value
(Intercept)	1	318	184.12372	<.0001
C	2	318	1.46545	0.2325
V	1	318	17.31473	<.0001 ***
âge_chronologique	1	18	0.27063	0.6092
C:V	2	318	1.36729	0.2563
C:âge_chronologique	2	318	2.26261	0.1058
V:âge_chronologique	1	318	0.06098	0.8051
C:V:âge_chronologique	2	318	0.40962	0.6642

Table 5.45 : tableau de résultats au test statistique sur les données des enfants NH (effets de groupe, lieu d'articulation, contexte vocalique et âge chronologique sur la distance entre VOT des consonnes voisées et non-voisées)

2.4.3. Résumé : effet de l'âge chronologique (enfants NH)

La table 5.46 présente un résumé des effets de l'âge chronologique sur les deux mesures de VOT étudiées dans ce paragraphe.

Effets	Mesures		
	VOT		Distance entre VOT des consonnes voisées et non-voisées
	Non-voisées	Voisées	
âge chronologique (NH)	n.s.		n.s.
âge chronologique (NH)*Lieu d'articulation	n.s.		n.s.
âge chronologique (NH) * Contexte vocalique	n.s.		n.s.
âge chronologique (NH) * Lieu d'articulation * Contexte vocalique	n.s.		n.s.

Table 5.46 : résumé des effets de l'âge chronologique seul et en interactions avec le groupe, le contexte vocalique et le lieu d'articulation sur le VOT et la distance entre VOT des consonnes voisées et non-voisées chez les enfants NH

Les résultats que nous obtenons indiquent que l'âge chronologique chez cette population d'enfants NH de 5;7 à 10;6 ans n'affecte pas la réalisation du voisement. Une explication peut être avancée : conformément à l'étude de Whalen et al. (2007) qui mettait en évidence l'acquisition du prévoisement des occlusives voisées dès la première année chez des bébés francophones (à la différence du contraste d'aspiration chez des bébés anglophones du même âge), nos résultats nous laisseraient penser que la coordination des gestes oro-laryngés nécessaire pour produire le prévoisement caractéristique des consonnes voisées du français et distinctif des consonnes non-voisées est acquise précocement par les enfants francophones et que les enfants NH de notre étude ont déjà une réalisation stabilisée du contraste de voisement des occlusives ; une deuxième explication possible nous semble au contraire moins pertinente : cette stabilisation pourrait n'être pas encore terminée, et ne pas évoluer dans notre échantillon de population, si l'on se base sur l'étude de Whiteside et al. (2003) qui montre que le contraste de voisement n'est acquis qu'à un âge compris entre 9 et 11 ans. Cette étude est cependant réalisée auprès d'enfants anglophones, pour lesquels le contraste de voisement se réalise entre consonnes ayant toutes un VOT positif, et c'est pour cette raison que nous aurions plutôt tendance à penser que le contraste de voisement est déjà globalement acquis par les enfants NH de notre étude. Une étude complémentaire incluant des productions d'adultes et d'adolescents pourrait nous permettre d'affiner nos résultats et de replacer nos résultats dans un contexte plus large d'évolution de la réalisation du voisement et de la distinction entre lieux d'articulation au cours de la vie.

2.5. Enfants CI : effet de l'âge à l'implantation et de la durée d'utilisation de l'implant

Comme pour les enfants NH, nous nous intéressons aux effets de l'âge sur la réalisation du VOT chez les enfants CI. Plutôt que de prendre en compte l'âge chronologique, nous avons choisi de considérer conjointement l'âge à l'implantation et la durée d'utilisation de l'implant (ou âge auditif).

2.5.1.VOT

La figure 5.37 ci-dessous présente l'évolution du VOT en fonction de l'âge à l'implantation et de l'âge auditif (ou durée d'utilisation de l'implant) chez les enfants CI. Une observation visuelle de ces deux figures semble indiquer une diminution du VOT des consonnes non-voisées et une augmentation du VOT des consonnes voisées avec l'âge à l'implantation, et une tendance inverse en fonction de l'âge à l'implantation. Les résultats présentés au paragraphe précédent §2.2 faisaient état d'une tendance chez les enfants CI à produire des consonnes non-voisées avec un VOT plus court que les enfants NH et à produire des consonnes voisées avec un prévoisement plus long. Il semblerait que cette tendance soit guidée par l'âge à l'implantation et l'âge auditif : plus les enfants ont reçu l'implant tôt, plus ils produisent des consonnes non-voisées avec un VOT et un prévoisement courts et approchant ceux des enfants NH, et plus ils ont une durée d'utilisation de l'implant prolongée, plus leurs occlusives ont des caractéristiques de VOT et de prévoisement proches de celles des enfants NH.

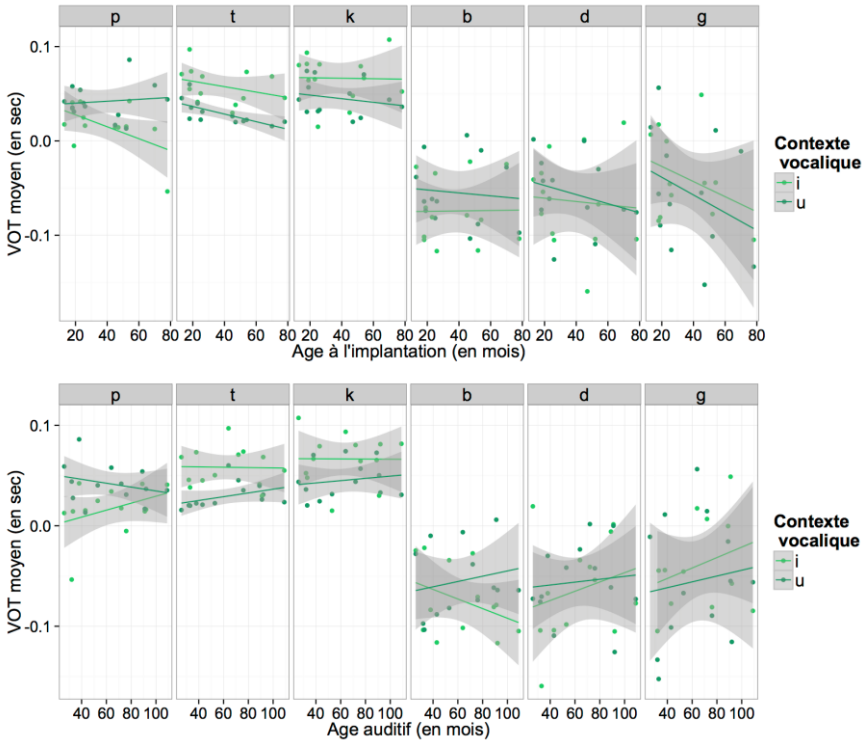


Figure 5.37 : évolution du VOT (en sec) en fonction de l'âge à l'implantation (haut) et de l'âge auditif (bas), en mois, par contexte vocalique et lieu d'articulation, pour les enfants CI

Comme pour l'étude de l'effet de l'âge chronologique chez les enfants NH, nous avons donc ajouté les deux variables « âge à l'implantation » et « âge auditif » au modèle utilisé précédemment, pour en étudier les effets sur le VOT. Les résultats de ce test sont présentés dans la table 5.47 ci-dessous.

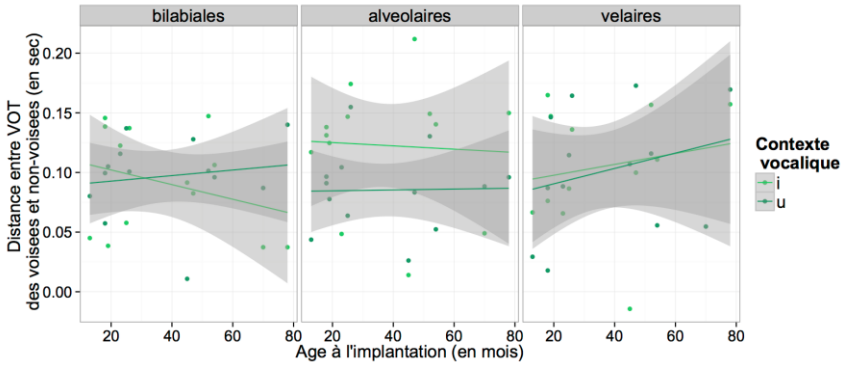
	numDF	denDF	F-value	p-value
(Intercept)	1	767	1.51909	0.2181
C	5	767	195.84850	<.0001
V	1	767	0.01127	0.9155
âge_implantation	1	9	1.55065	0.2445
âge_auditif	1	9	0.18977	0.6734
C:V	5	767	8.46098	<.0001 ***
C:âge_implantation	5	767	1.86624	0.0979
V:âge_implantation	1	767	0.04909	0.8247
C:âge_auditif	5	767	0.53076	0.7531
V:âge_auditif	1	767	0.17724	0.6739
âge_implantation:âge_auditif	1	9	0.19134	0.6721
C:V:âge_implantation	5	767	1.26331	0.2778
C:V:âge_auditif	5	767	3.80320	0.0021 **
C:âge_implantation:âge_auditif	5	767	3.62274	0.0030 **
V:âge_implantation:âge_auditif	1	767	0.08726	0.7678
C:V:âge_implantation:âge_auditif	5	767	3.41108	0.0047 **

Table 5.47 : tableau de résultats au test statistique sur les données des enfants CI (effets de groupe, lieu d'articulation, contexte vocalique, âge auditif et âge à l'implantation sur le VOT)

Les résultats indiquent que l'âge auditif ($F(1,9)=0,19$, $p>.05$) et l'âge à l'implantation ($F(1,9)=1,55$, $p>.05$) seuls n'ont pas d'effet significatif sur le VOT des occlusives, mais que l'âge auditif a un effet en interaction avec le lieu d'articulation et le contexte vocalique ($F(5,767)=3,80$, $p<.01$). Des tests supplémentaires par lieu d'articulation indiquent que seul l'âge auditif a un effet sur /p/ ($F(1,118)=10,57$, $p<.01$) et /t/ en interaction avec le contexte vocalique ($F(1,9)=5,53$, $p<.05$). Nous observons enfin un effet de l'interaction entre lieu d'articulation, contexte vocalique, âge à l'implantation et âge auditif indiquant une évolution différente du VOT avec l'âge auditif selon l'âge à l'implantation. Les tests post-hoc réalisés par lieu d'articulation ne permettent cependant pas d'identifier précisément quelles consonnes sont concernées par cet effet d'interaction.

2.5.2. Distance entre VOT des occlusives voisées et non-voisées

La figure 5.38 ci-dessous présente l'évolution de la distance entre consonnes voisées et non-voisées en fonction de l'âge à l'implantation et en fonction de l'âge auditif, chez les enfants CI. Les trois lieux d'articulation semblent suivre des trajectoires d'évolution différentes en fonction de l'âge à l'implantation et de l'âge auditif : la distance entre les vélaires semble augmenter plus fortement lorsque les enfants ont reçu l'implant le plus tard, mais ne semble pas évoluer avec l'âge auditif, au contraire des bilabiales et alvéolaires, qui semblent plus sensibles à l'âge auditif qu'à l'âge à l'implantation.



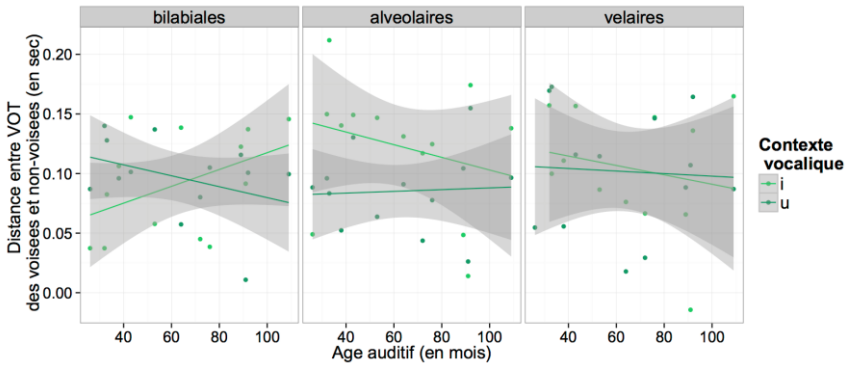


Figure 5.38 : évolution de la distance entre VOT des consonnes voisées et non-voisées (en sec) en fonction de l'âge à l'implantation (haut) et de l'âge auditif (bas), en mois, par contexte vocalique et lieu d'articulation, pour les enfants CI

La table 5.48 ci-dessous présente les résultats du test statistique des effets de l'âge à l'implantation et de l'âge auditif sur la distance entre VOT consonnes voisées et non-voisées.

	numDF	denDF	F-value	p-value
(Intercept)	1	182	151.34521	<.0001
C	2	182	0.44160	0.6437
V	1	182	4.16239	0.0428 *
âge_auditif	1	9	0.10773	0.7502
âge_implantation	1	9	0.02661	0.8740
C:V	2	182	2.87242	0.0591
C:âge_auditif	2	182	0.71353	0.4913
V:âge_auditif	1	182	0.32436	0.5697
C:âge_implantation	2	182	1.91630	0.1501
V:âge_implantation	1	182	3.61657	0.0588
âge_auditif:âge_implantation	1	9	1.41043	0.2654
C:V:âge_auditif	2	182	4.61389	0.0111 *
C:V:âge_implantation	2	182	1.45906	0.2352
C:âge_auditif:âge_implantation	2	182	0.55325	0.5760
V:âge_auditif:âge_implantation	1	182	0.04330	0.8354
C:V:âge_auditif:âge_implantation	2	182	5.73457	0.0038 *

Table 5.48 : tableau de résultats au test statistique sur les données des enfants CI (effets de groupe, lieu d'articulation, contexte vocalique, âge à l'implantation et âge auditif sur la distance entre VOT des consonnes voisées et non-voisées)

Les résultats indiquent un effet de l'âge auditif sur la distance entre VOT des consonnes voisées et non-voisées, en interaction avec le contexte vocalique et le lieu d'articulation, mais pas d'effet de l'âge auditif seul ($F(1,9)=0,11$, $p>.05$), ni de l'âge à l'implantation seul ($F(1,9)=0,03$, $p>.05$) ou en interaction avec d'autres variables.

Des tests supplémentaires par lieu d'articulation indiquent un effet de l'âge auditif en interaction avec le contexte vocalique pour les seules bilabiales ($F(1,54)= 7,73$, $p<.01$) et un effet de l'âge à l'implantation en interaction avec le contexte vocalique pour les seules alvéolaires ($F(1,60)=4,50$, $p<.05$) : lorsque la durée d'utilisation de l'implant augmente, la différence entre les VOT des alvéolaires voisées et non voisées diminue en contexte /i/, pour se rapprocher de celle des alvéolaires en contexte /u/, qui reste stable. Aucun effet d'âge n'est observé pour la distance entre /k/ et /g/.

Comme pour les analyses présentées au paragraphe §2.5.1, nous observons un effet de l'interaction entre lieu d'articulation, contexte vocalique, âge à l'implantation et âge auditif. Des

analyses supplémentaires par lieu d'articulation et contexte vocalique ne permettent pas de préciser cet effet conjoint de l'âge à l'implantation et de l'âge auditif sur la distance entre consonnes voisées et non-voisées.

2.5.3. Résumé : effets de l'âge à l'implantation et de la durée d'utilisation de l'implant (enfants CI)

La table 5.49 ci-dessous présente un résumé des effets de l'âge à l'implantation et de l'âge auditif sur les mesures de VOT et de distances entre VOT des consonnes voisées et non-voisées, testés dans ce paragraphe.

Effets	Mesures		
	VOT		Distance entre VOT des consonnes voisées et non-voisées
	Non-voisées	Voisées	
âge auditif (CI)	n.s.		n.s.
âge à l'implantation (CI)	n.s.		n.s.
âge auditif (CI)*Lieu d'articulation	n.s.		n.s.
âge auditif (CI) * Contexte vocalique	n.s.		n.s.
âge à l'implantation (CI)*Lieu d'articulation	n.s.		n.s.
âge à l'implantation (CI) * Contexte vocalique	n.s.		n.s.
âge à l'implantation (CI) * Contexte vocalique * Lieu d'articulation	n.s.	n.s.	/p-b/ : diminution avec /i/, augmentation avec /u/
âge auditif (CI) * Contexte vocalique * Lieu d'articulation	/p/ et /t/ : diminution avec /i/ et augmentation avec /u/	n.s.	n.s.

Table 5.49 : résumé des effets de l'âge à l'implantation et de l'âge auditif, seuls et en interactions avec le groupe, le contexte vocalique et le lieu d'articulation sur le VOT et la distance entre VOT des consonnes voisées et non-voisées chez les enfants CI

Nos résultats indiquent que l'âge auditif (qui correspond à la durée d'utilisation de l'implant) et l'âge à l'implantation influencent légèrement la production des consonnes occlusives et la réalisation du contraste de voisement. L'âge à l'implantation a un effet sur le contraste de voisement des bilabiales, différent selon le contexte vocalique, et l'âge auditif a un effet sur la réalisation des bilabiales et alvéolaires non-voisées, sensible également au contexte vocalique. Les résultats ne permettent en revanche pas de mettre en évidence des effets globaux et systématiques de l'âge à l'implantation et de l'âge auditif sur la réalisation des contrastes consonantiques et du contraste de voisement.

2.6. Effet de l'input auditif immédiat

2.6.1. VOT

La figure 5.39 ci-dessous présente le VOT par lieu d'articulation, par tâche (tâches de répétition et de dénomination), pour chaque groupe d'enfant (CI et NH) et pour le modèle audio adulte utilisé lors de la tâche de répétition. L'observation de cette figure semble indiquer que les VOT des consonnes des deux groupes sont similaires dans les deux tâches (tâches de répétition et de dénomination) et que les consonnes non-voisées sont réalisées de façon plus proche du modèle que les consonnes voisées. Ces dernières sont assez éloignées du modèle adulte présenté. Celui-ci semble avoir été réalisé avec un prévoisement important, qui témoigne d'une hyperarticulation qui ne se répercute pas sur les productions des enfants des deux groupes (CI et NH).

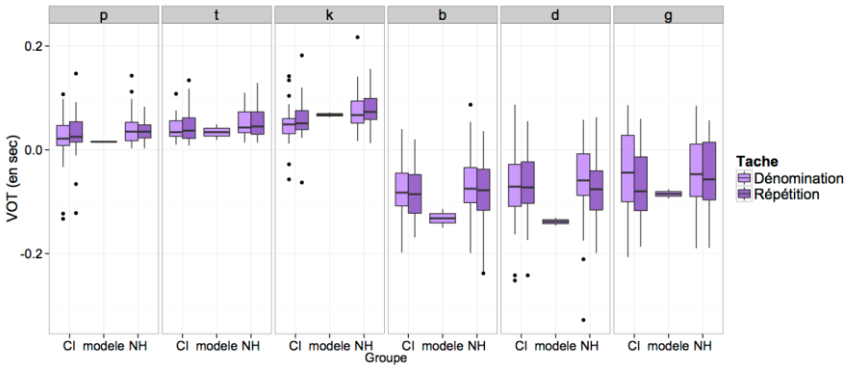


Figure 5.39 : VOT (en sec) par lieu d'articulation, par groupe et par tâche (répétition ou dénomination)

La table 5.50 ci-dessous présente les résultats du test statistique sur l'effet de la tâche sur la réalisation du VOT des consonnes occlusives.

	numDF	denDF	F-value	p-value
(Intercept)	1	1411	5.2555	0.0220
C	5	1411	433.3632	<.0001 ***
groupe	1	31	1.8496	0.1836
tâche	1	1411	1.3462	0.2461
C:groupe	5	1411	1.1874	0.3130
C:tâche	5	1411	1.4114	0.2172
groupe:tâche	1	1411	0.5413	0.4620
C:groupe:tâche	5	1411	1.0295	0.3986

Table 5.50 : tableau de résultats au test statistique sur les toutes données (effets de groupe, lieu d'articulation, et tâche sur le VOT)

Les tests réalisés indiquent que ni le VOT du groupe des enfants NH ni celui des enfants CI ne semblent affectés par la tâche ($F(1,1411)=1,35$, $p>.05$) : les VOT sont similaires chez les deux groupes dans les deux tâches. Des tests similaires réalisés séparément sur les consonnes voisées et non-voisées indiquent un effet de la tâche sur le VOT ($F(1,695)=5,39$,$p<.05$) pour les consonnes voisées et de l'interaction entre groupe et tâche sur le VOT ($F(1,677)=4,07$, $p<.05$) pour les consonnes non-voisées.

Pour les consonnes non-voisées, un test post-hoc de cette interaction entre groupe et tâche met en évidence un effet significatif de la tâche sur le VOT pour les enfants CI mais pas pour les enfants NH. La table 5.51 ci-dessous propose un résumé de ce test post-hoc.

Groupe	VOT
	Dénomination vs Répétition
CI	z=-2,80, p<.05 *
NH	z=-0,22, p>.05 (n.s.)

Table 5.51 : résumé du test post-hoc sur l'effet d'interaction entre groupe et tâche sur le VOT des consonnes non-voisées /p/, /t/ et /k/

2.6.2. Distance entre VOT des occlusives voisées et non-voisées

La figure 5.40 propose une représentation de la distance entre VOT des consonnes voisées et non-voisées pour chaque lieu d'articulation et chaque tâche (tâches de répétition et dénomination), ainsi que la distance entre VOT du modèle adulte proposé lors de la tâche de répétition. L'observation semble indiquer des réalisations du contraste de voisement similaires

dans les deux tâches, pour chaque lieu d'articulation à l'exception des alvéolaires chez les enfants NH, et des vélaires chez les enfants CI, pour lesquelles la distance entre VOT est plus marquée en répétition et plus proche du modèle adulte proposé lors de cette tâche.

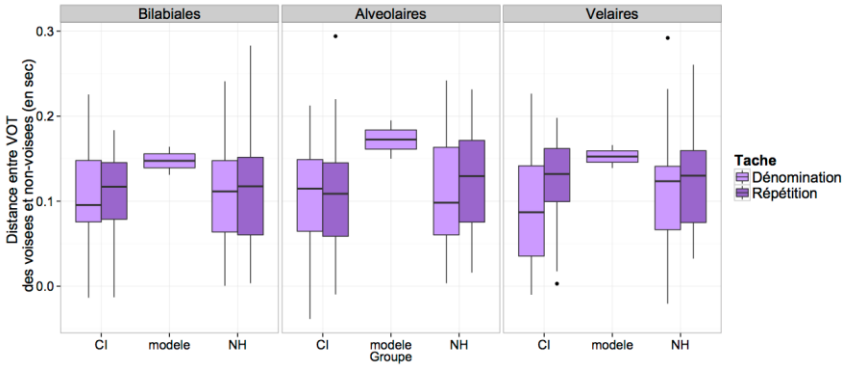


Figure 5.40 : Distance entre VOT des consonnes voisées et non-voisées (en sec) par lieu d'articulation, par groupe et par tâche (répétition ou dénomination)

La table 5.52 ci-dessous présente les résultats du test statistique réalisé pour étudier l'effet de la tâche sur la distance entre VOT des consonnes voisées et non-voisées.

	numDF	denDF	F-value	p-value
(Intercept)	1	345	298.31882	<.0001
tâche	1	345	6.13175	0.0138*

Table 5.52 : tableau de résultats au test statistique sur toutes les données (effets de groupe, lieu d'articulation et tâche sur la distance entre VOT des consonnes voisées et non-voisées)

Les résultats du test statistique indiquent un effet global de la tâche ($F(1,345)=6,13$, $p<.05$) sur la distance entre VOT des consonnes voisées et non-voisées. Des tests supplémentaires réalisés par groupe et lieu d'articulation permettent de mettre en évidence un effet sur les vélaires chez les enfants CI, qui produisent un contraste entre vélaires voisées et non-voisées atypique et moins fort lorsqu'ils ne disposent pas du modèle adulte, mais qui sont capables de s'adapter au modèle et de s'en approcher lorsque celui-ci est présenté. La table 5.53 ci-dessous propose un résumé de ces tests par groupe et lieu d'articulation.

Groupe	Lieu d'articulation	Distance entre VOT des consonnes voisées et non-voisées
		Dénomination vs Répétition
CI	bilabiales	$F(1,32)=0,05$, $p>.05$ (n.s.)
	alvéolaires	$F(1,38)=0,23$, $p>.05$ (n.s.)
	vélaires	$F(1,35)=5,03$, $p<.05$ *
NH	bilabiales	$F(1,58)=0,32$, $p>.05$ (n.s.)
	alvéolaires	$F(1,55)=1,87$, $p>.05$ (n.s.)
	vélaires	$F(1,56)=1,16$, $p>.05$ (n.s.)

Table 5.53 : résumé des tests de l'effet de la tâche sur la distance entre VOT des consonnes voisées et non-voisées, par groupe et lieu d'articulation

2.6.3. Résumé : effet de l'input audio immédiat

La table 5.54 propose un résumé de tous les effets testés, seuls ou en interaction, dans ce paragraphe.

Effets	Mesures		
	VOT		Distance entre VOT des consonnes voisées et non-voisées
	Non-voisées	Voisées	
Tâche	n.s.		D < R *
Tâche*groupe	CI : D < R *	n.s.	CI : /k-g/ : D < R *
Tâche*Lieu d'articulation			n.s.

Table 5.54 : résumé de l'effet de la tâche (Dénomination ou Répétition), seule et en interaction avec le groupe et le lieu d'articulation sur le VOT et la distance entre VOT des consonnes voisées et non-voisées

Les résultats présentés dans ce paragraphe indiquent certains effets de la tâche sur la réalisation du VOT : le groupe des enfants CI subit une influence de la tâche et produit des VOT de consonnes non-voisées plus courts lorsqu'ils ne disposent pas du modèle adulte, et une différence entre vélaires voisées et non-voisées plus marquée lorsqu'ils disposent du modèle adulte. On peut noter que les modèles de mots, en particulier ceux utilisés dans la production de consonnes voisées, ont été hyperarticulés par l'adulte et sont produits avec un prévoisement plus important que chez les deux groupes d'enfants. De façon générale nous aurions pu anticiper un prévoisement plus marqué dans la production des deux groupes d'enfants si ceux-ci calquaient leurs productions sur le modèle adulte. Comme ce n'est pas le cas chez les enfants NH, nous pouvons envisager que l'acquisition des occlusives est stabilisée pour ce groupe d'enfants, chez qui le contraste de voisement n'est pas réalisé différemment s'ils disposent d'un input auditif immédiatement disponible ou s'ils produisent les occlusives en se basant sur une représentation stabilisée du contraste de lieu d'articulation et de voisement. Chez les enfants CI cependant, le voisement des vélaires est sensible à la tâche proposée. Ceci pourrait indiquer que les enfants CI sont toujours dans un processus de stabilisation de la catégorisation voisée-non voisée pour les vélaires. Cette spécificité des vélaires est cohérente avec la littérature sur le développement phonologique (Jakobson, 1941, 1949, Ingram, 1992, 1999, Locke & Pearson, 1992, Kent, 1992).

2.7. Effet de la complexité du mot

2.7.1.VOT

La figure 5.41 ci-dessous présente le VOT par groupe et lieu d'articulation en fonction du degré de complexité du mot (codé « 0 » pour les mots simples et « 1 » pour les mots complexes). Une première observation de cette figure semble indiquer un traitement différent des consonnes non-voisées qui ne semblent pas affectées par la complexité du mot, et les consonnes voisées, pour lesquelles nous observons une différence entre mots simples et complexes : pour les consonnes voisées, le prévoisement est moins marqué lorsque les consonnes se trouvent dans les mots complexes, ce qui témoignerait d'une hyperarticulation lorsque les mots sont plus simples (mots mono- ou bisyllabiques, caractérisés par des structures syllabiques sans clusters consonantiques) ou d'une hypoarticulation lorsque les mots sont complexes.

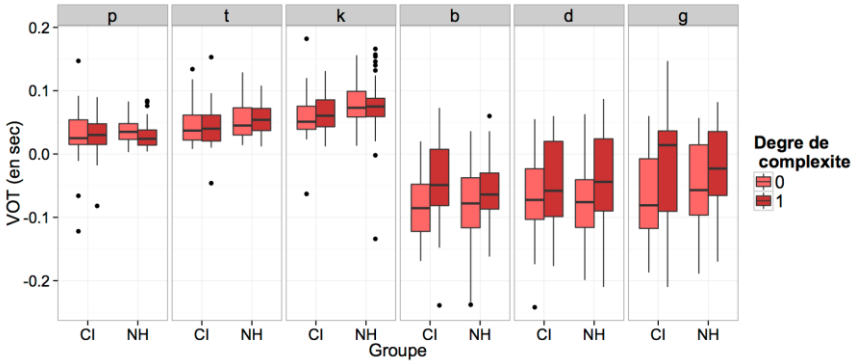


Figure 5.41 : VOT (en sec) par lieu d'articulation, par groupe et par degré de complexité du mot

La table 5.55 ci-dessous présente les résultats du test statistique réalisé pour étudier l'effet de la complexité du mot sur la réalisation du VOT.

	numDF	denDF	F-value	p-value
(Intercept)	1	1408	0.2947	0.5873
C	5	1408	394.7888	<.0001 ***
complexité_mot	1	1408	30.2241	<.0001 ***
groupe	1	31	0.5516	0.4632
C:complexité_mot	5	1408	7.4147	<.0001 ***
C:groupe	5	1408	1.6690	0.1390
complexité_mot:groupe	1	1408	0.7505	0.3865
C:complexité_mot:groupe	5	1408	0.7305	0.6006

Table 5.55 : tableau de résultats au test statistique sur toutes les données (effets de groupe, lieu d'articulation, contexte vocalique et degré de complexité sur le VOT)

Les résultats indiquent un effet du degré de complexité du mot sur le VOT seul (F(1,1408)=30,22, p<.001) et en interaction avec le lieu d'articulation (F(5,1408)=7,41, p<.001). Un test post-hoc nous permet de caractériser plus précisément cette interaction : le VOT des consonnes voisées est affecté par le degré de complexité du mot. Les résultats de ce test sont résumés dans la table 5.56 ci-dessous.

Consonne	VOT
	Degré de complexité : 0 vs 1
/p/	z=0,68, p>.05 (n.s.)
/t/	z=-0,04, p>.05 (n.s.)
/k/	z=-0,15, p>.05 (n.s.)
/b/	z=-3,94, p<.001 ***
/d/	z=-4,15, p<.001 ***
/g/	z=-5,90, p<.001 ***

Table 5.56 : résumé du test post-hoc sur l'effet d'interaction entre degré de complexité du mot et lieu d'articulation sur le VOT

Des tests supplémentaires par groupe confirment cet effet de la complexité du mot sur le VOT pour les enfants NH, seule (F(1,885)=14,95, p<.001) et en interaction avec le lieu d'articulation (F(5,885)=4,86, p<.001) et pour les enfants CI, seule (F(1,523)=14,67, p<.001) et en interaction avec le lieu d'articulation (F(5,523)=3,27, p<.01).

Des tests post-hoc réalisés séparément chez chaque groupe d'enfants indiquent une différence dans les lieux d'articulation affectés par la complexité du mot : pour les enfants NH, seules /d/ et

/g/ sont influencées par la complexité du mot, alors que pour les enfants CI, seules /b/ et /g/ le sont.

2.7.2. Distance entre VOT des occlusives voisées et non-voisées

La figure 5.42 ci-dessous présente la distance entre VOT des occlusives voisées et non-voisées par lieu d'articulation, groupe et degré complexité du mot. Une première observation nous indique que cette distance entre consonnes est plus marquée lorsque le mot est simple (codé « 0 » sur la figure) que lorsque celui-ci est complexe (codé « 1 » sur la figure). Cette tendance est constatée pour les deux groupes, aux trois lieux d'articulation.

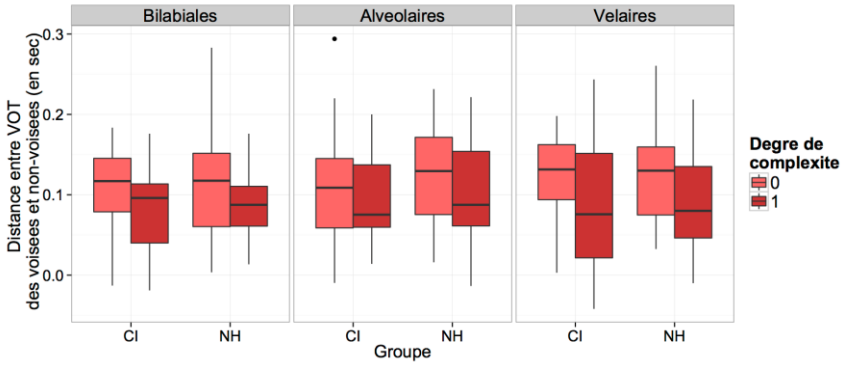


Figure 5.42 : Distance entre VOT des consonnes voisées et non-voisées (en sec) par lieu d'articulation, par groupe et par degré de complexité (mots simples ou mots complexes)

La table 5.57 ci-dessous présente les résultats du test statistique de l'effet de la complexité du mot sur la distance entre VOT des consonnes voisées et non-voisées.

	numDF	denDF	F-value	p-value
(Intercept)	1	334	298.26973	<.0001
C	2	334	2.20764	0.1116
complexité_mot	1	334	30.33152	<.0001 ***
groupe	1	31	0.37676	0.5438
C:complexité_mot	2	334	0.39207	0.6760
C:groupe	2	334	0.07407	0.9286
complexité_mot:groupe	1	334	0.01155	0.9145
C:complexité_mot:groupe	2	334	0.02930	0.9711

Table 5.57 : tableau de résultats au test statistique sur toutes les données (effets de groupe, lieu d'articulation et degré de complexité du mot sur la distance entre VOT des consonnes voisées et non-voisées)

Le test indique un effet de la complexité du mot seule ($F(1,334)=30,33$, $p<.001$), ce qui confirme l'impression visuelle : les deux groupes sont affectés de la même façon par ce facteur, et ce aux trois lieux d'articulation, ce qui témoigne d'un contraste entre consonnes voisées et non-voisées plus marqué lorsque les mots sont plus simples. Ceci concorde avec l'étude sur le VOT du paragraphe précédent, qui montrait que les enfants ont tendance à produire des consonnes voisées avec un prévoisement plus marqué lorsque les mots sont plus simples.

2.7.3. Résumé : effet de la complexité du mot

La table 5.58 ci-dessous présente un résumé de l'effet de la complexité du mot, testé dans ce paragraphe, seul et en interaction avec le groupe et le lieu d'articulation.

Effets	Mesures		
	VOT		Distance entre VOT des consonnes voisées et non-voisées
	Non-voisées	Voisées	
complexité	n.s.	0 > 1	0 > 1
complexité*groupe	n.s.		n.s.
complexité*Lieu d'articulation	n.s.	/b/ : 0 > 1 ***	n.s.
		/d/ : 0 > 1 ***	
		/g/ : 0 > 1 ***	

Table 5.58 : résumé de l'effet de la complexité du mot, seule et en interaction avec le groupe et le lieu d'articulation sur le VOT et la distance entre VOT des consonnes voisées et non-voisées

Les résultats indiquent que le degré de complexité du mot affecte les consonnes voisées, qui sont produites avec un prévoisement plus important lorsque le mot est simple, et ce qui se traduit par une distance entre consonnes voisées et non-voisées plus marquée lorsque les mots sont simples. Nous n'observons pas de différence entre les deux groupes d'enfants, CI et NH, ce qui semble indiquer que cet effet n'est pas lié à l'audition des enfants, mais plutôt à la coordination de gestes articulatoires qui sont acquis de la même façon par les deux groupes d'enfants : les enfants auraient tendance à hypoarticuler les occlusives voisées placées dans les mots qui leur sont les moins familiers ou hyperarticuler celles qui se trouvent dans les mots qui leur sont les plus familiers. Ces résultats pourraient indiquer que l'apprentissage de la production des consonnes se fait au niveau du mot, et n'est pas généralisé au phonème, puisque nous observons des différences entre consonnes produites dans les mots familiers et les mots non familiers, en particulier pour les consonnes voisées.

Nous notons en outre que cette variation des occlusives voisées dans les mots complexes vs simples ne peut pas être attribuée à un effet de la prosodie, puisque tous les mots étudiés dans ce paragraphe ont été produits dans la tâche de répétition, dont la prosodie était strictement contrôlée.

2.8. Discussion

La table 5.59 ci-dessous présente un résumé de tous les effets testés dans ce paragraphe sur la production d'occlusives par les enfants NH et CI de notre étude.

Effets	Mesures		
	VOT		Distance entre VOT des consonnes voisées et non-voisées
	Non-voisées	Voisées	
Groupe	n.s.		n.s.
Lieu d'articulation	/p/ < /t/ < /k/ ***	/g/ < /d/ < /b/ (n.s.)	n.s.
Groupe*Lieu d'articulation	/p/ : n.s.	/b/ : n.s.	n.s.
	/t/ : n.s.	/d/ : n.s.	n.s.
	/k/ : CI < NH **	/g/ : n.s.	n.s.
Contexte vocalique	n.s.		n.s.
Contexte vocalique * Groupe	n.s.		n.s.
Contexte vocalique * Lieu d'articulation	/p/ : i < u *	/b/ : i > u ***	/p-b/ : n.s.
	/t/ : i > u ***	/d/ : n.s.	/t-d/ : i > u ***
	/k/ : i > u ***	/g/ : n.s.	/k-g/ : n.s.
âge chronologique (NH)	n.s.		n.s.
âge chronologique (NH)*Lieu d'articulation	n.s.		n.s.
âge chronologique (NH) * Contexte vocalique	n.s.		n.s.
âge chronologique (NH) * Lieu d'articulation * Contexte vocalique	n.s.		n.s.
âge auditif (CI)	n.s.		n.s.
âge à l'implantation (CI)	n.s.		n.s.
âge auditif (CI)*Lieu d'articulation	n.s.		n.s.
âge auditif (CI) * Contexte vocalique	n.s.		n.s.
âge à l'implantation (CI)*Lieu d'articulation	n.s.		n.s.
âge à l'implantation (CI) * Contexte vocalique	n.s.		n.s.
âge à l'implantation (CI) * Contexte vocalique * Lieu d'articulation	n.s.	n.s.	/p-b/ : diminution avec /i/, augmentation avec /u/
âge auditif (CI) * Contexte vocalique * Lieu d'articulation	/p/ et /t/ : diminution avec /i/ et augmentation avec /u/	n.s.	n.s.
Tâche	n.s.		D < R *
Tâche*groupe	CI : D < R *	n.s.	CI : /k-g/ : D < R *
Tâche*Lieu d'articulation	n.s.		n.s.
complexité	n.s.	0 > 1	0 > 1
complexité*groupe	n.s.		n.s.
complexité*Lieu d'articulation	n.s.	/b/ : 0 > 1 ***	n.s.
		/d/ : 0 > 1 ***	
		/g/ : 0 > 1 ***	

Table 5.59 : résumé de tous les effets étudiés dans le cadre de notre étude de production d'occlusives par les enfants CI et NH (« n.s. » indique une différence non significative)

2.8.1. Production d'occlusives chez les enfants NH

Notre étude montre que les enfants NH produisent des occlusives similaires à celles des adultes francophones de France dont la production a été décrite dans la littérature (Caramazza & Yeni-Komshian, 1974, Abdelli-Beruh, 2009 et Nearey & Rochet, 1994) : les occlusives /p, t, k/ non-voisées des enfants NH ont un VOT positif, ce qui correspond à une apparition du voisement consécutive à l'explosion consonantique, et les occlusives /b, d, g/ un VOT négatif, qui correspond à un prévoisement initié avant l'explosion consonantique. Nous observons également une différence de durée du VOT significative entre les trois lieux d'articulation pour les consonnes non-voisées mais pas pour les consonnes voisées. La distance observée entre les VOT moyens par paires de consonnes voisées et non-voisées aux trois lieux d'articulation n'est pas

significative, ce qui indique une coordination entre contrôle laryngé et articulateurs similaire pour les trois lieux d'articulation.

Il ressort de notre étude que même les enfants NH les plus jeunes réalisent un contraste de voisement de façon déjà stable à cet âge, ce qui est en accord avec l'étude de Whalen et al. (2007). En effet, celle-ci montrait que chez les enfants francophones, le contraste de voisement qui se caractérise par un prévoisement des occlusives voisées, est en voie d'acquisition dès la fin de leur première année, contrairement à ce qui est observé chez des enfants anglophones du même âge, pour qui le voisement, caractérisé en anglais par un VOT positif pour consonnes non-voisées et voisées, se met en place plus tardivement. Ces résultats sont renforcés par l'absence d'effet de l'âge chronologique sur la réalisation des VOT, que ce soit pour les consonnes voisées que pour les non-voisées, et ce à chaque lieu d'articulation considéré.

Par ailleurs, notre étude montre que la coordination des gestes oro-laryngées par les enfants NH est acquise de façon robuste, puisque le VOT ne subit aucune influence de la tâche proposée (répétition vs dénomination). Par contre, nous observons une influence de la complexité du mot sur les consonnes voisées : les VOT des consonnes voisées sont légèrement plus longs lorsque les mots sont plus fréquents et simples que lorsqu'ils sont moins fréquents et plus complexes. Deux hypothèses peuvent expliquer cet effet : la familiarité du mot ou sa complexité articulatoire. Dans notre étude, les mots simples sont aussi les mots familiers puisqu'ils correspondent à des mots fréquents dans le lexique des enfants. On peut formuler l'hypothèse que l'acquisition du voisement des occlusives est un processus descendant, à partir du mot vers le phonème et que les mots moins familiers sont acquis plus tardivement. Ceci expliquerait que les occlusives voisées placées dans les mots les moins familiers, qui sont également les mots complexes sont plus sujettes à la variation. Une autre hypothèse serait que les mots « simples » de notre corpus sont les mots qui ont ont un degré plus faible de complexité articulatoire, puisqu'ils sont plus courts et ne comportent pas de clusters consonantiques. Dans une étude sur le VOT de consonnes produites dans des logatomes, Yu et al. (2014) ont montré un effet du nombre de syllabes sur la durée du VOT (le VOT de /p/ est plus long dans /pa/ que dans /pataka/). Nous pouvons rapprocher cet effet du nombre de syllabes dans le logatome sur la durée du VOT (Yu et al., 2014) à celui du degré de complexité que nous observons dans le cadre de notre étude, puisque les mots simples sont mono- ou bisyllabiques et les mots complexes sont bi- ou tri-syllabiques.

La coarticulation que nous observons est relativement inattendue puisqu'elle est liée à l'antériorité de la voyelle alors que la littérature (Nearey & Rochet, 1994) indique que la coarticulation anticipatoire pour les occlusives serait liée à la hauteur de la voyelle. Le contexte vocalique dans notre étude étant alternativement les deux voyelles hautes /i/ et /u/, nous nous attendions donc à obtenir des durées de VOT indépendantes du contexte vocalique. Nous observons pour /p/ un VOT plus long et donc plus proche de celui de /t/ lorsque celui-ci est suivi de /u/, alors que pour /t/ et /k/ le VOT est plus court lorsque ces consonnes sont suivies de /i/, ce qui indiquerait qu'elles sont produites de façon plus antérieure sous l'influence de /i/. Nous notons également un prévoisement plus important pour /b/ lorsque celle-ci est suivie de /i/ et une distance plus importante pour la paire /t-d/ lorsque ces deux occlusives sont suivies de /i/ également.

Les études utilisant l'équation de locus comme corrélat de la coarticulation anticipatoire (par exemple Noiray et al., 2013) indiquent que la coarticulation anticipatoire des alvéolaires est la plus

tardivement stabilisée chez l'enfant, en raison d'une acquisition du contrôle articulatoire de ces consonnes plus complexe et plus tardive chez l'enfant. Ces études utilisent la transition du deuxième formant (F2) pour étudier ces phénomènes, mais les résultats de notre étude sur le VOT peuvent nous laisser envisager que la coarticulation se manifeste également par des différences de VOT selon le contexte vocalique. Cette étude du voisement des occlusives sera complétée par une étude spécifique de la coarticulation de séquences occlusive-voyelle au moyen de l'équation de locus (voir ci-après, paragraphe §4, ce chapitre). Outre l'intérêt de l'utilisation du VOT pour comprendre le processus d'acquisition du voisement par les enfants normo-entendants, cette étude a mis en évidence l'utilisation du VOT comme corrélat de la coarticulation anticipatoire.

2.8.2. Production d'occlusives chez les enfants CI

Notre étude met en évidence des productions d'occlusives et du contraste de voisement relativement similaires chez les deux groupes d'enfants NH et CI, à la différence des études acoustiques de Horga & Liker (2006) et Uchanski & Geers (2003) qui se concentraient sur les seules alvéolaires /t/ et /d/. La seule occlusive pour laquelle nous observons une différence significative est la vélaire non-voisée /k/ : les enfants CI ont tendance à produire cette seule occlusive avec un VOT plus court que les enfants NH. Pour les autres consonnes, nous n'observons aucune différence significative entre les deux groupes d'enfants.

A la différence des enfants NH chez lesquels aucune influence de la tâche (répétition vs dénomination) n'était observée, les enfants CI ont une production qui varie en fonction de la tâche proposée, ce qui serait dû principalement aux vélaires qui sont les consonnes dont la réalisation du contraste de voisement est la plus variable. En effet, la tendance généralement observée sur toutes les consonnes pour ce groupe d'enfants indique que le VOT est plus court lorsque ceux-ci ne disposent pas du modèle audio adulte (ce qui laisse penser à une hyperarticulation lorsque le modèle est disponible ou à une hypoarticulation lorsqu'il ne l'est pas), mais les vélaires sont le seul lieu d'articulation pour lequel la différence entre consonnes voisées et non-voisées varie en fonction de la tâche, avec une différence entre VOT des voisées et non-voisées plus faible lorsque le modèle audio adulte n'est pas disponible.

En revanche, nos résultats indiquent que la complexité du mot affecte de façon identique les enfants CI et les enfants NH, puisqu'aucun effet d'interaction entre groupe et degré de complexité du mot n'était observé : les consonnes voisées sont les consonnes les plus affectées par cet effet, puisque les deux groupes produisent /b/, /d/ et /g/ avec un prévoisement plus important lorsque les mots sont simples.

Par ailleurs, ni l'étude de Horga & Liker (2006), ni celle de Uchanski & Geers (2003) n'étudiaient l'effet de l'âge à l'implantation ou de la durée d'utilisation de l'implant (ou âge auditif). Peut-être en raison du nombre trop limité d'enfants enregistrés lors de notre étude, nos résultats ne permettent pas de mettre en évidence de façon robuste l'effet de l'un de ces deux facteurs : 1) l'âge à l'implantation affecte seulement le contraste entre bilabiales voisées et non-voisées, qui est plus marqué lorsque la consonne est suivie de /u/ mais moins marqué lorsque celle-ci est suivie de /i/ chez les enfants ayant reçu l'implant le plus tardivement et 2) l'évolution des VOT de /p/ et /t/ avec l'âge auditif est différente selon le contexte vocalique (/i/ ou /u/), la différence entre les deux contextes vocaliques observée initialement ayant tendance à se réduire avec une utilisation plus longue de l'implant. Pour ces deux lieux d'articulation, il existerait donc un lien

entre degré de coarticulation et durée de l'expérience de l'oral chez les enfants CI alors que chez les enfants NH, l'effet de coarticulation est présent mais n'évolue pas avec l'âge chronologique.

Il semble donc que la coordination des gestes oro-laryngés nécessaire pour produire le contraste de voisement et le contraste entre les trois lieux d'articulation (bilabiales, alvéolaires et vélaires) est maitrisée de façon similaire par les enfants CI et les enfants NH. La seule différence entre les productions des deux groupes concerne les vélaires, produites avec plus de variation et plus de sensibilité à la tâche proposée que les bilabiales et alvéolaires par les enfants CI.

Ainsi, nos résultats sont en accord avec les études de Warner-Czyz & Davis (2008), Kim & Chin (2008), Chin & Pisoni (2000) en anglais et de Gaul-Bouchard (2007) en français québécois, qui mettaient en évidence, à partir de jugements perceptifs, une trajectoire d'acquisition des occlusives par les enfants CI similaire à celle des enfants NH, mais plus tardive : dans ces études, les consonnes labiales et coronales sont acquises avant les vélaires par les enfants CI, et les voisées avant les non-voisées, et une durée d'utilisation de l'implant plus longue permet aux enfants CI d'acquérir une production des occlusives approchant celle des enfants NH.

Le VOT est une caractéristique temporelle de la parole, et notre étude a mis en évidence le fait que cette caractéristique phonétique est assez peu affectée dans notre groupe d'enfants CI. Ce résultat est assez cohérent avec le fait que la dégradation du son due à l'implant cochléaire affecte essentiellement la structure spectrale du signal, en préservant sa structure temporelle. Il reste que, comme ce que les résultats de notre étude sur les voyelles orales indiquaient, la réalisation des occlusives est plus difficile pour les enfants CI, lorsque le lieu d'articulation ne peut pas être identifié visuellement (pour les vélaires par exemple).

3. Fricatives

3.1. Données analysées

Dans un premier temps, nous excluons des analyses les productions de mots non-exploitables : dans certains cas, les mots n'ont pas été produits ou ont été substitués par d'autres mots par les enfants. Nous excluons aussi huit fricatives, pour lesquelles nous avons des valeurs de skewness et de kurtosis aberrantes. Nos analyses sont donc réalisées à partir de 1096 fricatives exploitables. 99,2% des mots produits par les enfants NH et 92,7% des mots produits par les enfants CI ont pu être inclus dans les analyses ; un test t (test de Welch) indique une différence significative entre les deux groupes : $t(12,674)=2,99$, $p<0.05$. Des productions de mots exploitables, nous excluons ensuite toutes les productions de fricatives non-exploitables : les fricatives sont substituées par des occlusives ou d'autres consonnes, ou ne sont pas produites par les enfants. Parmi les mots exploitables, 96,7% des consonnes produites par les enfants NH sont des fricatives exploitables et 97,2% des consonnes produites par les enfants CI sont des fricatives exploitables ; un test t (test de Welch) indique que la différence entre les productions des deux groupes d'enfants n'est pas significative : $t(30,484)=-0,49$, $p>0,05$.

Le compte détaillé des productions de fricatives est donné dans la table 5.60 ci-dessous qui présente les valeurs moyennes par groupe, consonne et contexte vocalique, pour chaque moment spectral étudié dans le cadre de notre expérience de production. Dans la suite de cette thèse, nous avons choisi de désigner le coefficient de dissymétrie (ou troisième moment spectral) par le terme de « skewness » et le coefficient d'aplatissement (ou quatrième moment spectral) par le terme de « kurtosis ».

Groupe	Consonne	Contexte vocalique	Nombre total de fricatives	Nombre de fricatives par enfant (max=6)	Centre de gravité moyen (en Hz)	Ecart-type moyen (en Hz)	Skewness moyen (coefficient)	Kurtosis moyen (coefficient)
CI	f	i	64	4,92	3825 (1843)	2617 (872)	1,38 (1,32)	5,99 (7,64)
CI	f	u	71	5,46	3296 (1625)	2754 (802)	1,39 (1,66)	5,99 (11,87)
CI	s	i	75	5,77	5411 (1233)	2541 (782)	0,36 (0,89)	2,66 (4,76)
CI	s	u	68	5,23	4663 (1491)	2522 (497)	0,63 (1,11)	2,60 (6,24)
CI	ʃ	i	63	4,85	4484 (696)	1900 (560)	1,11 (1,07)	5,66 (7,93)
CI	ʃ	u	65	5	3998 (826)	2135 (475)	0,87 (0,99)	2,88 (4,43)
NH	f	i	115	5,75	5092 (1766)	3075 (1092)	1,38 (0,99)	5,23 (8,59)
NH	f	u	116	5,8	5186 (1988)	3506 (880)	0,79 (1,05)	1,75 (3,54)
NH	s	i	119	5,95	7378 (1611)	2498 (710)	0,68 (0,96)	2,99 (4,14)
NH	s	u	113	5,65	5996 (1613)	2359 (774)	1,29 (1,49)	5,95 (10,07)
NH	ʃ	i	108	5,4	4700 (625)	1589 (418)	1,84 (0,90)	8,57 (7,13)
NH	ʃ	u	119	5,95	4426 (1143)	2040 (557)	1,00 (1,07)	4,32 (7,95)

Table 5.60 : valeurs moyennes (avec écart-type) par groupe, par consonne et par contexte vocalique pour chaque moment spectral étudié (centre de gravité, écart-type, skewness, kurtosis)

Les quatre mesures (ou quatre moments spectraux) proposées dans la table 5.60 ci-dessus sont couramment utilisées dans la littérature pour décrire acoustiquement les fricatives (voir chapitre 4, paragraphe §2.2.4). Cependant, comme la distinction /s/-/ʃ/ est décrite dans la littérature comme difficile à réaliser pour les enfants porteurs d'implant cochléaire (Bharadwaj et al., 2006, Mildner & Liker, 2008), en raison d'une perception partielle des sons dans les hautes fréquences, nous avons décidé d'étudier plus spécifiquement cette distinction entre fricatives alvéolaires et postalvéolaires. A cet effet, nous avons calculé une distance moyenne entre les centres de gravité

de /s/ et /ʃ/, par sujet, contexte vocalique (/i/ ou /u/), par tâche (répétition ou dénomination) et par degré de complexité (mots simples ou complexes). Les distances moyennes entre centres de gravité de /s/ et /ʃ/ sont données dans la table 5.61 ci-dessous.

Groupe	Contexte vocalique	Nombre total de données	Nombre de données par enfant (max=3)	Distance s-ʃ moyenne (en Hz)
CI	i	33	2,54	1284 (1414)
CI	u	39	3	558 (1340)
NH	i	58	2,9	2406 (1219)
NH	u	60	3	1523 (1796)

Table 5.61 : distance moyenne (en Hz) entre les centres de gravité de /s/ et /ʃ/, par groupe et par contexte vocalique

Pour analyser ces fricatives, nous avons fait le choix d'utiliser des modèles linéaires mixtes qui prennent en compte à la fois des effets fixes et un effet aléatoire du sujet. Les variables à expliquer (ou variables dépendantes) ainsi que les variables explicatives (ou variables indépendantes) produisant les effets fixes sont détaillées dans la table 5.62 ci-dessous.

Variables dépendantes	Centre de gravité (1er moment spectral), mesuré en Hz
	Distance des centres de gravité moyens de /s/ et /S/, en Hz
	Ecart-type (2ème moment spectral), mesuré en Hz
	Skewness (3ème moment spectral), coefficient
	Kurtosis (4ème moment spectral), coefficient
Variables indépendantes	Groupe : CI ou NH
	C : Lieu d'articulation : labiodental /f/, alvéolaire /s/, post-alvéolaire /ʃ/
	V : Contexte vocalique : _i ou _u
	Age chronologique des enfants (NH)
	Age à l'implantation (enfants CI)
	Age auditif (qui correspond à la durée d'utilisation de l'implant pour les enfants CI)
	Tâche : tâche de répétition (R) ou tâche de dénomination (D)
	Complexité lexicale : mots simples (0) et mots complexes (1)
Variable ayant un effet aléatoire	Sujet

Table 5.62 : variables dépendantes, variables indépendantes et variable ayant un effet aléatoire, pris en compte dans l'analyse des fricatives

3.2. Effets du groupe et du lieu d'articulation

Dans les tables figurant dans ce chapitre, nous désignons par « C » le lieu d'articulation (ou consonne) et « V » le contexte vocalique.

3.2.1. Centre de gravité

La figure 5.43 ci-dessous présente l'effet du lieu d'articulation et du groupe sur le centre de gravité. Nous avons choisi de présenter ces deux effets conjointement, puisqu'ils peuvent donner lieu également à un effet d'interaction entre groupe et lieu d'articulation.

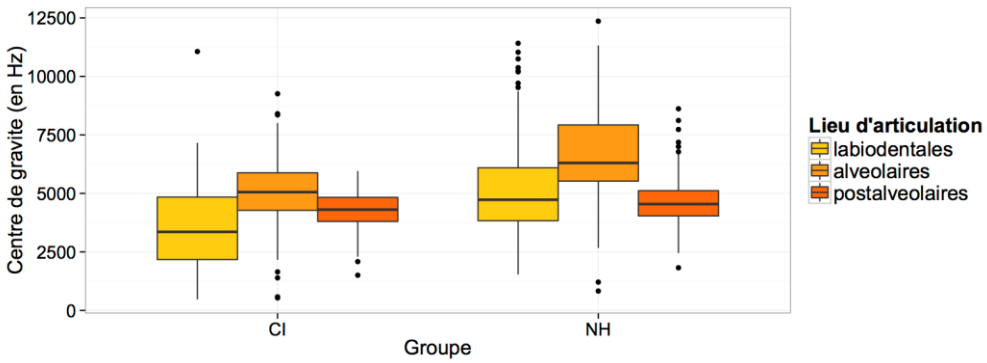


Figure 5.43 : centre de gravité (en Hz) par groupe et lieu d'articulation

Les résultats présentés dans la table 5.63 ci-dessous montrent un effet global du groupe sur le centre de gravité : les enfants CI ont en effet tendance à produire les fricatives avec un centre de gravité plus bas que les enfants NH (F(1,31)=35,56, p<.001).

	numDF	denDF	F-value	p-value
(Intercept)	1	1059	2530.1207	<.0001
C	2	1059	156.8116	<.0001 ***
groupe	1	31	35.5609	<.0001 ***
C:groupe	2	1059	23.2012	<.0001 ***

Table 5.63 : tableau de résultats au test statistique sur toutes les données (effets de groupe, lieu d'articulation sur le centre de gravité)

Ces résultats indiquent par ailleurs un effet du lieu d'articulation sur le centre de gravité (F(2,1059)=156,81, p<.001), différent selon le groupe (pour l'interaction entre groupe et lieu d'articulation : F(2,1059)=23,20, p<.001) : pour les enfants NH /s/ a le centre de gravité le plus haut, suivi de /f/ puis /ʃ/ alors que pour les enfants CI /s/ a le centre de gravité le plus haut, suivi de /ʃ/, puis de /f/ : un test post-hoc de comparaisons multiples (table 5.64, ci-dessous) indique que les différences pour chaque groupe entre paires de lieux d'articulations sont toutes significatives, ce qui correspond à une distinction entre les catégories de fricatives.

Groupe	Centre de gravité		
	Paire de consonnes testées		
	/f/-/s/	/f/-/ʃ/	/s/-/ʃ/
NH	z=-12,22, p<.001 ***	z=4,47, p<.001 ***	z=16,62, p<.001 ***
CI	z=-9,00, p<.001 ***	z=-4,07, p<.001 ***	z=4,75, p<.001 ***

Table 5.64 : résumé du test post-hoc sur l'effet d'interaction entre lieu d'articulation et groupe sur le premier moment spectral (centre de gravité)

Par ailleurs, un test post-hoc supplémentaire indique que les labiodentales et alvéolaires sont produites avec un centre de gravité plus bas chez les enfants CI et les postalvéolaires avec des centres de gravité similaires chez les deux groupes. Les résultats du test post-hoc de comparaisons multiples (CI vs NH) sont présentés dans la table 5.65 ci-dessous.

Consonne	Centre de gravité
	CI vs NH
/f/	z=-6,72, p<.001 ***
/s/	z=-7,08, p<.001, ***
/ʃ/	z=-1,42, p>.05 (n.s.)

Table 5.65 : résumé du test post-hoc sur l'effet d'interaction entre groupe et lieu d'articulation sur le premier moment spectral (centre de gravité)

3.2.2.Distance entre centres de gravité de /s/ et /ʃ/

La figure 5.44 ci-dessus présente la distance moyenne (en Hz) entre les centres de gravité de /s/ et /ʃ/ pour chaque groupe et la table 5.66 présente les valeurs correspondantes, par groupe et par contexte vocalique.

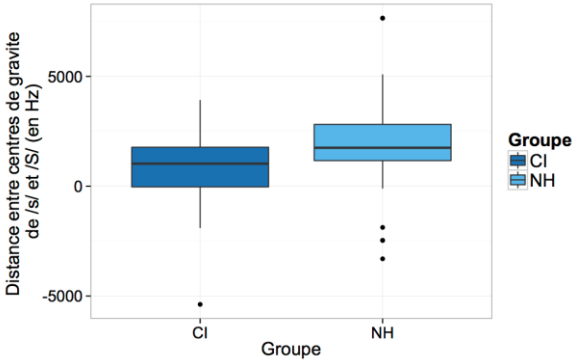


Figure 5.44 : distance entre les centres de gravité de /s/ et /ʃ/ (en Hz) par groupe

Les résultats statistiques de la table 5.66 ci-dessous indiquent un effet du groupe (F(1,31)=9,56, p<.01) : la différence en Hz entre /s/ et /ʃ/ est plus faible chez les enfants CI qui ont en effet tendance à produire des fricatives /s/ et /ʃ/ de façon moins distincte que les enfants NH.

	numDF	denDF	F-value	p-value
(Intercept)	1	157	89.80827	<.0001
groupe	1	31	10.28098	0.0031 ***

Table 5.66 : tableau de résultats au test statistique sur toutes les données (effets du groupe sur la distance entre centres de gravité de /s/ et /ʃ/)

Ces résultats sur le premier moment spectral des fricatives aux trois lieux d'articulation ont mis en évidence un effet du groupe sur le centre de gravité des labiodentales et des alvéolaires, ainsi que sur la distinction entre alvéolaires et postalvéolaires.

3.2.3.Ecart-type

Nous avons ensuite choisi, de la même façon, d'analyser le deuxième moment spectral ou écart-type, qui correspond à l'étendue du bruit de la fricative sur la bande de fréquence. La figure 5.45 ci-dessous présente l'effet du lieu d'articulation et du groupe sur l'écart-type. Comme pour le centre de gravité, nous avons choisi de présenter ces deux effets conjointement, puisqu'ils peuvent également donner lieu à un effet d'interaction entre groupe et lieu d'articulation.

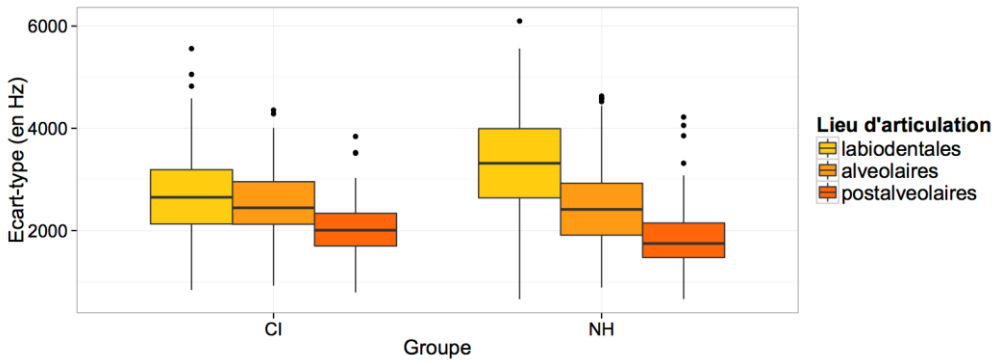


Figure 5.45 : Ecart-type (en Hz) par groupe et lieu d'articulation

Les résultats présentés dans la table 5.67 ci-dessous ne montrent pas d'effet global du groupe sur le deuxième moment spectral, ou écart-type : les enfants CI ont en effet tendance à produire les fricatives avec des valeurs d'écart-type similaire à celui des enfants NH (F(1,31)=0,88, p>.05).

	numDF	denDF	F-value	p-value
(Intercept)	1	1059	2206.8218	<.0001
C	2	1059	252.6070	<.0001 ***
groupe	1	31	0.8825	0.3548
C:groupe	2	1059	33.4594	<.0001 ***

Table 5.67 : tableau de résultats au test statistique sur toutes les données (effets de groupe, lieu d'articulation, contexte vocalique sur l'écart-type)

En revanche, le test statistique indique un effet du lieu d'articulation sur l'écart-type (F(2,1059)=252,61, p<.001) : la labiodentale /f/ est produite avec un écart-type supérieur à celui de l'alvéolaire /s/ puis de la postalvéolaire /ʃ/. Cet ordre est le même pour les deux groupes d'enfants, NH et CI. Nous observons également une plus grande distinction des valeurs d'écarts-types des trois lieux d'articulation chez les enfants NH, par rapport aux enfants CI, qui se traduit au niveau statistique par un effet significatif de l'interaction entre groupe et lieu d'articulation (F(2,1059)=33,46, p<.001) : un test post-hoc (table 5.68, ci-dessous) permet de mettre en évidence une distinction significative de l'écart-type entre toutes les paires de consonnes sauf pour la paire labiodentales-alvéolaires chez les enfants CI.

Groupe	Ecart-type		
	Paire de consonnes testées		
	/f/-/s/	/f/-/ʃ/	/s/-/ʃ/
NH	z=13,50, p<.001 ***	z=23,25, p<.001 ***	z=9,84, p<.001 ***
CI	z=1,93, p>.05 (n.s.)	z=7,73, p<.001 ***	z=5,94, p<.001 ***

Table 5.68 : résumé du test post-hoc sur l'effet d'interaction entre lieu d'articulation et groupe sur le deuxième moment spectral (écart-type)

Cet effet de l'interaction entre groupe et lieu d'articulation sur l'écart-type se traduit également par des valeurs d'écart-type plus élevées pour les labiodentales chez les enfants NH par rapport aux enfants CI : un test post-hoc (table 5.69, ci-dessous) indique que la différence entre les deux groupes n'est pas significative pour les alvéolaires et pour les postalvéolaires, alors que celle-ci est significative pour les labiodentales.

Consonne	Ecart-type
	CI vs NH
/f/	z=-4,91, p<.001 ***
/s/	z=0,70, p>.05 (n.s.)
/ʃ/	z=1,76, p>.05 (n.s.)

Table 5.69 : résumé du test post-hoc sur l'effet d'interaction entre groupe et lieu d'articulation sur le deuxième moment spectral (écart-type)

3.2.4.Skewness (coefficient de dissymétrie)

Nous avons également choisi de proposer une analyse du troisième moment spectral ou skewness (coefficient de dissymétrie), qui correspond à la répartition de la fricative et à son orientation (vers les hautes ou les basses fréquences) à partir du centre de gravité, sur la bande de fréquence. La figure 5.46 ci-dessous présente l'effet du lieu d'articulation et du groupe sur le skewness. Comme pour nos analyses précédentes, nous avons choisi de présenter ces deux effets conjointement, puisqu'ils peuvent donner lieu également à un effet d'interaction entre groupe et lieu d'articulation.

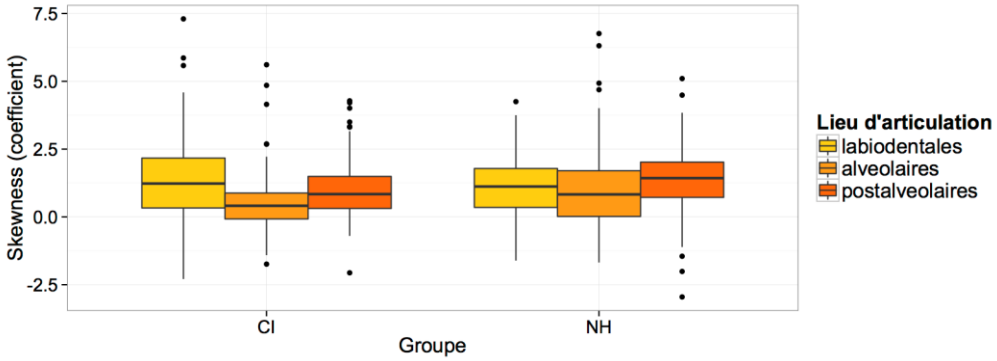


Figure 5.46 : skewness (coefficient) par groupe et lieu d'articulation

Les résultats présentés dans la table 5.70 ci-dessous ne montrent pas d'effet global du groupe sur le troisième moment spectral, ou skewness : les enfants CI ont en effet tendance à produire les fricatives des valeurs de skewness similaires à celles des enfants NH (F(1,31)=1,40, p>.05).

	numDF	denDF	F-value	p-value
(Intercept)	1	1059	185.53892	<.0001
C	2	1059	18.66926	<.0001 ***
groupe	1	31	1.39888	0.2459
C:groupe	2	1059	13.76664	<.0001 ***

Table 5.70 : tableau de résultats au test statistique sur toutes les données (effets de groupe, lieu d'articulation sur le skewness)

Cependant, le test statistique indique un effet de l'interaction entre lieu d'articulation et groupe sur le troisième moment spectral (F(1,1059)=13,77, p<.001), et un test post-hoc, résumé dans la table 5.71 ci-dessous, nous permet de préciser cette interaction : les enfants NH produisent les postalvéolaires avec un troisième moment spectral plus élevé que les labiodentales, lui-même similaire à celui des alvéolaires, alors que chez les enfants CI, les labiodentales ont un troisième moment spectral plus élevé que celui des postalvéolaires, lui-même plus élevé que celui des alvéolaires.

Groupe	Skewness		
	Paire de consonnes testées		
	/f/-/s/	/f/-/ʃ/	/s/-/ʃ/
NH	z=-1,43, p>.05 (n.s.)	z=-3,03, p<.05 *	z=-4,00, p<.001 ***
CI	z=6,87, p<.001 ***	z=2,95, p<.05 *	z=-3,79, p<.001 ***

Table 5.71 : résumé du test post-hoc sur l'effet d'interaction entre lieu d'articulation et groupe sur le troisième moment spectral (skewness)

Cet effet de l'interaction entre groupe et lieu d'articulation sur le troisième moment spectral se traduit également par des valeurs plus élevées pour les alvéolaires chez les enfants NH par rapport aux enfants CI : un test post-hoc, résumé dans la table 5.72 ci-dessous, indique que la différence entre les deux groupes n'est pas significative pour les labiodentales et pour les postalvéolaires, alors qu'elle est significative pour les alvéolaires.

Consonne	Skewness
	CI vs NH
/f/	z=1,67 (n.s.)
/s/	z=-2,63, p<.05
/ʃ/	z=-2,07 (n.s.)

Table 5.72 : résumé du test post-hoc sur l'effet d'interaction entre groupe et lieu d'articulation sur le troisième moment spectral (skewness)

3.2.5. Kurtosis (coefficient d'aplatissement)

La figure 5.47 ci-dessous présente l'effet du lieu d'articulation et du groupe sur le quatrième moment spectral ou kurtosis, qui correspond à une concentration (kurtosis élevé) ou à un étalement (kurtosis faible) du bruit de la fricative autour de son centre de gravité. Comme pour les trois premiers moments spectraux, nous avons choisi de présenter ces deux effets conjointement, puisqu'ils peuvent donner lieu également à un effet d'interaction entre groupe et lieu d'articulation.

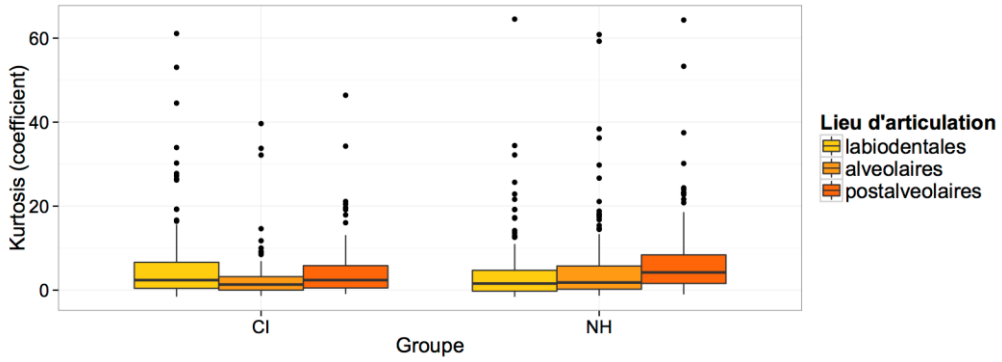


Figure 5.47 : kurtosis (coefficient) par groupe et lieu d'articulation

Les résultats présentés dans la table 5.73 ci-dessous ne montrent pas d'effet global du groupe sur le quatrième moment spectral, ou kurtosis : les enfants CI ont en effet tendance à produire les fricatives avec des valeurs de kurtosis globalement similaires à celles des enfants NH ($F(1,31)=0.25$, $p>.05$).

	numDF	denDF	F-value	p-value
(Intercept)	1	1059	154.62228	<.0001
C	2	1059	5.46663	0.0043 ***
groupe	1	31	0.24929	0.6211
C:groupe	2	1059	10.67029	<.0001 ***

Table 5.73 : tableau de résultats au test statistique sur toutes les données (effets de groupe, lieu d'articulation sur le kurtosis)

Le test statistique indique en outre un effet du lieu d'articulation seul (F(2,1059)=5,67, p<.01) ou en interaction avec le groupe (F(2,1059)=10,67, p<.001) : le test post-hoc réalisé, résumé dans la table 5.74 ci-dessous, indique une différence significative pour les paires /f/-/ʃ/ et /f/-/s/ chez les enfants NH, et pour la paire /f/-/s/ uniquement, chez les enfants CI.

Groupe	Kurtosis		
	Paire de consonnes testées		
	/f/-/s/	/f/-/ʃ/	/s/-/ʃ/
NH	z=-1,43, p>.05 (n.s.)	z=-4,14, p<.001 ***	z=-2,72, p<.05 *
CI	z=3,84, p<.001 ***	z=2,11, p>.05 (n.s.)	z=-1,75, p>.05 (n.s.)

Table 5.74 : résumé du test post-hoc sur l'effet d'interaction entre lieu d'articulation et groupe sur le kurtosis

Les enfants NH ont donc des valeurs de kurtosis plus élevées pour /ʃ/ que pour /f/ et /s/ dont les valeurs de kurtosis sont similaires, alors que les enfants CI ont des valeurs de kurtosis plus élevées pour /f/ mais proches de celles de /ʃ/ et différentes de celles de /s/ : les fricatives /ʃ/ et /s/ sont donc produites par les enfants NH avec une amplitude plus élevée autour du centre de gravité, par rapport à /s/, alors que chez les enfants CI, les fricatives /f/ et /ʃ/ sont celles dont l'amplitude est la plus élevée autour du centre de gravité.

Un test post-hoc inter-groupes (voir table 5.75 ci-dessous) indique par ailleurs que la différence entre les valeurs des kurtosis des deux groupes d'enfants est significative pour /f/ seulement : pour /f/, le kurtosis est plus élevé chez les enfants CI, qui produisent donc cette fricative avec une amplitude plus grande autour du centre de gravité.

Consonne	Kurtosis
	CI vs NH
/f/	z=2,63, p<.05 *
/s/	z=-1,76, p>.05 (n.s.)
/ʃ/	z=-2,02, p>.05 (n.s.)

Table 5.75 : résumé du test post-hoc sur l'effet d'interaction entre groupe et lieu d'articulation sur le kurtosis

3.2.6. Résumé : effets du groupe et du lieu d'articulation

Les effets du groupe, du lieu d'articulation et de l'interaction entre groupe et lieu d'articulation sur les quatre moments spectraux sont résumés dans la table 5.76 ci-dessous.

Effets	Moments spectraux				
	Centre de gravité	Distance entre centres de gravité de /s/ et /ʃ/	Ecart-type	Skewness	Kurtosis
Groupe	CI<NH	CI<NH	n.s.	n.s.	n.s.
Lieu d'articulation	s>ʃ>f	NA	f>s>ʃ	f≃ʃ>s	ʃ>f≃s
Groupe*Lieu d'articulation	NH : s>f>ʃ	NA	NH : f≃s>ʃ	NH : ʃ>f>s	NH : f≃s>ʃ
	CI : s>ʃ>f	NA	CI : f>s>ʃ	CI : f>ʃ>s	CI : f≃ʃ et f>s

Table 5.76 : résumé des effets des variables groupe, lieu d'articulation et de l'interaction entre groupe et lieu d'articulation sur chaque moment spectral (le symbole ≃ indique une absence de différence significative)

Les analyses réalisées sur les quatre moments spectraux et présentées dans ce paragraphe ont permis de mettre en évidence des réalisations différentes selon le groupe (NH et CI) : elles indiquent que les enfants CI produisent des fricatives au centre de gravité plus bas que les enfants NH et marquées par une distinction entre alvéolaires et postalvéolaires moins forte que les enfants NH, avec un deuxième moment ou écart-type similaire (sauf pour les labiodentales, pour lesquelles l'écart-type est plus bas chez les enfants CI), avec un troisième moment ou skewness similaire (sauf pour les alvéolaires, pour lesquelles ce troisième moment est plus élevé chez les enfants CI), et un quatrième moment ou kurtosis similaire (sauf pour les labiodentales, pour lesquelles ce quatrième moment est plus élevé chez les enfants CI).

Ces analyses mettent également en évidence une différence de réalisation selon le lieu d'articulation : les enfants NH produisent des fricatives dont le centre de gravité est plus haut pour les alvéolaires, puis pour les labiodentales, suivies des postalvéolaires, alors que les enfants CI produisent des fricatives dont le centre de gravité est plus haut pour les alvéolaires, puis pour les postalvéolaires, suivies des labiodentales ; les enfants NH ont des productions de labiodentales plus diffuses que les alvéolaires, et que les postalvéolaires (plus compactes), alors que les enfants CI produisent des labiodentales et alvéolaires similaires et plus diffuses que les postalvéolaires (plus compactes) ; les enfants NH produisent des fricatives postalvéolaires avec un coefficient de dissymétrie (ou skewness) plus élevé que les labiodentales, puis que les alvéolaires, alors que les enfants CI produisent des fricatives labiodentales avec un coefficient de dissymétrie plus élevé que les post-alvéolaires puis les alvéolaires ; enfin, les enfants NH produisent des fricatives post-alvéolaires plus intenses autour du centre de gravité que les labiodentales et les alvéolaires, alors que les enfants CI produisent des fricatives labiodentales d'intensité similaire aux post-alvéolaires mais plus intenses autour du centre de gravité que les alvéolaires.

Li et al. (2009) proposent une correspondance entre caractéristiques spectrales et articulatoires des fricatives alvéolaires et post-alvéolaires : le premier moment spectral (centre de gravité) qui correspond au centre de la distribution spectrale de la fricative, est corrélé négativement avec la longueur de la cavité de résonance, en avant de la constriction, ce qui explique par exemple que les fricatives alvéolaires ont un centre de gravité plus élevé que les postalvéolaires ; le deuxième moment spectral (écart-type) qui correspond à l'étendue de la distribution spectrale de la fricative, est lié à la configuration de la langue, et permet de différencier fricatives apicales et laminales ; le troisième moment spectral (skewness ou coefficient de dissymétrie) qui correspond à l'asymétrie de la distribution spectrale, est également corrélé négativement avec la longueur de la cavité de résonance, devant la langue ; le quatrième moment spectral (kurtosis ou coefficient

d'aplatissement) qui correspond à l'amplitude du pic spectral de la fricative, est également lié à la configuration de la langue et permet de distinguer fricatives apicales et laminales.

Nous pouvons donc déduire des différences inter-groupes mises en évidence par nos analyses que les alvéolaires sont produites par les enfants CI avec une cavité de résonance plus grande que les enfants NH, puisque les centres de gravité et skewness de leurs alvéolaires sont plus bas. En revanche, pour les post-alvéolaires, nous n'observons pas de différences pour ces deux moments spectraux entre les deux groupes d'enfants. Cette différence acoustique entre la réalisation des deux lieux d'articulation (alvéolaires et post-alvéolaires) résulte donc d'un contrôle moteur différent chez les deux groupes d'enfants et pas d'une différence de taille globale du conduit vocal : si la différence observée était liée à la physiologie des enfants (en particulier à la taille de leur conduit vocal, qui évolue avec l'âge) les post-alvéolaires ne seraient pas produites par les enfants CI et NH de façon identique et nous devrions observer une différence entre les deux groupes, à l'image de ce que nous observons pour les alvéolaires. Il nous semble donc que les post-alvéolaires sont produites avec le même contrôle articulatoire par les deux groupes d'enfants, mais que les alvéolaires sont plus postérieures chez les enfants CI, ce qui témoignerait d'un contrôle articulatoire différent chez les enfants CI.

Pour les labiodentales, nous observons une différence significative entre les enfants CI et NH pour le deuxième moment spectral (écart-type). D'après Johnson (2012) le centre de gravité est l'indice acoustique le plus saillant qui caractérise fricatives alvéolaires et post-alvéolaires, mais l'écart-type est celui qui est le plus pertinent pour décrire la production des labiodentales, puisqu'il correspond à la diffusion du flux d'air et à sa perturbation par l'obstacle que constitue les dents et les lèvres. Ainsi, pour ce seul lieu d'articulation des labiodentales, nous observons une différence d'écart-type, qui est plus bas chez les enfants CI que chez les enfants NH, qui pourrait correspondre à des fricatives labio-dentales plus compactes et donc une constriction plus forte chez les enfants CI.

3.3. Effet du contexte vocalique

Pour analyser l'effet du contexte vocalique, nous avons ajouté la variable « contexte vocalique » (désignée « V » dans les tables de ce paragraphe) aux modèles utilisés pour les tests statistiques réalisés au paragraphe §3.2.1 qui ne prenaient en compte que les effets du lieu d'articulation et du groupe sur la réalisation des quatre moments spectraux des fricatives de notre étude. Comme pour les analyses précédentes, nous étudions l'effet du contexte vocalique seul mais également en interaction avec le groupe d'une part et le lieu d'articulation d'autre part.

3.3.1. Centre de gravité

La figure 5.48 ci-dessous propose une représentation du centre de gravité, par lieu d'articulation, en fonction du groupe et du contexte vocalique.

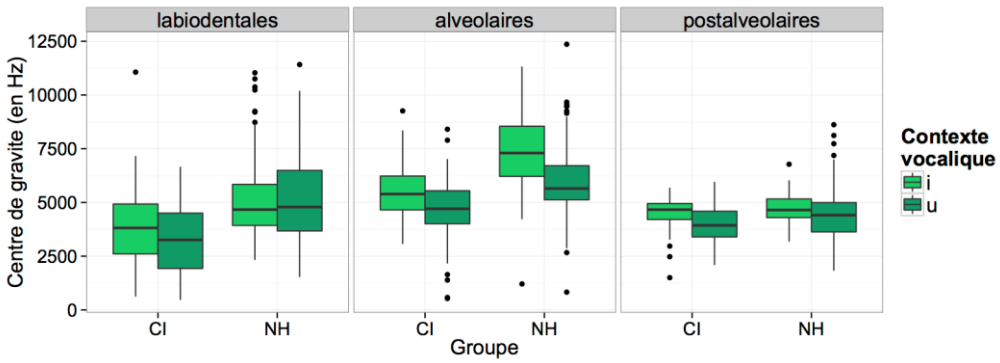


Figure 5.48 : centre de gravité (en Hz) par lieu d'articulation, par groupe et par contexte vocalique

La table 5.77 ci-dessous présente les résultats du test statistique réalisé sur toutes les données, qui nous permet d'analyser l'effet du contexte vocalique sur le centre de gravité des fricatives.

	numDF	denDF	F-value	p-value
(Intercept)	1	1053	2461.8729	<.0001
C	2	1053	168.6525	<.0001 ***
groupe	1	31	34.6556	<.0001 ***
V	1	1053	45.3179	<.0001 ***
C:groupe	2	1053	24.4951	<.0001 ***
C:V	2	1053	14.6591	<.0001 ***
groupe:V	1	1053	0.1098	0.7404
C:groupe:V	2	1053	5.1849	0.0057 **

Table 5.77 : tableau de résultats au test statistique sur toutes les données (effets de groupe, lieu d'articulation, contexte vocalique sur le centre de gravité)

Le test statistique réalisé sur toutes les données montre un effet global du contexte vocalique sur le centre de gravité (F(1,1053)=45,32, p<.001) mais également un effet de l'interaction du contexte vocalique avec le lieu d'articulation (F(2,1053)=14,66, p<.001) ainsi que de l'interaction du contexte vocalique avec le lieu d'articulation et le groupe (F(2,1053)=5,28, p<.01). L'observation graphique indique que le centre de gravité est plus haut lorsque la fricative est suivie de /i/ que lorsqu'elle est suivie de /u/. Un test post-hoc (table 5.78, ci-dessous) testant l'effet du contexte vocalique sur le centre de gravité en fonction du lieu d'articulation indique que seules les alvéolaires sont influencées par le contexte vocalique.

Consonne	Centre de gravité
	Contexte vocalique : /i/ vs /u/
/f/	z=0,98, p>.05 (n.s.)
/s/	z=8,18, p<.001 ***
/ʃ/	z=2,33, p>.05 (n.s.)

Table 5.78 : résumé du test post-hoc sur l'effet d'interaction entre contexte vocalique et lieu d'articulation sur le centre de gravité

Un test statistique (modèle mixte à effet aléatoire du sujet) sur le seul groupe NH indique un effet significatif du contexte vocalique seul (F(1,662)=24,50, p<.005) et de l'interaction entre contexte vocalique et lieu d'articulation (F(2,662)=18,79, p<.001) alors que la même analyse statistique sur le seul groupe CI indique un effet du contexte vocalique seul (F(1,377)=22,12, p<.001) mais pas de l'interaction entre contexte vocalique et lieu d'articulation (F(1,377)=0,55, p>.05). Un test post-hoc réalisé sur les centres de gravité des enfants NH indique que seules les alvéolaires sont influencées par le contexte vocalique (voir Table 5.79, ci-dessous).

Consonne	Centre de gravité
	Contexte vocalique : /i/ vs /u/
/f/	z=-6,45, p>.05 (n.s.)
/s/	z=7,72, p<.001 ***
/ʃ/	z=1,51, p>.05 (n.s.)

Table 5.79 : résumé du test post-hoc sur l'effet d'interaction entre contexte vocalique et lieu d'articulation sur le centre de gravité, chez les enfants NH

Il ressort donc de ces analyses statistiques un effet du contexte vocalique sur le centre de gravité qui varie selon les groupes et les lieux d'articulation : le centre de gravité des alvéolaires est le seul à être influencé par le contexte vocalique pour les enfants NH, alors que ceux de toutes les fricatives sont influencés par le contexte vocalique pour les enfants CI, ce qui témoigne d'une coarticulation anticipatoire plus importante chez le groupe des enfants CI. Ainsi, les enfants NH ont une production de labiodentales et postalvéolaires stable selon le contexte vocalique, alors que les enfants CI ont une production beaucoup plus variable : /si/ et /ʃi/ d'une part et /su/ et /ʃu/ d'autre part sont produites de façon beaucoup plus proche chez les enfants CI que chez les enfants NH, chez lesquels la production de /ʃ/ dans /ʃi/ et /ʃu/ est stable. Cet effet du contexte vocalique, observé acoustiquement, est également perceptible, puisque chez les enfants CI, la production de /si/ est proche de celle de /ʃi/ et celle de /su/ est proche de celle de /ʃu/.

3.3.2. Distance entre centres de gravité de /s/ et /ʃ/

Comme pour les analyses du centre de gravité en fonction du groupe et du lieu d'articulation, nous avons choisi de nous intéresser à l'effet du contexte vocalique sur la distance entre centres de gravité de /s/ et /ʃ/, seul et en interaction avec le groupe. La figure 5.49 ci-dessous propose une représentation de la distance entre centres de gravité de /s/ et /ʃ/ en fonction du groupe et du contexte vocalique (/i/ vs /u/).

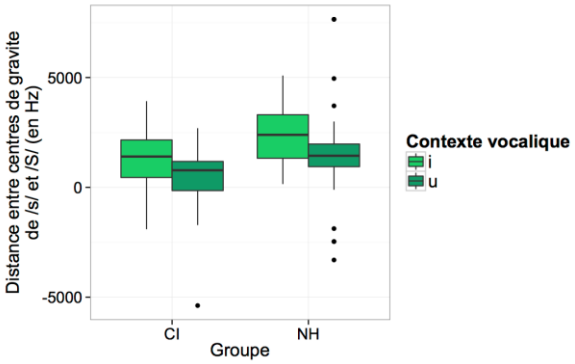


Figure 5.49 : distance moyenne (en Hz) entre les centres de gravité de /s/ et /ʃ/ par groupe et contexte vocalique

Une première observation de cette représentation montre une différence selon le contexte vocalique (la distance entre les centres de gravité est plus basse en contexte /u/) et selon le groupe (la distance entre les centres de gravité est plus basse pour le groupe CI).

La table 5.80 ci-dessous présente les résultats du test statistique étudiant l'effet du contexte vocalique sur la distance entre les fricatives alvéolaires et postalvéolaires.

	numDF	denDF	F-value	p-value
(Intercept)	1	156	89.11399	<.0001
V	1	156	21.57497	<.0001 ***
groupe	1	31	9.56233	0.0042 **

Table 5.80 : tableau de résultats au test statistique sur toutes les données (effets de groupe, contexte vocalique sur la distance entre centres de gravité de /s/ et /ʃ/)

Les tests statistiques que nous avons réalisés indiquent un effet du contexte vocalique sur la différence /s/-/ʃ/ (F(1,156)=21,57, p<.001), mais pas d'effet d'interaction, ce qui indique que la distinction /s/-/ʃ/ subit une influence similaire du contexte vocalique pour les deux groupes.

3.3.3.Ecart-type

La figure 5.50 ci-dessus présente l'écart-type pour chaque lieu d'articulation, en fonction du groupe et du contexte vocalique.

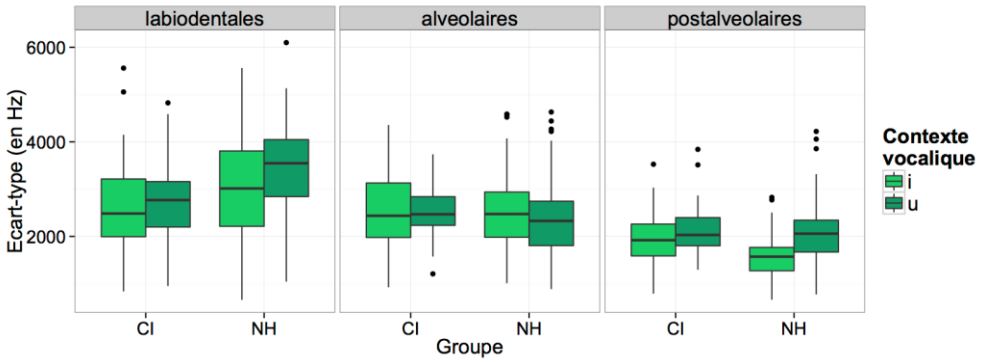


Figure 5.50 : écart-type (en Hz) par lieu d'articulation, par groupe et par contexte vocalique

La table 5.81 présente les résultats du test statistique étudiant l'effet du contexte vocalique sur l'écart-type.

	numDF	denDF	F-value	p-value
(Intercept)	1	1053	2154.6183	<.0001
C	2	1053	265.2844	<.0001 ***
groupe	1	31	0.8615	0.3605
V	1	1053	21.4716	<.0001 ***
C:groupe	2	1053	35.8376	<.0001 ***
C:V	2	1053	14.2991	<.0001 ***
groupe:V	1	1053	1.9420	0.1637
C:groupe:V	2	1053	2.4380	0.0878

Table 5.81 : tableau de résultats au test statistique sur toutes les données (effets de groupe, lieu d'articulation, contexte vocalique sur l'écart-type)

Les résultats indiquent un effet du contexte vocalique seul (F(1,1053)=21,47, p<.001) et en interaction avec le lieu d'articulation (F(2,1053)=14,30, p<.001). Cependant ces résultats n'indiquent pas d'effet du contexte vocalique en interaction avec le groupe (F(1,1053)=1,94, p>.05) ni en interaction avec le groupe et le lieu d'articulation (F(1,1053)=2,44, p>.05). Un test post-hoc (table 5.82 ci-dessous) sur l'interaction entre contexte vocalique et lieu d'articulation indique que un effet du contexte vocalique (/i/ ou /u/) sur l'écart-type des labiodentales et des postalvéolaires mais pas sur celui des alvéolaires : l'écart-type est en effet plus élevé lorsque la fricative est suivie de /u/ que de /i/, et ce pour les deux groupes.

Consonne	Ecart-type
	Contexte vocalique : /i/ vs /u/
/f/	z=-4,34, p<.001 ***
/s/	z=1,71, p>.05 (n.s.)
/ʃ/	z=-5,00, p<.001 ***

Table 5.82 : résumé du test post-hoc sur l'effet d'interaction entre contexte vocalique et lieu d'articulation sur l'écart-type

3.3.4.Skewness

La figure 5.51 ci-dessous propose une représentation du troisième moment spectral (skewness) par lieu d'articulation, en fonction du groupe et du contexte vocalique.

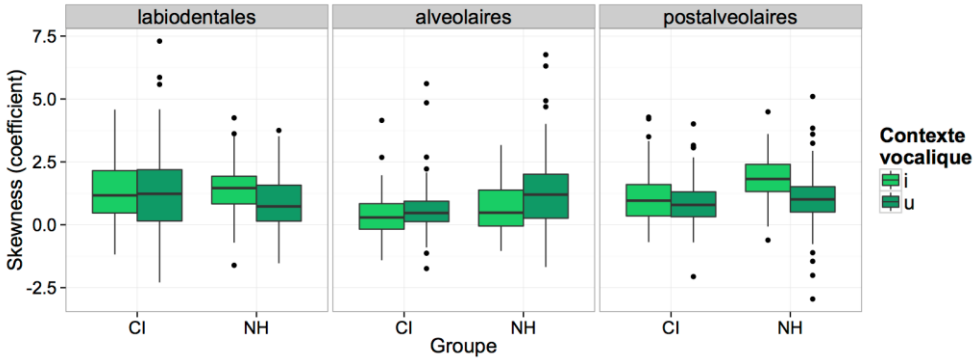


Figure 5.51 : skewness (coefficient) par lieu d'articulation, par groupe et par contexte vocalique

Les résultats du test statistique sont présentés dans la table 5.83 ci-dessous.

	numDF	denDF	F-value	p-value
(Intercept)	1	1053	184.62191	<.0001
C	2	1053	19.91080	<.0001 ***
groupe	1	31	1.38778	0.2477
V	1	1053	5.37229	0.0207 *
C:groupe	2	1053	14.91241	<.0001 ***
C:V	2	1053	27.08067	<.0001 ***
groupe:V	1	1053	4.09055	0.0434 *
C:groupe:V	2	1053	6.38173	0.0018 **

Table 5.83 : tableau de résultats au test statistique sur toutes les données (effets de groupe, lieu d'articulation, contexte vocalique sur le skewness)

Les analyses statistiques indiquent un effet du contexte vocalique sur le troisième moment spectral seul (F(1,1053)=5,37, p<.01), mais également en interaction avec le lieu d'articulation (F(2,1053)=27,08, p<.001), en interaction avec le groupe (F(2,1053)=6,38, p<.05) et en interaction avec le groupe et le lieu d'articulation (F(2,1053)=6,38, p<.01). Un test post-hoc sur l'ensemble des données (table 5.84 ci-dessous), prenant en compte l'interaction entre contexte vocalique et lieu d'articulation montre une différence significative entre skewness en contexte vocalique /i/ et /u/ pour les labiodentales, pour les alvéolaires et pour les postalvéolaires.

Consonne	Skewness
	Contexte vocalique : /i/ vs /u/
/f/	z=2,99, p<.01 **
/s/	z=-4,64, p<.001 ***
/ʃ/	z=4,00, p<.001 ***

Table 5.84 : résumé du test post-hoc sur l'effet d'interaction entre contexte vocalique et lieu d'articulation sur le deuxième moment spectral (skewness)

Un second test post-hoc (Table 5.85 ci-dessous, gauche), réalisé sur les valeurs de skewness des fricatives produites par les enfants NH indique un effet du contexte vocalique sur chaque lieu d'articulation, alors que le même test post-hoc (Table 5.85 ci-dessous, droite) réalisé sur les valeurs de skewness des fricatives produites par les enfants CI n'indique aucun effet du contexte vocalique pour ce groupe. Pour les enfants NH, les labiodentales et postalveolaires sont donc produites de façon plus symétrique, en contexte /i/ par rapport à /u/, alors que les alvéolaires sont moins symétriques en contexte /u/ par rapport au contexte /i/. Chez les enfants CI, les valeurs de skewness sont similaires dans des contextes vocaliques différents.

Consonne	Skewness
	Contexte vocalique : /i/ vs /u/
/f/	z=4,55, p<.001 ***
/s/	z=-4,83, p<.001 ***
/ʃ/	z=6,27, p<.001 ***

Consonne	Skewness
	Contexte vocalique : /i/ vs /u/
/f/	z=-0,74, p>.05 (n.s.)
/s/	z=-1,31, p>.05 (n.s.)
/ʃ/	z=1,06, p>.05 (n.s.)

Table 5.85 : résumé des tests post-hoc sur l'effet d'interaction entre contexte vocalique et lieu d'articulation sur le deuxième moment spectral (skewness) chez les enfants NH (gauche) et les enfants CI (droite)

3.3.5. Kurtosis

La figure 5.52 ci-dessous présente les valeurs de kurtosis par lieu d'articulation, en fonction du groupe et du contexte vocalique.

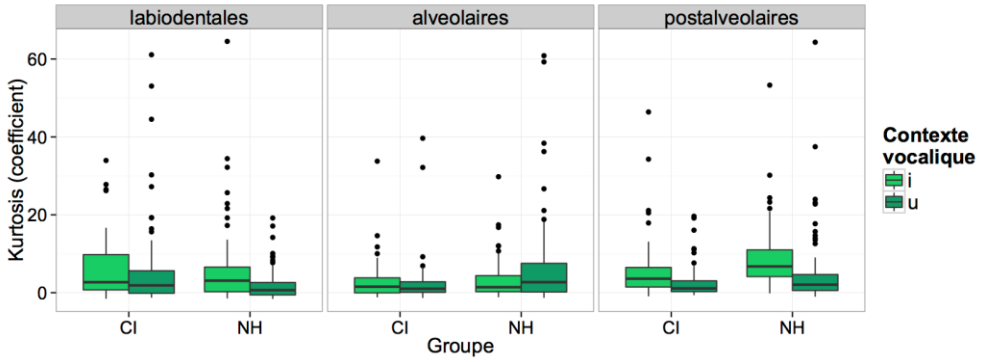


Figure 5.52 : kurtosis (coefficient) par lieu d'articulation, par groupe et par contexte vocalique

Les résultats du test statistique présentés dans la Table 5.86 ci-dessous indiquent un effet du contexte vocalique sur le quatrième moment spectral (ou kurtosis) seul ($F(1,1053)=8,35$, $p<.01$), en interaction avec le lieu d'articulation ($F(2,1053)=14,49$, $p<.001$) et en interaction avec le groupe et le lieu d'articulation ($F(2,1053)=4,81$, $p<.01$).

	numDF	denDF	F-value	p-value
(Intercept)	1	1053	153.62878	<.0001
C	2	1053	5.67375	0.0035 **
groupe	1	31	0.24355	0.6251
V	1	1053	8.34680	0.0039 **
C:groupe	2	1053	11.34244	<.0001 ***
C:V	2	1053	14.49015	<.0001 ***
groupe:V	1	1053	0.38274	0.5363
C:groupe:V	2	1053	4.80056	0.0084 **

Table 5.86 : tableau de résultats au test statistique sur toutes les données (effets de groupe, lieu d'articulation, contexte vocalique sur le kurtosis)

Un test post-hoc (Table 5.87 ci-dessous) indique une différence significative entre les valeurs de kurtosis en contexte /i/ et /u/ pour les labiodentales.

Consonne	Kurtosis
	Contexte vocalique : /i/ vs /u/
/f/	z=2,77, p<.01 **
/s/	z=-2,54, p<.01 **
/ʃ/	z=4,67, p<.001 ***

Table 5.87 : résumé du test post-hoc sur l'effet d'interaction entre contexte vocalique et lieu d'articulation sur le troisième moment spectral (kurtosis)

Un test statistique (modèle mixte) sur les données de kurtosis des enfants NH seuls indiquent un effet du contexte vocalique seul (F(1,680)=8,10, p<.01) et en interaction avec le lieu d'articulation (F(2,680)=18,41, p<.001) sur le quatrième moment spectral (ou kurtosis), qui est plus élevé en contexte /i/ avec les labiodentales et les postalvéolaires mais plus élevé en contexte /u/ avec les alvéolaires (les résultats du test post-hoc de l'interaction entre contexte vocalique et lieu d'articulation chez les enfants NH est présenté dans la table 5.88 ci-dessous).

Consonne	Kurtosis
	Contexte vocalique : /i/ vs /u/
/f/	z=3,74, p<.001 ***
/s/	z=-3,27, p<.01 **
/ʃ/	z=4,50, p<.001 ***

Table 5.88 : résumé du test post-hoc sur l'effet d'interaction entre contexte vocalique et lieu d'articulation sur le troisième moment spectral (kurtosis) chez les enfants NH

En revanche, le même test statistique sur les données de kurtosis des enfants CI seuls ne montrent pas d'effet du contexte vocalique sur les valeurs de kurtosis (F(1,373)=1,58, p>.05). Les enfants NH produisent donc des labiodentales et postalvéolaires de façon plus intense autour du centre de gravité en contexte /i/ mais des alvéolaires plus intense autour du centre de gravité en contexte /u/. Cette différence n'est pas observée chez les enfants CI.

3.3.6. Résumé : effet du contexte vocalique

La table 5.89 ci-dessous propose un résumé des effets observés du contexte vocalique sur chacun des moments spectraux étudiés dans ce paragraphe.

Effets	Moments spectraux				
	Centre de gravité	Distance entre centres de gravité de /s/ et /ʃ/	Ecart-type	Skewness	Kurtosis
Contexte vocalique	/i/ > /u/	/i/ > /u/	/i/ < /u/	/i/ > /u/	/i/ > /u/
Contexte vocalique * Groupe	NH : /i/ > /u/	n.s.	n.s.	NH : /i/ > /u/	NH : /i/ > /u/
	CI : /i/ > /u/		CI : n.s.	CI : n.s.	CI : n.s.
Contexte vocalique * Lieu d'articulation	f : n.s.	NA	f : /i/ < /u/	f : /i/ > /u/	f : /i/ > /u/
	s : /i/ > /u/		s : /i/ > /u/	s : /i/ < /u/	s : n.s.
	ʃ : n.s.		ʃ : /i/ < /u/	ʃ : /i/ > /u/	ʃ : /i/ > /u/

Table 5.89 : résumé des effets de la variable contexte vocalique et de l'interaction entre contexte vocalique et groupe et entre contexte vocalique et lieu d'articulation sur chaque moment spectral (n.s. indique une absence de différence significative)

Les analyses réalisées sur les quatre moments spectraux et présentées dans ce paragraphe ont permis de mettre en évidence des réalisations différentes selon le contexte vocalique (/i/ et /u/) : les analyses indiquent que les fricatives sont produites avec un centre de gravité plus haut lorsque la fricative est suivie de /i/. Cette différence n'est cependant significative que pour les alvéolaires, mais elle affecte de la même façon les deux groupes d'enfants (NH et CI).

Les résultats indiquent un écart-type globalement plus bas lorsque la fricative est suivie de /i/, mais cette différence liée au contexte vocalique n'est pas significative lorsque les données de chaque groupe sont analysées séparément. Une analyse par lieu d'articulation indique un effet qui varie selon la fricative : pour les labiodentales et postalvéolaires, l'écart-type est plus bas en contexte /i/, ce qui correspond à des fricatives plus compactes, et il est plus haut en contexte /u/, ce qui correspond à des fricatives plus diffuses.

Les troisième et quatrième moments spectraux sont quant à eux plus élevés lorsque la voyelle suivante est un /i/. Cette différence de réalisation en contexte vocalique /i/ ou /u/ est significative pour les enfants NH mais pas pour les enfants CI. Comme pour les premier et deuxième moments spectraux, les alvéolaires ont des résultats différents de ceux des autres fricatives : les skewness et kurtosis des labiodentales et postalvéolaires sont plus élevés lorsque celles-ci sont suivies du /i/, alors que le skewness des alvéolaires est plus bas lorsque la voyelle suivante est un /i/, et aucune différence significative n'est observée pour le kurtosis des alvéolaires.

3.4. Enfants NH : effet de l'âge chronologique

Comme nous cherchons à étudier l'effet de l'implantation sur la production des fricatives, nous avons décidé d'étudier l'effet de l'âge chronologique indépendamment chez les deux groupes d'enfants : pour les enfants NH qui ne sont pas concernés par l'implantation, l'âge chronologique correspond à l'âge auditif, alors que pour les enfants CI, l'âge chronologique correspond à l'âge auditif combiné à l'âge à l'implantation. Nous allons donc dans un premier temps étudier l'effet de l'âge chronologique sur les quatre moments spectraux chez les enfants NH.

3.4.1. Centre de gravité

La figure 5.53 ci-dessous présente l'évolution du premier moment spectral ou centre de gravité en fonction de l'âge chronologique des enfants NH, pour chaque lieu d'articulation (labiodentales, alvéolaires et postalvéolaires) et chaque contexte vocalique (/i/ et /u/).

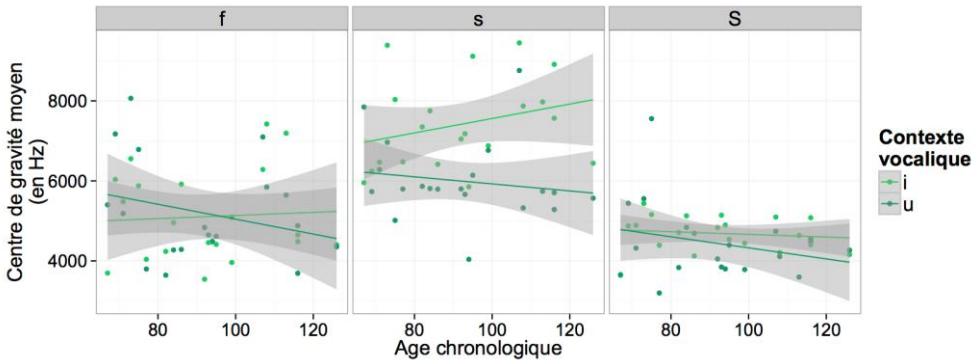


Figure 5.53 : centre de gravité (en Hz) en fonction de l'âge chronologique (en mois) par contexte vocalique et lieu d'articulation, pour les enfants NH

Nous avons testé avec un modèle mixte à effet aléatoire du sujet l'effet de l'âge chronologique sur la réalisation du centre de gravité par les enfants NH, seul et en interaction avec le lieu d'articulation et le contexte vocalique. La table 5.90 ci-dessous présente les résultats à ce test statistique.

	numDF	denDF	F-value	p-value
(Intercept)	1	662	1418.4040	<.0001
C	2	662	146.3509	<.0001 ***
V	1	662	24.4923	<.0001 ***
âge_chronologique	1	18	0.1677	0.6870
C:V	2	662	18.7092	<.0001 ***
C:âge_chronologique	2	662	2.3348	0.0976
V:âge_chronologique	1	662	11.0521	0.0009 ***

Table 5.90 : tableau de résultats au test statistique sur les données des enfants NH (effets de l'âge chronologique, du lieu d'articulation, du contexte vocalique sur le centre de gravité)

Les résultats indiquent que l'âge chronologique seul n'a aucun effet sur la hauteur du centre de gravité (F(1,18)=0,17, p>.05) : le centre de gravité est stable chez notre population d'enfants NH, dont l'âge va de 5;7 ans à 10;6 ans. Cependant, l'âge chronologique a un effet sur le centre de gravité en interaction avec le contexte vocalique (F(1,662)=10,5, p<.001) : l'observation de la figure 5.53 indique que le centre de gravité des labiodentales et alvéolaires suivies de /i/ augmente avec l'âge chronologique, alors qu'il diminue lorsque ces fricatives sont suivies de /u/. Le centre de gravité des postalvéolaires est quant à lui stable lorsque celles-ci sont suivies de /i/ et diminue après /u/. Des tests supplémentaires par voyelle indiquent cependant que cet effet de l'âge chronologique n'est pas significatif (pour le contexte /i/, F(1,18)=0,99, p>.05 et pour le contexte /u/ : F(1,18)=2,05, p>.05), ce qui est peut-être dû à un nombre trop faible de données.

3.4.2. Distance entre centres de gravité de /s/ et /ʃ/

La figure 5.54 ci-dessous présente l'évolution de la distance entre les centres de gravité moyens de /s/ et /ʃ/ en fonction de l'âge chronologique et du contexte vocalique, pour les enfants NH. L'observation de cette figure semble indiquer une stabilité de cette distance avec l'âge chronologique lorsque les fricatives sont suivies de /u/ mais une augmentation légère lorsqu'elles sont suivies de /i/.

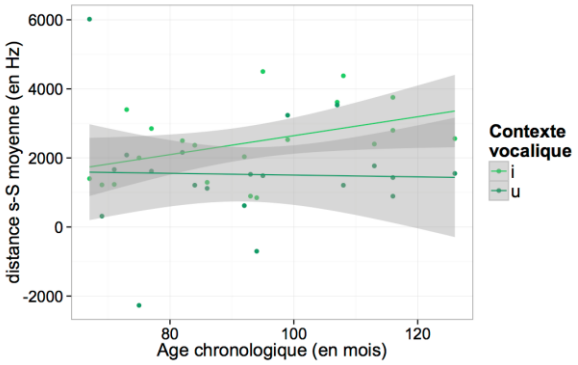


Figure 5.54 : distance entre les centres de gravité de /s/ et /ʃ/ (en Hz) en fonction de l'âge chronologique (en mois) par contexte vocalique, pour les enfants NH

Les résultats du test statistique réalisé sur les données des enfants NH sont proposés dans la table 5.91 ci-dessous.

	numDF	denDF	F-value	p-value
(Intercept)	1	96	76.41116	<.0001
V	1	96	15.32412	0.0002 ***
âge_chronologique	1	18	0.93291	0.3469
V:âge_chronologique	1	96	5.51632	0.0209 *

Table 5.91 : tableau de résultats au test statistique sur les données des enfants NH (effets de l'âge chronologique et du contexte vocalique sur la distance entre les centres de gravité de /s/ et /ʃ/)

Les résultats du test statistique indiquent que l'âge chronologique seul n'a pas d'effet sur la distance entre les centres de gravité de /s/ et /ʃ/ ($F(1,18)=0,93$, $p>.05$) mais que l'interaction entre âge chronologique et contexte vocalique a un effet significatif sur cette distance ($F(1,96)=5,52$, $p<.05$). L'observation visuelle de la figure 5.54 indique une augmentation de cette distance en contexte /i/ et une stagnation en contexte /u/ : cette évolution de la distance entre les centres de gravité de /s/ et /ʃ/ en fonction de l'âge chronologique témoigne d'une coarticulation plus forte en contexte /i/ qui ne s'est pas encore stabilisée chez les enfants NH, âgés de 5;7 ans à 10;6 ans. Cette évolution de la coarticulation est confirmée par un test statistique qui met en évidence un effet significatif de l'âge chronologique en contexte /i/ ($F(1,18)=4,83$, $p<.05$) mais pas en contexte /u/ ($F(1,18)=0,02$, $p>.05$).

3.4.3.Ecart-type

La figure 5.55 ci-dessous présente l'évolution du deuxième moment spectral ou écart-type en fonction de l'âge chronologique des enfants NH, pour chaque lieu d'articulation (labiodentales, alvéolaires et postalvéolaires) et chaque contexte vocalique (/i/ et /u/).

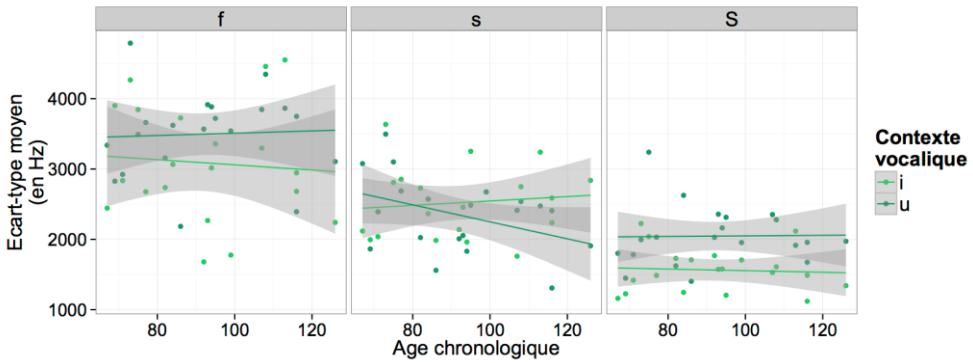


Figure 5.55 : écart-type (en Hz) en fonction de l'âge chronologique (en mois) par contexte vocalique et lieu d'articulation, pour les enfants NH

Les résultats du test statistique réalisé à partir des données du groupe des enfants NH sur l'écart-type sont présentés dans la table 5.92 ci-dessous.

	numDF	denDF	F-value	p-value
(Intercept)	1	660	1026.2652	<.0001
C	2	660	260.8945	<.0001 ***
V	1	660	21.4787	<.0001 ***
âge_chronologique	1	18	0.1384	0.7142
C:V	2	660	15.0227	<.0001 ***
C:âge_chronologique	2	660	0.5284	0.5898
V:âge_chronologique	1	660	0.7708	0.3803
C:V:âge_chronologique	2	660	4.5312	0.0111 *

Table 5.92 : tableau de résultats au test statistique sur les données des enfants NH (effets de l'âge chronologique, du lieu d'articulation, du contexte vocalique sur l'écart-type)

Ce test n'indique pas d'effet de l'âge chronologique seul ($F(1,18)=0,14$, $p>.05$), ni en interaction avec le lieu d'articulation ($F(1,660)=0,53$, $p>.05$) ou le contexte vocalique ($F(1,660)=0,77$, $p>.05$) sur la réalisation de l'écart-type. Cependant on note un effet de l'interaction entre âge chronologique, lieu d'articulation et contexte vocalique sur l'écart-type. Un test par lieu d'articulation et contexte vocalique n'atteint pas le niveau de significativité ($p<.05$) mais n'atteint que le niveau d'une tendance pour les alvéolaires en contexte /u/ seulement ($F(1,18)=3,65$, $p=0.0722$.

3.4.4.Skewness

La figure 5.56 ci-dessous présente l'évolution du troisième moment spectral ou skewness en fonction de l'âge chronologique des enfants NH, pour chaque lieu d'articulation (labiodentales, alvéolaires et postalvéolaires) et chaque contexte vocalique (/i/ et /u/).

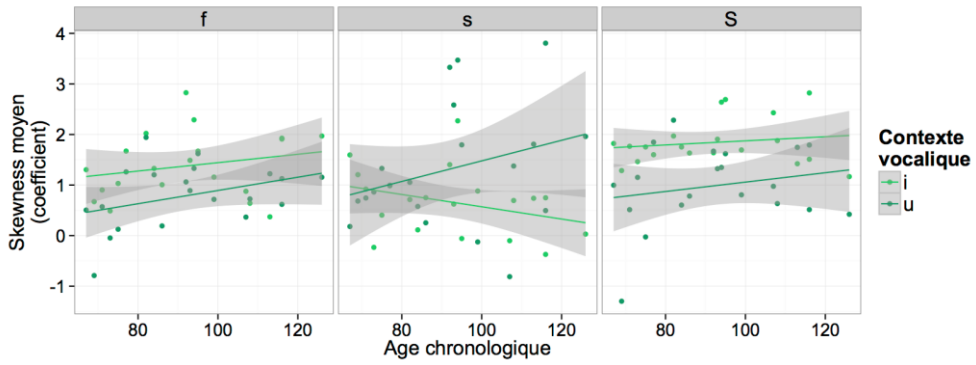


Figure 5.56 : skewness (coefficient) en fonction de l'âge chronologique (en mois) par contexte vocalique et lieu d'articulation, pour les enfants NH

Les résultats du test statistique réalisé à partir des données du groupe des enfants NH sur le skewness sont présentés dans la table 5.93 ci-dessous.

	numDF	denDF	F-value	p-value
(Intercept)	1	660	119.87242	<.0001
C	2	660	10.53393	<.0001 ***
V	1	660	11.91072	0.0006 ***
âge_chronologique	1	18	1.15471	0.2968
C:V	2	660	36.39941	<.0001 ***
C:âge_chronologique	2	660	1.06380	0.3457
V:âge_chronologique	1	660	11.89530	0.0006 ***
C:V:âge_chronologique	2	660	4.27000	0.0144 *

Table 5.93 : tableau de résultats au test statistique sur les données des enfants NH (effets de l'âge chronologique, du lieu d'articulation, du contexte vocalique sur le skewness)

Ces résultats n'indiquent pas d'effet de l'âge chronologique seul ($F(1,18)=1,15$, $p>.05$) mais un effet significatif de l'interaction entre contexte vocalique et âge chronologique ($F(1,660)=11,90$, $p<.001$), ainsi que de l'interaction entre contexte vocalique, lieu d'articulation et âge chronologique ($F(2,660)=4,27$, $p<.05$) : l'observation visuelle de la figure 5.56 indique une augmentation du skewness pour les labiodentales en contexte /i/ et /u/, une augmentation du skewness en contexte /u/ et une diminution en contexte /i/ pour les alvéolaires, ainsi qu'une augmentation légère dans les deux contextes pour les postalvéolaires. Des tests supplémentaires par lieu d'articulation et contexte vocalique ne confirment cependant pas ces observations.

3.4.5. Kurtosis

La figure 5.57 ci-dessous présente l'évolution du quatrième moment spectral ou kurtosis en fonction de l'âge chronologique pour les enfants NH, par lieu d'articulation (labiodentales, alvéolaires et postalvéolaires) et contexte vocalique (/i/ ou /u/)

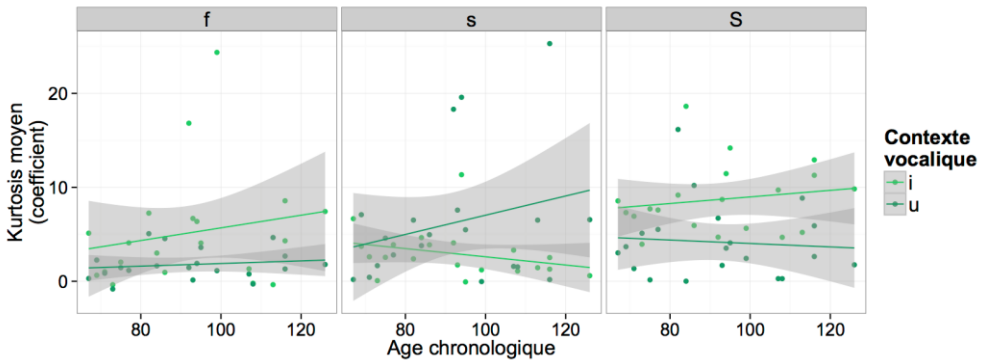


Figure 5.57 : kurtosis (coefficient) en fonction de l'âge chronologique (en mois) par contexte vocalique et lieu d'articulation, pour les enfants NH

Les résultats du test statistique réalisé à partir des données du groupe des enfants NH sur le kurtosis sont présentés dans la table 5.94 ci-dessous.

	numDF	denDF	F-value	p-value
(Intercept)	1	660	101.27921	<.0001
C	2	660	9.47552	0.0001 ***
V	1	660	8.10460	0.0046 **
âge_chronologique	1	18	0.51773	0.4810
C:V	2	660	18.40815	<.0001 ***
C:âge_chronologique	2	660	0.58754	0.5560
V:âge_chronologique	1	660	0.23469	0.6282
C:V:âge_chronologique	2	660	3.92557	0.0202 *

Table 5.94 : tableau de résultats au test statistique sur les données des enfants NH (effets de l'âge chronologique, du lieu d'articulation, du contexte vocalique sur le kurtosis)

Ces résultats indiquent que l'âge chronologique, seul (F(1,18)= 0,52, p>.05), en interaction avec le lieu d'articulation (F(2,660)=0,59, p>.05) ou en interaction avec le contexte vocalique (F(1,660)=0,23, p>.05) n'a pas d'effet significatif sur ce quatrième moment spectral. En revanche, l'interaction entre âge chronologique, contexte vocalique et lieu d'articulation est a un effet significatif sur la réalisation du quatrième moment spectral des enfants NH (F(2,660)=3,93, p<.05). L'observation de la figure 5.57 ci-dessus indique que le kurtosis augmente pour chaque lieu d'articulation et chaque contexte vocalique, sauf pour les alvéolaires en contexte /i/. Des tests statistiques supplémentaires ne permettent cependant pas de mettre en évidence un effet significatif de l'âge chronologique.

3.4.6. Résumé : effets de l'âge chronologique (enfants NH)

L'ensemble des effets de l'âge chronologique, observés dans ce paragraphe, sont résumés dans la table 5.95 ci-dessous.

Effets	Moments spectraux				
	Centre de gravité	Distance entre centres de gravité de /s/ et /ʃ/	Ecart-type	Skewness	Kurtosis
âge chronologique (NH)	n.s.	n.s.	n.s.	n.s.	n.s.
âge chronologique (NH)*Lieu d'articulation	n.s.	NA	n.s.	n.s.	n.s.
âge chronologique (NH) * Contexte vocalique	/i/ : augmentation (n.s.) /u/ : diminution (n.s.)	/i/ : augmentation (p<.05) /u/ : stabilité (n.s.)	n.s.	n.s.	n.s.
âge chronologique (NH) * Lieu d'articulation * Contexte vocalique	n.s.	NA	n.s.	f & ʃ : augmentation (n.s.) s : augmentation avec i (n.s) s : diminution avec u (n.s)	n.s.

Table 5.95 : résumé des effets chez les enfants NH de la variable « âge chronologique », seule et en interaction avec le groupe, le lieu d'articulation et contexte vocalique sur chaque moment spectral (« n.s. » indique une différence non significative)

Comme le montre la table 5.95, les analyses que nous avons effectuées n'ont pas permis de mettre en évidence un effet principal de l'âge chronologique sur les variables étudiées dans le cadre de notre étude. Des influences ponctuelles mais non significatives de l'âge chronologique sont observées mais seule la distance entre les centres de gravité des alvéolaires et postalvéolaires est affectée par l'âge chronologique, lorsque les fricatives sont suivies de /i/ : en effet, le centre de gravité de /s/ dans des séquences /si/ a tendance à s'élever, c'est-à-dire qu'il subit une coarticulation plus grande du /i/, pour s'éloigner de plus en plus de celui de /ʃ/ à mesure que l'enfant grandit, ce qui d'après la littérature (Nissen & Fox, 2005, Nittrouer, 1995 et Körkkö, 2015) était un résultat attendu. Nous notons par ailleurs que les second (écart-type), troisième (skewness) et quatrième (kurtosis) moments spectraux ne sont pas influencés par l'âge chronologique chez les enfants NH.

3.5. Enfants CI : effet de l'âge à l'implantation et de la durée d'utilisation de l'implant

Les analyses réalisées au paragraphe précédent n'ont pas mis en évidence un effet global de l'âge chronologique sur les fricatives produites par les enfants NH, mais quelques effets ponctuels, pour certains lieux d'articulation et dans certains contextes vocaliques seulement. Les analyses présentées dans la suite de ce paragraphe s'intéressent à deux facteurs : l'âge à l'implantation et la durée d'utilisation de l'implant (appelée « âge auditif » dans la suite de ce chapitre).

3.5.1.Centre de gravité

La figure 5.58 ci-dessous présente l'évolution du centre de gravité des fricatives produites par les enfants CI, en fonction de ces deux facteurs, pour chaque contexte vocalique. Une première observation de cette figure nous indique que les postalvéolaires semblent être les seules fricatives dont le centre de gravité n'est pas influencé par ces deux facteurs, à la différence des labiodentales et alvéolaires, dont l'évolution semble différer selon le contexte vocalique.

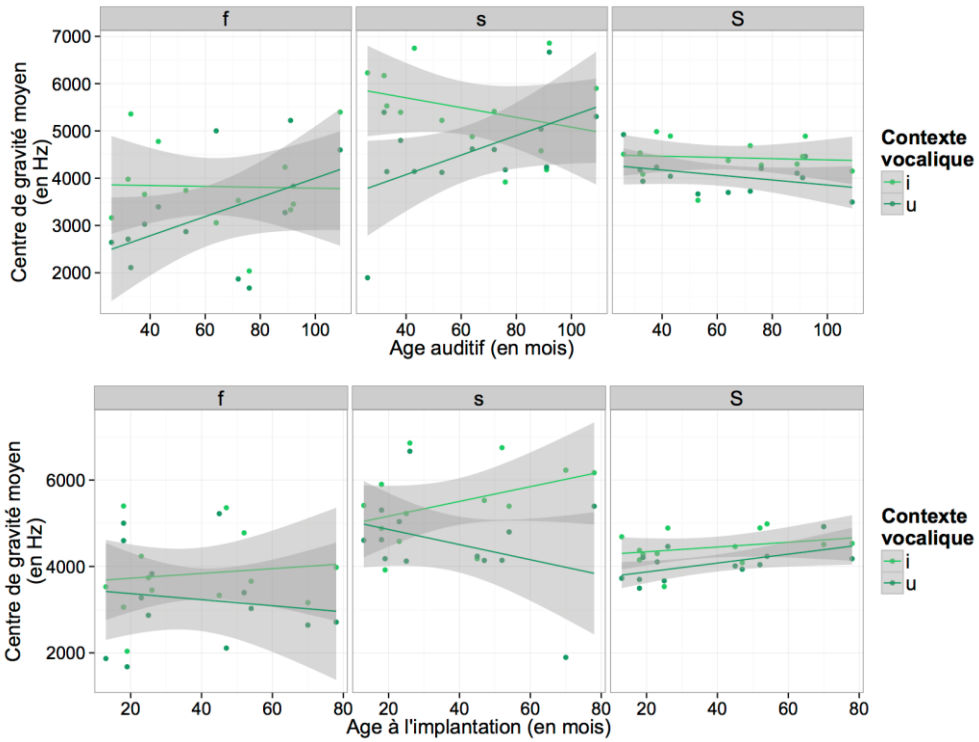


Figure 5.58 : Centre de gravité (en Hz) en fonction de l'âge auditif (haut, en mois) et âge à l'implantation (bas, en mois) par contexte vocalique et lieu d'articulation, pour les enfants CI

Ces observations semblent confirmées par le test statistique, dont les résultats sont présentés dans la table 5.96 ci-dessous. En effet, il n'y a pas d'effet significatif de l'âge auditif seul ($F(1,9)=1,13$, $p>.05$) ni de l'âge à l'implantation seul ($F(1,9)=2;63$, $p>.05$). Cependant nous observons un effet de l'âge auditif en interaction avec le lieu d'articulation ($F(2,377)=0,1$, $p<.05$) et avec le contexte vocalique ($F(1,377)=9,65$, $p<.01$).

	numDF	denDF	F-value	p-value
(Intercept)	1	377	1754.8819	<.0001
C	2	377	49.1561	<.0001 ***
V	1	377	22.1195	<.0001 ***
âge_auditif	1	9	1.1250	0.3165
âge_implantation	1	9	3.6295	0.0892
C:V	2	377	0.5496	0.5776
C:âge_auditif	2	377	3.3601	0.0358 *
V:âge_auditif	1	377	9.6453	0.0020 **
C:âge_implantation	2	377	0.9967	0.3701
V:âge_implantation	1	377	0.0740	0.7858
âge_auditif:âge_implantation	1	9	0.1235	0.7333
C:V:âge_auditif	2	377	2.8886	0.0569
C:âge_auditif:âge_implantation	2	377	2.1810	0.1144
V:âge_auditif:âge_implantation	1	377	3.8830	0.0495 *

Table 5.96 : tableau de résultats au test statistique sur les données des enfants CI (effets de l'âge auditif et de l'âge à l'implantation, du lieu d'articulation, du contexte vocalique sur le centre de gravité)

Testé séparément par lieu d'articulation, l'effet de l'âge auditif n'atteint cependant pas le niveau de significativité fixé à $p<.05$, peut-être en raison d'un nombre insuffisant de données. En revanche, les résultats des tests séparés par contexte vocalique n'atteignent pas ce niveau de significativité fixé à $p<.05$ lorsque la fricative est suivie de /i/ mais l'effet de l'âge auditif sur les centres de

gravité est significatif lorsque la fricative est suivie de /u/ (F(1,11)=7,07, p<.05). Globalement, l'effet de l'âge auditif sur le centre de gravité semble donc être une tendance due à certains lieux d'articulation et certains contextes vocaliques qu'il faudrait confirmer avec un nombre plus conséquent de données. L'antériorité de la voyelle /i/ ou /u/ varie donc au fur et à mesure de l'utilisation de l'implant mais l'âge à l'implantation seul n'influence pas la réalisation du centre de gravité des trois fricatives de notre étude.

Nous observons en outre un effet de l'interaction entre contexte vocalique, âge auditif et âge à l'implantation sur le centre de gravité des fricatives : la durée d'utilisation de l'implant a donc un effet sur le centre de gravité des fricatives qui est différent selon l'âge auquel les enfants ont reçu l'implant : les tests par lieu d'articulation et par contexte vocalique de cette interaction ne permettent cependant pas d'identifier les séquences fricative-voyelle qui sont influencées par cette interaction de l'âge à l'implantation et de l'âge auditif.

3.5.2. Distance entre centres de gravité de /s/ et /ʃ/

La figure 5.59 ci-dessous présente l'évolution de la distance entre les centres de gravité de /s/ et /ʃ/ qui correspond à la distance entre les deux catégories de fricatives par les enfants CI. Une première observation indique que l'âge auditif (gauche) et l'âge à l'implantation (droite) semblent avoir une influence inverse et différente selon le contexte vocalique : augmentation de la distance entre les centres de gravité en fonction de l'âge auditif et diminution en fonction de l'âge à l'implantation assez marquées lorsque les fricatives sont suivies de /u/, et diminution en fonction de l'âge auditif et augmentation en fonction de l'âge à l'implantation plus modérées lorsque les fricatives sont suivies de /i/. Les analyses de l'évolution du centre gravité en fonction de l'âge auditif et de l'âge à l'implantation présentées au paragraphe précédent (§3.5.1) indiquaient une augmentation du centre de gravité des alvéolaires en fonction de l'âge auditif avec la voyelle /u/ mais une diminution avec /i/, qui s'éloignait petit à petit du centre de gravité des postalvéolaires. Cette distinction des deux lieux d'articulation en contexte /u/ semble donc confirmée par l'observation de la figure 5.59, puisque celle-ci indique une augmentation avec l'âge auditif de la distance entre centres de gravité de /s/ et /ʃ/ lorsque les fricatives sont suivies de /u/ mais une stabilité apparente lorsqu'elles sont suivies de /i/.

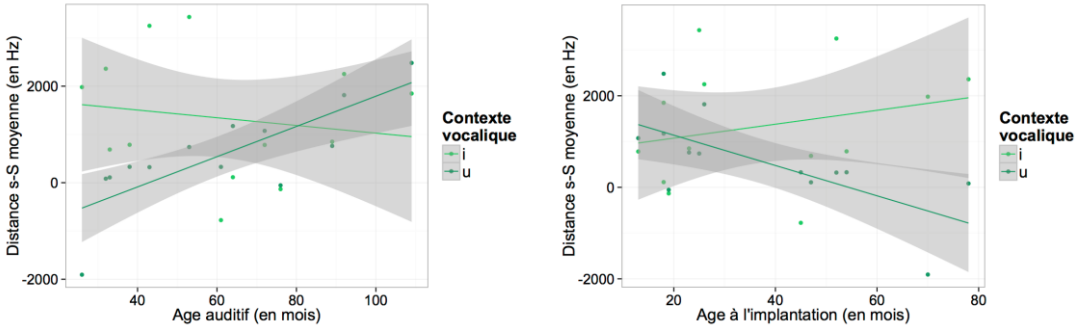


Figure 5.59 : distance entre les centres de gravité de /s/ et /ʃ/ (en Hz) en fonction de l'âge auditif (gauche, en mois) et âge à l'implantation (droite, en mois) par contexte vocalique, pour les enfants CI

La table 5.97 ci-dessous présente les résultats du test statistique effectué sur les distances entre centres de gravité de /s/ et /ʃ/ chez les enfants CI, qui prend en compte âge auditif et âge à l'implantation.

	numDF	denDF	F-value	p-value
(Intercept)	1	55	19.293508	0.0001
V	1	55	7.921734	0.0068 **
âge_auditif	1	9	2.528093	0.1463
âge_implantation	1	9	0.662511	0.4367
V:âge_auditif	1	55	15.305217	0.0003 ***
V:âge_implantation	1	55	3.475450	0.0676
âge_auditif:âge_implantation	1	9	0.089307	0.7718
V:âge_auditif:âge_implantation	1	55	2.477700	0.1212

Table 5.97 : tableau de résultats au test statistique sur les données des enfants CI (effets de l'âge auditif, de l'âge à l'implantation et du contexte vocalique sur la distance entre les centres de gravité de /s/ et /ʃ/)

Les résultats statistiques indiquent un effet de l'âge auditif en interaction avec le contexte vocalique ($F(1,55)=15,31$, $p<.001$), ce qui confirme la première observation visuelle de la figure 5.59. Cependant, l'évolution liée à l'âge à l'implantation n'est pas confirmée par le test statistique ($F(1,55)=3,48$, $p>.05$) : nous observons également que la distance entre /s/ et /ʃ/ augmente légèrement avec l'âge auditif, ce qui signifie que plus les enfants CI ont une utilisation prolongée de l'implant cochléaire, plus ils sont en mesure de différencier les deux fricatives /s/ et /ʃ/. Un test supplémentaire indique que cette augmentation de la distance entre /s/ et /ʃ/ avec l'âge auditif est significative lorsque la voyelle suivante est un /u/ ($F(1,11)=19,78$, $p<.001$).

3.5.3.Ecart-type

La figure 5.60 présente l'évolution de l'écart-type en fonction de l'âge auditif (gauche) et de l'âge à l'implantation pour chaque lieu d'articulation et chaque contexte vocalique. L'observation de cette figure indique une augmentation de l'écart-type des labiodentales en fonction de l'âge auditif et une diminution de celui des postalvéolaires en fonction de l'âge à l'implantation. L'écart-type des alvéolaires ne semble pas être affecté par ces deux facteurs.

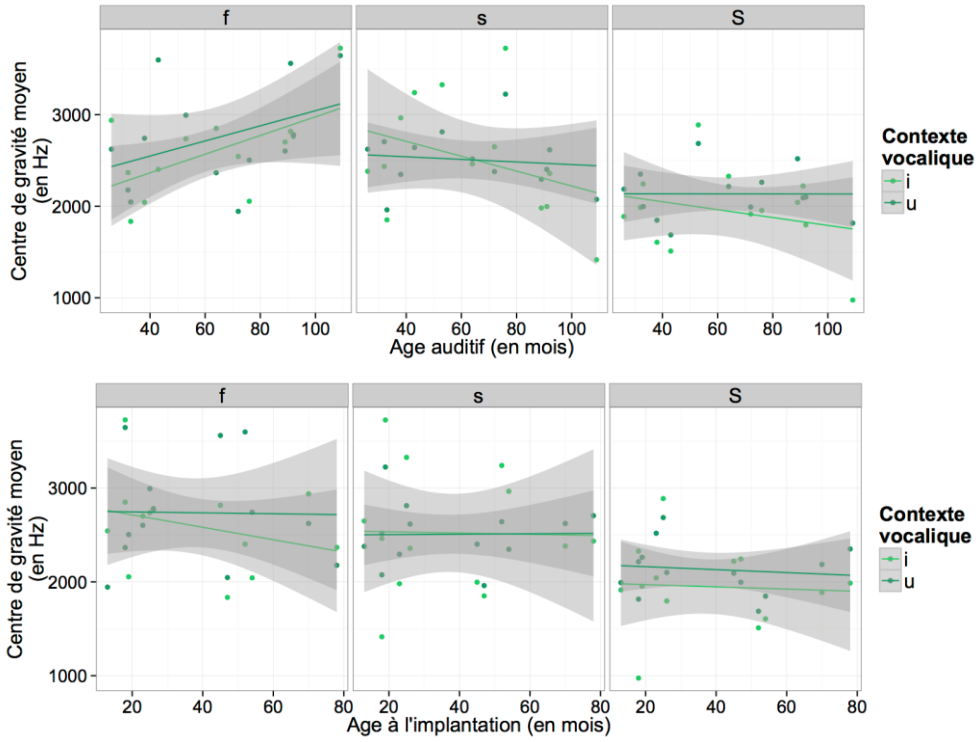


Figure 5.60 : Ecart-type (en Hz) en fonction de l'âge auditif (haut, en mois) et âge à l'implantation (bas, en mois) par contexte vocalique et lieu d'articulation, pour les enfants CI

La table 5.98 ci-dessous présente les résultats au test statistique sur l'écart-type des fricatives produites par les enfants CI. Nous observons comme pour le centre de gravité que ni l'âge auditif (F(1,9)=0,09,p>.05), ni l'âge à l'implantation (F(1,9)=0,34, p>.05) n'ont d'effet sur l'écart-type.

	numDF	denDF	F-value	p-value
(Intercept)	1	373	1774.6700	<.0001
C	2	373	40.5433	<.0001 ***
V	1	373	3.6464	0.0570
âge_auditif	1	9	0.0936	0.7666
âge_implantation	1	9	0.3432	0.5724
C:V	2	373	1.4981	0.2249
C:âge_auditif	2	373	18.0441	<.0001 ***
V:âge_auditif	1	373	1.2636	0.2617
C:âge_implantation	2	373	8.9593	0.0002 ***
V:âge_implantation	1	373	5.8480	0.0161 *
âge_auditif:âge_implantation	1	9	1.6660	0.2290
C:V:âge_auditif	2	373	1.9076	0.1499
C:V:âge_implantation	2	373	0.2645	0.7678
C:âge_auditif:âge_implantation	2	373	1.0068	0.3664
V:âge_auditif:âge_implantation	1	373	0.0083	0.9275
C:V:âge_auditif:âge_implantation	2	373	6.0687	0.0025 **

Table 5.98 : tableau de résultats au test statistique sur les données des enfants CI (effets de l'âge auditif et de l'âge à l'implantation, du lieu d'articulation, du contexte vocalique sur l'écart-type)

Comme pour le centre de gravité, l'interaction entre âge auditif et lieu d'articulation a un effet significatif (F(2,373)=18,04, p<.001). Testé séparément par lieu d'articulation, l'effet de l'âge auditif n'est significatif que pour les labiodentales (F(1,11)=5,77, p<.05), comme la représentation graphique (Figure 5.60) semblait l'indiquer. En revanche, testé séparément par lieu d'articulation, puis par contexte vocalique, nous n'observons aucun effet significatif de l'âge à l'implantation sur ces deux variables.

3.5.4.Skewness

La figure 5.61 présente l'évolution du troisième moment spectral ou skewness chez les enfants CI en fonction de l'âge auditif (haut) et de l'âge à l'implantation (bas), par lieu d'articulation et par contexte vocalique : l'observation de cette figure semble indiquer un effet marqué de l'âge auditif sur ce moment spectral qui diffère nettement selon le contexte vocalique, les valeurs de skewness diminuant avec l'âge auditif en contexte /u/ et augmentant en contexte /i/. Cette tendance semble cependant se réaliser en sens inverse avec l'âge à l'implantation : pour les alvéolaires et postalvéolaires, les valeurs de skewness augmentent avec l'âge à l'implantation en contexte /u/ et diminuent en contexte /i/. Comme le skewness est corrélé négativement avec la longueur de la cavité de résonance, en avant constriction (Li et al., 2009), l'augmentation avec l'âge à l'implantation en contexte /u/ correspondrait à une cavité de résonance plus grande c'est à dire une production des alvéolaires et postalvéolaires plus postérieures pour les enfants implantés tardivement. En revanche, la diminution du skewness des alvéolaires en contexte /i/ avec l'âge à l'implantation correspond à une production plus antérieure chez les enfants implantés tardivement.

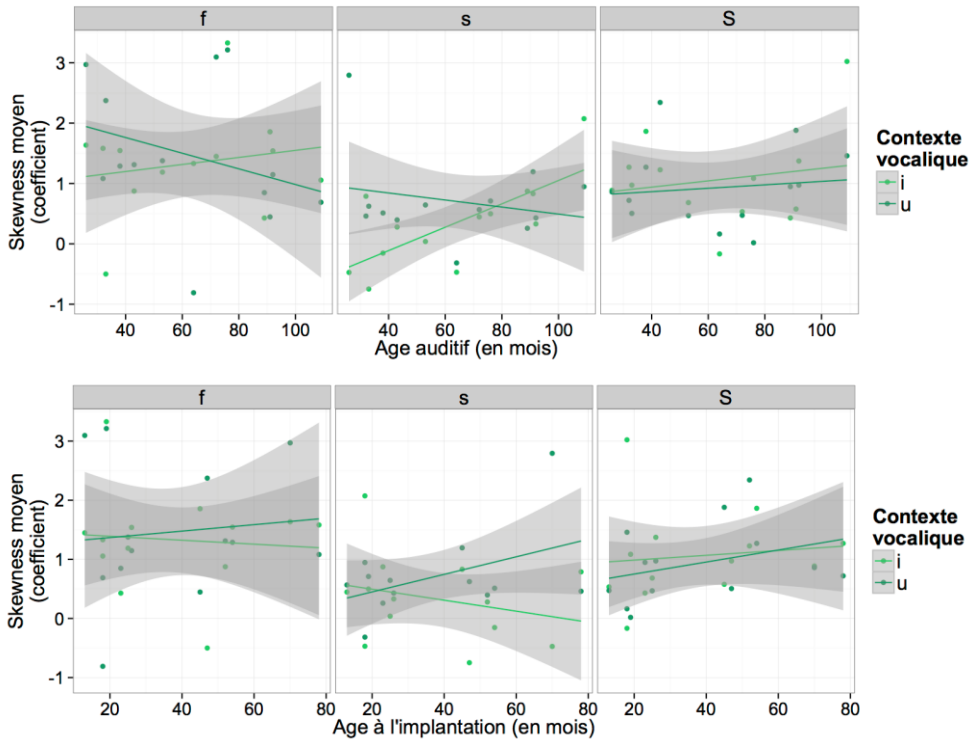


Figure 5.61 : Skewness (coefficient) en fonction de l'âge auditif (haut, en mois) et âge à l'implantation (bas, en mois) par contexte vocalique et lieu d'articulation, pour les enfants CI

Les résultats au test statistique de l'effet de l'âge auditif et de l'âge à l'implantation présentés dans la table 5.99 ci-dessous semblent confirmer cette tendance observée visuellement : il n'y a pas d'effet global de l'âge auditif (F(1,9)=0,47, p>.05) ni de l'âge à l'implantation (F(1,9)=2,96, p>.05).

	numDF	denDF	F-value	p-value
(Intercept)	1	373	102.23320	<.0001
C	2	373	22.93019	<.0001 ***
V	1	373	0.05363	0.8170
âge_auditif	1	9	0.47150	0.5096
âge_implantation	1	9	2.95620	0.1197
C:V	2	373	1.44210	0.2377
C:âge_auditif	2	373	3.79961	0.0233 *
V:âge_auditif	1	373	10.02745	0.0017 **
C:âge_implantation	2	373	6.97189	0.0011 **
V:âge_implantation	1	373	0.47511	0.4911
âge_auditif:âge_implantation	1	9	2.56752	0.1435
C:V:âge_auditif	2	373	1.35870	0.2583
C:V:âge_implantation	2	373	0.83570	0.4344
C:âge_auditif:âge_implantation	2	373	0.08048	0.9227
V:âge_auditif:âge_implantation	1	373	1.67248	0.1967
C:V:âge_auditif:âge_implantation	2	373	4.62362	0.0104 *

Table 5.99 : tableau de résultats au test statistique sur les données des enfants CI (effets de l'âge auditif et de l'âge à l'implantation, du lieu d'articulation, du contexte vocalique sur le troisième moment, ou skewness)

En revanche, on note un effet de l'âge auditif en interaction avec le lieu d'articulation (F(2,373)=3,80, p<.05) et avec le contexte vocalique (F(1,373)=10,03, p<.01), et un effet de l'âge à l'implantation en interaction avec le lieu d'articulation (F(2,373)=6,97, p<.01) sur le troisième

moment spectral. Des tests réalisés séparément par lieu d'articulation, puis par contexte vocalique ne permettent pas d'affiner cette analyse, et nous ne pouvons pas identifier précisément quel lieu d'articulation et quel contexte vocalique sont impactés par ces deux effets d'âge sur le troisième moment spectral.

3.5.5. Kurtosis

Enfin, nous avons étudié l'effet de ces deux variables (âge auditif et âge à l'implantation) sur le quatrième moment spectral, et nous présentons sur la figure 5.62 ci-dessous l'évolution de ce moment en fonction de ces deux variables : une première observation nous indique que les labiodentales semblent être les seules fricatives dont le quatrième moment spectral subit un effet de l'âge auditif, alors que le quatrième moment spectral des alvéolaires et postalvéolaires semble plus fortement soumis à l'influence de l'âge à l'implantation.

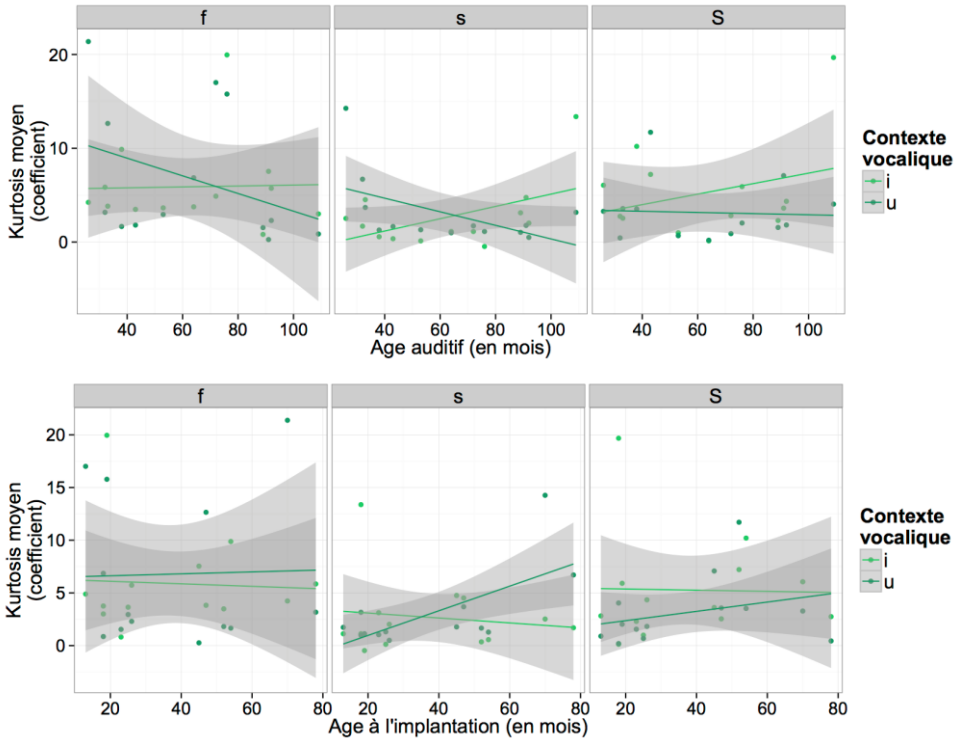


Figure 5.62 : Kurtosis (coefficient) en fonction de l'âge auditif (gauche, en mois) et âge à l'implantation (droite, en mois) par contexte vocalique et lieu d'articulation, pour les enfants CI

Les résultats des tests statistiques correspondant à la figure 5.62, pour le quatrième moment spectral ou kurtosis sont présentés dans la table 5.100 ci-dessous.

	numDF	denDF	F-value	p-value
(Intercept)	1	373	66.85531	<.0001
C	2	373	7.76520	0.0005 ***
V	1	373	1.57503	0.2103
âge_auditif	1	9	0.06544	0.8038
âge_implantation	1	9	0.43340	0.5268
C:V	2	373	1.55615	0.2123
C:âge_auditif	2	373	2.69561	0.0688
V:âge_auditif	1	373	10.98602	0.0010 ***
C:âge_implantation	2	373	6.98908	0.0010 ***
V:âge_implantation	1	373	1.55461	0.2132
âge_auditif:âge_implantation	1	9	3.63716	0.0889
C:V:âge_auditif	2	373	0.48209	0.6179
C:V:âge_implantation	2	373	1.30552	0.2723
C:âge_auditif:âge_implantation	2	373	0.32084	0.7257
V:âge_auditif:âge_implantation	1	373	0.14344	0.7051
C:V:âge_auditif:âge_implantation	2	373	5.32884	0.0052 **

Table 5.100 : tableau de résultats au test statistique sur les données des enfants CI (effets de l'âge auditif et de l'âge à l'implantation, du lieu d'articulation, du contexte vocalique sur le quatrième moment ou kurtosis)

L'âge auditif et l'âge à l'implantation seuls n'ont pas d'effet significatif sur la réalisation du quatrième moment spectral, cependant nous observons un effet de l'âge auditif en interaction avec le contexte vocalique ($F(1,373)=10,99$, $p<.001$) et de l'âge à l'implantation avec le lieu d'articulation ($F(2,373)=6,99$, $p<.001$). Comme pour ce que nous avions constaté pour le troisième moment spectral (skewness), les tests statistiques réalisés par lieu d'articulation et par contexte vocalique ne permettent pas de mettre en évidence un effet précis des deux variables d'âge sur un lieu d'articulation ou sur un contexte vocalique en particulier.

3.5.6. Résumé : effets de l'âge à l'implantation et de la durée d'utilisation de l'implant (enfants CI)

La table 5.101 ci-dessous présente un résumé des effets des variables « âge auditif » et « âge à l'implantation » étudiés chez les enfants CI, dans le cadre de notre étude sur la production des fricatives.

Effets	Moments spectraux				
	Centre de gravité	Distance entre centres de gravité de /s/ et /ʃ/	Ecart-type	Skewness	Kurtosis
âge auditif (CI)	n.s.	n.s.	n.s.	n.s.	n.s.
âge à l'implantation (CI)	n.s.	n.s.	n.s.	n.s.	n.s.
âge auditif (CI)*Lieu d'articulation	n.s.	NA	f : augmentation (p<.05)	n.s.	n.s.
âge auditif (CI) * Contexte vocalique	/i/ : stabilité (n.s.) /u/ : augmentation avec /f/ & /s/	n.s. /u/ : augmentation (p<.001)	n.s.	n.s.	n.s.
âge à l'implantation (CI)*Lieu d'articulation	n.s.	NA	n.s.	n.s.	n.s.
âge à l'implantation (CI) * Contexte vocalique	n.s.	n.s.	n.s.	n.s.	n.s.

Table 5.101 : résumé des effets chez les enfants CI des variables « âge auditif » et « âge à l'implantation », seules et en interaction avec le groupe, le lieu d'articulation et contexte vocalique sur chaque moment spectral (« n.s. » indique une différence non significative)

Comme nous le constatons dans la table 5.101, l'âge auditif et l'âge à l'implantation n'affectent pas directement les quatre moments spectraux étudiés dans notre étude. L'âge à l'implantation n'a d'ailleurs aucun effet non plus sur ces moments spectraux en interaction avec d'autres variables.

Cependant, nous notons un effet de l'âge auditif dans certains contextes vocaliques : les centres de gravité des labiodentales et alvéolaires augmentent au fur et à mesure de l'utilisation de l'implant par les enfants CI : l'analyse des effets du groupe et du lieu d'articulation proposée au paragraphe §3.2 avait mis en évidence d'une part des centres de gravité plus bas chez les enfants CI et d'autre part un centre de gravité des labiodentales plus bas que celui des postalvéolaires pour les enfants CI, contrairement à ce qui était observé pour les enfants NH (centre de gravité des labiodentales plus élevé que celui des postalvéolaires). Il semble donc qu'une durée post-implantation de deux ans n'est pas suffisante pour que les enfants CI atteignent une production de fricatives similaire à celle des enfants NH, mais que plus la durée d'utilisation de l'implant par les enfants CI augmente, plus leurs fricatives tendent à s'approcher de celles des enfants NH.

3.6. Effet de l'input auditif immédiat

Nous avons choisi d'étudier l'effet de l'input auditif immédiat, c'est à dire l'influence d'un modèle audio sur la production des fricatives et nous utilisons pour cela deux tâches différentes : répétition de mots d'après un modèle adulte puis dénomination des mêmes mots sans ce modèle. Les figures présentées dans ce paragraphe présentent les résultats des deux groupes pour chaque moment spectral, comparés au modèle proposé lors de la tâche de répétition. Nous excluons par ailleurs les mots complexes produits lors de la tâche de répétition, qui seront analysés plus précisément dans le paragraphe suivant §3.7.

3.6.1. Centre de gravité

La figure 5.63 présente les centres de gravité pour chaque lieu d'articulation, groupe et tâche comparés avec ceux du modèle audio utilisé lors de la tâche de répétition : une première observation nous permet de dire que les centres de gravité des labiodentales et des alvéolaires sont plus proches chez les enfants NH que ceux des mêmes fricatives chez les enfants CI. En revanche, les centres de gravité semblent similaires dans les deux tâches, exception faite de ceux des labiodentales produites par les enfants CI pour qui les centres de gravité lors de la tâche de répétition sont plus proches du modèle que lors de la tâche de dénomination sans modèle audio.

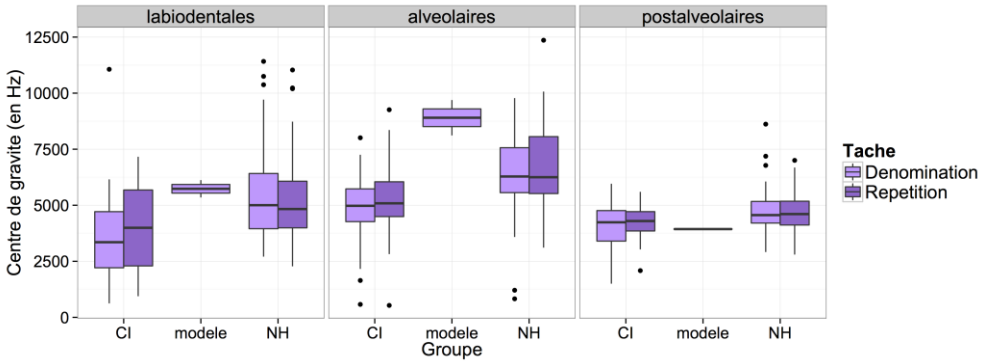


Figure 5.63 : centre de gravité (en Hz) par lieu d'articulation, par groupe et par tâche (répétition ou dénomination)

Les résultats du test statistique étudiant l'effet de la tâche (répétition vs dénomination) sur la réalisation du centre de gravité sont présentés dans la table 5.102 ci-dessous.

	numDF	denDF	F-value	p-value
(Intercept)	1	697	2010.5346	<.0001
C	2	697	89.7615	<.0001 ***
groupe	1	31	28.4875	<.0001 ***
tâche	1	697	0.3194	0.5721
C:groupe	2	697	11.0680	<.0001 ***
groupe:tâche	1	697	2.5905	0.1080

Table 5.102 : tableau de résultats au test statistique sur toutes les données (effets de groupe, lieu d'articulation, et tâche sur le centre de gravité)

Les résultats ne montrent pas d'effet de la tâche, ni seule (F(1,697)=0,32, p>.05), ni en interaction avec le groupe (F(1,697)=2,60, p>.05) : l'observation graphique initiale ne se traduit pas par un effet significatif de la tâche sur le centre de gravité.

3.6.2.Distance entre centres de gravité de /s/ et /ʃ/

La figure 5.64 ci-dessous présente la distance entre centres de gravité de /s/ et /ʃ/ pour chaque groupe et tâche, en la comparant avec le modèle présenté lors de la passation de l'expérience. Comme pour les lieux d'articulation, il semble que les enfants NH ont tendance à réaliser une distinction entre alvéolaires et postalvéolaires plus proche du modèle que les enfants CI.

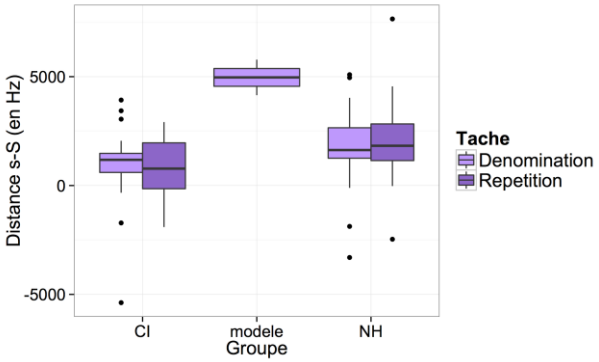


Figure 5.64 : distance moyenne (en Hz) entre les centres de gravité de /s/ et /ʃ/ par groupe et tâche (répétition ou dénomination)

Le test statistique présenté sur la table 5.103 ne permet pas de mettre en évidence un effet de la tâche, ni seule (F(1,92)=0,42, p>.05), ni en interaction avec le groupe (F(1,92)=0,41, p>.05).

	numDF	denDF	F-value	p-value
(Intercept)	1	92	79.37459	<.0001
groupe	1	31	8.42899	0.0067 **
tâche	1	92	0.41591	0.5206
groupe:tâche	1	92	0.41238	0.5224

Table 5.103 : tableau de résultats au test statistique sur toutes les données (effets de groupe et tâche sur la distance entre centres de gravité de /s/ et /ʃ/)

3.6.3.Ecart-type

La figure 5.65 présente les valeurs d'écart-type pour chaque groupe, lieu d'articulation et tâche, en les comparant au modèle audio utilisé lors de la tâche de répétition : l'observation visuelle nous laisse penser que les productions ne sont pas différentes d'une tâche à l'autre et que l'écart-type des labiodentales produites par les enfants NH est plus proche du modèle adulte que celui des enfants CI. En revanche, pour les alvéolaires et postalvéolaires, les écarts-types des deux groupes semblent éloignés du modèle de façon relativement similaire

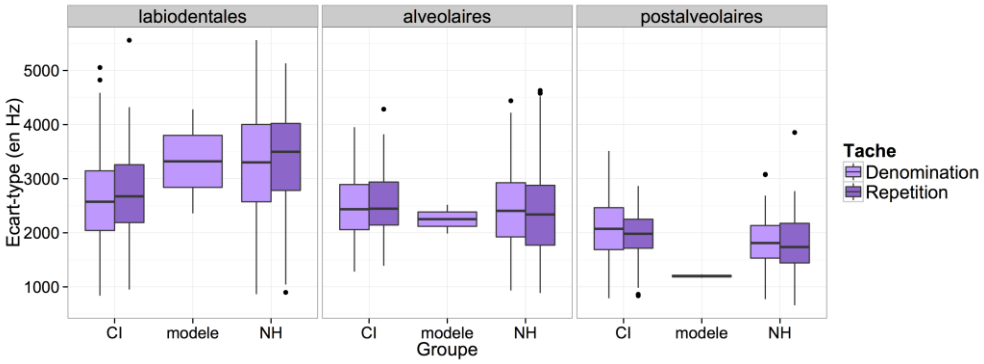


Figure 5.65 : écart-type (en Hz) par lieu d'articulation, par groupe et par tâche (répétition ou dénomination)

La procédure de sélection de variables effectuée lors de l'élaboration de notre modèle statistique a exclu la tâche du modèle, nous considérons donc que l'effet de la tâche sur l'écart-type n'a pas atteint le niveau de significativité fixé (p<.05) dans le cadre de notre étude.

3.6.4.Skewness

La figure 5.66 présente les valeurs de skewness pour chaque lieu d'articulation, pour chaque groupe et chaque tâche : les valeurs de skewness mesurées lors des deux tâches semblent relativement proches l'une de l'autres et du modèle adulte. Seules les valeurs des alvéolaires semblent différentes du modèle adulte.

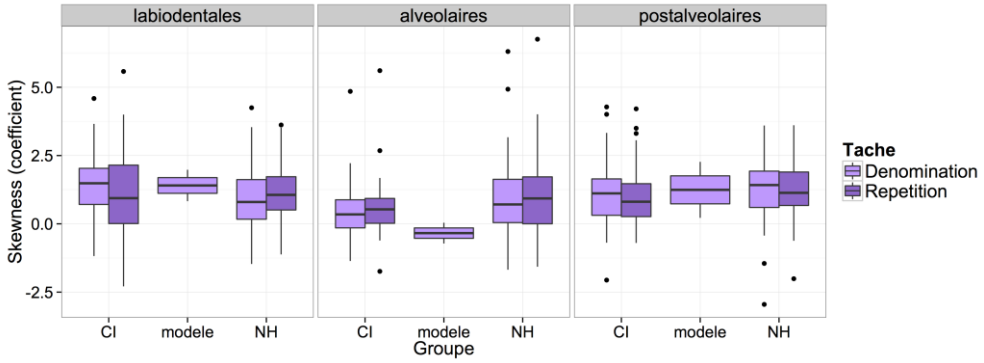


Figure 5.66 : skewness (coefficient) par lieu d'articulation, par groupe et par tâche (répétition ou dénomination)

Comme pour les valeurs d'écart-type, notre sélection des variables pertinentes pour la modélisation des données nous a conduit à exclure la tâche de l'analyse, et nous considérons que cet effet n'est pas significatif (p<.05) dans le cadre de notre étude.

3.6.5.Kurtosis

La Figure 5.67 présente les valeurs de kurtosis pour chaque lieu d'articulation, groupe et tâche, en les comparant au modèle adulte : une première observation de cette figure indique que le quatrième moment spectral n'est pas influencé par la tâche puisque les valeurs sont similaires dans les deux tâches pour chaque groupe et chaque lieu d'articulation. Ces valeurs sont également très proches du modèle, sauf pour les postalvéolaires, pour lesquelles le kurtosis du modèle est plus élevé que ceux des deux groupes.

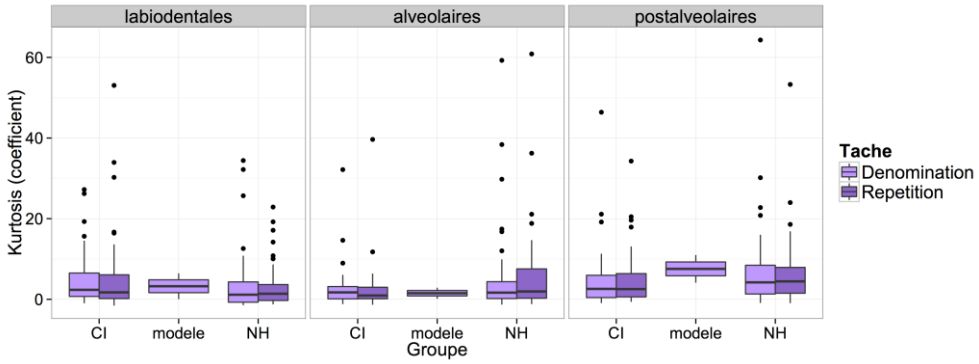


Figure 5.67 : kurtosis (coefficient) par lieu d'articulation, par groupe et par tâche (répétition ou dénomination)

A nouveau, à la suite de la procédure de sélection de variables, nous n'avons pas inclus la tâche dans notre modèle statistique, ce qui nous porte à dire que la tâche n'a pas un effet significatif sur le quatrième moment spectral ou kurtosis dans notre étude.

3.6.6. Résumé : effet de l'input auditif immédiat

La table 5.104 ci-dessous reprend tous les effets de la variable « tâche » étudiés dans ce paragraphe.

Effets	Moments spectraux				
	Centre de gravité	Distance entre centres de gravité de /s/ et /ʃ/	Ecart-type	Skewness	Kurtosis
Tâche	n.s.	n.s.	n.s.	n.s.	n.s.
Tâche*groupe	n.s.	n.s.	n.s.	n.s.	n.s.
Tâche*Lieu d'articulation	n.s.	NA	n.s.	n.s.	n.s.

Table 5.104 : résumé des effets de la variable tâche, seule et en interaction avec le groupe, le lieu d'articulation sur chaque moment spectral (« n.s. » indique une différence non significative)

Notre étude a clairement mis en évidence que les productions de fricatives ne sont absolument pas affectées par des tâches d'élicitation de mots différentes, ce qui nous laisse penser que ce type de consonnes reste identique, que l'enfant dispose d'un modèle audio de la fricative ou pas. Cet effet avait été testé dans l'optique de comprendre si la répétition immédiate d'un modèle adulte pouvait influencer la production des fricatives. Dans ce cas, les deux types de productions (avec ou sans modèle audio) semblent indiquer que le mode d'articulation et le contrôle articulatoire nécessaire à la production des fricatives sont stables chez chacun des groupes, même si des effets de l'implant sont observés : les enfants CI ont donc construit une représentation des fricatives en fonction des informations fournies par l'implant, cette représentation n'est pas tout à fait identique à celle des enfants NH qui ont un meilleur accès aux fréquences des sons de parole, mais elle est cependant déjà stable, deux ans après l'implantation.

3.7. Effet de la complexité du mot

Nous avons choisi d'étudier l'effet de la complexité du mot sur une production de mots simples (codés par « 0 » sur nos figures) et complexes (codés par « 1 » sur nos figures) dans la seule tâche de répétition. Nous excluons donc les mots produits lors de la tâche de dénomination, qui étaient tous simples, et sont analysés dans les paragraphes précédents.

3.7.1. Centre de gravité

La figure 5.68 ci-dessous présente les centres de gravité des fricatives (labiodentales, alvéolaires et postalvéolaires), par groupes et degrés de complexité lexicale. Une première observation

graphique semble indiquer que les centres de gravité des deux groupes et des trois lieux d'articulation ne sont pas affectés de la même façon par le degré de complexité du mot : les labiodentales des enfants CI semblent être les seules fricatives influencées par le degré de complexité du mot.

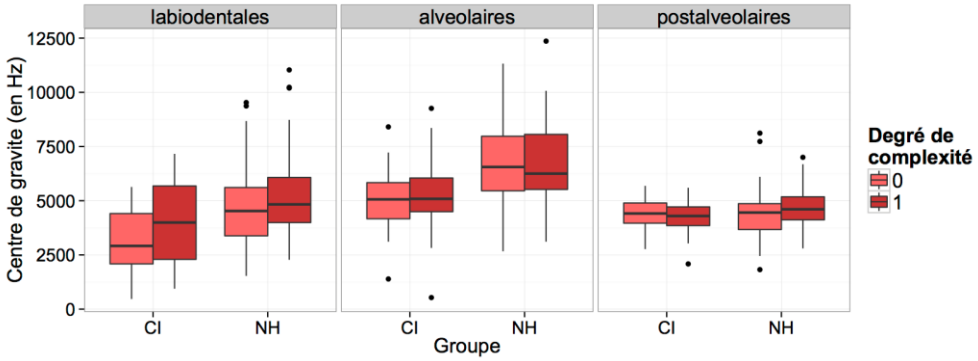


Figure 5.68 : centre de gravité (en Hz) par lieu d'articulation, par groupe et par degré de complexité du mot

La table 5.105 ci-dessus présente les résultats au test statistique correspondant, qui teste l'effet de la complexité du mot, seule ou en interaction avec le groupe et le lieu d'articulation sur le centre de gravité des fricatives.

	numDF	denDF	F-value	p-value
(Intercept)	1	694	2699.9544	<.0001
C	2	694	120.8504	<.0001 ***
groupe	1	31	31.6297	<.0001 ***
complexite_mot	1	694	4.5653	0.0330 *
C:groupe	2	694	17.4237	<.0001 ***
C:complexite_mot	2	694	4.4591	0.0119 *

Table 5.105 : tableau de résultats au test statistique sur toutes les données (effets de groupe, lieu d'articulation, et degré de complexité sur le centre de gravité)

Les résultats indiquent un effet du degré de complexité lexicale seule ($F(1,694)=4,57$, $p<.05$) et en interaction avec le lieu d'articulation ($F(2,694)=4,46$, $p<.05$), ce qui semble confirmer notre première observation visuelle. En effet, un test post-hoc (table 5.106, ci-dessous) le confirme : les fricatives labiodentales sont produites avec des centres de gravité différents selon le degré de complexité du mot, à la différence des alvéolaires et des postalvéolaires.

Consonne	Centre de gravité
	Complexité : mots simples (0) vs mots complexes (1)
/f/	z=-3,6, p<.001 ***
/s/	z=0,33, p>.05 (n.s.)
/ʃ/	z=-0,42, p>.05 (n.s.)

Table 5.106 : résumé du test post-hoc sur l'effet d'interaction entre complexité du mot et lieu d'articulation sur le centre de gravité

Des tests réalisés sur les données de chaque groupe indique en outre que les centres de gravité chez les enfants NH ne sont pas influencés par le degré de complexité lexicale ($F(1,439)=1,98$, $p>.05$), à l'inverse de ceux des enfants CI, pour lesquels nous observons un effet de l'interaction entre lieu d'articulation et complexité du mot ($F(2,252)=3,36$, $p<.05$). Un test post-hoc de cette

interaction chez les enfants CI indique que cet effet est localisé sur la labiodentale /f/ (z=-3,070, p<.01).

3.7.2. Distance entre centres de gravité de /s/ et /ʃ/

La figure 5.69 propose une représentation de la distance entre les centres de gravité de /s/ et /ʃ/ par groupe et degré de complexité lexicale. La distinction entre ces deux fricatives par les enfants NH semble influencée par le degré de complexité lexicale, à la différence des enfants CI.

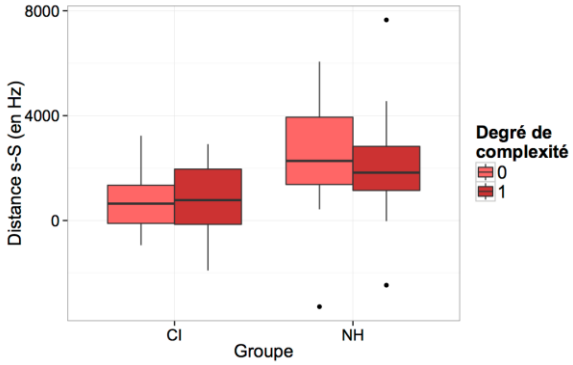


Figure 5.69 : distance moyenne (en Hz) entre les centres de gravité de /s/ et /ʃ/ par groupe et degré de complexité du mot

La table 5.107 ci-dessous propose le tableau des résultats au test statistique qui prend en compte l'effet de la complexité lexicale sur la distinction entre centres de gravité de /s/ et /ʃ/.

	numDF	denDF	F-value	p-value
(Intercept)	1	89	122.68530	<.0001
groupe	1	31	20.45498	0.0001 ***
complexite_mot	1	89	0.70857	0.4022
groupe:complexite_mot	1	89	1.12345	0.2920

Table 5.107 : tableau de résultats au test statistique sur toutes les données (effets de groupe et complexité lexicale sur la distance entre centres de gravité de /s/ et /ʃ/)

Les résultats indiquent que le degré de complexité lexicale n'a pas d'effet sur la distinction entre les centres de gravité de /s/ et /ʃ/, contrairement à ce que l'observation visuelle aurait pu laisser penser : les deux groupes d'enfants font la même distinction entre fricatives alvéolaires et postalvéolaires, que le mot soit simple ou complexe.

3.7.3. Ecart-type

La figure 5.70 ci-dessous présente les écarts-types des fricatives aux trois lieux d'articulation, par les deux groupes en fonction du degré de complexité lexicale. Une première observation nous indique que l'écart-type ne semble pas influencé par le degré de complexité lexicale.

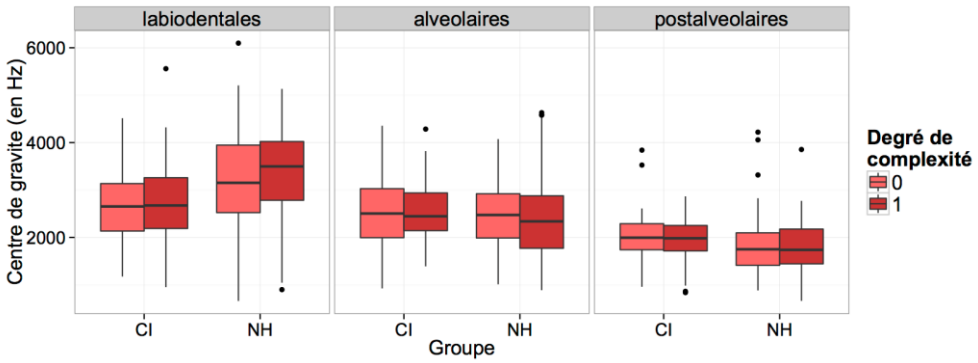


Figure 5.70 : écart-type (en Hz) par lieu d'articulation, par groupe et par degré de complexité du mot

Notre méthode de sélection des variables intégrées ensuite au modèle mixte comprenant effets fixes et effet aléatoire du sujet ne nous a pas conduite à incorporer le degré de complexité du mot au test statistique sur l'écart-type. Nous en concluons que ce degré de complexité du mot n'a pas d'effet significatif sur l'écart-type des fricatives produites par les deux groupes d'enfants NH et CI.

3.7.4.Skewness

La figure 5.71 ci-dessous présente les valeurs du troisième moment spectral ou skewness par les deux groupes d'enfants aux trois lieux d'articulation et pour chaque degré de complexité lexicale.

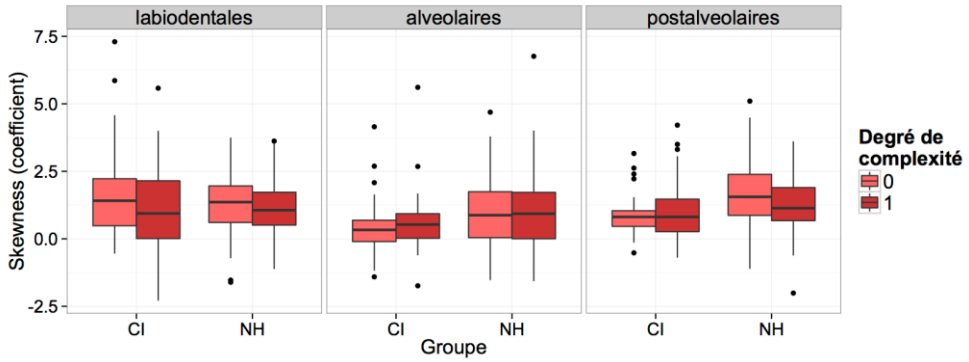


Figure 5.71 : skewness (coefficient) par lieu d'articulation, par groupe et par degré de complexité du mot

La table 5.108 ci-dessus présente les résultats au test statistique sur le troisième moment spectral ou skewness.

	numDF	denDF	F-value	p-value
(Intercept)	1	694	197.65761	<.0001
C	2	694	13.88726	<.0001 ***
groupe	1	31	2.44385	0.1281
complexite_mot	1	694	2.17372	0.1408
C:groupe	2	694	8.23952	0.0003 ***
C:complexite_mot	2	694	3.54933	0.0293 *

Table 5.108 : tableau de résultats au test statistique sur toutes les données (effets de groupe, lieu d'articulation et degré de complexité sur le troisième moment spectral ou skewness)

Le test statistique indique que le degré de complexité lexicale seul n'a pas d'effet sur le troisième moment spectral des fricatives produites par les deux groupes d'enfants ($F(1,694)=2,44$, $p>.05$). Cependant, l'interaction entre lieu d'articulation et degré de complexité lexicale a un effet significatif sur le troisième moment spectral ($F(2,694)=3,55$, $p<.05$) : l'observation de la figure

5.71 montre que pour les labiodentales, ce troisième moment spectral est plus élevé pour les labiodentales produites dans des mots simples, alors que pour les alvéolaires, il est similaire pour les deux degrés de complexité, et pour les postalvéolaires, il est similaire chez les enfants CI et plus élevé dans les mots simples chez les enfants NH. Un test post-hoc (table 5.109 ci-dessous) confirme l'observation visuelle pour les labiodentales : le degré de complexité a un effet significatif sur le troisième moment spectral des labiodentales, mais pas sur les alvéolaires, ni sur les postalvéolaires.

Consonne	Skewness
	Complexité : mots simples (0) vs mots complexes (1)
/f/	z=2,74, p<.05 *
/s/	z=-1,04, p>.05 (n.s.)
/ʃ/	z=0,82, p>.05 (n.s.)

Table 5.109 : résumé du test post-hoc sur l'effet d'interaction entre complexité du mot et lieu d'articulation sur le troisième moment spectral (skewness)

3.7.5. Kurtosis

La figure 5.72 ci-dessous présente le quatrième moment spectral ou kurtosis à chaque lieu d'articulation, pour chaque groupe et chaque degré de complexité lexicale.

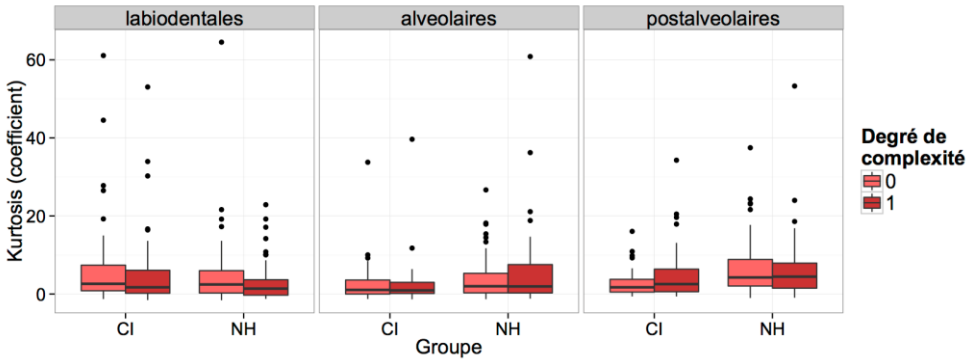


Figure 5.72 : kurtosis (coefficient) par lieu d'articulation, par groupe et par degré de complexité du mot

Notre méthode de sélection de variables intégrées ensuite au modèle mixte comprenant effets fixes et effet aléatoire du sujet ne nous a pas conduite à prendre en compte le degré de complexité du mot au test statistique sur le kurtosis. Nous en concluons que ce degré de complexité du mot n'a pas d'effet significatif sur le kurtosis des fricatives produites par les deux groupes d'enfants NH et CI.

3.7.6. Résumé : effet de la complexité du mot

La table 5.110 ci-dessous présente un résumé de l'effet de la variable « complexité lexicale » sur les différentes mesures de moments spectraux de notre étude sur les fricatives.

Effets	Moments spectraux				
	Centre de gravité	Distance entre centres de gravité de /s/ et /ʃ/	Ecart-type	Skewness	Kurtosis
complexité lexicale	0<1	n.s.	n.s.	n.s.	n.s.
complexité lexicale*groupe	NH : n.s. CI : 0<1	n.s.	n.s.	n.s.	n.s.
complexité lexicale*Lieu d'articulation	f : 0<1 s : n.s. ʃ : n.s.	NA	n.s.	f : 0<1 s : n.s. ʃ : n.s.	n.s.

Table 5.110 : résumé des effets de la variable « Complexité lexicale », seule et en interaction avec le groupe et le lieu d'articulation sur chaque moment spectral (« n.s. » indique une différence non significative)

Notre étude sur la production de fricatives par des enfants CI et NH a mis en évidence que le degré de complexité lexicale n'a que des effets très limités et localisés sur certains moments spectraux. Le centre de gravité est le seul moment spectral qui subit directement un effet global du degré de complexité lexicale : en effet, le centre de gravité des fricatives est plus bas lorsque le mot est simple que lorsqu'il est plus complexe. Cet effet est présent avant tout chez les enfants CI, et sur les labiodentales. Les enfants NH ne semblent pas subir cet effet, de même que les alvéolaires et postalvéolaires. La distance entre les centres de gravité de ces deux fricatives n'est d'ailleurs pas influencée par cette variable qui n'affecte pas non plus l'écart-type et le quatrième moment spectral (kurtosis). Le troisième moment spectral (skewness) des seules alvéolaires subit un effet de la complexité lexicale : les mesures de ce moment sont plus basses lorsque le mot est simple, ce qui traduit une concentration d'énergie de la fricative au dessus du centre de gravité dans ce cas.

Il apparaît donc que les enfants des deux groupes produisent des fricatives sensiblement identiques avec des mots plus ou moins complexes et que seules les labiodentales subissent un effet de la complexité lexicale.

3.8. Discussion

La table 5.111 ci-dessous reprend les effets de toutes les variables étudiés dans le cadre de cette étude de production de fricatives par des enfants NH et CI.

Effets	Moments spectraux				
	Centre de gravité	Distance entre centres de gravité de /s/ et /ʃ/	Ecart-type	Skewness	Kurtosis
Groupe	CI<NH	CI<NH	n.s.	n.s.	n.s.
Lieu d'articulation	s>ʃ>f	NA	f>s>ʃ	f≈ʃ>s	ʃ>f≈s
Groupe*Lieu d'articulation	NH : s>f>ʃ	NA	NH : f≈s>ʃ	NH : ʃ>f>s	NH : f≈s>ʃ
	CI : s>ʃ>f	NA	CI : f>s>ʃ	CI : f>ʃ>s	CI : f≈ʃ et f>s
Contexte vocalique	/i/ > /u/	/i/ > /u/	/i/ < /u/	/i/ > /u/	/i/ > /u/
Contexte vocalique * Groupe	NH : /i/ > /u/	n.s.	n.s.	NH : /i/ > /u/	NH : /i/ > /u/
	CI : /i/ > /u/		CI : n.s.	CI : n.s.	CI : n.s.
Contexte vocalique * Lieu d'articulation	f : n.s.	NA	f : /i/ < /u/	f : /i/ > /u/	f : /i/ > /u/
	s : /i/ > /u/		s : /i/ > /u/	s : /i/ < /u/	s : n.s.
	ʃ : n.s.		ʃ : /i/ < /u/	ʃ : /i/ > /u/	ʃ : /i/ > /u/
âge chronologique (NH)	n.s.	n.s.	n.s.	n.s.	n.s.
âge chronologique (NH)*Lieu d'articulation	n.s.	NA	n.s.	n.s.	n.s.
âge chronologique (NH) * Contexte vocalique	/i/ : augmentation (n.s.)	/i/ : augmentation (p<.05)	n.s.	n.s.	n.s.
	/u/ : diminution (n.s.)	/u/ : stabilité (n.s.)			
âge chronologique (NH) * Lieu d'articulation * Contexte vocalique	n.s.	NA	n.s.	f & ʃ : augmentation (n.s.)	n.s.
				s : augmentation avec i (n.s)	
				s : diminution avec u (n.s)	
âge auditif (CI)	n.s.	n.s.	n.s.	n.s.	n.s.
âge à l'implantation (CI)	n.s.	n.s.	n.s.	n.s.	n.s.
âge auditif (CI)*Lieu d'articulation	n.s.	NA	f : augmentation (p<.05)	n.s.	n.s.
âge auditif (CI) * Contexte vocalique	/i/ : stabilité (n.s.)	n.s.	n.s.	n.s.	n.s.
	/u/ : augmentation avec /f/ & /s/	/u/ : augmentation (p<.001)			
âge à l'implantation (CI)*Lieu d'articulation	n.s.	NA	n.s.	n.s.	n.s.
âge à l'implantation (CI) * Contexte vocalique	n.s.	n.s.	n.s.	n.s.	n.s.
Tâche	n.s.	n.s.	n.s.	n.s.	n.s.
Tâche*groupe	n.s.	n.s.	n.s.	n.s.	n.s.
Tâche*Lieu d'articulation	n.s.	NA	n.s.	n.s.	n.s.
complexité	0<1	n.s.	n.s.	n.s.	n.s.
complexité*groupe	NH : n.s.	n.s.	n.s.	n.s.	n.s.
	CI : 0<1				
complexité*Lieu d'articulation	f : 0<1	NA	n.s.	f : 0<1	n.s.
	s : n.s.			s : n.s.	
	ʃ : n.s.			ʃ : n.s.	

Table 5.111 : résumé de tous les effets étudiés dans le cadre de notre étude de production de fricatives par les enfants CI et NH (« n.s. » indique une différence non significative)

3.8.1.Production de fricatives par les enfants NH

Il ressort de nos analyses que les enfants NH réalisent des fricatives avec des centres de gravité similaires à ceux d'enfants et d'adultes NH décrits dans la littérature (centre de gravité de /s/ supérieur à celui de /f/ puis de /ʃ/, notamment chez des adultes francophones (Calliope, 1989)) : l'étude de Nissen & Fox (2005) d'une part mettait en évidence chez des enfants de trois à cinq ans de fricatives que les centres de gravité de /s/ et /ʃ/ devenaient de plus en plus distincts pour s'approcher de la production d'adultes, et observait une diminution forte à partir de quatre ans du centre de gravité de /ʃ/, et celle de Nittrouer (1995), d'autre part, faisait état de la même tendance chez des enfants de trois à sept ans. Nos résultats vont dans ce sens, puisque les enfants NH de

notre étude sont légèrement plus âgés que ceux de l'étude de Nittrouer (1995) mais ils produisent des fricatives avec des centres de gravité (caractérisés notamment par la distinction entre centres de gravité de /s/ et /ʃ/) similaires à celles d'adultes NH.

En revanche, Nissen & Fox (2005) montraient une distinction pour l'écart-type entre fricatives sibilantes /s/ et /ʃ/ et non-sibilantes /f/ et /θ/ que nos résultats contredisent : en effet, alors que dans l'étude de Nissen & Fox (2005) l'écart-type de /f/ et /θ/ était supérieur à celui de /s/ et /ʃ/ (ce qui signifie que /f/ et /θ/ sont plus diffuses que /s/ et /ʃ/), nos résultats pour les enfants NH indiquent que /f/ et /s/ ont des écarts-types similaires, et supérieurs à celui de /ʃ/ (ce qui signifie que /f/ et /s/ sont plus diffuses que /ʃ/). Nos résultats sont en revanche en accord avec ceux de Li et al. (2009) chez des enfants anglophones de deux et trois ans puisque cette étude met en évidence une différence d'écart-type entre /s/ et /ʃ/. Les troisième et quatrième moments spectraux des enfants NH de notre étude sont par ailleurs similaires à ceux de l'étude de Nissen & Fox (2005) et de Nittrouer (1995) qui étaient à notre connaissance les seules à en proposer l'analyse. Li et al. (2009) ne proposent pas d'analyse de ces moments spectraux. L'anglais est la langue des études de Nissen & Fox (2005), Nittrouer (1995) et Li et al. (2009), et la fricative /s/ de l'anglais est décrite comme une alvéolaire à la différence de celle du français qui est décrite comme une alvéo-dentale dans la base UPSID (UCLA Phonological Segment Inventory Database, 2014). Cette variation pourrait peut-être expliquer certaines différences avec la littérature que nous observons, en particulier pour l'écart-type qui distingue fricatives sibilantes et non-sibilantes dans l'étude de Nissen & Fox (2005), et /f/ et /s/ de /ʃ/ dans notre étude.

Nos résultats chez les enfants NH mettent par ailleurs en lumière un effet du contexte vocalique sur chacun des moments spectraux que nous avons étudiés. Cet effet est également visible sur le centre de gravité en interaction avec l'âge chronologique. Nos résultats sont une confirmation de ceux de l'étude de Körkkö (2015), qui montre que le centre de gravité de /s/ est plus haut lorsque la fricative est suivie de /i/ que de /u/. Nous observons par ailleurs que les centres de gravité des labiodentales et des postalvéolaires ne sont pas influencés par ce contexte vocalique. En revanche, les trois autres moments spectraux le sont : les labiodentales et postalvéolaires sont plus compactes lorsqu'elles sont suivies de /i/ et les alvéolaires de /u/ ; les labiodentales et postalvéolaires sont plus concentrées dans les fréquences inférieures au centre de gravité lorsqu'elles sont suivies de /i/ et les alvéolaires de /u/ (qui correspond à un skewness plus haut dans ces deux cas) ; enfin, les labiodentales et postalvéolaires ont une amplitude plus forte lorsqu'elles sont suivies de /i/ (kurtosis plus haut dans ce cas), et aucune différence n'est observée pour les alvéolaires. L'étude de Körkkö (2015), réalisée auprès d'enfants de six, neuf et onze ans, locuteurs du finnois met en évidence l'effet de la coarticulation anticipatoire : l'alvéolaire /s/ est influencée par la voyelle qui suit immédiatement la fricative. Nos résultats indiquent le même effet de coarticulation anticipatoire, même si la même fricative est décrite comme une alvéolaire en finnois et une alvéo-dentale en français dans la base de données UPSID (UCLA Phonological Segment Inventory Database, 2014) : la fricative alvéolaire ou alvéodentale a des caractéristiques qui se rapprochent de celles de la postalvéolaire /ʃ/ lorsque celle-ci est suivie de la voyelle /u/.

L'âge chronologique des enfants NH de notre étude n'a pas d'impact sur la production des fricatives (pour les quatre moments spectraux), ce qui confirme également les résultats de l'étude

de Nittrouer (1995), dans laquelle une différence inter-groupes entre adultes et enfants était observée, mais pas de différence intra-groupe, et ce pour les quatre moments spectraux. Li et al. (2009) et Nissen & Fox (2005) n'étudiaient quant à eux pas d'effet de l'âge chronologique sur les moments spectraux.

Enfin, il ressort de ces résultats que les deux dernières variables (tâche et degré de complexité lexicale) n'ont aucun effet sur la production de fricatives par les enfants NH : les fricatives produites par ce groupe d'enfants ne sont pas sensibles à la tâche proposée, et qu'elles ne sont pas influencées par la complexité du mot. L'absence de ces deux effets nous amène donc à penser que leur représentation est déjà stable chez des enfants normo-entendants de 5;7 ans à 10;6 ans.

La suite de cette discussion va s'intéresser aux fricatives produites par les enfants CI pour essayer d'en dégager les caractéristiques.

3.8.2. Production de fricatives par les enfants CI

Notre étude a mis en évidence une réalisation différente des fricatives des enfants CI par rapport aux enfants NH, appariés en âge chronologique. Cette différence est avant tout liée au premier moment spectral, ou centre de gravité, qui est en effet plus bas chez les enfants CI de notre étude pour chaque lieu d'articulation, et se manifeste également par une distinction plus faible entre alvéolaires et postalvéolaires chez les enfants CI. Ces différences entre les deux groupes d'enfants sont conformes aux résultats de l'étude de Todd et al (2011), dans laquelle les valeurs du premier moment spectral de /s/ des enfants CI étaient plus basses que celles des enfants NH, appariés en âge chronologique et les valeurs du centre de gravité de /s/ et /ʃ/ des enfants CI étaient plus basses que celles des enfants NH, appariés en âge auditif. Les valeurs du centre de gravité de /ʃ/ étaient similaires chez les enfants CI et les enfants NH appariés en âge chronologique. Notre étude ne permet malheureusement pas de comparer les fricatives des enfants CI avec celles d'un groupe d'enfants appariés en âge auditif, puisque notre groupe d'enfants CI a un âge auditif compris entre 2;2 ans et 9;1 ans (âge auditif moyen : 5;3 ans), alors que l'enfant NH le plus jeune de notre étude a 5;7 ans (âge chronologique et auditif moyen du groupe NH : 7;2 ans). En revanche, nos résultats sont en contradiction avec l'étude de Mildner & Liker (2008), dans laquelle les centres de gravité des alvéolaires et postalvéolaires du groupe CI sont plus bas que ceux du groupe NH lors des trois premières sessions d'enregistrement, puis plus hauts lors de la quatrième. La différence entre /s/ et /ʃ/ de notre étude est également plus marquée chez les enfants CI que dans celle de Mildner & Liker (2008), qui met en évidence des productions d'alvéolaires et de fricatives aux centres de gravité très proches. Enfin, nos résultats (centres de gravité plus bas chez les enfants CI que chez les enfants NH) diffèrent de ceux de l'étude de Uchanski & Geers (2003) qui montraient une réalisation de fricatives similaires chez les enfants CI et NH. Cependant, comme Uchanski et Geers (2003), nous observons un centre de gravité des alvéolaires plus élevé que celui des postalvéolaires. Les études de Uchanski & Geers (2003) et Todd et al. (2011) sont réalisées auprès d'enfants anglophones, celles de Mildner & Liker (2008) auprès d'enfants locuteurs du croate. La fricative /s/ du croate, comme celle de l'anglais est décrite comme une alvéolaire dans la base de données UPSID, à la différence du /s/ du français, décrit comme une alvéodentale. Comme pour les enfants NH, certains résultats des enfants CI de notre étude pourraient peut-être s'expliquer par une variation liée à la langue de notre étude, notamment au niveau du centre de gravité et de l'écart-type.

Nous observons également des différences significatives entre les deux groupes au niveau des deuxième (les labiodentales sont plus compactes chez les enfants CI), troisième (les alvéolaires ont un coefficient de dissymétrie plus élevé chez les enfants CI) et quatrième moments spectraux (les labiodentales ont un kurtosis plus élevé chez les enfants CI). Ces résultats diffèrent de ceux de Uchanski & Geers (2003) qui observaient des troisième et quatrième moments spectraux similaires chez les enfants CI et NH. Mildner & Liker (2008) et Todd et al. (2011) ne proposaient pas d'analyses de ces trois moments spectraux.

Nos résultats montrent par ailleurs un effet du contexte vocalique similaire chez les deux groupes d'enfants pour les deux premiers moments spectraux. En effet, comme les enfants NH, les enfants CI produisent des alvéolaires au centre de gravité plus élevé lorsqu'elles sont suivies de /i/ plutôt que de /u/, ce qui semble indiquer que les enfants CI sont sensibles au même effet de coarticulation que les enfants NH pour la production des alvéolaires. Cependant, la différence moyenne entre les centres de gravité de /s/ et /ʃ/ en contexte /i/ et /u/ est plus marquée chez les enfants NH : en effet, /s/ est plus variable chez les enfants CI qui produisent cette fricative avec un centre de gravité proche de celui de /ʃ/, lorsque ces fricatives sont suivies de /u/. Pour les troisième et quatrième moments spectraux, nous n'observons pas d'effet du contexte vocalique chez les enfants CI, à la différence des enfants NH.

Les études de Uchanski & Geers (2003), Todd et al. (2011) et Mildner & Liker (2008) ne prenaient en compte ni l'effet de l'âge à l'implantation, ni celui de la durée d'utilisation de l'implant (ou âge auditif). Mildner & Liker (2008) indiquaient seulement qu'à partir de 34 mois après le début de leur étude, les productions de fricatives avaient tendance à se stabiliser, ce qui était visible notamment au niveau de la distinction entre alvéolaires et postalvéolaires. Cette étude longitudinale était réalisée sur une période de 46 mois, mais commençait pour chaque enfant à un âge auditif différent. Notre étude met en évidence que l'âge à l'implantation n'a pas d'effet sur la production des fricatives (pour les quatre moments spectraux) alors que l'âge auditif affecte quelques caractéristiques des fricatives produites par les enfants CI, sous l'influence du contexte vocalique : le centre de gravité des labiodentales et alvéolaires, et la distance entre /s/ et /ʃ/ augmentent avec l'âge auditif lorsque les fricatives sont suivies de /u/, et l'écart-type des labiodentales augmente avec l'âge auditif.

Comme ce que nous observions chez les enfants NH, les fricatives produites par les enfants CI ne sont pas influencées par des tâches différentes de dénomination de mots, ce qui nous laisse penser que plus de deux ans après implantation, les enfants CI ont déjà pu construire une représentation stable des différentes catégories de fricatives, indépendante du modèle audio proposé, même si celle-ci continue à évoluer légèrement pour se rapprocher de celle des enfants NH. Cette production similaire de fricatives lors de deux tâches différentes caractérisées par une mise à disposition ou non d'un modèle audio peut être comparée aux résultats de l'étude de Bharadwaj et al. (2006) qui obtenaient des résultats similaires pour les quatre moments spectraux, pour trois enfants sur les quatre enfants CI de leur étude, tous âgés de 7 à 10;9 ans avec implant activé et désactivé : dans cette étude, l'accès à un feedback influençait plus les enfants que les adultes, mais la distinction entre alvéolaires et postalvéolaires n'était que marginalement affectée par cette absence de feedback auditif.

Cette stabilisation des catégories de fricatives chez les enfants CI de notre étude semble confirmée par le fait que seul un lieu d'articulation (labiodentales) est influencé par le degré de

complexité du mot. En effet, cet effet très localisé (sur le centre de gravité et le coefficient de dissymétrie) nous laisse penser que les productions des enfants CI ne sont que peu différentes lorsque le mot est moins fréquent ou possède des propriétés phonotactiques plus complexes.

Ces résultats sur les fricatives mettent en évidence des difficultés persistantes de production des enfants CI qui n'ont pas accès à la même information fréquentielle que les enfants NH, du fait de l'utilisation de l'implant cochléaire. Ils contrastent avec ceux de notre étude sur la production d'occlusives puisque les indices de durée nécessaires à la mise en place de la coordination de gestes articulatoires qui permettent de réaliser le contraste de voisement sont relativement bien utilisés par les enfants CI, pour produire des consonnes occlusives proches de celles des enfants NH.

Comme nous l'avons évoqué précédemment (§3, ce chapitre), le VOT peut être utilisé pour caractériser l'évolution du degré de coarticulation anticipatoire de séquences occlusive-voyelle chez les enfants. Nous pouvons faire le même constat pour les quatre moments spectraux utilisés pour caractériser la production des fricatives dans cette étude : le contexte vocalique joue un rôle important sur la production des fricatives et permet d'attester d'une évolution de cette production au cours du temps. Les moments spectraux (centre de gravité, écart-type, skewness et kurtosis) sont donc de bons indicateurs du degré de coarticulation anticipatoire. Ils permettent entre autres de caractériser la taille de la cavité de résonance devant la langue pour les alvéolaires et postalvéolaires ou l'amplitude de la turbulence créée au niveau des dents et des lèvres pour les labiodentales.

Dans la suite de ce chapitre, nous allons présenter les résultats de notre étude sur la réalisation de la coarticulation, qui nous permet de caractériser la coordination des gestes articulatoires lors de l'enchainement de segments.

4. Coarticulation

4.1. Données analysées

Pour cette étude de coarticulation dans des séquences CV, nous avons fait le choix d'enregistrer un corpus de locuteurs adultes, de façon à pouvoir comprendre dans quelle mesure les enfants CI et NH ont acquis des caractéristiques de coarticulation similaires à celles des adultes. En effet, la revue de littérature sur l'équation de locus présentée au chapitre 4 a mis en évidence que la transition CV est acquise tardivement par les enfants (par exemple Noiray et al., 2013), et que la mesure d'équation de locus subit une influence de la tâche utilisée pour recueillir la parole (lecture ou parole spontanée dans l'étude de Duez, 1992 et production de mots isolés ou de parole spontanée dans l'étude de Krull, 1989). Nous avons donc enregistré, selon la procédure décrite au chapitre 4 pour les deux groupes d'enfants, huit locuteurs adultes : trois hommes (âge moyen : 33 ans (12;1)) et cinq femmes (âge moyen : 30;9 ans (5;8)), tous locuteurs francophones monolingues, résidant à Grenoble au moment des enregistrements. Ce corpus de parole adulte nous permettra donc de pouvoir localiser chez les deux groupes d'enfants la variation liée à l'âge et à une acquisition non stabilisée, et la variation liée à l'utilisation de l'implant cochléaire.

Comme pour les voyelles et les consonnes, nous basons notre étude de la coarticulation sur des productions exploitables : nous excluons donc des analyses les mots non produits ou non identifiables. 99% des mots produits par les enfants NH et 96,4% des mots produits par les enfants CI ont pu être inclus dans ces analyses : un test t (test de Welch) met en évidence une absence de différence significative entre les deux groupes ($t(13,995)=1,762$, $p>.05$). Parmi ces productions de mots exploitables, nous avons exclu toutes les productions de voyelles non-exploitables (car dévoisées, creaky, élidées…) : nos analyses sont donc réalisées à partir de 2395 voyelles exploitables. 73,1% des voyelles produites par les enfants NH et 72,9% des voyelles produites par les enfants CI sont donc incluses dans ces analyses : un test t (test de Welch) indique que la différence entre le nombre de voyelles incluses dans les analyses pour chaque groupe d'enfants n'est pas significative ($t(1,25,248)=0,0199$, $p>.05$).

Le compte détaillé des productions incluses dans notre étude sur la coarticulation est donné dans la table 5.112 ci-dessous. A partir des mesures acoustiques réalisées sur les voyelles des séquences CV, nous avons calculé pour chaque sujet et chaque tâche (dénomination et répétition) une équation de locus selon la procédure décrite précédemment (paragraphe §2.3.1, chapitre 4).

Groupe	Consonne	Nombre total de voyelles	Nombre de voyelles par sujet (max=20)	Pente moyenne de l'équation de locus (coefficient)
AD	b	160	20	0,73 (0,15)
AD	d	160	20	0,35 (0,18)
AD	g	160	20	1,18 (0,19)
CI	b	248	19,08	0,65 (0,22)
CI	d	247	19,00	0,54 (0,27)
CI	g	232	17,85	1,02 (0,35)
NH	b	400	20,00	0,75 (0,13)
NH	d	395	19,75	0,61 (0,15)
NH	g	393	19,65	1,13 (0,16)

Table 5.112 : valeurs moyennes des pentes d'équation de locus (avec écart-type) par groupe et par lieu d'articulation

Les pentes par groupe et lieu d'articulation sont également détaillées dans la table 5.112 ci-dessus. Comme pour l'analyse des voyelles et des consonnes, nous utilisons des modèles linéaires mixtes qui prennent en compte à la fois des effets fixes de certaines variables, et des effets aléatoires d'autres variables du sujet. Toutes les variables (dépendantes, indépendantes et ayant un effet aléatoire) utilisées pour ces analyses sont détaillées dans la table 5.113 ci-dessous.

Variables dépendantes	Pente de l'équation de locus (coefficient)
Variables indépendantes	Groupe : CI, NH, AD
	C : Lieu d'articulation : bilabial /b/, alvéolaire /d/ et vélaire /g/
	Age chronologique des enfants (NH)
	Age à l'implantation (enfants CI)
	Age auditif (qui correspond à la durée d'utilisation de l'implant, pour les enfants CI)
	Tâche : tâche de répétition (R) ou tâche de dénomination (D)
Variable ayant un effet aléatoire	Sujet

Table 5.113 : variables dépendantes, variables indépendantes et variable ayant un effet aléatoire, pris en compte dans l'analyse des équations de locus

4.2. Effets du groupe et du lieu d'articulation

La figure 5.73 ci-dessous présente les pentes d'équation de locus de chaque groupe, à chaque lieu d'articulation : pour les bilabiales et les vélaires, les pentes chez les adultes et les enfants NH semblent similaires, alors que celles des enfants CI semblent plus basses (ce qui serait le signe d'une plus faible coarticulation chez les enfants CI) et pour les alvéolaires, les deux groupes d'enfants semblent avoir des pentes d'équation de locus similaires, et plus élevées que celles des adultes (ce qui indiquerait une plus grande coarticulation chez les deux groupes d'enfants, par rapport aux adultes).

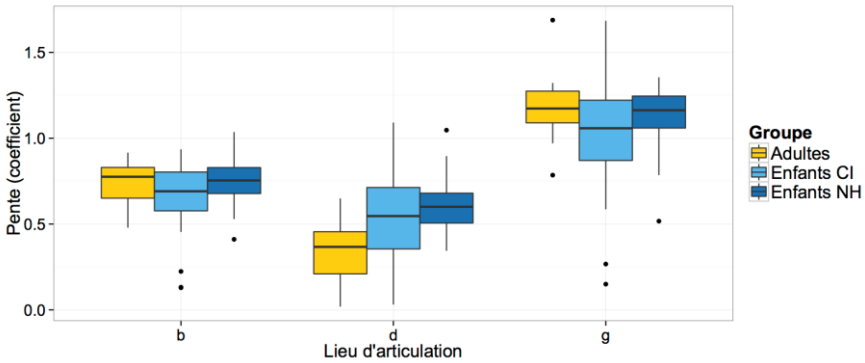


Figure 5.73 : pentes d'équation de locus (coefficient) par groupe et lieu d'articulation

Dans un premier temps, nous avons testé l'effet du lieu d'articulation et du groupe sur la réalisation de la pente de l'équation de locus (table 5.114) chez les trois groupes (adultes, enfants NH et enfants CI). Les résultats indiquent un effet du lieu d'articulation seul ($F(2, 199)=203{,}70$, $p<.001$), et de l'interaction entre lieu d'articulation et groupe ($F(4,199)=5{,}59$, $p<.001$) qui confirment la première observation graphique des données.

	numDF	denDF	F-value	p-value
(Intercept)	1	199	1916.5049	<.0001
C	2	199	203.7003	<.0001 ***
groupe	2	38	2.8274	0.0717
C:groupe	4	199	5.5891	0.0003 ***

Table 5.114 : tableau de résultats au test statistique sur toutes les données, adultes et enfants (effets de groupe et lieu d'articulation sur la pente de l'équation de locus)

Pour affiner ces résultats, nous avons réalisé un test post-hoc de comparaisons multiples, pour mieux comprendre où est localisé cet effet d'interaction. Ce test, dont les résultats sont résumés dans la table 5.115 ci-dessous, nous permet notamment de comparer les résultats de chaque groupe d'enfant à ceux du groupe d'adultes.

Groupes	Pentes d'équation de locus		
	Consonne		
	/b/	/d/	/g/
AD vs CI	z=1,11, p>.05 (n.s.)	z=-2,79, p<.05 *	z=2,36, p>.05 (n.s.)
AD vs NH	z=-0,30, p>.05 (n.s.)	z=-3,96, p<.001 ***	z=0,76, p>.05 (n.s.)
CI vs NH	z=-1,75, p>.05 (n.s.)	z=-1,12, p>.05 (n.s.)	z=-2,08, p>.05 (n.s.)

Table 5.115 : résumé du test post-hoc testant l'effet inter-groupe par lieu d'articulation sur la pente d'équation de locus

Les tests inter-groupes indiquent que les deux groupes d'enfants ont des réalisations similaires, et différentes de celles des adultes pour les alvéolaires, qui est le seul lieu d'articulation pour lequel les différences entre les pentes des adultes et des enfants NH d'une part et celles des adultes et des enfants CI d'autre part sont significatives. Nous avons également effectué un test post-hoc inter-lieux pour comprendre si les patrons sont les mêmes pour les deux groupes d'enfants et le groupe d'adultes. Les résultats de ce test sont présentés dans la table 5.116 ci-dessous.

Groupe	Pentes d'équation de locus		
	Paire de consonnes testée		
	/b/ - /d/	/b/ - /g/	/d/ - /g/
AD	z=5,91, p<.001 ***	z=-7,04, p<.001 ***	z=-12,95, p<.001 ***
CI	z=2,16, p>.05 (n.s.)	z=-7,27, p<.001 ***	z=-9,43, p<.001 ***
NH	z=3,53, p<.01 *	z=-9,46, p<.001 ***	z=-12,99, p<.001 ***

Table 5.116 : résumé du test post-hoc testant l'effet du lieu d'articulation par groupe sur la pente d'équation de locus

Le test statistique nous indique que les enfants NH ont une réalisation de la transition de F2 similaire à celle des adultes : la pente d'équation de locus est plus élevée pour les vélaires, suivies des bilabiales, puis des alvéolaires. En revanche, la réalisation de la transition de F2 chez les enfants CI est différente de celle des enfants NH et des adultes : la pente d'équation de locus est plus élevée pour les vélaires, puis pour les bilabiales et alvéolaires, mais ces deux derniers lieux d'articulation ont une pente dont la différence n'est pas significative, à la différence de ce qui est observé pour le groupe d'enfants NH et le groupe d'adultes.

4.2.1. Résumé : effets du groupe et du lieu d'articulation

Les résultats des tests de l'effet du groupe sur la pente de locus sont résumés dans la table 5.117 ci-dessous.

Effets	Pentes de l'équation de locus
Groupe	n.s.
Lieu d'articulation	AD : /g/ > /b/ > /d/
	NH : /g/ > /b/ > /d/
	CI : /g/ > /b/ ≃ /d/
Groupe*Lieu d'articulation	/b/ : n.s.
	/d/ : NH/CI > AD
	/g/ : n.s.

Table 5.117 : résumé des effets des variables « Groupe » et « Lieu d'articulation », seules et en interaction sur la pente de l'équation de locus (« n.s. » et ≃ indiquent une absence de différence significative)

Les analyses réalisées sur la pente de l'équation de locus chez les trois groupes (adultes, enfants NH et enfants CI) ont mis en évidence une différence de réalisation des transitions CV entre adultes et enfants, en particulier au niveau des alvéolaires qui sont réalisées avec un degré plus important de coarticulation pour les enfants. Des analyses supplémentaires de l'effet de l'âge chronologique chez les enfants NH vont permettre d'affiner cette analyse, de façon à comprendre si cet effet évolue chez ce groupe d'enfants de 5;7 à 10;6 ans vers une coarticulation similaire à celle des adultes ou s'il témoigne d'une acquisition plus tardive de la coarticulation de séquences alvéolaire-voyelle.

4.3. Enfants NH : effet de l'âge chronologique

La figure 5.74 ci-dessous présente l'évolution de la pente d'équation de locus en fonction de l'âge chronologique chez les enfants NH. Nous observons une légère diminution de cette pente pour les alvéolaires, et une stabilité pour les bilabiales et vélaires.

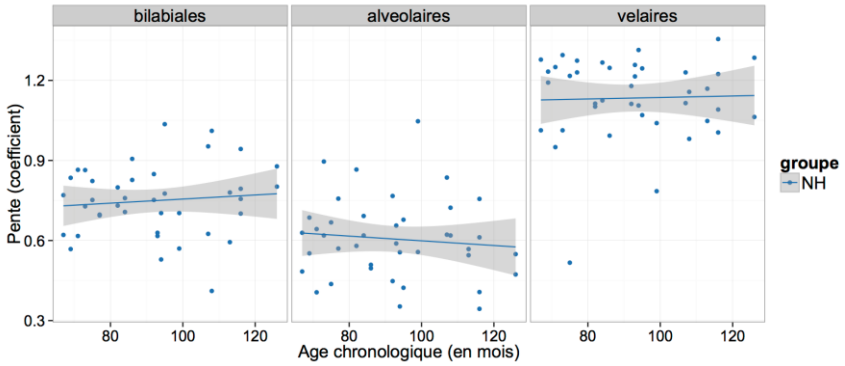


Figure 5.74 : pente de l'équation de locus (coefficient) en fonction de l'âge chronologique (en mois) par lieu d'articulation, pour les enfants NH

Les résultats du test statistique de l'effet de l'âge chronologique sur la pente d'équation de locus sont présentés dans la table 5.118 ci-dessous.

	numDF	denDF	F-value	p-value
(Intercept)	1	96	3758.887	<.0001
C	2	96	135.646	<.0001 ***
âge_chronologique	1	18	0.005	0.9449
C:âge_chronologique	2	96	0.388	0.6792

Table 5.118 : tableau de résultats au test statistique sur les données des enfants NH (effets de l'âge chronologique, du lieu d'articulation sur la pente d'équation de locus)

Le test statistique indique que l'âge chronologique n'a pas d'effet significatif sur la réalisation de la pente d'équation de locus, ni seul ($F(1,18)=0,01$, $p>.05$) ni en interaction avec le lieu d'articulation ($F(2,96)=0,38$, $p>.05$).

4.3.1. Résumé : effet de l'âge chronologique (enfants NH)

La table 5.119 ci-dessous propose un résumé des résultats obtenus dans ce paragraphe, de l'effet de l'âge chronologique sur la pente d'équation de locus chez les enfants NH.

Effets	Pentes de l'équation de locus
âge chronologique (NH)	n.s.
âge chronologique (NH)*Lieu d'articulation	n.s.

Table 5.119 : résumé des effets de l'âge chronologique seul et en interaction avec le lieu d'articulation sur la pente de l'équation de locus

4.4. Enfants CI : effet de l'âge à l'implantation et de la durée d'utilisation de l'implant

La figure 5.75 ci-dessous présente l'évolution de la pente de l'équation de locus en fonction de l'âge à l'implantation et de l'âge auditif chez les enfants CI.

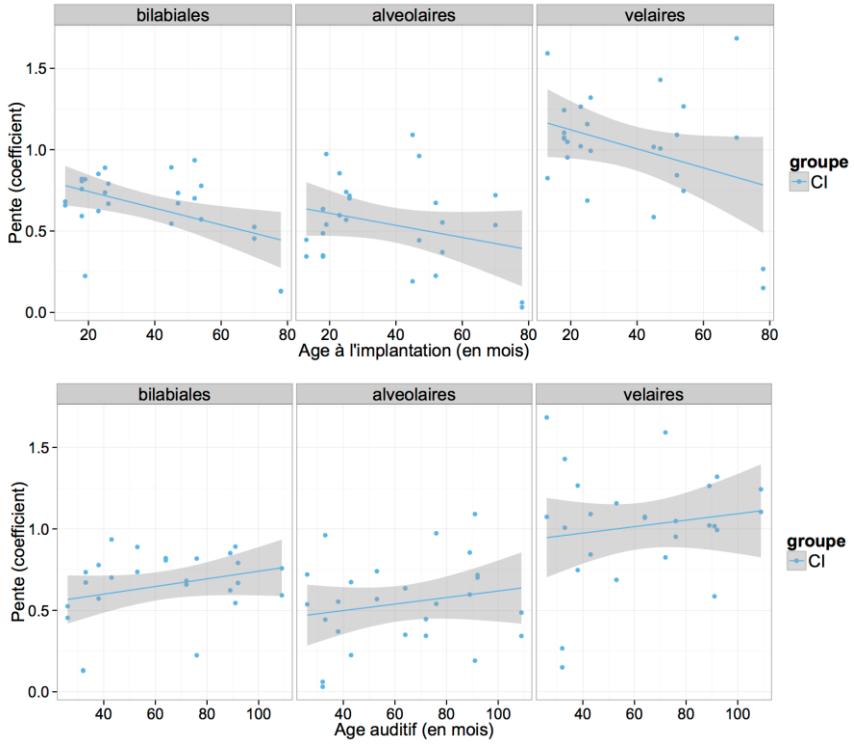


Figure 5.75 : pente de l'équation de locus (coefficient) en fonction de l'âge à l'implantation (haut) et de l'âge auditif (bas), en mois par lieu d'articulation, pour les enfants CI

L'observation de la figure 5.76 indique une diminution de la pente de l'équation de locus en fonction de l'âge à l'implantation : il semble que les enfants ayant reçu un implant le plus tardivement ont une réalisation de la transition CV avec un degré de coarticulation plus faible que ceux ayant reçu un implant tôt. En revanche l'évolution en fonction de l'âge auditif est inverse : plus les enfants ont une longue expérience de leur implant, plus la pente d'équation de locus est élevée, et plus la coarticulation est forte.

Les résultats du test statistique des effets de l'âge auditif et de l'âge à l'implantation sur la pente de l'équation de locus sont présentés dans la table 5.120 ci-dessous.

	numDF	denDF	F-value	p-value
(Intercept)	1	57	277.25518	<.0001
C	2	57	29.29855	<.0001 ***
âge_auditif	1	9	1.56808	0.2420
âge_implantation	1	9	4.01066	0.0762 .
C:âge_auditif	2	57	0.01349	0.9866
C:âge_implantation	2	57	0.50079	0.6087
âge_auditif:âge_implantation	1	9	0.62891	0.4482
C:âge_auditif:âge_implantation	2	57	2.39297	0.1005

Table 5.120 : tableau de résultats au test statistique sur les données des enfants CI (effets de l'âge auditif, de l'âge à l'implantation et du lieu d'articulation sur la pente d'équation de locus)

Les résultats du test statistique indiquent que la pente d'équation de locus a tendance à être affectée par l'âge à l'implantation (F(1,9)=4,01, p=0,072), mais cet effet n'atteint pas le niveau de significativité fixé pour notre étude (p<.05). L'âge auditif n'a en revanche aucun effet sur la réalisation de la pente d'équation de locus (F(1,9)=1,57, p>.05).

4.4.1. Résumé : effets de l'âge à l'implantation et de la durée d'utilisation de l'implant (enfants CI)

La table 5.121 ci-dessus propose un résumé des effets de l'âge (âge à l'implantation et âge auditif) sur la pente de l'équation de locus étudiés dans ce paragraphe.

Effets	Pentes de l'équation de locus
âge auditif (CI)	n.s.
âge à l'implantation (CI)	n.s.
âge auditif (CI)*Lieu d'articulation	n.s.
âge à l'implantation (CI)*Lieu d'articulation	n.s.

Table 5.121 : résumé des effets de l'âge à l'implantation et de l'âge auditif seuls et en interaction avec le lieu d'articulation sur la pente de l'équation de locus

Ces résultats laissent penser que la coarticulation anticipatoire chez les enfants CI est avant tout liée à une coordination de gestes articulatoires qui, comme chez les enfants NH n'évolue pas en fonction de l'âge des enfants entre 5 et 11 ans, mais à un âge plus tardif, et que cette coordination est liée au développement du contrôle moteur des articulateurs plus qu'à un retour auditif plus ou moins partiel de la production de parole par les enfants.

4.5. Effet de l'input auditif immédiat

La figure 5.76 ci-dessous présente les pentes d'équation de locus pour les trois groupes (adultes, et enfants NH et CI) par lieu d'articulation en fonction de l'input auditif immédiatement disponible (lié à la tâche : dénomination vs répétition). L'observation de la figure 5.76 ne montre pas d'effet systématique et clair de la tâche sur la pente de l'équation de locus, à l'exception de celle des vélaires, chez les adultes et les enfants CI.

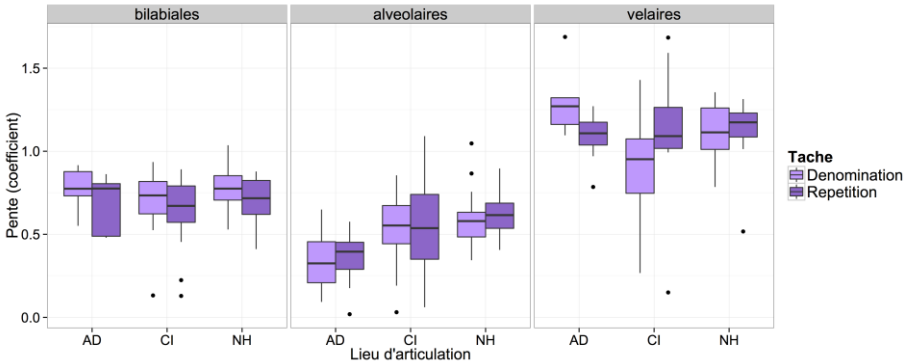


Figure 5.76 : pentes d'équation de locus (coefficient) par groupe (Adultes, enfants CI et enfants NH), par tâche et lieu d'articulation

La table 5.122 ci-dessous présente les résultats du test statistique de l'effet de la tâche sur la réalisation de la pente d'équation de locus chez les trois groupes.

	numDF	denDF	F-value	p-value
(Intercept)	1	194	1876.0725	<.0001
C	2	194	208.9153	<.0001 ***
groupe	2	38	2.7678	0.0755
tâche	1	194	0.0872	0.7680
C:groupe	4	194	5.7322	0.0002 ***
C:tâche	2	194	2.1166	0.1232
groupe:tâche	2	194	2.4513	0.0888

Table 5.122 : tableau de résultats au test statistique sur toutes les données, adultes et enfants (effets de groupe, tâche et lieu d'articulation sur la pente de l'équation de locus)

Ces résultats indiquent que la réalisation de la pente de l'équation de locus n'est pas affectée par la tâche seule (F(1,194)=0,09, p>.05) ou en interaction avec le lieu d'articulation (F(2,194)=2,12, p>.05). La différence constatée visuellement sur les vélaires ne se traduit donc pas par une différence significative entre les pentes d'équation de locus des deux tâches.

4.5.1. Résumé : effet de l'input auditif immédiat

La table 5.123 ci-dessous propose un résumé des effets de la tâche (répétition ou dénomination) sur la pente de l'équation de locus, testés dans ce paragraphe.

Effets	Pentes de l'équation de locus
Tâche	n.s.
Tâche*groupe (AD vs NH vs CI)	n.s.
Tâche*groupe (NH vs CI)	n.s.
Tâche*Lieu d'articulation	n.s.

Table 5.123 : résumé des effets de la tâche seule et en interaction avec le groupe et le lieu d'articulation sur la pente de l'équation de locus

Notre étude montre que la différence de tâche (répétition d'après un modèle audio adulte ou dénomination sans modèle) n'influence pas de façon significative la réalisation de la coarticulation par les trois groupes de locuteurs. Les différences inter-tâches relevées pour les vélaires chez les adultes et les enfants NH se sont révélées non-significatives, ce qui nous laisse penser que l'input auditif ne joue qu'un rôle marginal dans la réalisation de la coarticulation de séquences CV : les

trois groupes ont donc des productions qui traduisent des caractéristiques articulatoires indépendantes de l'input auditif proposé.

4.6. Discussion

La table 5.124 ci-dessous présente un résumé de tous les effets étudiés dans le cadre de cette étude sur la coarticulation en utilisant la pente de l'équation de locus pour modéliser la transition du deuxième formant.

Effets	Pentes de l'équation de locus
Groupe	n.s.
Lieu d'articulation	AD : /g/ > /b/ > /d/
	NH : /g/ > /b/ > /d/
	CI : /g/ > /b/ ≃ /d/
Groupe*Lieu d'articulation	/b/ : n.s.
	/d/ : NH/CI > AD
	/g/ : n.s.
âge chronologique (NH)	n.s.
âge chronologique (NH)*Lieu d'articulation	n.s.
âge auditif (CI)	n.s.
âge à l'implantation (CI)	n.s.
âge auditif (CI)*Lieu d'articulation	n.s.
âge à l'implantation (CI)*Lieu d'articulation	n.s.
Tâche	n.s.
Tâche*groupe (AD vs NH vs CI)	n.s.
Tâche*groupe (NH vs CI)	n.s.
Tâche*Lieu d'articulation	n.s.

Table 5.124 : résumé de tous les effets étudiés dans le cadre de notre étude de production de la coarticulation dans des séquences CV par les enfants CI et NH, en comparaison avec un groupe d'adultes

4.6.1.Coarticulation chez les enfants NH

Nos analyses mettent en évidence un patron de réalisation de la transition du deuxième formant globalement similaire chez les adultes et les enfants NH : la pente d'équation de locus est la plus élevée pour les vélaires, puis pour les bilabiales, puis pour les alvéolaires, ce qui est un résultat cohérent avec une partie seulement de la littérature sur l'acquisition de la coarticulation par les enfants. En effet, la revue méthodologique sur l'utilisation de l'équation de locus dans des études développementales, présentée au chapitre 4, paragraphe §2.3.2 avait mis en évidence une grande disparité de résultats d'une étude à l'autre.

Les études de Sussman et al. (1999) et de Gibson & Ohde (2007) qui étudient la coarticulation dans des production en anglais de babillage canonique et de premiers mots d'enfants de sept à 40 mois (Sussman et al., 1999) ou d'enfants de 17 à 22 mois (Gibson & Ohde, 2007) obtenaient des résultats similaires : les pentes d'équation de locus des alvéolaires sont les plus faibles dans les deux études, suivies de celles des bilabiales puis des vélaires pour Sussman et al. (1999) et Gibson & Ohde (2007). Chez des enfants plus grands (enfants de trois et cinq ans pour l'étude de Turnbaugh et al. (1985), de trois, cinq et sept ans pour celle de Nittrouer (1993) et de quatre et cinq ans pour celle de Noiray et al., 2013), les alvéolaires ont également systématiquement la pente d'équation de locus la plus faible. Elles sont suivies des bilabiales puis des vélaires pour les enfants de trois ans de l'étude de Turnbaugh et al. (1985), dont la différence a tendance à s'inverser peu à peu (les pentes des bilabiales et des vélaires sont similaires chez des enfants de cinq ans) pour atteindre un ordre alvéolaires-vélaires-bilabiales chez les adultes de cette étude. Les

analyses comparatives de l'évolution du F2 dans des séquences /ətV/ et /əkV/ de l'étude de Nittrouer (1993) chez des enfants de trois, cinq et sept ans montre que les vélaires sont plus sensibles au contexte vocalique que les alvéolaires, et que le F2 varie selon l'âge des enfants, la consonne et le contexte vocalique. Cette étude n'inclut pas les bilabilales. Enfin, l'étude de Noiray et al. (2013) met en évidence chez des enfants de quatre et cinq ans une coarticulation plus faible pour les alvéolaires, suivies des vélaires puis des bilabiales, avec une grande variabilité inter-sujets, en particulier pour les alvéolaires, ce qui conduit les auteurs de cette étude à envisager que les synergies entre gestes articulatoires utilisées dans la production de séquences alvéolaires-voyelles ne sont pas encore stabilisées chez les enfants de quatre et cinq ans.

Les résultats de notre étude comparative de la coarticulation chez les enfants NH et les adultes sont donc en accord pour l'ordre des pentes d'équation de locus pour chaque lieu d'articulation (/d/ < /b/ < /g/) avec ceux des études où les enfants étaient les plus jeunes (enfants de trois ans chez Turnbaugh et al., 1985, enfants de 17 à 22 mois chez Gibson & Ohde, 2007 et un enfant de 7 à 40 mois chez Sussman et al., 1999) mais se distinguent des études avec des enfants plus âgés (ordre des pentes d'équation de locus : /d/ < /b/-/g/ pour les enfants de cinq ans de l'étude de Turnbaugh et al., 1985, et de quatre-cinq ans chez Noiray et al., 2013).

Nos résultats pour les vélaires et bilabiales sont similaires à ceux de Sussman et al. (1999) chez l'adulte : une plus grande coarticulation pour les vélaires que pour les bilabiales. Nous pourrions donc envisager qu'à la suite du babillage et des premiers mots, les enfants de 5;7 ans à 10;6 ans de notre étude ont une production de séquences vélaire-voyelle et bilabiale-voyelle qui commence à s'approcher de celle de l'adulte.

Notre étude confirme en outre les résultats de l'étude de Noiray et al. (2013) pour laquelle les enfants de quatre-cinq ans n'ont pas encore atteint une coarticulation similaire à celles des adultes pour les alvéolaires, dont la production nécessite une coordination articulatoire que les enfants ne semblent pas encore maîtriser à cet âge. Ceci se manifeste notamment dans l'étude de Noiray et al. (2013) par une grande variabilité inter-sujets et dans notre étude par une différence entre adultes et enfants NH significative pour les alvéolaires.

Nos résultats n'ont cependant pas permis de mettre en évidence un effet de l'âge chronologique chez les enfants NH. En effet, entre cinq et onze ans, la coarticulation des alvéolaires est plus marquée que chez les adultes et il semblerait que la réalisation de la transition du F2 ne varie que peu et que la diminution de la coarticulation pour atteindre un niveau comparable à celui des adultes n'interviendrait que plus tard. La variabilité de la pente de l'équation de locus au cours du développement est mise en avance par plusieurs études, réalisées auprès d'enfants d'âges différents (Sussman et al., 1999 et Noiray et al., 2013, par exemple). Nos résultats pour les vélaires et les bilabiales, différents de ceux de Noiray et al. (2013), pourtant également réalisés auprès d'adultes et d'enfants francophones, sont peut-être liés à cette évolution non linéaire des patrons de coarticulation chez l'enfant, avec une succession de périodes de stabilité plus ou moins longues. Sussman et al. (1999) distinguent en effet la coarticulation lors de la première année de vie, puis lors des trois années suivantes. Nous pouvons faire l'hypothèse que les résulats observés chez Noiray et al. (2013), puis ceux deux notre étude, pourraient correspondre à deux paliers différents dans la mise en place d'une coarticulation similaire à celle de l'adulte : les bilabiales seraient plus sensibles à la coarticulation avant l'âge de quatre-cinq ans, puis seraient produites de façon plus stables à partir de cinq ans. Des analyses complémentaires de la coarticulation chez

une population d'enfants plus âgés et de jeunes adultes nous permettraient de comprendre plus précisément à quel âge la mise en place du patron observé chez l'adulte intervient.

Notre étude montre enfin que la tâche n'influence pas les pentes d'équation de locus des groupes d'enfants NH et d'adultes, ce qui est un résultat différent de ceux de Krull (1989) et Duez (1992) qui montraient un effet du style de parole (parole spontanée vs mots isolés pour Krull, 1989, et parole spontanée vs lue pour Duez, 1992). La différence entre parole spontanée et production de mots isolés est plus grande que celle qui distingue les deux tâches de notre étude (répétition vs dénomination de mots), ce qui pourrait expliquer que nous n'observons pas de différence entre les deux tâches. Une étude de la coarticulation à partir du corpus de parole spontanée que nous avons constitué en parallèle avec les mêmes enfants pourrait nous permettre de poursuivre cette caractérisation de la coarticulation de l'enfant normo-entendant.

4.6.2. Coarticulation chez les enfants CI

Notre étude de la coarticulation a mis en évidence des résultats similaires en de nombreux points chez les enfants NH et les adultes NH, la principale différence entre ces deux groupes étant localisée sur les alvéolaires. Les résultats des enfants CI indiquent la même tendance que pour les enfants NH. En effet, les enfants CI réalisent des séquences bilabiale-voyelle et vélaire-voyelle avec une coarticulation similaire à celle des enfants NH et des adultes NH. En revanche, ils réalisent des séquences alvéolaire-voyelle avec une pente d'équation de locus plus haute, donc un degré de coarticulation plus grand que les enfants et adultes NH. Nous n'observons pas de différence significative entre les pentes d'équation de locus des alvéolaires chez les deux groupes d'enfants.

Cette réalisation de la coarticulation similaire chez les deux groupes d'enfants de notre étude est assez proche de celui de l'étude de Monsen (1976) réalisée auprès d'adolescents sourds de 14 à 18 ans, non appareillés, chez qui les mesures de F2 au début de la transition CV et au milieu de la voyelle indiquaient des transitions similaires à ceux d'adolescents normo-entendants du même âge.

Si pour les bilabiales, nous n'observons pas de différences entre les productions des deux groupes d'enfants et le groupe d'adultes, notre étude indique cependant un degré de coarticulation plus proche pour les bilabiales et les alvéolaires (absence de différence significative entre ces deux lieux d'articulation) chez les enfants CI, alors que celui ci est bien distinct chez les enfants NH et les adultes NH de notre étude.

A la différence de l'étude de McCaffrey-Morrison (2012), nous n'avons pas pu comparer la coarticulation CV des enfants CI de notre étude avec une population d'enfants NH appariés en âge auditif. Dans cette étude chez des enfants CI de 11 à 50 mois (0;11 à 4;2 ans), des enfants CI réalisaient des transitions CV similaires à celles d'enfants NH appariés en âge chronologique pour les bilabiales et les vélaires, et similaires à celles d'enfants appariés en âge auditif pour les alvéolaires, ce qui indiquait une stabilisation légèrement plus tardive pour ce dernier lieu d'articulation.

Enfin, nos résultats indiquent que l'âge (âge chronologique chez les enfants NH, et âge à l'implantation ou âge auditif chez les enfants CI) n'a pas d'effet sur la réalisation de la pente d'équation de locus : l'évolution de la coarticulation CV est donc probablement plus fortement liée à des caractéristiques articulatoires en lien principalement avec la maturité du contrôle moteur

chez les enfants, et dans une moindre mesure avec le retour auditif qui ne perturbe que peu la réalisation des séquences CV chez les enfants CI. Ces résultats vont donc également dans le sens de l'étude de Lane et al. (2001) qui ne constatait pas de différence entre les pentes des équations de locus dans des productions de séquences CV chez des adultes sourds post-linguaux porteurs d'implants cochléaires et chez des adultes normo-entendants : la perte d'audition ne semble donc pas affecter de façon importante la réalisation de la transition CV, qui reste stable et comparable à celle d'adultes NH pour Lane et al. (2001) et à celle d'enfants NH pour notre étude.

Cet argument peut être renforcé par le fait que nous n'avons pas constaté de différence significative entre les pentes d'équation de locus des trois groupes (adultes NH, enfants NH et enfants CI) lorsqu'ils disposaient d'un modèle audio ou pas : en effet, la présence ou l'absence de l'input immédiat n'ont eu aucune influence sur les réalisations des séquences CV de notre étude.

5. Discussion générale des résultats de notre étude de production de parole

Les analyses que nous avons réalisées dans ce chapitre ont pu mettre en évidence des différences inter-groupes de réalisation de certaines catégories et contrastes phonologiques, ainsi que d'importantes similitudes entre ces réalisations qui nous permettent de tirer un certain nombre de conclusions sur les facteurs du développement phonologique et sur l'impact de l'implant.

D'une part, on trouve une grande proximité entre les productions des enfants porteurs d'implants et celles des enfants typiques, ce qui nous montre qu'il est possible d'acquérir la phonologie d'une langue malgré un manque d'accès à l'audio pendant la première, voire les premières années de vie. Cette observation amène à questionner la notion de « période critique » pendant laquelle il serait indispensable de construire des représentations perceptives conformes à la langue maternelle (par exemple, Aslin & Pisoni, 1980 (cit. in Best, 1994), ou Kuhl, 2000). La proximité entre nos deux groupes d'enfants montre également que le développement phonologique peut se faire malgré une information auditive dégradée.

D'autre part, les différences observées entre enfants CI et enfants NH mettent en évidence les effets des différents facteurs de développement phonologique : contraintes articulatoires, poids des modalités perceptives, effet de l'input langagier.

Ainsi, nos résultats sont en accord avec la littérature sur l'acquisition du langage et le développement phonologique, puisqu'ils suivent les trajectoires de développement décrites par Jakobson (1941, 1949), Ingram (1992, 1999), Locke & Pearson (1992) ou encore Kent (1992) qui indiquent que les voyelles sont acquises avant les consonnes, et les occlusives avant les fricatives, mais également par Davis & McNeilage (2004) qui, dans une analyse articulatoire de la production de parole, mettent en évidence que les fricatives sont plus tardivement stabilisées que les occlusives dans l'inventaire des sons de parole des enfants, puisque les fricatives nécessitent un contrôle plus précis des articulateurs que les occlusives. Nous observons chez les enfants implantés le même poids des contraintes articulatoires sur l'ordre d'acquisition des phonèmes que chez les enfants typiques.

La production des voyelles semble globalement acquise par nos deux populations d'enfants, puisque les quatre degrés de hauteur et les trois degrés d'antériorité sont bien distincts chez les deux groupes. Cependant nous obtenons une différence inter-groupes pour les voyelles antérieures arrondies : les enfants CI produisent cette série de voyelles de façon similaire à celle des voyelles postérieures. Une explication de cette spécificité pourrait s'appuyer sur le statut particulier de ces voyelles antérieures arrondies dans la typologie des langues du monde : rappelons que (voir chapitre 4, paragraphe §2.1.1.1) d'après la base de données UPSID (UCLA Phonological Segment Inventory Database, 2014) et Vallée (1994), seuls 0,3 % des langues du monde comportent la série des voyelles antérieures arrondies /y/, /ø/ et /œ/. Ces voyelles sont donc typologiquement rares, ce qui pourrait être dû à une plus grande complexité de production. Ceci permettrait d'expliquer également qu'elles sont plus difficilement acquises par les enfants CI. Il faudrait pour prouver cette hypothèse observer leur acquisition chez les enfants NH plus jeunes que ceux de notre échantillon.

Les difficultés de perception de l'information audio, mais également visuelle des sons de parole peuvent expliquer certaines différences inter-groupes mises en lumière lors de notre étude de production. Nos résultats indiquent tout d'abord un impact différent de l'implantation sur le

traitement des informations fréquentielles des sons de parole et sur celui des informations de durée. En effet, nous avons étudié dans ce travail certaines caractéristiques phonétiques liées avant tout au traitement perceptif d'informations fréquentielles (les formants pour les voyelles, et les quatre moments spectraux pour les fricatives) et d'autres caractéristiques qui sont associées aux informations de durée (le VOT pour les occlusives). Notre étude a montré une différence d'acquisition selon que les contrastes étudiés reposaient sur les informations de fréquence ou de durée : les informations de durée sont plus facilement maîtrisées par les enfants CI, qui comme les enfants NH produisent un voisement des occlusives proche de celui des adultes francophones de la littérature, alors que les informations fréquentielles, en particulier celles qui caractérisent les fricatives, donnent lieu à de plus grandes différences entre les deux groupes d'enfants de notre étude. Ceci peut s'expliquer par les caractéristiques de l'implant qui propose un codage assez précis des informations de durée mais qui, du fait d'un nombre de canaux limités (un maximum de 22 électrodes sont disponibles pour recomposer et transmettre l'information captée par l'implant), impose forcément une dégradation de la résolution fréquentielle.

Pour ce qui concerne la production des voyelles antérieures arrondies, l'explication de la spécificité des enfants CI peut également être liée à leur perception de l'input langagier, puisqu'il s'agit d'une série de voyelles qui présente une ambiguïté, i.e. leur confusion possible avec les voyelles postérieures, qui ne peut être résolue par la seule modalité visuelle. Ceci pourrait expliquer que seule cette série de voyelles est affectée par le port d'un implant cochléaire.

Nous observons par ailleurs que l'ambiguïté visuelle des sons de parole affecte indifféremment voyelles et consonnes puisque 1) les voyelles antérieures non-arrondies sont acquises avant les voyelles antérieures, arrondies, qui visuellement ne se distinguent pas des postérieures, également arrondies, 2) les fricatives alvéolaires et post-alvéolaires sont plus difficilement produites que les labiodentales par les enfants CI, en comparaison avec celles des enfants NH, qui ont des caractéristiques proches de celles d'adultes de la littérature et 3) la production des occlusives vélaires est relativement proche de celle des alvéolaires chez les enfants CI, par rapport aux enfants NH. Notre étude a également mis en évidence que les occlusives vélaires, qui visuellement ne se distinguent pas des alvéolaires, sont les plus variables chez les enfants CI. Ces analyses sont notamment en accord avec l'étude de Geffner (1980) qui montrait que la production des enfants sourds non-appareillés était plus difficile pour les contrastes dont la perception visuelle peut être ambiguë. L'implant cochléaire permet donc à l'enfant qui l'utilise de construire une phonologie proche de celle de l'enfant normo-entendant mais plus tardive, et qui conserve des caractéristiques d'enfants sourds.

Enfin, nous pouvons expliquer les différences inter-groupes lors de notre étude par un effet de l'input (à la fois retour perceptif interne et bain langagier) auquel les enfants CI sont exposés.

La plus grande dispersion sur F1-F2 et F2-F3 de certaines catégories vocaliques peut être attribuée elle aussi à la perception de l'input auditif, mais s'explique peut-être mieux encore par le fait que l'implant cochléaire joue également sur le retour qu'a l'enfant sur sa propre production de parole. Ce retour auditif dégradé induit sans doute chez les enfants CI une difficulté à s'appuyer sur la boucle audio-phonatoire pour construire des catégories vocaliques conformes à celles de la langue adulte. Ces résultats sont en cohérence avec l'approche de Bybee (2001) et Pierrehumbert (2003) qui analysent le développement phonologique par l'interaction de la boucle audio-

phonatoire de l'enfant avec l'exposition à la langue maternelle et avec la fréquence des sons de la langue maternelle.

Pour ce qui concerne les voyelles antérieures arrondies, on pourrait également invoquer leur fréquence d'occurrence dans la langue maternelle pour expliquer leur spécificité, mais cet aspect ne semble pas pertinent au regard des données disponibles. En effet, selon les chiffres de l'étude de Monnin (2010), basés sur les bases de données Lexique 2 et 3 (New et al., 2007) ainsi que sur une étude originale sur du français adressé à l'enfant (Child-Directed Speech ou CDS), cette série de voyelles n'a pas une fréquence d'apparition particulièrement faible dans la langue française, en tout cas par rapport à la série des voyelles postérieures arrondies par exemple.

Enfin, nos résultats sur la coarticulation de séquences occlusive-voyelle n'ont pas mis en évidence de différence majeure entre les deux groupes d'enfants, ce qui est le reflet de gestes articulatoires similaires, mais une différence essentiellement visible sur les consonnes alvéolaires entre les enfants et le groupe d'adultes, ce qui indique que la coarticulation n'est pas encore stabilisée à l'âge de 11 ans. L'étude de Noiray et al. (2013), basée sur des analyses acoustiques (équation de locus) et articulatoire (ultrason) de la coarticulation anticipatoire, mettait en évidence un mouvement différent de la langue par des enfants de quatre-cinq ans et des adultes, tous canadiens francophones, pour produire des alvéolaires. Notre étude semble donc montrer que le développement des stratégies de coarticulation ne dépend pas des informations perceptives, mais d'une évolution du contrôle articulatoire non encore terminée à l'âge de 11 ans. Une étude complémentaire chez des enfants plus âgés pourra confirmer cette hypothèse.

Notre étude a mis en évidence un certain nombre de difficultés persistantes dans la production de parole des enfants implantés, et nous avons donc réalisé une étude visant à évaluer l'impact de ces difficultés sur l'intelligibilité de ces enfants, qui fera l'objet des chapitres suivants.

TROISIÈME PARTIE :
Intelligibilité de la parole
(2ème étude expérimentale)

CHAPITRE 6 : Étude d'intelligibilité (2ème étude expérimentale)

Parallèlement à notre étude de production (chapitres 4 et 5), qui reposait sur une analyse acoustique des productions de parole des enfants porteurs d'implant cochléaire, nous avons choisi de proposer une analyse perceptive de cette production de parole. Notre but est d'essayer de confronter une évaluation objective de la production et des difficultés persistantes rencontrées par cette population d'enfants avec une évaluation subjective de ces difficultés. Dans les sections suivantes, nous allons présenter des méthodologies disponibles dans la littérature et fréquemment utilisées pour juger de l'intelligibilité de la parole d'enfants CI, de façon à expliquer nos propres choix méthodologiques, avant de proposer notre propre étude ainsi que nos résultats et nos analyses des facteurs de l'intelligibilité de la parole d'enfants porteurs d'implant cochléaire.

Dans cette deuxième étude expérimentale, nous cherchons en particulier à caractériser l'intelligibilité des enfants porteurs d'implants cochléaires et normo-entendants : nous allons donc nous intéresser aux effets du groupe (CI et NH), de l'âge chronologique des enfants (CI et NH), de l'âge à l'implantation (enfants CI) et de la durée d'utilisation de l'implant (enfants CI) sur l'intelligibilité de la parole continue. Cette évaluation subjective va nous permettre de mieux comprendre si l'évaluation subjective de l'intelligibilité (Deuxième partie, ce chapitre) est sujette aux effets des mêmes facteurs que les caractéristiques acoustiques de la parole des enfants (Première partie, chapitres 4 et 5).

1. Méthodologie d'études d'intelligibilité dans la littérature

1.1. Introduction

L'évaluation de l'intelligibilité de locuteurs, qu'ils aient une audition dégradée ou non, est une question complexe, puisqu'elle repose obligatoirement sur un jugement perceptif. Les études d'intelligibilité présentées au chapitre 3, ci-dessus, tentent de comprendre quelles sont les caractéristiques objectivement observables qui permettent d'expliquer des différences perçues et qui causent une moindre intelligibilité de la parole par un interlocuteur. Ces caractéristiques peuvent être à la fois des caractéristiques liées au locuteur (âge, genre, degré d'audition, type d'aide auditives, etc), ou des caractéristiques quantifiables dans sa production de parole (caractéristiques acoustiques, aérodynamiques de sons de parole, etc). Le début de ce chapitre est consacré aux méthodes de jugement de l'intelligibilité disponibles dans une sélection d'études expérimentales, qui vont ensuite guider nos choix méthodologiques pour étudier l'intelligibilité des enfants CI et NH.

1.2. Revue de méthodologies disponibles

Les études sur l'intelligibilité consistent généralement à éliciter et à recueillir des productions de la part d'un groupe d'individus, puis à soumettre ces productions au jugement d'auditeurs naïfs ou spécialistes de la parole. On peut ensuite éventuellement corréler le score d'intelligibilité aux caractéristiques phonétiques des productions. Nous nous intéresserons dans un premier temps aux différentes méthodes d'élicitation de parole utilisées dans le cadre de ces études, qui permettent à la fois de caractériser la parole des enfants et de construire les stimuli utilisés pour les jugements d'intelligibilité. Dans un deuxième temps, nous passerons en revue les méthodes et les critères utilisés pour juger l'intelligibilité de ces productions de parole. Enfin, nous nous intéresserons aux méthodes d'analyse des corrélats de l'intelligibilité.

Auteurs	Date	Langue	Participants	Âge des participants	Type d'étude / nombre de sessions	Type de corpus	Test d'intelligibilité standardisé
Blamey et al.	2001	Anglais (Australie)	9 enfants CI	3;9 ans (âge impl)	étude longitudinale : 3m avt impl., puis 9 sessions sur 6a	conversation	
Calmels et al.	2004	Français	63 enfants CI (35G, 27F)	1-10ans (âge impl.)	étude longitudinale : 3, 6, 9, 12, 18m, 2, 3, 5a post-impl.	conversation / parole spontanée	Speech Intelligibility Rating (Allen et al. 1998)
Chuang et al.	2012	Mandarin (Nord de Taïwan)	24 enfants CI 24 enfants NH	93,8 mois (72-133m) même âge	1 session	lecture de phrases, lecture de mots, voyelles tenues	
Flipsen et al.	2006	Anglais (US)	6 enfants CI	5a (3;9a à 6;2a)	étude longitudinale, tous les 3 mois, sur une période de 12 à 21 mois selon les enfants	conversation	Intelligibility Index (Shriberg 1986) : Echantillons de 90 mots par enfant par session
Habib et al.	2010	Anglais (US)	37 enfants CI	2;5-18a (âge impl. : 8-36m)	1 session	répétitions de phrases	Beginner's Intelligibility Test (Osberger et al. 1994) : 2 répétitions*4 listes de phrases (2-6 mots mono- et bisyllabliques par phrase), imitation d'un modèle audio adulte
Hassanzadeh	2012	Farsi	7 enfants CI (parents sourds) 7 enfants CI (parents NH)	106 mois (52-153) 106 mois (51-147)	4 sessions : - 1er réglage - 1er réglage + 6m - 1er réglage + 12m - non spécifié	répétition de phrases	Beginner's Intelligibility Test (Osberger et al. 1994) : production de phrases
Horga et Liker	2006	Croate	10 enfants CI 10 enfants NH 10 enfants HA	9;7-15;2 9;6-15;1 7;4-15;1	1 session	répétition de mots et phrases (audio + image ou texte)	
Khwaileh et Flipsen	2010	Anglais (US)	17 enfants CI (13F, 4M)	4;8a-11;1a	1 session	répétition de mots et de phrases	Beginner's Intelligibility Test : répétition de phrases, Children's Speech Intelligibility Measure (CSIM) : répétition de mots
Löfqvist et al.	2010	Suédois	12 enfants CI 11 enfants NH	184 mois (144-229) 108 mois (103-118)	1 session	Production de mots + conversation	
Löhle et al.	1999	Allemand	83 enfants HA (4 groupes selon niveau d'audition) 33 enfants CI (3 groupes d'âge d'impl.)	âge exact des enfants non spécifié	Comparaison avant/après impl.	reconnaissance de mots (audiogramme), parole spontanée, lecture de mots	
Miyamoto et al.	1997	Anglais (US)	Enfants CI Enfants HA (3 groupes selon niveau d'audition) Nombre d'enfants pas spécifiés	5;0 (âge d'impl.) moyen	étude longitudinale, 1 enregistrement pré-impl., puis tous les 6m post-impl. (jusqu'à 7;5a post-impl.)	répétition de phrases	Enfants de moins de 6a : Beginner's Intelligibility Test (Osberger et al. 1994) Enfants de plus de 6a : Monsen Sentences Test (Monsen 1981)

Montag et al.	2014	Anglais (US)	63 enfants CI 63 enfants NH	19 enf. (7-9a), 30 enf. (10-14a), 14 enf (15+a) pour chaque groupe	1 session	répétition de phrases (avec support écrit)	McGarr Sentence Intelligibility Test (McGarr 1981) : phrases de 3, 5, 7 syllabes avec un mot cible (6 mots fréquents, 6 mots peu fréquents)
Peng et al.	2004	Anglais (US)	24 enfants CI (18 enfants avec la stratégie de codage SPEAK, 6 enfants avec la stratégie de codage MPEAK)	147;5m (113;3-217;9), âge d'impl. 61;3 (30;9-132;5)	1 session	répétition de phrases (modèle adulte)	Test "Iowa Children's Cochlear Implant Project" (Tye-Murray 1998), partie "Short-Long Sentence Test" -> 14 phrases de 7,43 mots (6-10 mots)
Perrin et al.	1999	Français	4 enfants CI 4 enfants NH	(9-14a)	1 session	Lecture de "La Bise et le Soleil"	
Poissant et al.	2006	Anglais (US)	6 enfants CI	7;1a (5;3-10;2)	2 sessions : implant-ON vs OFF	Production de mots isolés et des mêmes mots dans phrases porteuses	
Svirsky	2000	Anglais (US)	Enfants CI Enfants HA (3 groupes selon niveau d'audition) Nombre d'enfants non spécifiés	âge exact des enfants non spécifié	1 session	répétition de phrases	Enfants de moins de 6a : Beginner's Intelligibility Test (Osberger et al. 1994) Enfants de plus de 6a : Monsen Sentences Test (Monsen 1981)
Van Lierde et al.	2005	Néerlandais (Belgique)	9 enfants CI 6 enfants NA	8;10a (5;10-12;9) 10;6a (6;9-13;8)	1 session	conversation + lecture/répétition de phrases	

Table 6.1 : résumé des procédures d'élicitation de parole, utilisées dans des études récentes sur l'intelligibilité de la parole d'enfants porteurs d'implants cochléaires (classement par ordre alphabétique du premier auteur)

La table 6.1 ci-dessus présente les méthodes d'élicitation de parole et de construction des stimuli utilisés dans les jugements d'intelligibilité de ces études, ainsi que les populations étudiées et les langues dans lesquelles ces études sont réalisées. Les tests standardisés sont des procédures définies et dont les résultats peuvent être comparés à un résultat normé qui est le plus souvent calibré en fonction des participants (adultes vs enfants), de leur âge, etc. et qui peuvent être disponibles dans plusieurs langues, ce qui permet des comparaisons inter-langues.

Les études d'intelligibilité se basent sur des productions de parole, qui peuvent être évaluées simultanément d'un point de vue quantitatif (par des évaluations de nombre de mots ou segments correctement produits, par exemple) et qualitatif (par des jugements d'auditeurs, par exemple). Cette double analyse permet ensuite aux différents auteurs d'étudier les relations entre la production et l'intelligibilité de la parole.

Plusieurs méthodes d'élicitation de parole et plusieurs types de corpus sont ainsi présents dans la littérature. Dans ces études, les productions peuvent être enregistrées soit en suivant la méthodologie d'un test existant, soit en élaborant une méthodologie propre à l'étude en question,

qui permet aux auteurs d'étudier parallèlement un ou plusieurs corrélats, notamment acoustiques, de l'intelligibilité dans les productions de parole.

1.2.1. Méthodes d'élicitation de la parole

1.2.1.1. Enregistrement des productions de parole

Ce chapitre se concentre sur une sélection d'études récentes, dont nous avons présenté les résultats au chapitre 3.

- **Répétition de mots ou de phrases**

Parmi les tests standardisés utilisés, le BIT, « Beginner's Intelligibility Test » (Osberger et al., 1994), utilisé par Miyamoto et al. (1997), Svirsky (2000), Habib et al. (2010), Khwaileh et Flipsen (2010) et Hassanzadeh (2012), est l'un des plus répandus. Il s'agit d'un test de répétition de phrases comprenant de deux à six mots mono- et bisyllabiques, à partir d'un modèle audio adulte. Ce test est notamment utilisé lorsque les enfants ne lisent pas encore : au début des études longitudinales de Miyamoto et al. (1997) et Svirsky (2000), les enfants ont un âge moyen de 5;0 ans, dans celle de Habib et al. (2010) les enfants les plus jeunes ont 2;8 ans, dans celle de Khwaileh et Flipsen (2010) les enfants les plus jeunes ont 4;8 ans et dans celle de Hassanzadeh (2012) les enfants les plus jeunes ont 4;4 ans. Cependant, dans aucune de ces études, il n'est précisé si les répétitions se font avec un modèle audio préalablement enregistré ou avec une production de l'expérimentateur présent au moment de l'enregistrement et qui pourrait varier d'une session à une autre.

Khwaileh et Flipsen (2010) complètent ce test par un deuxième test standardisé, le « Children's Speech Intelligibility Measure » (CSIM, Wilcox and Morris, 1999), qui consiste en une répétition de mots par les enfants d'après un modèle adulte. Il n'est pas précisé non plus de quelle façon le stimulus audio est proposé.

Miyamoto et al. (1997) et Svirsky (2000) utilisent le « Monsen Sentences Test » (Monsen, 1981) pour enregistrer les productions d'enfants de plus de 6 ans, en complément du BIT qu'ils utilisent avec les enfants de moins de 6 ans. Le « Monsen Sentences Test » est un test similaire au BIT (Osberger et al., 1994), mais l'élicitation de la parole se fait par la lecture de phrases.

De la même façon, le test « McGarr Sentence Intelligibility Test » (McGarr, 1981) permet à Montag et al. (2014) d'enregistrer des productions de parole à partir de la lecture sans modèle audio de phrases de trois longueurs différentes (trois, cinq et sept syllabes par phrase) qui contiennent chacune un mot-cible parmi une liste de 12 mots (six mots fréquents et six mots peu fréquents).

Peng et al. (2004) utilisent le « Short-Long Sentence Test », qui est extrait du « Iowa Children's Cochlear Implant Project » (Tye-Murray, 1998) ; il s'agit d'un test similaire au BIT (Osberger et al., 1994)) dans lequel il est demandé aux enfants de répéter des phrases (14 phrases comprenant entre six et dix mots) d'après un modèle audio adulte.

- **Elicitation de parole spontanée**

Les méthodologies utilisées par Calmels et al. (2004) et Flipsen et al. (2006) pour construire et analyser les stimuli de leurs études d'intelligibilité sont également tirées de protocoles standardisés : des productions de parole spontanée (conversation des enfants avec un adulte) sont enregistrées puis analysées avec des grilles de notation : « Speech Intelligibility Rating » (Allen et

al., 1998) pour Calmels et al. (2004) et « Intelligibility Index » (Shriberg, 1986) pour Flipsen et al. (2006).

Par ailleurs, si la plupart des études utilisent des procédures standardisées existantes, un certain nombre d'autres études sont réalisées avec une méthodologie propre qui peut être similaire à celle d'un ou plusieurs tests existants. Ainsi, les études de Van Lierde et al. (2005) et Horga et Liker (2006) utilisent une procédure similaire au BIT (Osberger et al., 1994) de répétition de phrases avec des supports écrits ou des images.

Blamey et al. (2001), Löfqvist et al. (2010) et Löhle et al. (1999) enregistrent des productions de parole spontanée dont ils évaluent ensuite l'intelligibilité avec une grille qu'ils ont élaborée pour répondre à leur problématique. Quelques études utilisent la lecture de phrases ou de mots sans modèle audio : Chuang et al. (2012) et Löhle et al. (1999) enregistrent la lecture de phrases et de mots, et Perrin et al. (1999) la lecture du texte « La Bise et le Soleil ». Enfin, Poissant et al. (2006) enregistrent des productions de mots isolés et des mêmes mots dans une phrase porteuse à partir d'images présentées aux enfants.

Le choix d'un style de parole (lue, spontanée, répétée, etc.) est déterminant pour une étude d'intelligibilité : les résultats et analyses obtenus par des méthodes d'élicitation de parole ne permettent pas tous de juger les mêmes caractéristiques de l'intelligibilité. Ainsi, la répétition de mots isolés ou la conversation avec un proche ne permettent pas de juger de la même façon de la capacité de l'enfant à produire un discours intelligible.

1.2.1.2.Populations

Les études peuvent être soit longitudinales soit transversales, ce qui va jouer un rôle dans les résultats obtenus par leurs auteurs. Ces caractéristiques sont également présentées dans la Table 6.1 ci-dessus.

Les études longitudinales permettent d'évaluer l'intelligibilité en fonction de l'expérience d'utilisation de l'implant par les enfants. Löhle et al. (1999), Blamey et al. (2001), Hassanzadeh (2012) et Miyamoto et al. (1997) enregistrent les productions des enfants avant l'implantation, puis suivent ces mêmes enfants sur des périodes plus ou moins longues (jusqu'à sept ans et demi pour l'étude de Miyamoto et al., 1997). L'étude de Calmels et al. (2004) démarre trois mois après l'implantation et s'étend sur une période de cinq ans, alors que celle de Flipsen et al. (2006) est une étude à moyen terme, elle commence trois ans après l'implantation et s'étend sur une période de 12 à 21 mois selon les enfants. Ces études longitudinales permettent de suivre l'évolution de l'intelligibilité à partir de la première session d'enregistrement des productions et ne prennent en compte que l'intelligibilité des enfants porteurs d'implant cochléaire, sans référence à des enfants normo-entendants.

Les études transversales quant à elles se déroulent en une session d'enregistrement avec un groupe ou une condition contrôle dans la majorité des cas. Les études de Habib et al. (2010) et de Khwaileh et Flipsen (2010) se déroulent en une seule session d'enregistrement et sans groupe ni condition contrôle. Peng et al. (2004) proposent une étude transversale sans groupe contrôle normo-entendant mais avec deux groupes d'enfants dont les implants fonctionnent avec deux stratégies différentes de codage (stratégie SPEAK vs MPEAK). Leur but est d'étudier l'effet de la stratégie de codage sur l'intelligibilité. Poissant et al. (2006) comparent l'intelligibilité des mêmes enfants dans deux conditions d'audition (implant activé vs implant désactivé).

Les études de Perrin et al. (1999), Svirsky et al. (2000), Van Lierde et al. (2005), Löfqvist et al. (2010), Chuang et al (1012) et Montag et al. (2014) comparent l'intelligibilité des enfants porteurs d'implants à celle d'enfants normo-entendants appariés en âge. L'étude de Horga et Liker (2006) compare l'intelligibilité de deux groupes d'enfants sourds (porteurs d'implants et utilisateurs d'aides auditives) avec celle d'un groupe d'enfants normo-entendants.

Cette revue des méthodes d'élicitation de parole et de constitution de corpus pour étudier l'intelligibilité des enfants porteurs d'implants a mis en évidence une diversité des méthodes, avec une utilisation de tests standardisés ou de méthodologies propres aux études et liées aux caractéristiques de production de parole que les auteurs souhaitent étudier en lien avec l'intelligibilité. On peut noter également que le choix d'une méthode d'élicitation n'est pas lié au type d'étude (longitudinale ou transversale), ni aux groupes d'enfants de ces études.

1.2.2.Méthodes de jugement de l'intelligibilité

Les caractéristiques relatives aux jugements d'intelligibilité (caractéristiques des juges, méthodes d'évaluation de l'intelligibilité des productions) des études de ce chapitre sont présentées dans la Table 6.2 ci-dessous.

Auteurs	Date	Juges de l'intelligibilité	Critères d'évaluation de l'intelligibilité	Corrélats testés de l'intelligibilité	Méthode d'évaluation des productions
Blamey et al.	2001	Les 4 auteurs de l'étude	Nombre de mots intelligibles par phrase Nombre de syllabes intelligibles par phrase Pourcentage de segments corrects (réalisation vs cible)	Correction des : - segments (monophtongues, diphtongues, consonnes) - syllabes - mots	Transcription orthographique de phrases Transcription phonétique large de phrases
Calmels et al.	2004	Orthophonistes qui suivent les enfants	Jugements d'intelligibilité par les orthophonistes qui suivent les enfants : Speech Intelligibility Rating (Allen et al. 1998) - Note de 1 (inintelligible) à 5 (intelligible pour tout interlocuteur)	Tests de perception par les enfants : Test for the Evaluation of Voice Perception (Vieu et al. 1999)	
Chuang et al.	2012	4 orthophonistes expérimentés	Notes d'intelligibilité : échelle de 1 (problèmes sévères) à 5 (aucun problème), évaluation de la correction des voyelles, consonnes, tons et mots : pourcentage de sons produits correctement (transcription des mots produits)	Correction des voyelles, débit de parole : analyses acoustiques (formants des voyelles, débit de parole)	
Flipsen et al.	2006	Plusieurs étudiants	2 mesures différentes de l'« Intelligibility Index » (Shriberg 1986) : - nb de mots intelligibles, phrase par phrase ou par échantillon de parole -> Comparaison du nb de mots inintelligibles avec résultats d'une norme (enfants typiques)	âge chronologique âge d'implantation	Transcription phonétique étroite des échantillons par plusieurs étudiants
Habib et al.	2010	3 locuteurs naïfs différents par phrase	Pourcentage de mots-cibles correctement transcrits	âge chronologique âge d'implantation	Transcription orthographique de

			Score d'intelligibilité : 1 (non intelligible) à 5 (tout à fait intelligible)	nombre de L1 nombre d'implants audition résiduelle	chaque phrase
Hassanzadeh	2012	1 orthophoniste	Score d'intelligibilité : 1 (non intelligible) à 5 (tout à fait intelligible)	- Persian Language Development Test (Sentence imitation) - Persian Auditory Perception Test for the Hearing Impaired	
Horga et Liker	2006	Plusieurs phonéticiens	Jugements perceptifs de la qualité de voix (échelle GRBAS) : scores de 1 (max) à 7 (min)	Segments (voyelles, voisement des consonnes, VOT), accent de mot, accent de phrase, qualité de voix	Analyses acoustiques, analyses perceptives (qualité de voix, qualité de la prononciation, intelligibilité des voyelles)
Khwaileh et Flipsen	2010	51 étudiants pour le test BIT et 51 étudiants pour le test CSIM -> locuteurs naïfs : trois juges par item	Pour chaque test : pourcentage de mots correctement identifiés, moyenne des trois juges	âge chronologique âge d'implantation	Phrases : transcription orthographique Mots : transc. ortho et QCM (identification du mot dans une liste)
Löfqvist et al.	2010	20 étudiants en orthophonie (intelligibilité des phrases) 12 étudiants en orthophonie (intelligibilité des voyelles)	Phrases tirées du corpus de conversation, présentées dans le bruit (SNR 50%), jugements par 20 étudiants en orthophonie + test d'identification de voyelles par 12 étudiants	Caractéristiques acoustiques de 9 voyelles, taille de l'espace vocalique	Analyses acoustiques
Löhle et al.	1999	Plusieurs étudiants	Jugements perceptifs de la production de mots et de la prosodie de la parole spontanée, par des étudiants qui ne sont pas familiarisés avec la parole pathologique	Étude du rythme et de l'intonation (parole spontanée) : Analyses acoustiques de la prosodie (rythme et F0)	
Miyamoto et al.	1997	3 locuteurs naïfs	Scores d'intelligibilité : moyennes du nombre de mots correctement perçus par les trois locuteurs	âge chronologique expérience auditive audition résiduelle	Transcription orthographique des phrases
Montag et al.	2014	171 étudiants (locuteurs naïfs)	intelligibilité du mot-cible : 0 (mot inintelligible) - 1 (mot intelligible), calculée d'après la transcription ortho, puis moyenne de 3 juges. Intelligibilité totale de la phrase : nb de mots correctement transcrits pour chaque phrase, puis moyenne de 3 juges	âge chronologique âge d'implantation mode de communication revenu de la famille	Transcription orthographique du mot cible et de la phrase
Peng et al.	2004	72 adultes naïfs (25M + 47F)	Transcriptions : pourcentage de mots correctement transcrits (moyenne de 3 juges) Echelle d'intelligibilité : 1 (pas intelligible) à 5 (tout à fait intelligible)	âge d'implantation stratégie de codage (SPEAK vs MPEAK)	Transcription orthographique de chaque phrase / Score d'intelligibilité : 2 écoutes de chaque phrase -> 2 transcriptions / 2 scores d'intelligibilité
Perrin et al.	1999	1 orthophoniste	Jugements de plusieurs critères (perçus) par un orthophoniste (note de 0 (min) à 5 (max)) : pitch (niveau et variation),	11 voyelles (F0, 3 formants), débit d'articulation (durée de production de la phrase)	Comparaison des résultats avec des valeurs standard (norme)

			intensité (niveau et variation), fluency (level et variation), pause (nombre et durées), articulation et caractère agréable de la production.		
Poissant et al.	2006	6 étudiants en orthophonie	Jugements d'intelligibilité par 6 autres étudiants en orthophonie : 0 (mot incompréhensible) - 10 (tout à fait intelligible)	5 mots monosyllabiques, 3 mots bisyllabiques -> voyelles (F0, F1, F2, F3), durées de mots	Analyses acoustiques, transcriptions (phonétiques ?), par 6 étudiants en orthophonie
Svirsky	2000	3 locuteurs naïfs	Scores d'intelligibilité : moyennes du nombre de mots correctement perçus par les trois locuteurs	durée de rééducation orthophonique âge lors du test	Transcription orthographique des phrases
Van Lierde et al.	2005	2 orthophonistes	2 jugements par échantillon de parole : échelle de 0 (intelligibilité normale) à 3 (intelligibilité dégradée)	Articulation, phonation, résonance,	analyses aérodynamiques, acoustiques (F0, Jitter), étendue de voix (intensité, F0), évaluation (GRBAS)

Table 6.2 : résumé des méthodes de jugement de l'intelligibilité de quelques études récentes sur l'intelligibilité de la parole d'enfants porteurs d'implants cochléaires (classement par ordre alphabétique du premier auteur)

1.2.2.1.Juges d'intelligibilité

Les auteurs des études présentées dans ce chapitre, font appel à plusieurs profils d'auditeurs qui jugent l'intelligibilité des productions de parole. Les juges de l'intelligibilité sont soit des phonéticiens, auteurs des études (Blamey et al., 2001 ou Horga et Liker, 2006), soit un ou plusieurs orthophonistes (Calmels et al., 2004, Chuang et al., 2012, Hassanzadeh, 2012 et Perrin et al., 1999), des étudiants en orthophonie (Poissant et al., 2006 et Löfqvist et al., 2010), ou des locuteurs et étudiants naïfs (Miyamoto et al., 1997, Löhle et al., 1999, Svirsky, 2000, Peng et al., 2004, Flipsen et al., 2006, Habib et al., 2010, Khwaileh et Flipsen, 2010 et Montag et al., 2014). Il est intéressant de noter que ces études font le choix d'un profil d'auditeurs auxquels on demande un jugement de l'intelligibilité de la parole des enfants, mais aucune d'entre elles ne compare les résultats de plusieurs groupes d'auditeurs (auditeurs naïfs vs orthophonistes, etc).

1.2.2.2.Méthodes d'évaluation de l'intelligibilité

Certaines études font appel à des jugements objectifs au moyen de transcriptions (orthographiques ou phonétiques), d'autres études font appel à des évaluations subjectives de la perception de l'intelligibilité au moyen de notes et quelques unes combinent les deux types de jugements. Enfin, certaines études ne précisent pas quelle est la méthode utilisée. La Table 6.3 ci-dessous propose un résumé des différentes méthodes utilisées, en complément des informations données sur ces mêmes études, dans la Table 6.2.

Auteurs	Date	Test	Pourcentage de mots correctement transcrits (orthographiquement)	Pourcentage de segments correctement transcrits (phonétiquement)	Note	Autre méthode
Habib et al.	2010	BIT (Osberger et al., 1994)	oui		1 (non intelligible) à 5 (tout à fait intelligible)	
Hassanzadeh	2012	BIT (Osberger et al., 1994)	oui		1 (non intelligible) à 5 (tout à fait intelligible)	
Khwaileh et Flipsen	2010	BIT (Osberger et al., 1994)	oui		1 (non intelligible) à 5 (tout à fait intelligible)	
Miyamoto et al.	1997	BIT (Osberger et al., 1994) et Monsen Sentence Test (Monsen, 1981)	oui		1 (non intelligible) à 5 (tout à fait intelligible)	
Peng et al.	2004	Test "Iowa Children's Cochlear Implant Project" (Tye-Murray, 1998)	oui		1 (non intelligible) à 5 (tout à fait intelligible)	
Svirsky	2000	BIT (Osberger et al., 1994) et Monsen Sentence Test (Monsen, 1981)	oui		1 (non intelligible) à 5 (tout à fait intelligible)	
Montag et al.	2014	McGarr Sentence Intelligibility Test (McGarr, 1981)	oui			
Blamey et al.	2001		oui	oui		
Flipsen et al.	2006	Intelligibility Index (Shriberg, 1986)		oui		
Calmels et al.	2004	SIR (Allen et al.,1998)			1 (non intelligible) à 5 (tout à fait intelligible)	
Chuang et al.	2012				1 (problèmes sévères) à 5 (aucun problème)	
Perrin et al.	1999				0 (min) à 5 (max)	
Poissant et al.	2006				0 (mot incompréhensible) à 10 (mot tout à fait compréhensible)	
Van Lierde et al.	2005				0 (intell. normale) à 3 (intell. dégradée)	
Horga et Liker	2006	méthode non précisée			Note ?	jugements d'intelligiblité des voyelles
Löfqvist et al.	2010	méthode non précisée			Note ?	identification de voyelles
Löhle et al.	1999	méthode non précisée			Note ?	

Table 6.3 : comparaison des méthodes de jugement (subjectives et objectives) utilisées par les auteurs de quelques études récentes sur l'intelligibilité de la parole d'enfants porteurs d'implants cochléaires.

- **Evaluations objectives : transcriptions**

Deux études (Blamey et al., 2001 et Flipsen et al., 2006) utilisent une transcription phonétique, ce qui permet d'obtenir un pourcentage de segments corrects (en comparant la réalisation effective à la cible) ou un nombre de mots intelligibles phrase par phrase ou par échantillon de parole. Dans l'étude de Blamey et al. (2001), les juges effectuent également une transcription orthographique des phrases, qui permet d'obtenir un pourcentage des mots et syllabes intelligibles par phrase. Montag et al. (2014) utilisent également une évaluation objective de l'intelligibilité au moyen d'une note binaire calculée à partir d'une transcription orthographique de mot-cibles : 0 (mot inintelligible) ou 1 (mot intelligible). Flipsen et al. (2006) utilisent deux mesures d'intelligibilité à partir d'une transcription phonétique étroite des échantillons de parole : un nombre de mots intelligibles phrase par phrase, ou par échantillon de parole ; ces deux mesures sont ensuite comparées aux résultats calibrés établis à partir des résultats d'enfants typiques normo-entendants et faisant partie du protocole du test standardisé.

- **Evaluations subjectives : notes d'intelligibilité**

Calmels et al. (2004) utilisent l'évaluation du test SIR (Allen et al., 1998), qui permet de noter les productions sur une échelle de 1 (non intelligible) à 5 (intelligible pour toute personne). Dans l'étude de Chuang et al. (2012), l'évaluation de l'intelligibilité se fait au moyen d'une note d'intelligibilité de 1 (problèmes sévères) à 5 (aucun problème) à plusieurs niveaux, qui permet d'obtenir un pourcentage d'éléments produits correctement pour chaque niveau de l'étude (voyelles, consonnes, tons et mots). Dans l'étude de Perrin et al. (1999), un orthophoniste effectue des jugements perceptifs de plusieurs critères, avec une note de 0 (note minimale) à 5 (note maximale) : pitch (niveau et variation), intensité (niveau et variation), fluence (niveau et variation), pauses (nombre et durées), articulation et caractère agréable de la production. Ces critères sont également quantifiés acoustiquement. Dans l'étude de Poissant et al. (2006), l'intelligibilité des mots est évaluée une échelle de 0 (mot incompréhensible) à 10 (tout à fait compréhensible). Dans l'étude de Van Lierde et al. (2005) une note de 0 (intelligibilité normale) à 3 (intelligibilité dégradée) est donnée à chaque échantillon de parole.

- **Combinaison d'évaluations objectives et subjectives : transcriptions et notes**

Pour les études qui utilisent le test BIT (Miyamoto et al., 1997, Svirsky, 2000, Habib et al., 2010, Khwaileh et Flipsen, 2010 et Hassanzadeh, 2012), l'intelligibilité est évaluée au travers d'une transcription orthographique des phrases entendues et d'une note d'intelligibilité attribuée à la même phrase. A partir des transcriptions orthographiques obtenues, un pourcentage du nombre de mots correctement transcrits est calculé et moyenné lorsque plusieurs juges effectuent cette transcription. La note d'intelligibilité est une note de 1 (non intelligible) à 5 (tout à fait intelligible), qui est également moyennée si plusieurs juges effectuent cette évaluation. Peng et al. (2004) utilisent deux types d'évaluations : à partir d'une transcription orthographique effectuée par trois juges, ils calculent la moyenne des pourcentages des mots correctement transcrits par chaque juge, et une note d'intelligibilité pour chaque phrase sur une échelle de 1 (non intelligible) à 5 (tout à fait intelligible).

- **Autres méthodes**

Dans l'étude de Horga et Liker (2006), l'intelligibilité des voyelles est notée mais la méthode d'évaluation n'est pas précisée. Dans l'étude de Löfqvist (2010), deux évaluations sont utilisées : l'intelligibilité des phrases du corpus présentées dans le bruit est notée par un groupe d'étudiants et un test d'identification de voyelles est effectué par un autre groupe d'étudiants. L'évaluation de l'intelligibilité dans l'étude de Löhle et al. (1999) est réalisée par des jugements perceptifs de la production de mots et de la prosodie de phrases tirées du corpus de parole. Les critères exacts de l'évaluation ne sont pas précisés.

1.3. Conclusion

La comparaison des différentes méthodologies utilisées dans les études sur l'intelligibilité de la parole des enfants porteurs d'implants cochléaires a mis en évidence plusieurs techniques d'élicitation de la parole des enfants, de construction de stimuli, et de jugements perceptifs d'intelligibilité. La parole est élicitée au moyen de conversations, de répétitions de mots ou de phrases, de lecture de phrases ou de textes ou par dénomination d'images. Deux techniques d'évaluation de l'intelligibilité sont privilégiées : la transcription orthographique à partir de laquelle un nombre de segments ou de mots correctement produits est calculé et un système de notes (avec des échelles de notation variables).

Peng et al. (2004) et Habib et al. (2010) ont mis en évidence une forte corrélation entre les deux méthodes (corrélations : 0.91 dans l'étude de Peng et al., 2004 et 0.79 dans l'étude de Habib et al., 2010), qui donnent des résultats comparables. Peng et al. (2004) en concluent qu'en raison de cette très forte corrélation, la méthode de jugement par notes est la plus appropriée pour juger de l'intelligibilité, puisqu'elle est plus rapide et moins fastidieuse que le recours à une transcription.

2. Méthodologie de notre étude d'intelligibilité

Pour élaborer la méthodologie de notre étude, nous nous sommes inspirées à la fois de la revue des méthodes de jugement de l'intelligibilité disponibles dans la littérature et d'une étude pilote (Aubert, 2016), réalisée dans le but de tester plusieurs de ces méthodes, à partir du corpus dont nous disposions.

2.1. Etude pilote

Parallèlement à notre travail, une étude pilote a été réalisée dans le cadre d'un stage et d'un mémoire de fin d'études (Aubert, 2016), réalisés au sein du Gipsa-Lab par Corinne Aubert, étudiante en orthophonie à l'UFR de Médecine et de Pharmacie de Besançon (Université de Franche-Comté). Cette étude pilote a permis de tester en parallèle plusieurs méthodes de jugement de l'intelligibilité de trois énoncés par enfant, tirés de la tâche de narration de notre propre étude de production (voir chapitre 4, paragraphe §1.2.1.4) : deux questions fermées à réponses multiples, et un jugement par transcription orthographique. La procédure était la suivante : les participants à l'étude pilote répondaient à une première question pour chaque énoncé, puis à une deuxième question pour tous les stimuli qui était représentés lors d'une deuxième phase d'écoute, puis il leur était demandé de transcrire orthographiquement une partie des énoncés après une troisième écoute. La première question posée demandait aux auditeurs un jugement de l'intelligibilité des énoncés, c'est-à-dire un jugement lié à la compréhension du message de l'enfant par l'auditeur, alors que la deuxième question posée avait pour but d'obtenir un jugement de la qualité de production de l'enfant par l'auditeur. La perspective était donc différente pour le premier (réception du message) et le deuxième jugement (émission du message).

Deux groupes d'auditeurs ont été recrutés et ont participé à cette étude pilote : un groupe d'auditeurs francophones naïfs, et un groupe d'orthophonistes, ayant l'habitude de la parole pathologique et de la parole de l'enfant. Les résultats de l'étude pilote n'ont pas permis de mettre en évidence un effet de l'expertise, puisque pour les deux questions posées, les réponses des deux groupes n'étaient pas significativement différentes : $t(49.86)=0.06$, $p=0.95$ pour le test-t comparant les réponses des deux groupes pour la première question portant sur la compréhension du message et $t(50)=0.11$, $p=0.92$ pour le test-t comparant les réponses des deux groupes pour la deuxième question portant sur la qualité de production du message.

Pour chaque partie de l'étude pilote, les stimuli étaient présentés systématiquement et automatiquement deux fois à la suite, sans possibilité pour l'auditeur d'avoir un contrôle du lancement du stimulus, avec le logiciel Psychopy (Peirce, 2007). Les retours des participants à l'étude pilote ont permis de montrer que 1) deux écoutes n'étaient pas nécessaires pour chaque stimulus, 2) que les réponses à la deuxième question posée sur les mêmes stimuli pouvaient être biaisées par une familiarisation trop grande avec des stimuli déjà entendus, et 3) la transcription d'énoncés posait des problèmes de traitement.

Prenant en compte les résultats de l'étude pilote, ainsi que ceux des études disponibles dans la littérature, nous avons élaboré une étude qui permet à deux groupes d'auditeurs (naïfs et experts en parole pathologique de l'enfant) d'évaluer l'intelligibilité de la parole des 13 enfants CI ainsi que de 13 enfants NH appariés en âge chronologique de notre étude de production, présentée aux chapitres 4 et 5 ci-dessus. A la différence de l'étude pilote, nous avons fait le choix de poser

une seule question, qui évalue la compréhension du message par l'auditeur. Dans les paragraphes suivants, nous allons détailler le protocole expérimental, la méthodologie de l'analyse statistique et les résultats de notre étude.

2.2. Notre étude

2.2.1. Protocole

2.2.1.1. Question

A la différence de l'étude pilote, nous avons choisi de ne poser qu'une seule question, pour juger l'intelligibilité de la parole produite par les enfants, dans le sens proposé par Svirsky et al. (2000) qui définissent l'intelligibilité comme « […] the accuracy with which a speaker can produce speech that is understandable by others »[31] (cit. in Habib et al., 2010). Notre choix est motivé d'une part par les retours des participants à l'étude pilote (Aubert, 2016, voir ce chapitre, §2.1, ci-dessus) qui évoquent un effet de redondance et un effet de familiarisation trop grand aux énoncés utilisés qui sont les mêmes pour répondre aux deux questions, mais également par la revue des méthodologies utilisées dans la littérature et proposée précédemment, qui montre que les jugements subjectifs de l'intelligibilité se font systématiquement avec une question unique. Enfin, l'étude pilote (Aubert, 2016) ayant montré une forte corrélation entre les réponses aux deux questions (coefficients de corrélation de Pearson : 0.97 ($p<0.01$) pour le groupe des auditeurs experts et 0.96 ($p<0.01$) pour les auditeurs naïfs), il ne nous est pas paru pertinent d'avoir recours à deux questions différentes pour juger de l'intelligibilité de la parole des enfants.

Par ailleurs, nous avons fait le choix d'augmenter le nombre de stimuli proposés pour chaque enfant par rapport à l'étude pilote (Aubert, 2016) : là où l'étude pilote proposait trois stimuli par enfant avec deux questions, nous avons choisi de proposer cinq stimuli par enfant avec une seule question, de façon à garder une durée de passation similaire, ne dépassant pas une vingtaine de minutes, tout en obtenant un plus grand nombre de jugements pour chaque enfant. Nous disposons donc au total de 130 énoncés (5 énoncés pour chacun des 13 enfants CI et des 13 enfants NH).

Enfin, l'étude pilote (Aubert, 2016) comportait un troisième volet au cours duquel il était demandé aux participants une transcription orthographique de certains des énoncés déjà présentés. Il est ressorti de ce troisième jugement une impossibilité de traiter objectivement les transcriptions obtenues. En effet, les énoncés sont tirés d'un corpus de parole semi-spontanée enregistrés au moyen d'une tâche de narration d'un dessin animé (voir chapitre 4, paragraphe §1.2.1.4), qui sont par nature tous différents, et ne peuvent pas être comparés à une cible. La revue de littérature proposée précédemment indique que les types de tâches d'élicitation de la parole enregistrées pour ce type de jugement par transcription (phonétique ou orthographique) sont soit la lecture de phrases ou de mots, soit la répétition de mots d'après modèle. Les transcriptions réalisées dans l'étude pilote (Aubert, 2016) permettent une analyse comparative des différentes transcriptions entre elles et donc une évaluation de niveaux de consensus entre ces différentes transcriptions mais elles ne permettent pas de confronter les transcriptions réalisées avec une cible, qui n'est pas disponible du fait de la nature des productions jugées, assimilables à des productions de parole spontanée.

[31] « […] la précision avec laquelle un locuteur peut produire une parole compréhensible par d'autres » (cit. in Habib et al. (2010), notre traduction)

2.2.1.2. Echelle

Nous avons fait le choix de juger l'intelligibilité au moyen de la question suivante : « Vous devez évaluer l'intelligibilité de ce que vous avez entendu, c'est à dire la faculté de l'enfant à se faire comprendre. Merci de donner une note sur une échelle de 1 à 7 : 1 (pas du tout intelligible/compréhensible) - 7 (tout à fait intelligible/compréhensible) ». Il s'agit d'une question fermée, à choix multiples, fréquemment utilisée en psychologie ou en sociologie expérimentale (Myers & Hansen, 2007). Ce type de questions permet d'obtenir un jugement de chaque énoncé présenté. L'échelle utilisée pour la réponse est une échelle ordinale (Myers & Hansen, 2007), de type « échelle de Likert » (Likert, 1932) à 7 niveaux : ce type d'échelle nous permet d'obtenir un jugement relativement précis. Les échelles de Likert (Likert, 1932) peuvent avoir cinq, sept ou neuf niveaux (Hatch & Lazaraton, 1991), nous avons choisi d'utiliser sept niveaux, qui permettent aux participants de notre étude perceptive de limiter le choix tout en disposant d'une précision suffisante pour juger de l'intelligibilité de la parole des enfants.

2.2.1.3. Interface

A la différence de l'étude pilote (Aubert, 2016), réalisée sur le logiciel Psychopy (Peirce, 2007), nous avons choisi d'utiliser la fonction « ExperimentMFC » sur le logiciel Praat (Boersma & Weenink, 2015) pour réaliser notre propre étude d'intelligibilité : nous avons adapté un scénario d'expérience, disponible dans le manuel utilisateur Praat (Boersma & Weenink, 2015) aux besoins et contraintes de notre expérience d'intelligibilité. Les scénarios que nous avons élaborés pour les deux phases de notre étude d'intelligibilité sont donnés en annexe 6 (phase d'entrainement) et annexe 7 (phrase de test).

La figure 6.1 ci-dessous présente le premier écran de la phase d'entrainement et de la phase de test : le participant lit la question avant de commencer l'expérience.

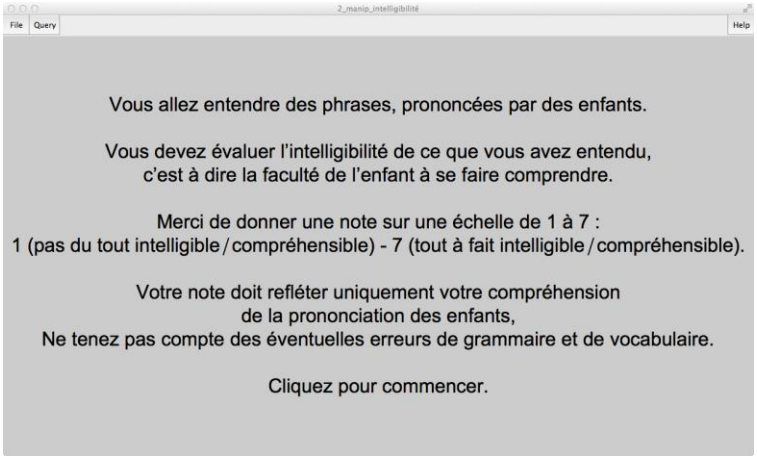


Figure 6.1 : écran d'accueil de l'expérience d'intelligibilité, proposée sur Praat (Boersma & Weenink, 2015)

La figure 6.2 ci-dessous présente l'interface de réponse de la phase d'entrainement et de la phase de test. Un rappel abrégé de la question posée est présenté en haut de la fenêtre, ainsi que l'échelle de notation. Deux boutons sont disponibles : le bouton « Réécouter » permet au participant de bénéficier de deux écoutes supplémentaires de chaque stimulus, et le bouton « noter » permet d'accéder à l'échelle de notation. Une fois la note choisie, un bouton « Suivant » apparaît, qui permet de passer au stimulus suivant.

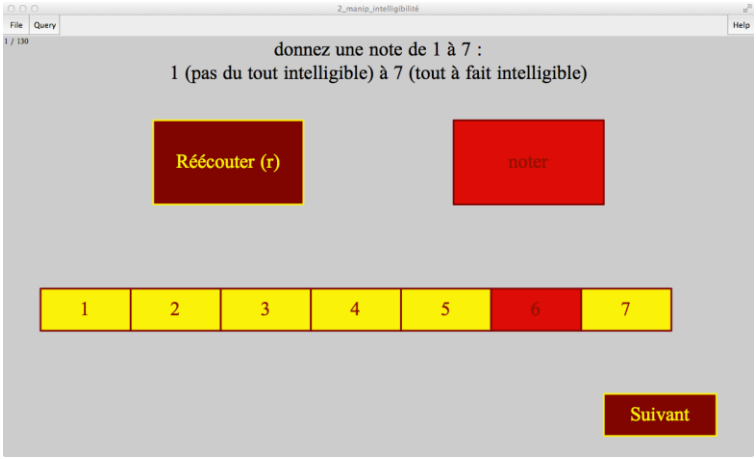


Figure 6.2 : interface de réponse de la phase d'entrainement et de la phase de test d'intelligibilité

Tant que le participant ne clique pas sur le bouton « Suivant », il peut réécouter le stimulus et modifier la note qu'il attribue au stimulus. Lorsqu'il clique sur « Suivant », le participant ne peut plus revenir en arrière, la note ne peut plus être modifiée et le stimulus suivant est présenté. Cette gestion des stimuli à l'initiative du participant nous a semblé plus adéquate pour notre question de recherche que la gestion des stimuli par le logiciel Psychopy (Peirce, 2007), utilisé lors de l'expérience pilote (Aubert, 2016). En effet, lors de cette étude pilote, deux écoutes successives obligatoires de chaque stimulus étaient proposées au participant, l'attribution de la note était définitive et provoquait la présentation du stimulus suivant. L'utilisation de Praat (Boersma & Weenink, 2015) donne une plus grande liberté au participant dans la gestion de la présentation des stimuli lors de l'étude perceptive d'intelligibilité.

Lors de la phase de test, les 130 stimuli étaient présentés en ordre aléatoire, en cinq blocs de 26 énoncés, séparés par une pause. Chaque stimulus n'était présenté qu'une seule fois. De même qu'il choisissait de valider sa réponse et de passer au stimulus suivant, le participant avait la possibilité de choisir la durée de chaque pause par un clic, qui lui permettait de continuer l'expérience.

Les résultats sont collectés sur Praat (Boersma & Weenink, 2015), directement après la passation de chaque participant et enregistrés au format .txt. Nous obtenons ainsi pour chaque stimulus entendu une note d'intelligibilité, et un temps de réaction que nous avons choisi de ne pas utiliser, puisque l'étude est réalisée sans aucune contrainte de temps.

2.2.1.4. Conditions de passation

L'expérience a été réalisée au domicile des participants ou sur leur lieu de travail, dans une pièce calme. Les participants ont tous passé l'expérience de perception sur un ordinateur portable, équipé d'un casque et d'une souris externe. Nous avons choisi de ne pas faire visionner le dessin animé par les participants, pour qu'ils puissent véritablement évaluer leur capacité à comprendre le message, sans qu'ils l'interprètent à partir de leur propre connaissance du dessin animé.

2.2.2. Stimuli

2.2.2.1. Groupes d'enfants

Pour cette expérience d'intelligibilité, nous avons utilisé les productions enregistrées d'enfants CI et NH issues de la tâche de narration de notre première étude de production (voir chapitre 4), enregistrées lors de l'étude de Scarbel (2012) pour la phase d'entrainement, et lors de notre

propre étude de production (voir chapitre 4, paragraphe §1.2.1.4) pour la phase de test. Il s'agit d'enregistrement de parole semi-spontanée : les enfants visionnaient un dessin animé divisé en huit séquences d'une vingtaine de secondes, puis devaient raconter à l'expérimentatrice tout ce qu'ils avaient vu dans chaque séquences. Pour que nos deux groupes de sujets soient les plus homogènes possibles, nous avons choisi les 13 enfants NH dont l'âge chronologique était le plus proche des 13 enfants du groupe CI. Les deux groupes ainsi constitués sont appariés en âge chronologique (t(23.954)=0.039, p=0.9692).

ENFANTS SOURDS, PORTEURS D'IMPLANT COCHLÉAIRE						**ENFANTS NORMO-ENTENDANTS**			
Enfant	Age (en mois)	Age à l'implantation (en mois)	Durée d'utilisation de l'implant (en mois)	Sexe	Classe	Enfant	Age (en mois)	Sexe	Classe
CI13	78	25	53	M	GSM	NH32	77	M	CP
CI19	80	47	33	M	CE1	NH25	82	M	CP
CI14	82	18	64	F	CP	NH16	84	F	CE1
CI21	85	13	72	F	CE1	NH30	86	M	CP
CI10	92	54	38	F	CE1	NH14	92	F	CE1
CI7	95	19	76	M	CE1	NH21	94	F	CE1
CI12	95	52	43	F	CE1	NH23	95	F	CE1
CI6	96	70	26	M	CE1	NH33	99	F	CE1
CI22	106	45	91	M	CM1	NH20	107	F	CE2
CI17	110	78	32	F	CE2	NH26	108	M	CM1
CI16	112	23	89	M	CE1	NH19	113	M	CM1
CI11	118	26	92	M	CM1	NH29	116	M	CM1
CI20	127	18	109	F	CM2	NH34	126	F	CM2

Table 6.4 : Enfants des deux groupes (CI et NH) dont les productions ont été sélectionnées pour l'étude d'intelligibilité

2.2.2.2.Choix des énoncés

Le corpus que nous utilisons est un corpus de parole semi-spontanée, qui nous permet de disposer de productions assimilables à une parole naturelle, mais les conditions contrôlées nous permettent d'obtenir un contenu similaire chez tous les enfants. Le choix des mots est différent pour chaque enfant, mais le contenu du message est globalement similaire chez tous.

- **Choix des énoncés pour la phase d'entrainement**

Notre phase de test est précédée d'un entrainement. Le but de cet entrainement est de familiariser le participant à la question posée, au style de parole, à la modalité de réponse et à l'interface de test. Il permet également au participant de régler le volume du son pour la phase de test. Pour cet entrainement, nous avons choisi des énoncés tirés de la tâche de narration de dessin animé et enregistrés lors de l'étude de Scarbel (2012), de façon à ne pas utiliser des enregistrements des mêmes enfants que dans la phase de test. Nous avons choisi deux énoncés produits par deux enfants sourds porteurs d'implant cochléaire et deux enfants normo-entendants. Cette phase d'entrainement comporte donc huit énoncés au total, qui ont été traités de la même façon que ceux qui sont utilisés dans la phase de test (voir la section « Traitement des stimuli » ci-dessous).

- **Choix des énoncés pour la phase de test**

Nous avons choisi cinq énoncés par enfant, de façon systématique : la tâche de narration étant divisée en huit séquences, nous avons sélectionné les énoncés sur les troisième, quatrième, cinquième, sixième et septième séquences. Ce choix nous permet d'éviter les premières séquences où les enfants ne sont pas encore à l'aise et peuvent hésiter à parler et d'éviter la fin de la tâche de narration, où les enfants peuvent être soumis à une certaine fatigue.

Le choix des énoncés se fait de la façon suivante : nous avons sélectionné le premier énoncé de la séquence, qui comportait minimalement un sujet (pronom, syntagme nominal), un verbe (un seul mot ou un syntagme verbal) et un complément (qui peut être un syntagme nominal, un syntagme adjectival ou une courte subordonnée). Nous définissons un énoncé comme une production isolable, précédée et suivie dans la mesure du possible d'une pause, et ayant un contour terminal assimilable à ceux de la finalité, de la continuation majeure, de l'implication ou de l'exclamation, décrits par Delattre (1966). L'utilisation de pauses pour segmenter le flux de parole étant différente d'un enfant à l'autre, nous avons dû, dans certains cas, sélectionner des énoncés qui n'étaient pas isolés par des pauses, mais que nous avons jugés indépendants de l'énoncé précédent et de l'énoncé suivant. Les énoncés sélectionnés correspondent donc au groupe intonatif (« Intonational Phrase » ou « IP »), décrit dans Jun & Fougeron (2000).

2.2.2.3.Traitement des stimuli

Les stimuli sont tous enregistrés au format .wav (échantillonnés à 44 100 Hz, mono, 16 bits). La procédure et le matériel utilisé sont détaillés au chapitre 4, ci-dessus (paragraphe §1.3). Nous avons gardé les caractéristiques des fichiers .wav, en normalisant l'intensité des fichiers sons que nous utilisons pour notre étude perceptive : nous avons utilisé la fonction « Scale intensity » sur Praat (Boersma & Weenink, 2015), en choisissant, pour chaque stimulus, une intensité à 60 dB au dessus du seuil normal d'audition.

2.2.3. Participants

Comme pour l'étude pilote, les participants de notre étude d'intelligibilité sont un premier groupe d'auditeurs experts (huit orthophonistes et une enseignante spécialisée, en classe CLIS), et un groupe d'auditeurs naïfs (17 auditeurs : neuf hommes et huit femmes). Tous les participants ont déclaré n'avoir aucun trouble de l'audition diagnostiqué. L'âge moyen de chaque groupe est donné dans la table 6.4 ci-dessous.

Groupe	Nombre d'auditeurs	Âge moyen en années (avec écart-type)
Auditeurs experts	9 F	30;3 (40)
Auditeurs naïfs	17 (8F, 9H)	28;8 (5;11)

Table 6.5 : âge moyen (avec écart-type) des participants à l'étude d'intelligibilité

Tous les participants du groupe des auditeurs naïfs sont de langue maternelle française : quatorze d'entre eux sont monolingues et trois ont déclaré avoir une deuxième L1 (Créole guadeloupéen, italien et tamoul). Ils étaient domiciliés en Isère (14), Val-de-Marne (1), Bas-Rhin (1) et Loir-et-Cher (1) au moment de la passation de l'expérience.

Tous les participants du groupe des auditeurs experts sont monolingues, de langue maternelle française, et étaient domiciliés en Isère (9) au moment de la passation de l'expérience.

2.3. Analyses statistiques

2.3.1. Analyses préliminaires : différence entre auditeurs naïfs et auditeurs experts et différences intra-sujet

Plusieurs références nous ont permis de sélectionner les tests statistiques les plus adaptés à nos données de jugements d'intelligibilité, dont les distributions ne sont pas normales et les variables peuvent être catégorielles et discrètes : Hatch & Lazaraton (1991), Howell et al. (2008), Grosjean & Dommergues (2011) et Welkowitz et al. (2012).

Dans un premier temps, nous cherchons à savoir s'il existe une différence inter-groupes entre les notes données par les auditeurs experts et les auditeurs naïfs.

La figure 6.3 ci-dessous présente la répartition des notes données par chaque groupe d'auditeurs.

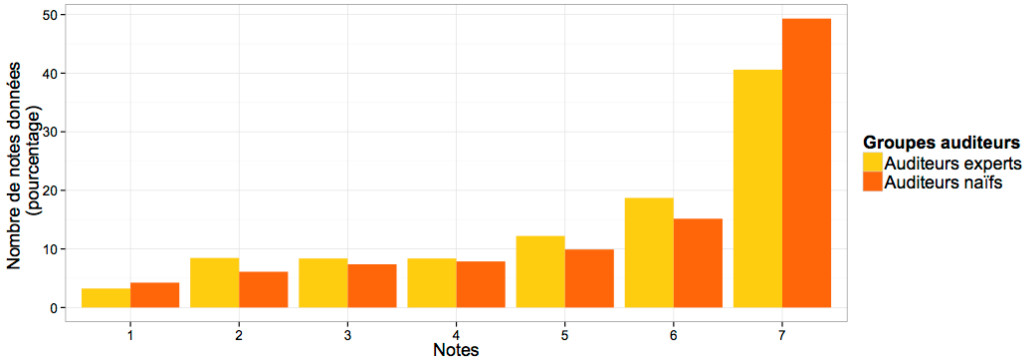


Figure 6.3 : choix des notes d'intelligibilité données par chaque groupe d'auditeurs (en pourcentage)

La figure 6.3 montre que les notes données pour chaque catégorie ne varient que très peu entre les deux groupes d'auditeurs : pour les réponses 1, 2, 3, 4, 5 et 6 en particulier, les pourcentages de notes données par les deux groupes d'auditeurs sont très proches les uns des autres.

La note attribuée est une variable ordinale, discrète mais dans cette analyse préliminaire, nous l'avons cependant traitée comme une variable continue, pour pouvoir comparer les notes données par les deux groupes. A cet effet, nous avons calculé une moyenne par stimulus pour chaque groupe d'auditeurs et nous avons étudié le lien entre ces deux moyennes. La moyenne n'est pas envisagée dans ce cadre comme une valeur numérique, mais un moyen de hiérarchiser les stimuli et de les classer selon le niveau d'intelligibilité établi par chaque groupe.

Nous avons utilisé un test de corrélation non-paramétrique de Kendall, qui crée pour chaque groupe un classement des stimuli du plus intelligible (moyenne la plus haute) au moins intelligible (moyenne la plus basse) et qui étudie la corrélation entre les deux groupes à partir de ce classement, à la différence d'une corrélation de Pearson qui teste la relation linéaire entre des moyennes. Nous avons préféré la corrélation des rangs de Kendall à la corrélation de Spearman, en raison de valeurs moyennes qui peuvent être ex-aequo et qui ne sont pas traitées par la corrélation de Spearman. Le test de Kendall donne un résultat τ (tau) qui est compris entre -1 et 1, où 0 correspond à une absence de corrélation et -1/+1 à une corrélation maximale, et la valeur p qui indique si la corrélation éventuellement observée est significative. La figure 6.4 ci-dessous présente cette relation entre les notes moyennes données par chaque groupe d'auditeurs.

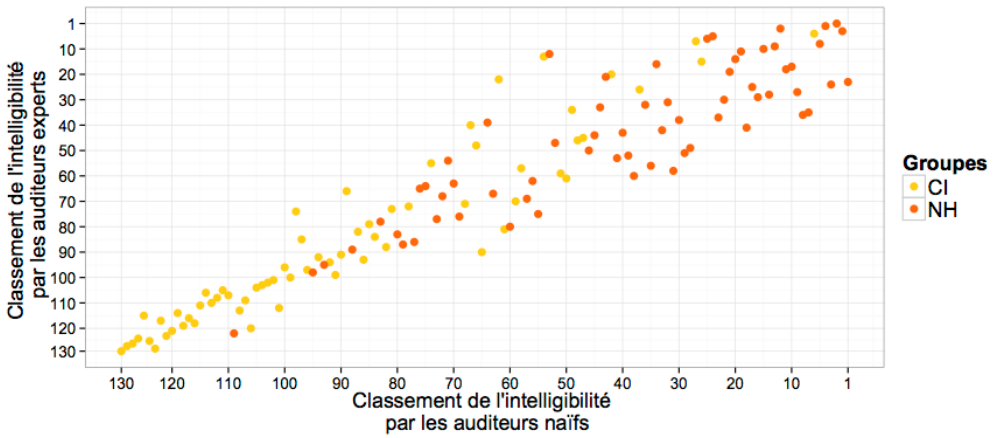


Figure 6.4 : classement des stimuli par niveau d'intelligibilité des groupes CI et NH pour chaque groupe d'auditeurs (1=énoncé le plus intelligible et 130=énoncé le moins intelligible)

Le test de corrélation de Kendall indique que les notes données par les deux groupes sont corrélées : $\tau = 0{,}803$, $p<.001$. Nous pouvons déduire de ce test préliminaire que les jugements ne sont pas différents d'un groupe à l'autre et nous allons donc traiter les résultats de notre étude sans distinguer les deux groupes d'auditeurs. Nous avons donc recalculé une note moyenne du groupe de tous les auditeurs (N=26) pour chaque stimulus.

Comme nous avons cinq stimuli et donc cinq moyennes par enfant, nous avons réalisé un test des rangs de Friedman, pour étudier la variance intra-enfant. Ce test indique que les notes moyennes par stimulus ne sont pas significativement différentes pour chaque enfant : $\chi^2(4)=5{,}96$, $p>.01$.

2.3.2. Analyses statistiques

Les auditeurs ont jugé l'intelligibilité de la production des enfants CI et NH en choisissant une note sur une échelle de type échelle de Likert, allant de 1 (pas du tout intelligible) à 7 (tout à fait intelligible). La réponse choisie est donc catégorielle, puisqu'elle correspond à un jugement, codé par un chiffre et non à une mesure continue. Une analyse de variance avec un test paramétrique de type t-test ou ANOVA n'est pas possible dans ce cas : les notes données et les notes moyennes calculées pour chaque enfant permettent d'établir si un stimulus ou un enfant est jugé plus intelligible qu'un autre, mais la distance entre des notes moyennes par groupe n'est pas pertinente.

Dans un premier temps, nous allons donc analyser les résultats obtenus en utilisant un test de khi-carré, pour étudier la distribution de chaque catégorie de réponse (1, 2, 3, 4, 5, 6 et 7), pour chaque groupe d'enfants (enfants CI ou NH) pour comprendre si l'utilisation de l'implant influence la note donnée et donc l'intelligibilité perçue par les auditeurs.

Dans un deuxième temps, nous nous intéresserons à l'effet de l'âge chronologique sur l'intelligibilité des enfants NH, en utilisant un test de corrélation de Kendall qui est un test non-paramétrique. Ce test de corrélation permet de classer les stimuli les uns par rapport aux autres en fonction de la note moyenne, et de classer les enfants en fonction de leur âge chronologique puis le test compare l'ordre des notes moyennes d'intelligibilité et l'ordre de l'âge chronologique. Ce test permet donc d'établir si la note d'intelligibilité et l'âge des enfants sont corrélés.

Enfin, nous étudierons l'effet de l'âge à l'implantation et de la durée d'utilisation de l'implant (ou âge auditif) sur les notes moyennes d'intelligibilité des enfants CI. Nous souhaitons faire le même

type d'analyse que pour l'effet de l'âge chronologique sur les notes moyennes d'intelligibilité chez les enfants NH. Nous utiliserons donc les mêmes tests de corrélation de Kendall pour tester la corrélation entre l'âge à l'implantation et l'intelligibilité d'une part, et la corrélation entre l'âge auditif et l'intelligibilité d'autre part.

3. Résultats

3.1. Effet du groupe

Les distributions des notes d'intelligibilité données par tous les auditeurs (N=26) pour les cinq stimuli de chaque enfant sont données dans la table 6.6 ci-dessous. Nous avons indiqué en couleur le mode de la distribution de chaque enfant.

Enfants	Distribution des notes des enfants CI						
	1	2	3	4	5	6	7
CI10	6	27	33	26	17	16	5
CI11	1	1	10	16	21	35	46
CI12	6	20	23	25	19	12	25
CI13	57	32	16	13	7	4	1
CI14	0	1	9	9	16	29	66
CI16	1	20	29	21	26	18	15
CI17	0	0	0	3	14	23	90
CI19	16	26	27	26	19	11	5
CI20	0	5	9	10	9	26	71
CI21	8	15	24	8	8	22	45
CI22	27	53	19	14	10	5	2
CI6	2	10	16	15	19	29	39
CI7	0	5	9	17	22	22	55

Enfants	Distribution des notes des enfants NH						
	1	2	3	4	5	6	7
NH14	0	0	0	0	1	16	113
NH16	0	4	8	10	12	25	71
NH19	0	0	2	3	9	10	106
NH20	1	2	2	3	15	26	81
NH21	0	0	0	0	7	16	107
NH23	0	0	1	0	10	17	102
NH25	0	2	6	10	19	41	52
NH26	0	1	6	10	17	25	71
NH29	7	9	2	6	6	18	82
NH30	0	0	1	0	12	21	96
NH32	0	0	4	10	16	27	73
NH33	0	1	5	9	16	38	61
NH34	0	0	0	8	15	22	85

Table 6.6 : distribution des notes attribuées pour chaque enfant, des groupes CI (gauche) et NH (droite), le mode pour chaque enfant est indiqué en couleur

La figure 6.5 présente la distribution des notes d'intelligibilité données pour chaque enfant par tous les auditeurs.

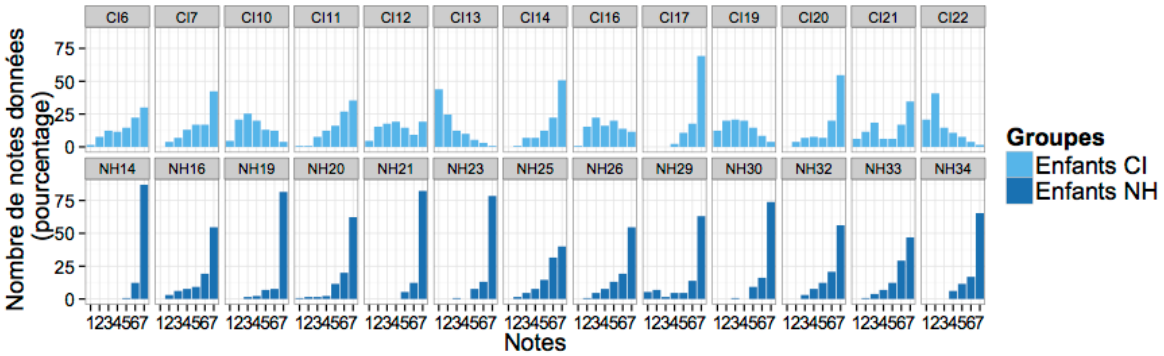


Figure 6.5 : distribution des notes d'intelligibilité pour chaque enfant du groupe CI (haut) et NH (bas)

Les distributions des notes d'intelligibilité données à chaque enfant présentées dans la table 6.6 et la figure 6.5 indiquent que la note de 7 est la plus fréquemment donnée pour juger de l'intelligibilité des enfants NH : dans les distributions données dans la table 6.6, tous les enfants NH ont un mode de 7, certaines catégories (1, 2, 3 et 4) étant très peu choisies par les auditeurs et certaines (5, 6 et 7) étant majoritairement choisies par les auditeurs. Ces résultats indiquent que les enfants NH sont clairement jugés intelligibles par les auditeurs lors de cette expérience. Si les

distributions de certains enfants (NH25, NH26 et NH33) sont plus aplaties que celles d'autres enfants, le mode est également de 7 pour ces trois enfants.

Les distributions des enfants CI sont beaucoup moins homogènes. En effet, trois patrons sont observés : les distributions de sept enfants CI sur treize ont un mode de 7, celle d'un seul enfant (CI12) a un mode de 4 qui correspond à une intelligibilité moyenne, et celles de cinq enfants (CI10, CI13, CI16, CI19 et CI22) ont un mode de 1, 2 ou 3, ces derniers étant donc jugés très peu intelligibles par les auditeurs lors de cette expérience.

La figure 6.6 ci-dessous présente la distribution des notes d'intelligibilité attribuées à chaque groupe. Une première observation de cette figure semble confirmer les répartitions des notes données à chaque enfant : la note de 7 est la plus fréquemment donnée pour les deux groupes d'enfants mais les notes données aux enfants NH sont quasiment toutes réparties entre les notes 5, 6 et 7, alors que celles des enfants CI sont partagées entre toutes les notes possibles.

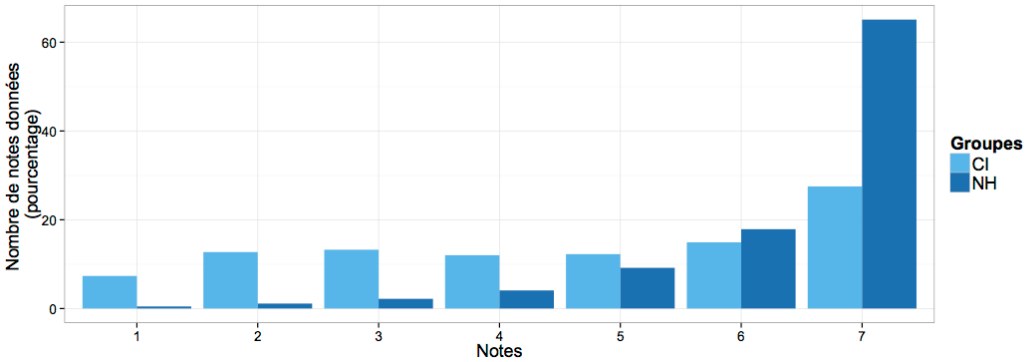


Figure 6.6 : distribution des notes d'intelligibilité par groupe (CI et NH)

Comme nous cherchons à mettre en évidence un éventuel effet de groupe entre les enfants CI et NH, nous avons réalisé un test de khi-carré avec la variable dépendante « nombre de mesures pour chaque catégorie de notes (de 1 à 7) » et la variable indépendante « groupe (CI vs NH) ». Le test indique que les distributions des deux groupes sont significativement différentes : $\chi^2(6) = 735{,}74$, $p<.001$, ce qui confirme l'observation visuelle de la figure 6.6 et la plus grande intelligibilité globale des enfants NH par rapport aux enfants CI.

Comme les notes supérieures à 4 correspondent à une bonne intelligibilité, nous avons calculé, pour chaque enfant, la proportion des notes de 5, 6 et 7 par rapport au total des notes reçues par cet enfant. Les résultats sont présentés dans la table 6.7 ci-dessous.

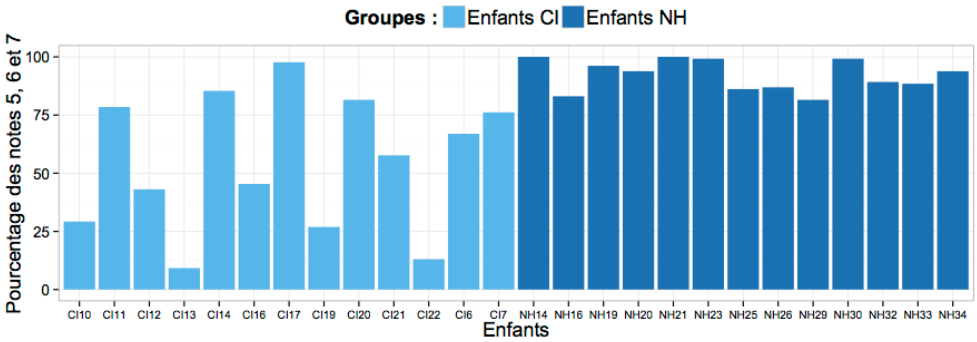


Figure 6.7 : pourcentage des notes 5, 6 et 7 données à chaque enfant des deux groupes (CI et NH)

Nous observons que tous les enfants NH ont reçu une majorité de notes supérieures à 4, puisque la proportion de ces notes est dépasse les 75% pour tous les enfants, et deux enfants (NH14 et NH21) atteignent même 100% de notes supérieures à 4. En revanche, seuls 7 enfants CI sur 13 obtiennent une majorité de notes supérieures à 4 (CI6, CI7, CI11, CI14, CI17, CI20 et CI21) et sont donc jugés intelligibles. Cette mesure de l'intelligibilité, calculée à partir de la proportion des notes strictement supérieures à 4 correspond à un indice d'intelligibilité, qui nous permet de compléter nos analyses.

3.2. Enfants NH : effet de l'âge chronologique

La figure 6.7 présente les résultats pour chaque enfant NH, par ordre d'âge chronologique.

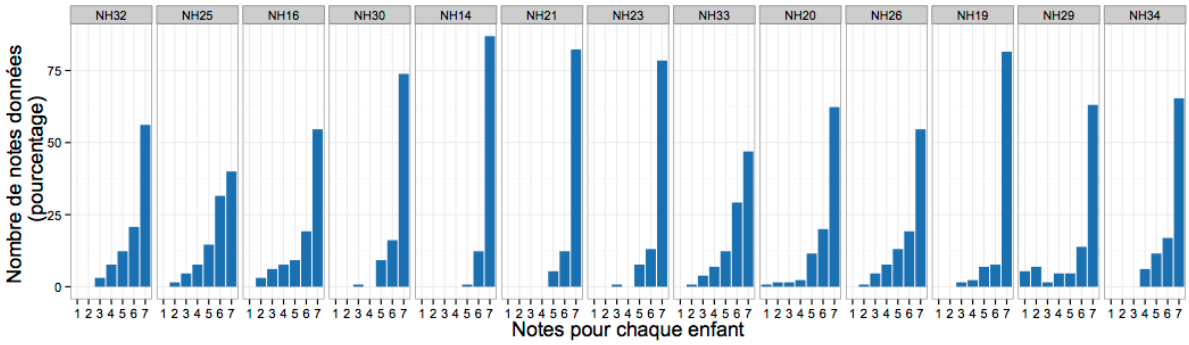


Figure 6.8 : notes d'intelligibilité obtenues par chaque enfant NH, par ordre d'âge chronologique (de gauche à droite : du plus jeune au plus âgé)

A première vue, la figure 6.8 ne semble pas indiquer d'effet global de l'âge : tous les enfants NH obtiennent une majorité de notes réparties entre 5, 6 et 7.

Le test statistique (corrélation de Kendall) confirme cette observation, puisque ce test indique une absence de corrélation entre âge chronologique et note moyenne donnée pour chaque stimulus : $\tau = 0{,}065$, $p>.05$.

Pour compléter l'analyse de d'effet de l'âge chronologique sur l'intelligibilité des enfants NH, nous avons ensuite utilisé un test de corrélation de Pearson réalisé à partir du pourcentage de notes supérieures à 4 (qui correspond une bonne intelligibilité), calculé précédemment (voir paragraphe §3.1) et de l'âge chronologique. Le test indique une absence de corrélation significative ($r=-0{,}02$, $p>.05$) entre la bonne intelligibilité des enfants NH et leur âge chronologique.

3.3. Enfants CI : effet de l'âge à l'implantation

La figure 6.9 présente les résultats pour chaque enfant CI, classés par ordre d'âge à l'implantation (du plus précoce au plus tardif).

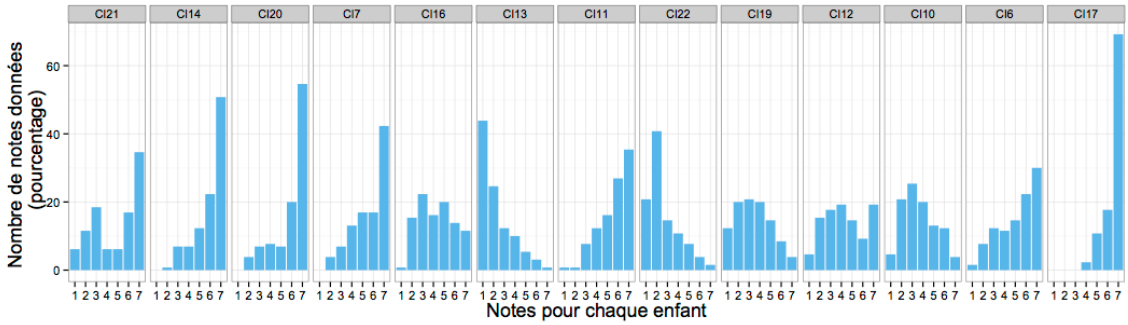


Figure 6.9 : notes d'intelligibilité obtenues par chaque enfant CI, par ordre d'âge à l'implantation (de gauche à droite : de l'âge à l'implantation le plus précoce au plus tardif)

La figure 6.9 semble indiquer que les quatre enfants qui ont reçu l'implant le plus tôt (CI21 à 1;1 ans, CI14 et CI20 à 1;6 ans et CI7 à 1;7 ans) obtiennent majoritairement les notes de 5, 6 et 7, alors que les enfants qui ont reçu l'implant plus tardivement (CI16 à 1;11 ans, CI13 à 2;1 ans, CI11 à 2;2 ans, CI22 à 3;9 ans, CI19 à 3;11 ans, CI12 à 4;4 ans, CI10 à 4;6 ans, CI6 5;10 ans à et CI17 à 6;6 ans) obtiennent des résultats beaucoup plus disparates, avec des distributions de notes réparties entre toutes les notes, ou localisées sur les notes les plus basses 1, 2 et 3. Seuls trois enfants (CI22, CI6 et CI17) sur les neuf enfants ayant reçu l'implant après 19 mois ont des distributions de notes localisées sur les notes de 5, 6 et 7.

Un test de corrélation de Kendall réalisé sur les notes moyennes par stimulus n'indique pas de corrélation entre la note moyenne par stimulus et l'âge à l'implantation : $\tau = -0{,}075$, $p>.05$.

Nous avons également réalisé un test de corrélation de Pearson entre le pourcentage des notes strictement supérieures à 4 calculé précédemment (paragraphe §3.1) et l'âge à l'implantation. Le résultat de ce test indique une absence de corrélation entre notre indice de bonne intelligibilité et l'âge à l'implantation ($r=-0{,}03$, $p>.05$).

L'âge de dix-huit mois à deux ans est régulièrement cité dans la littérature comme un âge critique pour l'implantation, après lequel les enfants auraient plus de difficultés à accéder au langage (voir entre autres Le Normand, 2004). A la lumière de l'apparente dichotomie entre enfants ayant reçu l'implant avant 20 mois et ceux l'ayant reçu plus tardivement, nous avons donc regroupé les enfants CI par âge à l'implantation, avec un premier groupe d'enfants ayant reçu l'implant avant 20 mois (ou 1;8 ans), et un deuxième groupe l'ayant reçu après 20 mois. Un test de khi-carré avec la variable dépendante « nombre de mesures pour chaque catégorie de notes (de 1 à 7) » et la variable indépendante « groupe d'enfants CI (implantation avant 20 mois vs après 20 mois) ». Le test indique que les distributions des deux groupes sont significativement différentes : $\chi^2(6) = 183{,}59$, $p<.001$, ce qui confirme l'observation visuelle : les enfants ayant reçu un implant de façon plus précoce obtiennent des résultats significativement meilleurs que ceux qui l'ont reçu plus tardivement.

3.4. Enfants CI : effet de la durée d'utilisation de l'implant

La figure 6.10 présente les distributions des notes données aux enfants CI, classés dans l'ordre de durée d'utilisation de l'implant (de la durée la plus courte à la plus longue).

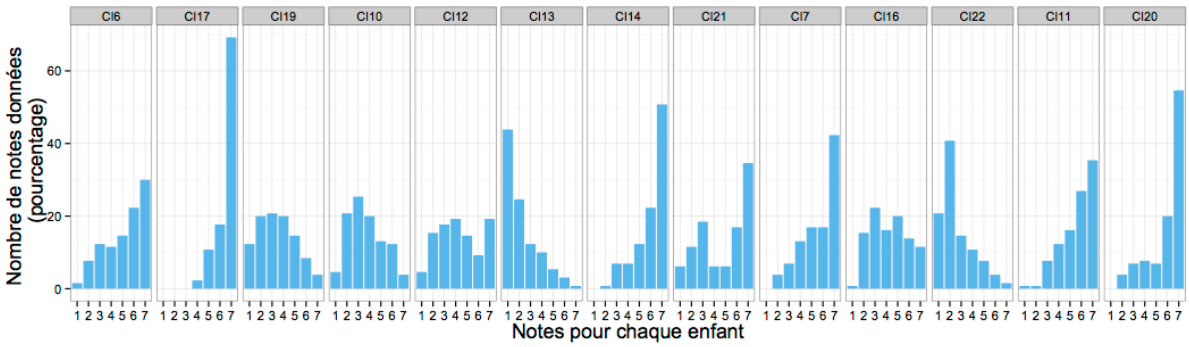


Figure 6.10 : notes d'intelligibilité obtenues par chaque enfant CI, par ordre de la durée d'utilisation de l'implant (de gauche à droite : de la durée la plus courte à la plus longue)

A la différence de l'âge à l'implantation, nous n'observons pas sur la figure 6.9 une distinction claire entre les enfants les plus intelligibles et les moins intelligibles en lien avec la durée d'utilisation de l'implant. En effet, les enfants les plus intelligibles ne sont pas ceux qui ont utilisé l'implant le plus longtemps et les moins intelligibles ne sont pas ceux qui l'ont utilisé moins longtemps : les enfants CI6 (durée d'utilisation de l'implant de 2;2 ans) et CI17 (2;8 ans) sont en effet jugés intelligibles avec des notes d'intelligibilité majoritairement de 5, 6 et 6, alors qu'ils utilisent leur implant depuis peu de temps, puis les enfants CI19 (2;11 ans), CI10 (3;2 ans) et CI12 (3 ;7 ans) sont jugés moyennement intelligibles, avec une majorité de notes comprises entre 2 et 5, puis l'enfant CI13 (4;5 ans) est jugé peu intelligible avec des notes entre 1 et 4, l'enfant CI14 (5;4 ans), CI21 (6;0 ans) et CI7 (6;4 ans) sont à nouveau jugés intelligibles, l'enfant CI16 (7;5 ans) est jugé moyennement intelligible, l'enfant CI22 (7 ;7 ans) peu intelligible, enfin les enfants CI11 (7;8 ans) et CI20 (9;1 ans) sont quant à eux jugés tout à fait intelligibles et sont les enfants qui utilisent leur implant depuis le plus longtemps.

Un test de corrélation de Kendall réalisé sur les notes moyennes par stimulus n'indique pas de corrélation entre la note moyenne par stimulus et la durée d'utilisation de l'implant : $\tau = 0{,}013$, $p>.05$.

Le test de corrélation de Pearson réalisé à partir de l'indice d'intelligibilité présenté au paragraphe §3.1 indique que le pourcentage des notes strictement supérieures à 4, obtenues par chaque enfant n'est pas corrélé avec la durée d'utilisation de l'implant ($r=0{,}13$, $p>.05$).

A la différence de l'âge à l'implantation, les patrons observés pour la durée d'utilisation de l'implant ne nous permettent pas de regrouper les enfants par groupes en fonction de la durée d'utilisation de l'implant, pour pouvoir préciser ces résultats et les observations visuelles des résultats.

3.5. Enfants CI : effet de l'âge chronologique

La figure 6.11 présente les distributions des notes données aux enfants CI, classés dans l'ordre de leur âge chronologique (de l'enfant le plus jeune au plus âgé au moment des enregistrements).

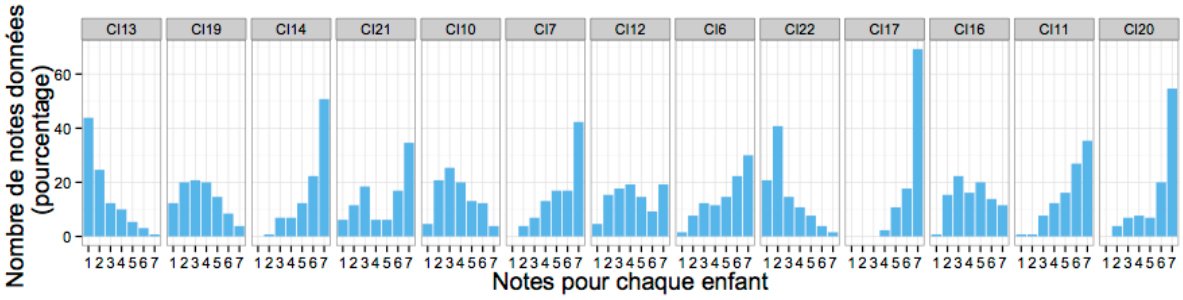


Figure 6.11 : notes d'intelligibilité obtenues par chaque enfant CI, par ordre d'âge chronologique (de gauche à droite : de l'enfant le plus jeune au plus âgé)

Comme pour la durée d'utilisation de l'implant, aucun patron clair d'intelligibilité se dégage : les enfants les plus intelligibles (CI14, CI21, CI7, CI6, CI17, CI11 et CI20) ont respectivement 6;10, 7;1, 7;11, 8;0, 9;2 ,9;1 et 10;7 ans alors que les enfants les moins intelligibles (CI13, CI19, CI10, CI12, CI22 et CI16) ont respectivement 6;6, 6;8, 7;8, 7;11, 8;1 et 9;4 ans.

Un test de corrélation de Kendall réalisé sur les notes moyennes par stimulus indique une corrélation faible (car proche plus de 0 que de 1), mais significative entre la note moyenne par stimulus et l'âge chronologique : $\tau = 0{,}2605$, $p<.01$.

Le résultat du test de corrélation de Pearson indique en revanche une absence de corrélation significative entre l'indice d'intelligibilité (correspondant au pourcentage de notes strictement supérieures à 4) et l'âge chronologique des enfants CI ($r=0{,}42$, $p>.05$).

L'intelligibilité des enfants CI semble donc être dans une certaine mesure liée à l'âge chronologique des enfants, ce qui n'était pas le cas chez les enfants NH, qui étaient tous jugés intelligibles. Nos résultats sur l'effet de l'âge à l'implantation sur l'intelligibilité et de la corrélation entre âge chronologique et intelligibilité semblent donc indiquer que ces deux facteurs peuvent expliquer la meilleure intelligibilité de certains enfants.

4. Discussion

La revue d'études d'intelligibilité présentée au chapitre 3 a montré que les jugements d'intelligibilité peuvent être réalisés par des groupes d'auditeurs naïfs ou experts (entre autres étudiants en orthophonie, orthophonistes, chercheurs) mais aucune étude ne s'est intéressée à la différence de jugement entre auditeurs naïfs et auditeurs experts. Nos résultats indiquent clairement que les jugements de l'intelligibilité ne sont pas liés au niveau d'expertise des participants, ce qui confirme les résultats de l'étude pilote (Aubert, 2016) réalisée préalablement à notre propre étude : les notes attribuées ne sont pas liées au groupe auquel l'auditeur appartient, ce qui confirme que les auditeurs ayant une grande habitude de la voix et de la parole pathologique n'ont pas une sensibilité plus grande du fait de cette expertise, puisque leur jugement de l'intelligibilité n'est pas différent de celui des auditeurs tout-venants.

4.1. Liens entre caractéristiques de l'implantation cochléaire et intelligibilité

Notre étude d'intelligibilité a mis en évidence que les enfants CI sont moins intelligibles que les enfants NH, ce qui est en accord avec l'étude de Flipsen et al. (2006) qui indiquaient que les enfants NH sont intelligibles à 100 % dès l'âge de quatre ans alors que les enfants CI n'atteignent ce niveau d'intelligibilité que plus tardivement et plus lentement. L'étude de Flipsen et al. (2006) utilisait des productions de parole spontanée, ce qui est également le cas de notre étude, mais Flipsen et al. (2006) s'appuyaient sur des transcriptions phonétiques d'échantillons de conversations par des étudiants, là où notre étude utilise des jugements perceptifs par des auditeurs naïfs et des auditeurs experts.

Nos résultats sont en accord avec ceux de Calmels et al. (2004) et Löhle et al. (1999) puisque, à partir d'échantillons de parole spontanée, ces études mettent en évidence un lien entre un âge à l'implantation précoce et les jugements d'intelligibilité, et avec ceux de Habib et al. (2010), Montag et al. (2014) et Peng et al. (2004) qui obtiennent des résultats équivalents avec des échantillons de parole obtenus lors d'une tâche de répétition de mots ou de phrases. Ils se distinguent en revanche de ceux de Blamey et al. (2001) et Khwaileh & Flipsen (2010), qui observent une absence de corrélation entre l'âge à l'implantation et l'intelligibilité des enfants CI. Par ailleurs, nos résultats semblent indiquer qu'il existe un âge à l'implantation critique compris entre un an et demi et deux ans, en dessous duquel les enfants arrivent à avoir un bon niveau d'intelligibilité, ce qui confirme les résultats de Löhle et al. (1999), puisque dans cette étude, les enfants ayant été implantés entre deux et quatre ans sont jugés plus intelligibles que ceux qui ont reçu l'implant entre cinq et huit ans, puis entre neuf et 14 ans.

En outre, nous ne mettons pas en évidence de corrélation entre durée d'utilisation de l'implant et intelligibilité, à la différence de toutes les études qui étudiaient l'effet de ce facteur sur l'intelligibilité : les études de Khwaileh & Flipsen (2010), Miyamoto et al. (1997), Svirsky (2000) qui utilisent des répétitions de mots ou de phrases et celles de Blamey et al. (2001) et Flipsen et al. (2006) des productions de parole spontanée.

4.2. Liens entre caractéristiques acoustiques et intelligibilité

Les résultats de notre étude acoustique de la production de parole par les enfants CI et NH (Deuxième Partie, chapitres 5 et 6) lors d'une tâche de répétition de mots ont mis en évidence des différences de production des voyelles, des consonnes et de séquences CV entre les deux groupes d'enfants. La figure 6.12 présente conjointement les résultats de l'étude d'intelligibilité ainsi que

certaines des caractéristiques acoustiques étudiées aux chapitres 4 et 5 (espace vocalique F1-F2, VOT des occlusives et centre de gravité des fricatives).

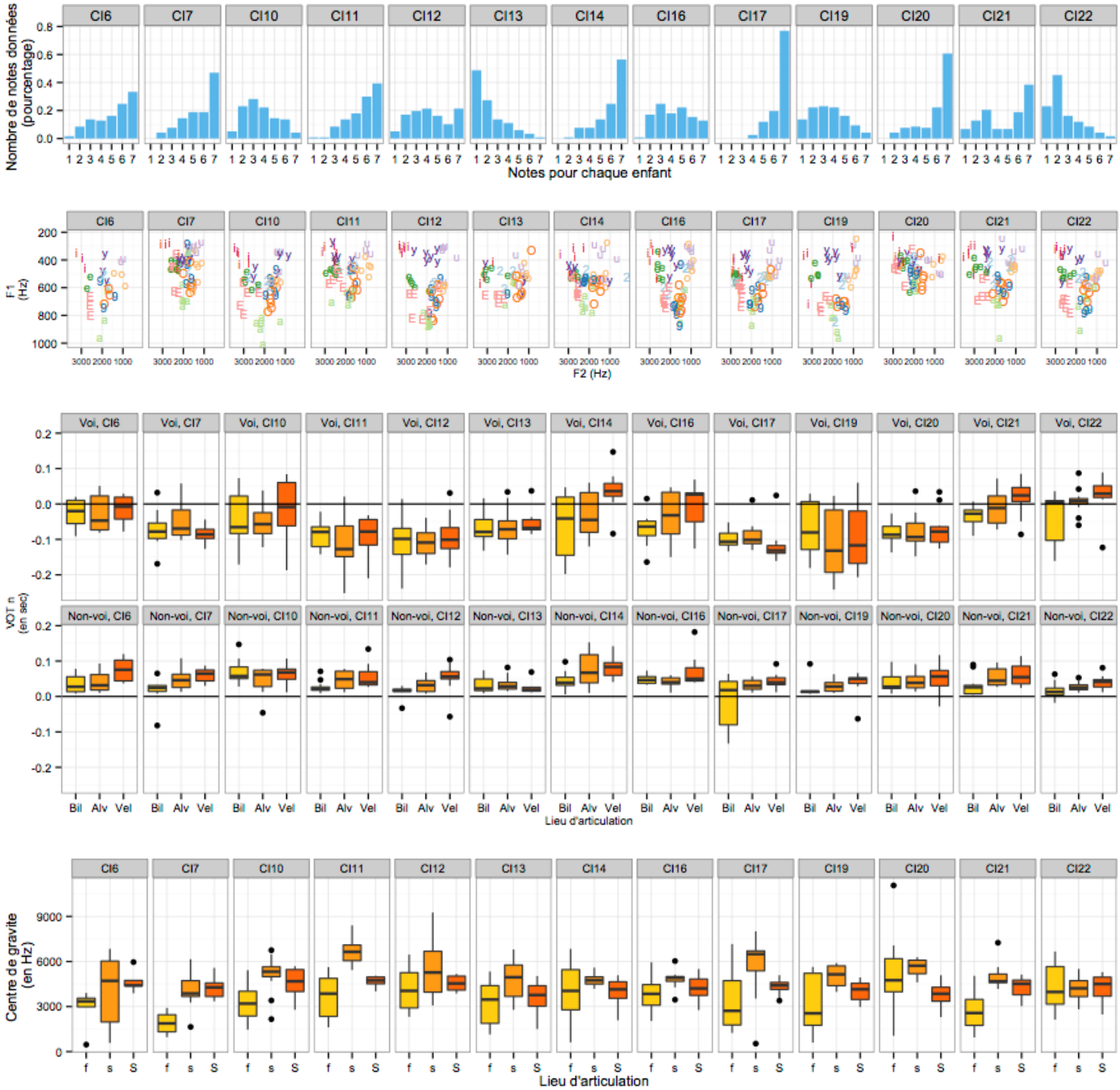


Figure 6.12 : Comparaison des notes d'intelligibilité obtenues par chaque enfant CI, et des caractéristiques acoustiques étudiées dans l'étude de production (de haut en bas : espace vocalique F1-F2, VOT des occlusives voisées et non-voisées, centre de gravité des fricatives)

Nous avons regroupé les sept enfants les plus intelligibles (CI6, CI7, CI11, CI14, CI17, CI20 et CI21, dont plus de 50% des notes sont 5, 6 et 7) et les six enfants les moins intelligibles (CI10, CI12, CI13, CI19, CI22 et CI16, dont moins de 50% des notes sont 5, 6 et 7), et nous avons comparé leurs productions, pour essayer de dégager des patrons de production qui peuvent

expliquer une bonne intelligibilité. Nous nous intéressons en particulier à la réalisation des voyelles antérieures arrondies, à la réalisation du prévoisement des occlusives voisées, à la distinction des alvéolaires et postalvéolaires par le VOT, ainsi qu'à la distinction entre centres de gravité des alvéolaires et postalvéolaires, puisque nous avons identifié des différences inter-groupes pour ces caractéristiques acoustiques lors de notre étude de production (chapitres 4 et 5).

Pour les voyelles, les différences inter-groupes (CI vs NH) observées lors de l'étude de production étaient localisées pour les voyelles à la fois au niveau du F1 de /a/, et au niveau du F2, ce qui témoignait d'une tendance à la centralisation des voyelles antérieures arrondies chez les enfants CI évoluant très légèrement avec l'âge auditif, notamment pour /y/ et /ø/. Seule une différence inter-groupe pour le F3 de /y/ était observée dans notre étude, ainsi que des différences inter-groupes de dispersion de plusieurs catégories vocaliques (/a/, /y/ et /ø/, plus dispersées chez les enfants CI) sur le plan F1-F2. Lorsque nous regardons la réalisation des voyelles par le groupe des enfants les plus intelligibles, nous observons que la plupart des sept enfants de ce groupe réalisent les voyelles /y/ et /u/ avec des degrés antériorité clairement distincts, mais seuls trois des six enfants les moins intelligibles marquent cette différence d'antériorité. Nous observons en outre que le contraste d'antériorité similaire de /ø/ et /o/ d'une part et de /œ/ et /ɔ/ d'autre part est réalisé de la même façon par les deux groupes d'enfants.

L'intelligibilité est liée à la dimension de hauteur (visible sur le F1) pour les études de Horga & Liker (2006) et Poissant et all. (2006), et d'antériorité (visible sur le F2) pour Chuang et al. (2012), Horga & Liker (2006) ou encore Poissant et al. (2006). En revanche, l'étude de Löfqvist et al. (2010) ne montre aucun lien entre F1 et F2, et intelligibilité d'enfants et d'adolescents CI, ni entre taille de l'espace vocalique et intelligibilité, à la différence de l'étude de Horga & Liker (2006) et celle de Bradlow et al. (1996) qui montrent une corrélation forte entre taille de l'espace vocalique et intelligibilité pour les locuteurs normo-entendants, les espaces vocaliques plus grands se traduisant par une intelligibilité moindre.

En revanche, les caractéristiques des consonnes, occlusives ou fricatives, ne sont, à notre connaissance, pas prises en compte par les études qui lient intelligibilité et caractéristiques acoustiques de la production de parole. Si notre étude de production mettait en évidence un voisement et une production des occlusives similaires pour les deux groupes d'enfants, à l'exception de la vélaire /k/, les résultats globaux de notre étude ne permettaient pas de constater une différence majeure entre la production des occlusives par les deux groupes d'enfants. L'âge à l'implantation n'avait aucun effet sur le voisement, mais l'âge auditif avait un effet limité aux bilabiales et alvéolaires non-voisées /p/ et /t/, dont le VOT diminue lorsque la consonne est suivie de /i/ mais augmente lorsque celle-ci est suivie de /u/.

La figure 6.12 indique que cinq des sept enfants les plus intelligibles réalisent des occlusives voisées avec un prévoisement qui correspond à un VOT négatif ou proche de zéro. Deux enfants de ce groupe réalisent les occlusives voisées avec un VOT positif. Le groupe des enfants les moins intelligibles réalisent quant à eux les occlusives voisées avec un prévoisement (VOT négatif ou proche de zéro) pour quatre d'entre eux et un VOT positif pour deux d'entre eux. Pour les occlusives non-voisées, nous observons que les enfants les plus intelligibles produisent des alvéolaires et vélaires avec un VOT distinct pour quatre d'entre eux, alors que trois d'entre eux ont une production similaire pour ces deux lieux d'articulation. De même, trois des enfants les moins intelligibles distinguent les alvéolaires et les vélaires par le VOT, et trois ont une réalisation

similaire. L'utilisation du VOT par les enfants CI ne semble pas clairement distincte entre nos deux sous-groupes de sujets, et ne semble donc pas pouvoir expliquer la meilleure intelligibilité de certains enfants.

Notre étude de production a montré que les fricatives étaient clairement différentes entre les enfants CI et NH, avec des centres de gravité plus bas chez les enfants CI, mais sans effet de l'âge à l'implantation ni de l'âge auditif. L'âge auditif a cependant un effet sur la distance entre centres de gravité de /s/ et /ʃ/ lorsqu'elles sont suivies de /u/. Le groupe des enfants les plus intelligibles produit des fricatives avec un centre de gravité de /s/ plus élevé que celui de /ʃ/ pour cinq d'entre eux, et avec des centres de gravité similaires pour deux d'entre eux. Le groupe des enfants les moins intelligibles produit des fricatives alvéolaires et postalvéolaires aux centres de gravité distincts pour quatre d'entre eux et similaires pour deux d'entre eux. Encore une fois, il ne semble pas y avoir de relation claire entre la production des fricatives par nos sujets, et leur intelligibilité.

Il nous semble donc impossible de dégager clairement un patron des caractéristiques segmentales de la production de parole qui puisse correspondre à une bonne intelligibilité de la parole continue. En effet, notre expérience de production a pu mettre en évidence la capacité des enfants NH et CI à produire des sons de parole avec plus ou moins de variabilité selon le contraste phonologique considéré mais nous pouvons nous interroger sur la capacité de ces enfants à reproduire ces contrastes dans un flux de parole continue, ce qui est une tâche plus complexe et qui ne repose pas uniquement sur les caractéristiques segmentales de la parole. Dans la littérature, l'intelligibilité des enfants CI était également liée à des caractéristiques supra-segmentales de la parole : la qualité de voix chez Perrin et al. (1999) et Van Lierde et al. (2005), débit de parole et pauses chez Chuang et al. (2012) et Perrin et al. (1999), et F0 chez Löhle et al. (1999) et Poissant et al. (2006). Notre étude étant focalisée exclusivement sur des caractéristiques segmentales de la parole, nos résultats ne permettent pas de confirmer ou infirmer ces corrélations entre caractéristiques supra-segmentales et intelligibilité de la parole.

Des analyses de l'utilisation de certaines caractéristiques acoustiques supra-segmentales dans la parole continue permettraient de compléter cette étude, notamment une analyse du débit de parole en parole spontanée, ou une analyse de la position de l'accent et de la réalisation de contours terminaux de groupes accentuels. Les liens entre caractéristiques segmentales et supra-segmentales de la parole des enfants pourraient également contribuer à approfondir notre étude : nous pourrions envisager notamment de comparer la réalisation des segments (consonnes et voyelles) dans des positions accentuelles différentes pour comprendre, par exemple, le rôle de l'accentuation dans les phénomènes de réduction vocalique ou consonantique chez les deux groupes d'enfants et l'effet de ces phénomènes sur l'intelligibilité de la parole.

CONCLUSION
ET PERSPECTIVES

De nombreux travaux en acquisition du langage et en développement phonologique ont été réalisés chez de très jeunes enfants, dès la naissance, et proposent des trajectoires d'acquisition jusqu'à 54 mois environ (MacLeod et al, 2011) mais relativement peu d'études sont réalisées chez des enfants plus âgés et à notre connaissance, aucune ne propose d'étude de la production des sons de parole et de la réalisation de contrastes phonologiques chez l'enfant de plus de cinq ans, francophone. Notre étude contribue ainsi à la description acoustique de la production de parole chez des enfants typiques, normo-entendants et francophones.

L'étude de la parole d'enfants sourds, porteurs d'implant cochléaire, permet à la fois de pouvoir identifier les difficultés persistantes de ces enfants à percevoir et reproduire des sons dont ils ne perçoivent les caractéristiques acoustiques que partiellement grâce à l'implant cochléaire, mais également de mieux comprendre les interactions entre les différentes modalités (auditives, visuelles, gestuelles) qui entrent en jeu dans la production de parole et dans son acquisition par l'enfant.

Le but de cette thèse était donc de proposer une description acoustique de la parole d'enfants sourds, porteurs d'implant cochléaire, en la comparant à celle d'enfants normo-entendants pour pouvoir localiser le plus précisément possible les difficultés qu'ils rencontrent, puis d'évaluer l'intelligibilité des deux groupes d'enfants. Nous avons choisi d'étudier la parole d'enfants de cinq à onze ans, puisque c'est un âge qui correspond à une durée d'utilisation prolongée de l'implant, ce qui nous permet d'étudier les effets à long terme de l'implantation cochléaire. Il s'agit en outre d'un âge peu décrit : la littérature se concentre essentiellement sur la production par le bébé normo-entendant ou sur les effets à très court terme de l'implantation cochléaire.

Lors de notre première étude expérimentale de la production de parole (Deuxième partie, chapitres 4 et 5), nous avons mis en évidence que l'acquisition des différents contrastes phonologiques n'est pas simultanée et suit des étapes précises, qui sont le reflet de la phonologie de la langue maternelle et qui obéissent également aux contraintes dictées par les capacités articulatoires des enfants, ainsi qu'aux contraintes liées à la perception, qui est partielle pour les enfants sourds porteurs d'implant cochléaire. Ainsi, les enfants normo-entendants et sourds porteurs d'implant cochléaire réalisent des voyelles aux niveaux de hauteur, d'antériorité et d'arrondissement similaires à celles d'adultes normo-entendants de la littérature. Les enfants sourds porteurs d'implant cochléaire ont cependant des voyelles plus variables, car plus dispersées que les enfants normo-entendants, et une série de voyelles antérieures arrondies plus postérieures, mais ayant tendance à s'antérioriser au fur et à mesure de l'utilisation de l'implant. Les caractéristiques de focalisation (F2-F1 et F3-F2) sont globalement similaires chez les deux groupes d'enfants et en accord avec les descriptions de l'adulte francophone dans la littérature. Le voisement des occlusives est réalisé de façon similaire chez les deux groupes d'enfants, avec quelques différences localisées notamment sur les vélaires, ce qui témoigne d'une acquisition plus tardive de ce lieu d'articulation chez les enfants CI, du fait notamment de l'ambiguïté visuelle et acoustique qui rend la perception du contraste entre vélaires et alvéolaires plus difficile. L'étude des fricatives a mis en évidence que les enfants normo-entendants francophones ont une réalisation similaire à celle des adultes de la littérature, avec un effet du contexte vocalique attendu sur les quatre moments spectraux. Les enfants porteurs d'implant cochléaire ont quant à eux des fricatives aux centres de gravité plus bas, ainsi que des différences entre centres de gravité des alvéolaires et post-alvéolaires moins marquées que celles des enfants normo-entendants, en

étant sensibles de la même manière aux effets de la coarticulation. Enfin, notre étude de la coarticulation à travers l'analyse de la pente de l'équation de locus a pu mettre en évidence un degré de coarticulation de séquences occlusive-voyelle similaire chez les deux groupes d'enfants, et distincte des adultes pour les alvéolaires, ce qui témoigne d'une utilisation des gestes articulatoires différente de celle des adultes et plus impactée par le contexte vocalique que ceux-ci.

Notre étude a mis en évidence que l'acquisition de contrastes phonologiques tels que les contrastes de hauteur et d'antériorité des voyelles, du voisement et du lieu d'articulation des occlusives, du lieu d'articulation des fricatives est plus compliquée pour les enfants CI lorsque ceux-ci dépendent de caractéristiques spectrales que lorsqu'il s'agit de caractéristiques temporelles : en effet, le voisement qui repose sur une coordination successive de plusieurs gestes articulatoires est globalement acquis par les enfants CI (avec une différence inter-groupes observée pour les vélaires), alors que le contraste d'antériorité des voyelles et celui des fricatives alvéolaires et postalvéolaires est plus difficile à mettre en place et intervient plus tardivement que chez les enfants NH. Ceci est à mettre en rapport avec le traitement du signal par l'implant, qui transmet les caractéristiques temporelles des sons de façon plus précise que leurs caractéristiques spectrales.

Les résultats de notre thèse confirment la complexité des paramètres qui entrent en jeu dans le parcours de l'enfant sourd dans l'accès au langage oral : si la précocité est largement décrite dans la littérature comme le principal prédicteur du succès de l'implantation cochléaire, nos résultats sont plus nuancés. L'intelligibilité de la parole spontanée semble liée à la précocité de l'implantation, mais la durée de l'utilisation de l'implant peut également expliquer les différences constatées avec la production des enfants normo-entendants.

Plus généralement, notre thèse a montré que l'enfant qui acquiert la phonologie de sa langue maternelle fait appel conjointement à plusieurs niveaux de compétence : articulatoire, perceptive ou encore statistique. Les résultats de notre étude de production ont montré que ces différents niveaux ne sont pas sollicités dans les mêmes proportions et aux mêmes moments du développement phonologique.

La maturation du contrôle moteur permet à l'enfant de développer et coordonner des gestes articulatoires de plus en plus précis et de plus en plus proches de ceux de l'adulte : les consonnes occlusives et leur voisement sont acquis avant les fricatives, dont la production nécessite un contrôle plus fin des articulateurs. Les occlusives ne sont elles-mêmes pas toutes acquises au même rythme, les vélaires étant les consonnes les plus variables dans notre étude sur les occlusives, elles sont probablement les dernières à être maitrisées par les enfants. L'enchaînement des gestes articulatoires utilisés pour produire des séquences occlusive-voyelle s'est révélé similaire à celui des adultes, à l'exception des séquences alvéolaire-voyelle, qui nécessitent une coordination plus complexe et plus précise des muscles de la langue que pour les séquences bilabiale-voyelle et vélaire-voyelle, et qui sont acquises plus tardivement par les enfants, probablement à l'adolescence, puisque les enfants les plus âgés de notre étude (10;7 ans pour le plus âgé) ne produisent pas encore des séquences alvéolaire-voyelle similaires à celles des adultes.

Les capacités perceptives de l'enfant jouent également un rôle important dans l'acquisition de la langue maternelle, puisque l'enfant fait appel conjointement à l'input auditif disponible autour de lui et à sa familiarité avec cet input, mais également au retour perceptif sur ses propres productions, qui lui permet de produire des catégories vocaliques et consonantiques de plus en

plus précises et de plus en plus identifiables par rapport aux catégories adjacentes. Notre travail a pu montrer que ces capacités perceptives ne permettent pas une utilisation similaire de toutes les caractéristiques acoustiques des sons de parole et que ces capacités perceptives ne reposent pas que sur la modalité auditive mais également sur la modalité visuelle qui peut permettre de désambiguiser des sons proches. Cette ambiguïté visuelle de certains sons de parole va influencer leur acquisition, qui sera plus tardive que celle des sons non ambigus. Ainsi les sons dont le contraste n'est perçu qu'auditivement, comme les voyelles antérieures arrondies et les fricatives alvéolaires et postalvéolaires, sont acquis plus tardivement par l'enfant.

Enfin, un niveau de traitement statistique des mots contribue également à la variation de la production de parole chez l'enfant et lui permet de produire les sons présents dans les mots qui lui sont plus familiers avant ceux qui lui sont moins familiers : Vihman & Croft (2007) ont montré que l'enfant produit des mots simples, avant d'arriver à des productions plus complexes, nous pouvons donc faire l'hypothèse que l'acquisition de la langue maternelle se fait par le mot, dont le patron global est acquis avant la production isolée du phonème, puisqu'une plus grande variabilité de la production des sons de parole est observée dans les mots les moins familiers.

Notre deuxième étude expérimentale (Troisième partie, chapitre 6) a mis en évidence une intelligibilité de la parole semi-spontanée plus faible des enfants porteurs d'implant cochléaire par rapport aux enfants normo-entendants. L'intelligibilité des enfants porteurs d'implant cochléaire est liée à l'âge à l'implantation, et les enfants ayant reçu l'implant avant l'âge de 20 mois obtiennent de meilleurs résultats. Nous pouvons nous interroger sur les liens entre le niveau d'intelligibilité des enfants CI et la variation observée entre les deux groupes d'enfants : une réalisation plus variable de certains contrastes phonologiques chez les enfants porteurs d'implant cochléaire, notamment en ce qui concerne l'antériorité de certaines voyelles, le voisement des occlusives vélaires et des fricatives, en particulier la plus grande difficulté de ce groupe d'enfants à réaliser des fricatives alvéolaires et post-alvéolaires bien distinctes.

Ce travail apporte un ensemble de mesures acoustiques de production de parole chez une population d'enfants normo-entendants francophones, à un âge relativement peu étudié dans la littérature, qui se concentre généralement sur les premières années de vie de l'enfant. Il apporte également des mesures acoustiques de la production de parole chez les enfants implantés.

Par les trois tâches proposées aux enfants lors de ces deux études expérimentales, nous disposons donc à la fois de données de parole contrôlée, qui permettent de caractériser des contrastes phonologiques précis, et de données de parole semi-spontanée, qui s'apparentent à des productions de parole naturelle, permettant d'étudier les caractéristiques de la parole continue de l'enfant au développement typique.

Cette thèse nous a enfin permis de constituer une base de données conséquente incluant à la fois un corpus pouvant être utilisé comme corpus contrôle dans le cadre d'autres études d'acquisition de la langue maternelle par des populations d'enfants présentant des troubles de la parole, ainsi qu'un corpus de parole semi-spontanée qui peut être utilisé pour des analyses complémentaires de la production de sons de parole dans de la parole continue. Le corpus des enfants porteurs d'implant cochléaire peut également être utilisé dans le cadre d'autres études chez cette population d'enfants.

Cette thèse nous a donc permis d'étudier la réalisation de plusieurs contrastes segmentaux en production, mais nous pouvons regretter de ne pas avoir étudié des contrastes supplémentaires comme celui de nasalité par exemple. Lors de la constitution du corpus et lors des enregistrements que nous avons réalisés, nous avons inclus des voyelles nasales. Cependant l'analyse acoustique du trait de nasalité s'est révélée être une tâche très compliquée, du fait du peu de références et de normes acoustiques de la nasalisation des voyelles du français, en particulier chez l'enfant, et nous avons dû laisser de côté cette étude. Une étude spécifique articulatoire et aérodynamique de la nasalité aurait été très intéressante, mais implique des contraintes plus fortes sur l'enfant participant à l'étude, et nous ne souhaitions pas réaliser une étude invasive chez des enfants.

Notre étude s'est concentrée sur la réalisation de plusieurs contrastes phonologiques par des enfants normo-entendants et des enfants porteurs d'implant cochléaire. Nous avons fait le choix de contrôler le plus possible la situation d'élicitation de parole, de façon à éviter des interférences avec des effets qui masquent les phénomènes que nous souhaitions étudier : contrôler l'accentuation des mots produits nous permettait en particulier d'éviter des phénomènes de réduction vocalique et consonantique. Nous pouvons cependant regretter un contrôle trop important de ces productions de mots isolés, qui ne nous permet pas de comprendre quelles caractéristiques de la production sont corrélées avec une bonne intelligibilité de la parole des enfants.

Par ailleurs, si notre étude a pu mettre en évidence des réalisations des contrastes phonologiques chez l'enfant similaires à ceux d'adultes francophones de la littérature, elle pourrait être complétée par une confrontation des productions des enfants normo-entendants avec des productions similaires d'adultes francophones, de façon à mieux comprendre les chronologies de l'acquisition des voyelles et des consonnes chez l'enfant de 5 à 11 ans.

Les résultats encourageants que nous avons obtenus dans le cadre de cette thèse ainsi que des lectures complémentaires nous ont conduites à une réflexion plus générale sur les mécanismes d'acquisition de la phonologie de la langue maternelle par l'enfant et plusieurs pistes restent à explorer.

Il nous semble pertinent, pour poursuivre cette étude de la production et de l'intelligibilité de la parole d'enfants sourds, porteurs d'implant cochléaire, de prendre en compte conjointement les caractéristiques segmentales et supra-segmentales de la parole spontanée. Une étude des corrélats segmentaux de l'accentuation dans la parole spontanée pourrait en effet contribuer à comprendre quels mécanismes permettent à l'enfant d'avoir une production de plus en plus proche de celle de l'adulte et de comprendre quelles sont les étapes tardives du développement de la parole chez l'enfant.

Notre étude ouvre des perspectives d'études non-invasives, notamment celle de caractéristiques supra-segmentales telles que la qualité de voix, ou de la prosodie. En effet, notre deuxième étude expérimentale a mis en évidence que les caractéristiques segmentales à elles seules ne permettent pas d'expliquer une bonne intelligibilité des enfants, qu'ils soient normo-entendants ou porteurs d'implant cochléaire. Les caractéristiques supra-segmentales de la parole spontanée doivent notamment jouer un rôle important dans l'intelligibilité de la parole spontanée et nous avions envisagé une étude de caractéristiques de la prosodie des enfants. Lors d'une étude préliminaire réalisée sur les données collectées lors de l'étude pilote de Scarbel (2012), nous avons souhaité

étudier la réalisation de l'accentuation chez les deux groupes d'enfants. Nous avons cherché en particulier à caractériser la position de l'accent primaire du français, qui se manifeste par un allongement systématique de la voyelle de la dernière syllabe de groupe accentuel (entre autres Lacheret-Dujour & Beaugendre, 1999, DiCristo, 1998, Jun & Fougeron, 2000). Cette étude préliminaire, réalisée en partenariat avec M.J.-Martinez et A. Samson (laboratoire LJK, UMR 5224, Grenoble) semblait indiquer que la réalisation de l'allongement des voyelles en position finale de syntagme était similaire chez les deux groupes d'enfants. Ces résultats se basent cependant sur des annotations de la morphologie du mot (nom, verbe, adjectif) et de la syntaxe des énoncés, qui pourraient être plus précises, en incluant notamment des informations sémantico-pragmatiques, pour pouvoir être approfondies. En parallèle, une étude de la réalisation de la F0 des contours finaux de groupes accentuels pourrait permettre de caractériser l'intonation des énoncés chez les enfants en parole spontanée, puisque la F0 est principalement liée à la réalisation du contour global de l'énoncé.

Pour l'enfant sourd, porteur d'implant cochléaire, l'acquisition de la langue maternelle est donc un processus plus complexe que pour l'enfant normo-entendant et doit nécessairement être accompagnée d'une rééducation orthophonique adaptée, qui prenne en compte les difficultés spécifiques des enfants porteurs d'implant cochléaire. Ces difficultés spécifiques étant probablement dues en grande partie au manque de visibilité de ces catégories, elles pourraient être compensées et désambiguïsées, par exemple, par l'utilisation de techniques d'imagerie telles que l'ultrason (Fabre et al., 2016), pour permettre à l'enfant de visualiser donc d'acquérir une compétence métaphonétique de leur articulation et ainsi systématiser les mouvements de ses articulateurs lors de la production de parole. Un accompagnement adapté, notamment orthophonique et scolaire, de l'enfant sourd, qu'il soit appareillé ou pas apparaît donc comme un enjeu crucial de son développement et de son accès à la parole et au langage, et le choix de l'implant cochléaire ne peut pas être envisagé comme la seule option possible pour permettre à l'enfant de communiquer avec son entourage et la société dans laquelle il évolue.

Références

Abdelli-Beruh, N. B. (2009), Influence of place of articulation on some acoustic correlates of the stop voicing contrast in Parisian French, *Journal of Phonetics, 37*, 66-78.

Adank, P., Smits, R. & Van Hout, R. (2004), A comparison of vowel normalization procedures for language variation research, *The Journal of the Acoustical Society of America, 116*, 3099.

Allen, M. C., Nikolopoulos, T. P., & O'Donoghue, G. M. (1998), Speech Intelligibility in Children After Cochlear Implanation, *Otology & Neurotology*, *19*(6), 742-746.

Amelot, A. (2004), *Etude aérodynamique, fibroscopique, acoustique et perceptive des voyelles nasales du français,* (Thèse de doctorat, Université Paris III).

Aronson, E., Rosenbloom, S. (1971), Space Perception in Early Infancy: Perception within a Common Auditory-Visual Space, *Sciences, 172 (3988),* 1161-1163.

Aslin, R. N., & Pisoni, D. B. (1980), Some developmental processes in speech perception, *Child phonology*, *2*, 67-96.

Aubert, C. (2016), *Intelligibilité de la parole des enfants sourds porteurs d'implants cochléaires*, Mémoire pour l'obtention du certificat de capacité d'orthophoniste, Université de Franche-Comté.

Baudonck, N., Dhooge, I., D'haeseleer, E. & Van Lierde, K. (2010), A comparison of the consonant production between Dutch children using cochlear implants and children using hearing aids, *International journal of pediatric otorhinolaryngology, 74*, 416-421.

Baudonck, N., Van Lierde, K.; Dhooge, I. & Corthals, P. (2011), A Comparison of Vowel Productions in Prelingually Deaf Children Using Cochlear Implants, Severe Hearing-Impaired Children Using Conventional Hearing Aids and Normal-Hearing Children, *Folia phoniatrica et logopaedica, 63*, 154.

Best, C. T. (1994), The emergence of native-language phonological influences in infant : A perceptual assimilation model, In Goodman, J. C. & H. C. Nusbaum (Eds.), *The development of speech perception : The transition from speech sounds to spoken words*, Cambridge, MA : MIT Press, 167-224.

Bharadwaj, S. V., Katz, W. F. & Tobey, E. A. (2007), Effects of auditory feedback deprivation on non-native French vowels produced by children with cochlear implants, *Audiological Medicine, 5*, 274-282.

Bharadwaj, S., Tobey, E., Assmann, P. & Katz, W. (2006), Effects of auditory feedback on fricatives produced by cochlear-implanted adults and children: Acoustic and perceptual evidence, *The Journal of the Acoustical Society of America, 119*, 1626.

Blamey, P., Barry, J., Bow, C., Sarant, J., Paatsch, L. & Wales, R. (2001), The development of speech production following cochlear implantation, *Clinical Linguistics & Phonetics, 15*, 363-382.

Boersma, P. & Weenink, D. (2015). Praat: doing phonetics by computer [Computer program]. Version 5.4.09, retrieved 1 June 2015 from http://www.praat.org/.

Bouton, S., Serniclaes, W., Bertoncini, J. & Cole, P. (2012), Perception of speech features by French-speaking children with Cochlear Implants, *Journal of Speech, Language and Hearing Research, 55*, 139.

Boysson-Bardies, B. (1996), *Comment la parole vient aux enfants*, Paris : Odile Jacob.

Bradlow, A. R., Torretta, G. M. & Pisoni, D. B. (1996), Intelligibility of normal speech I: Global and fine-grained acoustic-phonetic talker characteristics, *Speech communication,* 20, 255-272.

Briec, J. (2012), *Implant cochléaire et développement du langage chez les jeunes enfants sourds profonds,* Thèse de doctorat, Université Rennes 2.

Briec, J., Le Maner-Idrissi, G., Dardier, V., Rouxel, G., Tan-Bescond, G. & Godey, B. (2012), Echanges conversationnels avec un partenaire familier: étude comparative entre enfants sourds, implantés cochléaires depuis 2 ans et enfants entendants, *Psychologie Française, 112*, 17-48.

Brin, F., Courrier, C., Lederlé, E. & Masy, V. (2004), *Dictionnaire d'orthophonie,* Ortho édition.

Bronckart, J. P. (1977), *Théories du langage: une introduction critique (Vol. 70)*, Bruxelles : Mardaga.

Bybee, J. (2001), *Phonology and Language Use*, Cambridge Studies in Linguistics (Vol. 94), Cambridge : Cambridge University Press.

Bybee, J. (2002), Word frequency and context of use in the lexical diffusion of phonetically conditioned sound change*, Language variation and change, 14(03)*, 261-290.

Calliope (1989), *La Parole et son traitement automatique,* Paris : Masson.

Calmels, M.-N., Saliba, I., Wanna, G., Cochard, N., Fillaux, J., Deguine, O. & Fraysse, B. (2004), Speech perception and speech intelligibility in children after cochlear implantation, *International journal of pediatric otorhinolaryngology,* 68, 347-351.

Caramazza, A. & Yeni-Komshian, G. H. (1974), Voice onset time in two French dialects, *Journal of Phonetics, 2,* 239-245.

Chin, S. B. & Pisoni, D. B. (2000), A phonological system at 2 years after cochlear implantation, *Clinical linguistics & phonetics, 14*, 53-73.

Chistovich, L. A. & Lublinskaya, V. V. (1979), The « center of gravity » effect in vowel spectra and critical distance between the formants: psychoacoustical study of the perception of vowel-like stimuli, *Hearing research, 1*, 185-195.

Cho, T. & Ladefoged, P. (1999), Variation and universals in VOT: evidence from 18 languages, *Journal of Phonetics, 27*, 207-229.

Chomsky, N. (1959), A review of BF Skinner's Verbal Behavior, *Language, 35*, 26-58.

Chomsky, N. (1967), Recent contributions to the theory of innate ideas, *Synthese, 17,* 2-11.

Chomsky, N. (1969), *Language and Mind,* New York : Harcourt, Brace and World.

Chomsky, N. (2001), *Le langage et la pensée,* Paris : Payot, traduit de l'anglais par L.J. Calvet.

Chomsky, N. (1975), *Reflections on Language*, New York, NY : Pantheon Books.

Chomsky, N. (1999), *Reflexions sur le langage,* Paris : Flammarion, traduit de l'anglais par J. Milner, B. Vautherin & P. Fiala.

Chouard, C. H. (2009). *L'oreille musicienne: les chemins de la musique de l'oreille au cerveau,* Paris : Gallimard.

Chuang, H.-F., Yang, C.-C., Chi, L.-Y., Weismer, G. & Wang, Y.-T. (2012), Speech intelligibility, speaking rate, and vowel formant characteristics in Mandarin-speaking children with cochlear implant, *International journal of speech-language pathology 14*, 119-129.

CISIC (2014), *L'ambition d'entendre - Guide CISIC 2014*, version électronique disponible sur www.cisic.fr.

CISIC (2016), Site web du Centre d'Information sur la Surdité et l'Implant Cochléaire, www.cisic.fr (consulté le 16 mars 2016).

Davis, B. L., & MacNeilage, P. F. (2004), The frame/content theory of speech evolution: From lip smacks to syllables, *Primatologie*, *6*, 305-328.

Dehaene-Lambertz, G., & Houston, D. (1998), Faster orientation latencies toward native language in two-month-old infants. *Language and Speech*, *41*(1), 21-43.

Delais-Roussarie E. & H. Yoo (2014), Corpus and research in phonetics and phonology (I): Methodological and formal considerations. In: Durand J., Gut U. & G. Kristoffersen (eds): *The Oxford Handbook of Corpus Phonology,* Oxford : Oxford University Press.

Delais-Roussarie, E., Post, B., Avanzi, M., Buthke, C., Di Cristo, A., Feldhausen, I., ... & Sichel-Bazin, R. (2015), Intonational phonology of French: developing a ToBI system for French, *Intonation in Romance*, 63-100.

Delattre, P. (1948), Un triangle acoustique des voyelles orales du français, *French review, 21,* 477-484.

Delattre, P. (1966), Les dix intonations de base du français, *French review,* 1-14.

Delattre, P. C., Liberman, A. M. & Cooper, F. S. (1955), Acoustic loci and transitional cues for consonants, *The Journal of the Acoustical Society of America, 27*, 769-773.

Dettman, S., Barker, E., Dowell, R., Dawson, P., Blamey, P. & Clark, G. M. (1995), Vowel imitation tasks: results over time for 28 cochlear implant children under the age of eight years, *Scientific publications, vol. 8,* no. 733.

Di Cristo, A. (1998), Intonation in French, in Di Cristo, A. & Hirst, D. (Eds), *Intonation systems: A survey of twenty languages*, 195-218.

Di Cristo, A. (2000), Vers une modélisation de l'accentuation du français (seconde partie), *Journal of French Language Studies, Cambridge Univ Press, 10*, 27-4.

Dillon, C., Pisoni, D., Cleary, M. & Carter, A. (2004), Nonword imitation by children, with cochlear implants: Consonant analyses, *Archives of Otolaryngology- Head and Neck Surgery, 130*, 587.

Disner, S. F. (1980), Evaluation of vowel normalization procedures, *The Journal of the Acoustical Society of America, 67*, 253.

Dole, M., Loevenbruck, H., Pascalis, O., Schwartz, J.-L., Vilain, A. (en preparation), The role of production abilities in consonant invariance detection in infants.

Duez, D. (1992), Second formant locus-nucleus patterns: An investigation of spontaneouos French speech, *Speech communication, 11*, 417-427.

Eimas, P., Siqueland, E., Jusczyk, P., Vigorito, J. (1971), Speech perception in Infants, *Science, 171*, 303-306.

Ertmer, D. J. (2001), Emergence of a vowel system in a young cochlear implant recipient, *Journal of Speech, Language and Hearing Research, 44*, 803.

Fabre, D., Hueber, T., Canault, M., Bedoin, N., Acher, A., Bach, C., Lambourion, L. & Badin, P. (2016), Apport de l'échographie linguale à la rééducation orthophonique. In *XVIèmes Rencontres Internationales d'Orthophonie, Orthophonie et technologies innovantes* (N. Joyeux & S. Topouzkhanian, editors), Chapitre 11, pp. 1-27, Paris, France.

Fant, G. (1960), *Acoustic theory of speech production*, The Hague : Mouton.

Flipsen, P. & Colvard, L. G. (2006), Intelligibility of conversational speech produced by children with cochlear implants, *Journal of Communication Disorders, 39*, 93-108.

Fougeron, C. & Keating, P. A. (1997), Articulatory strengthening at edges of prosodic domains, *The journal of the acoustical society of America, 101*, 3728-3740.

Fowler, C. A. (1994), Invariants, specifiers, cues: An investigation of locus equations as information for place of articulation, *Perception & Psychophysics, 55*, 597-610.

Fuller, C. D., Galvin, J. J., Maat, B., Free, R. H., & Başkent, D. (2014), The musician effect: does it persist under degraded pitch conditions of cochlear implant simulations?, *Frontiers in neuroscience, 8,* 179.

Garabédian, E. & Loundon, N. (2010), L'implant cochléaire chez l'enfant *e-mémoires de l'Académie Nationale de Chirurgie,* 2010*, 9*, 47-51.

Gaul-Bouchard, M.-E., Le Normand, M.-T. & Cohen, H. (2007), Production of consonants by prelinguistically deaf children with cochlear implants, *Clinical linguistics & phonetics, 21*, 875-884.

Geffner, D. (1980), Feature characteristics of spontaneous speech production in young deaf children, *Journal of Communication Disorders, 13*, 443-454.

Georgeton, L. & Fougeron, C. (2014), Domain-initial strengthening on French vowels and phonological contrasts: Evidence from lip articulation and spectral variation, *Journal of Phonetics, 44*, 83-95.

Georgeton, L.; Paillereau, N.; Landron, S.; Gao, J. & Kamiyama, T. (2012), Analyse formantique des voyelles orales du français en contexte isolé: à la recherche d'une référence pour les apprenants de FLE, *Conférence conjointe JEP-TALN-RECITAL 2012,* 145-152.

Gerstman, L. (1968), Classification of self-normalized vowels, IEEE, Trans. Audio Electroacoust. AU-16, 78–80.

Gibson, T., & Ohde, R. (2007), F2 locus equations: Phonetic descriptors of coarticulation in 17- to 22-month-old children, *Journal of Speech, Language, and Hearing Research, 50,* 97–108.

Glasberg, B. R., and Moore, B. C. J. (1990), Derivation of auditory filter shapes from notched noise data, *Hear. Res.* 47, 103–138.

Grosjean, F. & Dommergues, J. Y. (2011), *La statistique en clair*, Paris : Ellipses.

Guerrero-Lopez, H. A. (2010), *Caractérisation de la voix de l'enfant sourd appareillé et implanté cochléaire: approches acoustique et perceptuelle et proposition de modélisation,* Thèse de doctorat, Montpellier 3.

Habib, M. G.; Waltzman, S. B.; Tajudeen, B. & Svirsky, M. A. (2010), Speech production intelligibility of early implanted pediatric cochlear implant users, *International journal of pediatric otorhinolaryngology, 74*, 855-859.

Hage, C., Charlier, B., & Leybaert, J. (2006), *Compétences cognitives, linguistiques et sociales de l'enfant sourd: pistes d'évaluation*, Editions Mardaga.

Hallé, P. A., De Boysson-Bardies, Bé. & Vihman, M. M. (1991), Beginnings of prosodic organization: Intonation and duration patterns of disyllables produced by Japanese and French infants, *Language and Speech, 34*, 299-318.

Hanna & Barbera (1949), *Tom & Jerry - Hatch Up Your Troubles.*

Harshman, R. A. (1970), Foundations of the PARAFAC procedure: Models and conditions for an "explanatory" multi-modal factor analysis, *University of California at Los Angeles Los Angeles.*

Hassanzadeh, S. (2012), Outcomes of cochlear implantation in deaf children of deaf parents: comparative study, *Journal of Laryngology and Otology, 126*, 989.

Hatch, E., & Lazaraton, A. (1991), *The research manual. Design and statistics for applied* linguistics, Boston : Heinle & Heinle Publishers.

Havy, M., Nazzi, T. & Bertoncini, J. (2013), Phonetic processing during the acquisition of new words in 3-to-6-year-old French-speaking deaf children with cochlear implants, *Journal of communication disorders, 46*, 181-192.

Hocevar-Boltezar, I., Boltezar, M. & Zargi, M. (2008), The influence of cochlear implantation on vowel articulation, *Wiener Klinische Wochenschrift, 120*, 228-233.

Horga, D. & Liker, M. (2006), Voice and pronunciation of cochlear implant speakers, *Clinical linguistics & phonetics, 20*, 211-217.

Howell, D., Bestgen, Y., Yzerbyt, V., & Rogier, M. (2008), *Méthodes statistiques en sciences humaines*, Bruxelles : De Boeck Supérieur.

Huber, J. E., Stathopoulos, E. T., Curione, G. M., Ash, T. A. & Johnson, K. (1999), Formants of children, women, and men: The effects of vocal intensity variation, *The Journal of the Acoustical Society of America, 106*, 1532.

Ingram, D. (1992), Early phonological acquisition: a cross-linguistic perspective, in Ferguson, C., Menn, L., Stoel-Gammon, C. (Eds) (1992), *Phonological development: Models, research, implications*, *423-435*.

Ingram, D. (1999), Phonological acquisition, in Barrett, M. (Ed) (1999), *The development of language*, 73-97.

Iverson, P. & Kuhl, P. K. (1995), Mapping the perceptual magnet effect for speech using signal detection theory and multidimensional scaling, *The Journal of the Acoustical Society of America, 97*, 553-562.

Jacewicz, E., Fox, R. A. & Salmons, J. (2007), Vowel space areas across dialects and gender, *Proceedings of the XVIth Int. Congress of Phonetic,* 1465-1468.

Jakobson, R. (1941), Kindersprache, aphasie und allgemeine lautgesetze, in Jakobson, R. (1971), *Selected writings*, vol. 2, *The Hague: Mouton,* 328-401.

Jakobson, R. (1949), Les lois phoniques du langage enfantin et leur place dans la phonologie générale, in Jakobson, R. (1971), *Selected writings*, vol. 2, *The Hague: Mouton,* 317-327.

Johnson, K. (2012), *Acoustic and auditory phonetics*, John Wiley & Sons.

Jun, S.-A. & Fougeron, Cé. (2000), A phonological model of French intonation, *Intonation,* Springer, 209-242.

Jusczyk, P. (1992), Developing Phonological Categories from the Speech Signal, in Ferguson, C., Menn, L., Stoel-Gammon, C. (Eds.) (1992), *Phonological Development, Models, Research, Implications,* York Press, 17-64.

Keating, P., Cho, T., Fougeron, C., & Hsu, C. S. (2004), Domain-initial articulatory strengthening in four languages. *Papers in laboratory phonology VI: Phonetic interpretation*, 145-163.

Kent, R. D. (1992), The biology of phonological development, in Ferguson, C., Menn, L., Stoel-Gammon, C. (Eds) (1992), *Phonological development: Models, research, implications*, 65-90.

Kent, R. & Forner, L, (1979), Developmental study of vowel formant frequencies in an imitation task, *The Journal of the Acoustical Society of America, 65*, 208-217.

Kent, R. D. & Forner, L. (1980), Speech segment duration in sentence recitations by children and adults. *Journal of Phonetics, 8,* 157-168.

Kern, S., & Dos Santos, C. (2011), Input et acquisition du lexique en français: rôle de la fréquence et de la densité de voisinage, *Travaux de didactique du français langue étrangère,* volume 66, 53-70..

Khomsi, A. (2001), *ELO: évaluation du langage oral.* ECPA, Les Editions du Centre de psychologie appliquée.

Khwaileh, F. A. & Flipsen Jr, P. (2010), Single word and sentence intelligibility in children with cochlear implants, *Clinical linguistics & phonetics, 24*, 722-733.

Kim, J. & Chin, S. B. (2008), Fortition and lenition patterns in the acquisition of obstruents by children with cochlear implant, *Clinical linguistics & phonetics, 22*, 233-251.

Konopczynski, G. (1990a), *Le langage émergent : caractéristiques rythmiques*, (Vol. 1), Hambourg : Buske Verlag.

Konopczynski, G. (1990b), *Le langage émergent : aspects vocaux et mélodiques*, (Vol. 2), Hambourg : Buske Verlag.

Konopczynski, G. (1991), L'intonation de question dans le langage emergent, *Actes du 12e Congrès International des Sciences Phonétiques,* 174-177.

Koopmans-van Beinum, F. J., & van der Stelt, J. M. (1986), Early stages in the development of speech movements, In Lindblom & Zetterström (Eds.), *Precursors of early speech*, 37-50, Palgrave Macmillan UK.

Körkkö, P. (2015), Spectral moments analysis of /s/ coarticulation development in Finnish-speaking children, In The Scottish Consortium for ICPhS 2015 (Ed.), *Proceedings of the 18th International Congress of Phonetic Sciences,* Glasgow, UK: the University of Glasgow.

Krull, D. (1989), Second formant locus patterns and consonant-vowel coarticulation in spontaneous speech, *Phonetic experimental research at the institute of linguistics university of Stockholm-PERILUS, 10*, 87-108.

Kuhl, P. K. (1991), Human adults and human infants show a "perceptual magnet effect" for the prototypes of speech categories, monkeys do not, *Perception & psychophysics, 50*, 93-107.

Kuhl, P. K. (2000), A new view of language acquisition, *Proceedings of the National Academy of Sciences, 97*, 11850-11857.

Kuhl, P. K., Meltzoff, A. (1982), The Bimodal Perception of Speech in Infancy, *Science, 218 (2477),* 1138-1141.

Kuhl, P. K., Williams, K. A., Lacerda, F., Stevens, K. N. & Lindblom, B. (1992). Linguistic experience alters phonetic perception in infants by 6 months of age, *Science, 255,* 606-608.

Kunisue, K., Fukushima, K., Nagayasu, R., Kawasaki, A. & Nishizaki, K. (2006), Longitudinal formant analysis after cochlear implantation in school-aged children, *International journal of pediatric otorhinolaryngology, 70*, 2033-2042.

Lacheret-Dujour, A., Beaugendre, F. (1999), *La Prosodie du Français*, Paris : Ed. du CNRS.

Ladefoged, P. & Johnson, K. (2011), *A course in Phonetics*, 6th international edition, Wadsworth Cengage Learning.

Lane, H.; Matthies, M., Perkell, J., Vick, J. & Zandipour, M. (2001), The effects of changes in hearing status in cochlear implant users on the acoustic vowel space and CV coarticulation, *Journal of Speech, Language and Hearing Research, 44*, 552.

Legendre, C. (2014), *Intelligibilité de la parole d'enfants sourds porteurs d'un implant cochléaire unilatéral et d'enfants normo-entendants* (Thèse de doctorat, Paris 3).

Legerstee, M. (1990), Infants Use Multimodal Information to Imitate Speech Sounds, *Infant Behavior and Development, 13,* 343-354.

Le Maner-Idrissi, G., Rouxel, G., Pajon, C., Dardier, V., Gavornikova-Baligand, Z., Tan-Bescond, G. & Godey, B. (2009), Cochlear implant and lexical diversity development in deaf children: intra-and interindividual differences. *Current psychology letters, 25*.

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.

Léon, P. (2007), *Phonétisme et prononciations du français (5ème édition)*, Paris : Armand Colin

Leybaert, J., Bayard, C., Colin, C., & LaSasso, C. (2015), Cued Speech and Cochlear Implants: A Powerful Combination for Natural Spoken Language Acquisition and the Development of Reading, in *The Oxford Handbook of Deaf Studies in Language*, 359.

Li, F., Edwards, J. & Beckman, M. E., (2009), Contrast and covert contrast: The phonetic development of voiceless sibilant fricatives in English and Japanese toddlers, *Journal of Phonetics, 37*, 111-124.

Liker, M., Mildner, V. & Šindija, B. (2007), Acoustic analysis of the speech of children with cochlear implants: A longitudinal study, *Clinical linguistics & phonetics, 21*, 1-11.

Likert, R. (1932), A technique for the measurement of attitudes, *Archives of psychology*, vol. 140, 1932, 1-55.

Lina-Granade, G. & Truy, E. (2005), Conduite à tenir devant une surdité de l'enfant, *EMC-Oto-rhino-laryngologie, 2*, 290-300.

Lindblom, B. (1963), Spectrographic study of vowel reduction, *The journal of the Acoustical society of America, 35*, 1773-1781.

Ligny, C., Simon, P., Renglet, T., Schepers, F. & Mansbach, A.-L. (2006), Chapitre 11. L'enfant porteur d'implant cochléaire: Intervention chirurgicale, suivi et évaluation, *Pratiques psychologiques*, 247-266.

Lisker, L. & Abramson, A. S. (1964), A Crosslanguage study of voicing in initial stops, *The Journal of the Acoustical Society of America, 35*, 1889-1890.

Lobanov, B. M. (1971), Classification of Russian vowels spoken by different speakers, *The Journal of the Acoustical Society of America, 49*, 606.

Locke, J. L., Pearson (1992), in D. M., Ferguson, C. A., Menn, L., & Stoel-Gammon, C. (Eds), (1992), Vocal learning and the emergence of phonological capacity: A neurobiological approach, *Phonological development: Models, research, implications*, 91-129.

Löfqvist, A. (1999), Interarticulator phasing, locus equations, and degree of coarticulation, *The Journal of the Acoustical Society of America, 106*, 2022.

Löfqvist, A., Sahlén, B. & Ibertsson, T. (2010), Vowel spaces in Swedish adolescents with cochlear implants, *The Journal of the Acoustical Society of America, 128*, 3064.

Löhle, E., Frischmuth, S., Holm, M., Becker, L., Flamm, K., Laszig, R., Beck, C. & Lehnhardt, E. (1999), Speech recognition, speech production and speech intelligibility in children with hearing aids versus implanted children, *International journal of pediatric otorhinolaryngology,* 47, 165-169.

MacLeod, A. A., Sutton, A., Trudeau, N., & Thordardottir, E. (2011), The acquisition of consonants in Québécois French: A cross-sectional study of pre-school aged children, *International Journal of Speech-Language Pathology*, *13*(2), 93-109.

MacNeilage, P. F. (1998), The frame/content theory of evolution of speech production, *Behavioral and brain sciences*, *21*(04), 499-511.

Martin, P. (2009) *Intonation du Français*, Paris : Armand Colin.

MATLAB and Statistics Toolbox Release 2014b (version 8.4.0), The MathWorks, Inc., Natick, Massachusetts, United States.

Maye, J., Werker, J. F., & Gerken, L. (2002), Infant sensitivity to distributional information can affect phonetic discrimination, *Cognition*, *82*(3), B101-B111.

McCaffrey Morrison, H. (2008), The locus equation as an index of coarticulation in syllables produced by speakers with profound hearing loss, *Clinical linguistics & phonetics, 22*, 726-740.

Mccaffrey Morrison, H. (2012), Coarticulation in early vocalizations by children with hearing loss: A locus perspective, *Clinical linguistics & phonetics, 26*, 288-309.

Meltzoff A., Borton, R. (1979), Intermodal matching by human neonates, *Nature, 282 (5737),* 403-404.

Ménard, L., Schwartz, J.-L., Boë, L.-J. & Aubin, Jé. (2007), Articulatory--acoustic relationships during vocal tract growth for French vowels: Analysis of real data and simulations with an articulatory model, *Journal of Phonetics, 35*, 1-19.

Menn, L., & Stoel-Gammon, C. (1995), Phonological development, in Fletcher, P. MacWhinney, B. (Eds), *The handbook of child language,* Wiley-Blackwell 335-359.

Meunier, C., & Espesser, R. (2011), Vowel reduction in conversational speech in French: The role of lexical factors, *Journal of Phonetics, 39(3),* 271-278.

Mildner, V. & Liker, M. (2008), Fricatives, affricates, and vowels in Croatian children with cochlear implants, *Clinical linguistics & phonetics, 22*, 845-856.

Miyamoto, R. T., Svirsky, M., Kirk, K. I., Robbins, A., Todd, S. & Riley, A. (1997), Speech intelligibility of children with multichannel cochlear implants, *The Annals of otology, rhinology & laryngology. Supplement,* 168, 35-36.

Mondain, M., Blanchet, C., Venail, F. & Vieu, A. (2005), Classification et traitement des surdités de l'enfant, *EMC-Oto-rhino-laryngologie, 2*, 301-319.

Monnin, J. (2010), *Influence de la langue ambiante sur l'acquisition phonologique: une comparaison du français et du drehu* (Thèse de Doctorat, Université de Grenoble & Université de Nouvelle Calédonie).

Monsen, R. B. (1976), Second formant transitions of selected consonant-vowel combinations in the speech of deaf and normal-hearing children, *Journal of speech and hearing research, 19*, 279.

Monsen, R. B. (1981), A usable test for the speech intelligibility of deaf talkers, *American Annals of the Deaf*, *156*(7), 845-852.

Montag, J. L., AuBuchon, A. M., Pisoni, D. B. & Kronenberger, W. G. (2014), Speech Intelligibility in Deaf Children After Long-Term Cochlear Implant Use, *Journal of Speech, Language, and Hearing Research, 57*, 2332-2343.

Most, T. & Peled, M. (2007), Perception of suprasegmental features of speech by children with cochlear implants and children with hearing aids, *Journal of deaf studies and deaf education, 12*, 350.

Myers, A., & Hansen, C. H. (2007), *Psychologie expérimentale, 2e éd,* Traduction française par Dieudonné Zelé, révision par Ludovic Ferrand, Bruxelles : De Boeck Supérieur.

Nazzi, T. (2006), Spécificité phonétique : de la perception précoce à l'acquisition des premiers mots, *Recherches linguistiques de Vincennes, 35*, 31-52.

Nazzi, T., Ramus, F. (2003), Perception and acquisition of linguistic rhythm by infants, *Speech Communication, 41,* 233-234.

Nearey, T. M. (1978), Phonetic feature systems for vowels, *Indiana University Linguistics Club, 77*.

Nearey, T. M., & Rochet, B. L. (1994), Effects of place of articulation and vowel context on VOT production and perception for French and English stops, *Journal of the International Phonetic Association*, *24*(01), 1-18.

Neumeyer, V., Harrington, J. & Draxler, C. (2010), An acoustic analysis of the vowel space in young and old cochlear-implant speakers, *Clinical linguistics & phonetics*, 1-8.

New, B., Brysbaert, M., Veronis, J., & Pallier, C. (2007), The use of film subtitles to estimate word frequencies, *Applied psycholinguistics, 28(04),* 661-677.

Nissen, S. L. & Fox, R. A. (2005), Acoustic and spectral characteristics of young children's fricative productions: A developmental perspectivea), *The Journal of the Acoustical Society of America, 118*, 2570-2578.

Nittrouer, S. (1993), The emergence of mature gestural patterns is not uniform : evidence from an acoustic study, *Journal of Speech, Language, and Hearing Research*, *36*(5), 959-972.

Nittrouer, S. (1995), Children learn separate aspects of speech production at different rates: Evidence from spectral moments, *The Journal of the Acoustical Society of America, 97*, 520-530.

Noiray, A.; Ménard, L. & Iskarous, K., (2013), The development of motor synergies in children: Ultrasound and acoustic measurements, *The Journal of the Acoustical Society of America, 133*, 444-452.

Oller, D. K., & Eilers, R. E. (1988), The role of audition in infant babbling, *Child development*, 441-449.

Osberger MJ, Robbins AM, Todd SL, Riley A. (1994), Speech intelligibility of children with cochlear implants, *Volta Review, 96,* 169–180.

Peirce, J.W. (2007), PsychoPy - Psychophysics software in Python, *Journal of Neuroscience Methods, 162 (1-2),* 8-13.

Peng, S., Weiss, A., Cheung, H. & Lin, Y. (2004), Consonant production and language skills in Mandarin-speaking children with cochlear implants, *Archives of Otolaryngology- Head and Neck Surgery, 130*, 592.

Peng, S.-C., Spencer, L. J. & Tomblin, J. B. (2004), Speech intelligibility of pediatric cochlear implant recipients with 7 years of device experience, *Journal of Speech, Language, and Hearing Research, 47*, 1227-1236.

Perrin, E., Berger-Vachon, C., Topouzkhanian, A., Truy, E. & Morgon, A. (1999), Evaluation of cochlear implanted children's voices, *International journal of pediatric otorhinolaryngology,* 47, 181-186.

Peterson, G. E. & Barney, H. L. (1952), Control methods used in a study of the vowels, *The Journal of the Acoustical Society of America, 24*, 175.

Piaget, J. (1989), *Le langage et la pensée chez l'enfant*, Delachaux & Niestlé (1ère édition publiée en 1923).

Piaget, J., Inhelder, B. (2008), *La psychologie de l'enfant*, Paris : PUF (1ère édition publiée en 1966).

Pierrehumbert, J. B. (2001), Exemplar dynamics: Word frequency, lenition and contrast, in Bybee, J. and Hopper, P. (Eds.) (2001), *Frequency and the Emergence of Linguistic Structure*, Typological studies in Language (45), John Benjamins, 137-158.

Pierrehumbert, J. B. (2003), Phonetic diversity, statistical learning, and acquisition of phonology, *Language and speech*, *46*(2-3), 115-154.

Pinker, S. (1994), *The Language Instinct,* New York, NY : Harper Perennial Modern Classics.

Pinker, S. (2002), *The Blank Slate*, New York, NY : Viking.

Pisoni, D. B. (2005), Speech perception in deaf children with cochlear implants, in Pisoni, D., & Remez, R. (Eds.) (2008), *The handbook of speech perception*. John Wiley & Sons, 494-523.

Plecy, A. (2013), Quelle voix pour les enfants sourds?, *Enfances & Psy,* 114-124.

Poissant, S. F., Peters, K. A. & Robb, M. P. (2006), Acoustic and perceptual appraisal of speech production in pediatric cochlear implant users, *International journal of pediatric otorhinolaryngology, 70*, 1195-1203.

Projet « Phonologie du Français Contemporain », site web http://www.projet-pfc.net.

Pujol, R. (2013), http://www.cochlea.org/ (grand public) http://www.cochlea.eu/ (étudiants et professionnels).

R Development Core Team (2012), R: A language and environment for statistical computing. R, Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL http://www.R-project.org/.

Ramus, F., Hauser, M. D., Miller, C., Morris, D., & Mehler, J. (2000), Language discrimination by human newborns and by cotton-top tamarin monkeys, *Science*, *288*(5464), 349-351.

Rondal, J. A. (1999), *Comment le langage vient aux enfants,* Bruxelles : Labor.

Rose, Y. (2014), Corpus-based Investigations of Child Phonological Development, in Durand, J., Gut, U., & Kristoffersen, G. (Eds.), *The Oxford handbook of corpus phonology*, Oxford University Press, Oxford.

Ryalls, J., Larouche, A. & Giroux, F. (2003), Acoustic comparison of CV syllables in French-speaking children with normal hearing, moderate-to-severe and profound hearing impairment, *Journal of Multilingual Communication Disorders, 1*, 99-114.

Saussure, F. (1995), *Cours de linguistique générale,* Paris : Payot (1ère édition publiée en 1916).

Scarbel, L., Vilain, A., Loevenbruck, H., Schmerber, S., (2012) An acoustic study of speech production by French children wearing cochlear implants. *3rd Early Language Acquisition Conference,* Lyon, France.

Scarbel, L., Vilain, A., & Loevenbruck, H. (2013), Phonetic characteristics of speech production by French children wearing cochlear implants, In *International Child Phonology Conference 2013 (ICPC 2013).*

Schenk, B. S.; Baumgartner, W.-D. & Hamzavi, J.-S. (2003), Effect of the loss of auditory feedback on segmental parameters of vowels of postlingually deafened speakers, *Auris Nasus Larynx, 30*, 333-339.

Schwartz, J.-L., Abry, C., Boë, L.-J., Ménard, L. & Vallée, N. (2004), Asymmetries in vowel perception, in the context of the Dispersion--Focalisation Theory, *Speech Communication, 45*, 425-434.

Schwartz, J.-L. & Escudier, P. (1989), A strong evidence for the existence of a large-scale integrated spectral representation in vowel perception, *Speech communication, 8*, 235-259.

Seifert, E., Oswald, M., Bruns, U., Vischer, M., Kompis, M. & Haeusler, R. (2002), Changes of voice and articulation in children with cochlear implants, *International journal of pediatric otorhinolaryngology, 66*, 115-123.

Skinner, B. F. (1957), *Verbal Behavior*, New York, NY : Appleton-Century-Crofts.

Shriberg, L. D. (1986), *PEPPER: Programs to examine phonetic and phonologic evaluation records*, Madison, WI: University of Wisconsin-Madison. [Computer Program].

Smolensky, P. (1996), On the comprehension/production dilemma in child language, *Linguistic inquiry, 27*, 720-731.

Sten, H. (1963), *Manuel de phonétique française*, Munksgaard.

Stoel-Gammon, C. (1992), Prelinguistic Vocal Development : Measurements and Predictions, in Ferguson, C.A., Menn, L., Stoel-Gammon, C. (Eds.) (1992) *Phonological Development, Models, Research, Implications,* York Press, 439-456.

Stokes, S. F., Kern, S., & Dos Santos, C. (2012). Extended statistical learning as an account for slow vocabulary growth, *Journal of Child Language, 39(01),* 105-129..

Sussman, H. M.; Duder, C.; Dalston, E. & Cacciatore, A. (1999), An acoustic analysis of the development of CV coarticulation: A case study, *Journal of Speech, Language, and Hearing Research, 42*, 1080-1096.

Sussman, H. M. & Shore, J. (1999), Locus equations as phonetic descriptors of consonantal place of articulation, *Perception & psychophysics, 58*, 936-946.

Svirsky, M. A. (2000), Speech intelligibility of pediatric cochlear implant users and hearing aid users, *Cochlear implants*, 312-314.

Svirsky MA, Chin SB, Miyamoto RT, Sloan RB, Caldwell M. (2000), Speech intelligibility of profoundly deaf pediatric cochlear hearing aid users, *Volta Review*, *102(4),* 175–198.

Todd, A. E.; Edwards, J. R. & Litovsky, R. Y. (2011), Production of contrast between sibilant fricatives by children with cochlear implants, *The Journal of the Acoustical Society of America, 130*, 3969.

Traunmüller, H. (1990), Analytical expressions for the tonotopic sensory scale, *The Journal of the Acoustical Society of America, 88*, 97-100.

Truy, E. & Lina, G. (2003), Implantation cochléaire de l'enfant: Technologie, bilan médical et sélection des candidats, réhabilitation, *Archives de pédiatrie, 10*, 554-564.

Tseng, S.-C.; Kuei, K. & Tsou, P.-C. (2011), Acoustic characteristics of vowels and plosives/affricates of Mandarin-speaking hearing-impaired children, *Clinical linguistics & phonetics, 25*, 784-803.

Turnbaugh, K. R.; Hoffman, P. R.; Daniloff, R. G. & Absher, R. (1985), Stop-vowel coarticulation in 3-year-old, 5-year-old, and adult speakers, *The Journal of the Acoustical Society of America, 77*, 1256-1257.

Tye-Murray N. (1998), Speech, language, and literacy development. In: Tye-Murray N, Clark W, editors. *Foundations of aural rehabilitation: Children, adults, and their family members*. Singular; San Diego, CA, 415-446.

Tye-Murray, N. & Kirk, K. I. (1993), Vowel and diphthong production by young users of cochlear implants and the relationship between the phonetic level evaluation and spontaneous speech, *Journal of Speech, Language, and Hearing Research, 36*, 488-502.

Uchanski, R. & Geers, A. (2003), Acoustic characteristics of the speech of young cochlear implant users: a comparison with normal-hearing age-mates, *Ear and hearing, 24*, 90S.

UCLA Phonological Segment Inventory Database (2014) Standard French sound inventory (UPSID). In: Moran, Steven & McCloy, Daniel & Wright, Richard (eds.) *PHOIBLE Online.* Leipzig: Max Planck Institute for Evolutionary Anthropology. (Available online at http://phoible.org/inventories/view/331, Accessed on 2016-04-18.).

Vaissière, J. (2007), Area functions and articulatory modeling as a tool for investigating the articulatory, acoustic and perceptual properties of sounds across languages, *Experimental approaches to phonology, OUP Oxford,* 54-71.

Vaissière, J. (2011), On the acoustic and perceptual characterization of reference vowels in a cross-language perspective, *The 17th International Congress of Phonetic Sciences (ICPhS XVII),* 52-59.

Vallée, N. (1994), *Systèmes vocaliques: de la typologie aux prédictions* (Thèse de doctorat, Grenoble 3).

Van Lierde, K.; Vinck, B.; Baudonck, N.; De Vel, E. & Dhooge, I. (2005), Comparison of the overall intelligibility, articulation, resonance, and voice characteristics between children using cochlear implants and those using bilateral hearing aids: a pilot study, *International Journal of Audiology,* 44, 452-465.

Ventureyra, V. A.; Pallier, C. & Yoo, H.-Y. (2004), The loss of first language phonetic perception in adopted Koreans, *Journal of Neurolinguistics, 17*, 79-91.

Verhoeven, J.; Hide, O.; De Maeyer, S.; Gillis, S. & Gillis, S. (2016), Hearing impairment and vowel production. A comparison between normally hearing, hearing-aided and cochlear implanted Dutch children, *Journal of communication disorders, 59*, 24-39.

Vieu, A., Mondain, M., Sillon, M., Piron, J.P., Uziel, A. (1999), A test for the evaluation of voice perception and production, *Rev. Laryngol. Otol. Rhinol.*, 120 (4), 219-225.

Vihman, M. M. (1996), *Phonological development: The origins of language in the child*, Blackwell Publishing.

Vihman, M., & Croft, W. (2007), Phonological development: Toward a "radical" templatic phonology, *Linguistics*, *45*(4), 683-725.

Virole, B. (Ed), (2006), *Psychologie de la surdité, 3ème édition augmentée,* Bruxelles : De Boeck Supérieur.

Vorperian, H. K. & Kent, R. D. (2007), Vowel acoustic space development in children: A synthesis of acoustic and anatomic data, *Journal of speech, language and hearing research, 50(6),* 1510-1545.

Vygotski, L. (1997), *Pensée et langage,* Paris : La Dispute.

Walter, H. (1977), *La phonologie du français*, Paris : Presses Universitaire de France.

Warner-Czyz, A. D. & Davis, B. L. (2008), The emergence of segmental accuracy in young cochlear implant recipients, *Cochlear Implants International, 9*, 143-166.

Warner-Czyz, A. D., Davis, B. L. & MacNeilage, P. F. (2010), Accuracy of consonant-vowel syllables in young cochlear implant recipients and hearing children in the single-word period, *Journal of Speech, Language, and Hearing Research, 53*, 2-17.

Watson, T. J. (1925), *Behaviorism*, New-York : Norton, Trad. : Watson, T.J. (1925), *Le Behaviorisme,* Paris : C.E.P.L. (Cité dans Bronckart, 1977).

Wauquier-Gravelines, S. (2003), Troncation et réduplication. Peut-on parler de gabarits morphologiques dans le lexique précoce, in Fradin, B., Dal, G., Hathout, M., Kerleroux, F., Roché, M. Plénat, M. (Eds.), *Les unités morphologiques. Silexicales*, *3*, Université de Lille III.

Wauquier-Gravelines, S. (2005), Acquisition et développement phonologiques, *Phonologie et phonétique, forme et substance,* Paris: Hermès, 325-345.

Welkowitz, J., Cohen, B. H., & Ewen, R. B. (2012), *Introductory statistics for the behavioral sciences*, John Wiley & Sons, 7ème édition.

Wells, J.C., (1997), SAMPA computer readable phonetic alphabet, In Gibbon, D., Moore, R. and Winski, R. (eds.), 1997. *Handbook of Standards and Resources for Spoken Language Systems*. Berlin and New York: Mouton de Gruyter. Part IV, section B.

Werker, J. F., & Tees, R. C. (1984), Cross-language speech perception: Evidence for perceptual reorganization during the first year of life, *Infant behavior and development*, *7*(1), 49-63.

Whalen, D. H., Levitt, A. G. & Goldstein, L. M. (2007), VOT in the babbling of French-and English-learning infants, *Journal of Phonetics, 35*, 341-352.

Whiteside, S. P., Dobbin, R. & Henry, L. (2003), Patterns of variability in voice onset time: A developmental study of motor speech skills in humans, *Neuroscience Letters, 347*, 29-32.

Wilcox, K.M., Morris, S. (1999), *Children's Speech Intelligibility Measure,* Orlando, FL : The Psychological Corporation.

Yu, V., De Nil, L. F., & Pang, E. W. (2014), Effects of Age, Sex and Syllable Number on Voice Onset Time: Evidence from Children's Voiceless Aspirated Stops, *Language and Speech*, 58, 152–167.

Liste des figures

Figure 1.1 : ordre d'apparition des phonèmes chez des enfants francophones (tiré de Rondal, 1999) 26
Figure 2.1 : schéma de l'oreille, tiré de Brin et al. (2004) 34
Figure 2.2 : schéma d'une coupe transversale de la cochlée, figure tirée de Pujol et al. (2013), www.cochlea.eu. 1. Canal cochléaire, 2. Rampe vestibulaire, 3. Rampe tympanique, 4. Ganglion cochléaire, 5. Nerf auditif. 35
Figure 2.3 : distribution tonotopique des fréquences à l'intérieur de la cochlée, schémas tirés de Pujol et al. (2013), www.cochlea.eu. 36
Figure 2.4 : Classification des types de surdités chez l'enfant, tirée de Mondain et al. (2005) 37
Figure 2.5 : composant externe (gauche) et composant interne (droite) d'un implant cochléaire, tiré du site du fabricant Med El, consulté le 19 mars 2016 : www.medel.com 40
Figure 2.6 : schéma représentant un implant cochléaire en fonctionnement, tiré du site internet du fabricant MedEl, consulté le 19 mars 2016. 41
Figure 2.7 : spectrogrammes de quatre mots en néerlandais (de gauche à droite : bus, souvent, stylo, vide), produits de façon naturelle (gauche) et avec un implant cochléaire (droite), tirés de Fuller et al. (2014) 42
Figure 2.8 : Présentation du code LPC utilisé en anglais américain (Langue Parlée Complétée ou Langage Parlé Complété) 48
Figure 3.1 : espace vocalique F1-F2 d'un enfant CI dans l'étude de Hocevar-Bolthezar et al. (2008), pré- vs post-implantation 54
Figure 3.2 : évolution de l'organisation de l'espace vocalique de l'enfant 1 et de celui de sa mère (dernière figure) dans l'étude de Kunisue et al. (2006) 55
Figure 3.3 : évolution de l'espace vocalique des enfants CI et comparaison avec l'espace vocalique d'enfants NH dans l'étude de Liker et al. (2007) 55
Figure 3.4 : comparaison de l'espace vocalique des enfants sourds profonds HA (triangles et ligne en pointillés), sourds modérés et sévères HA (cercles et ligne en pointillés) et NH (carrés et ligne pleine) dans l'étude de Ryalls et al. (2003) 62
Figure 3.5 : comparaison de l'espace vocalique des enfants CI (signe +), HA (losanges) et NH (cercles) dans l'étude de Verhoeven et al. (2016) 62
Figure 3.6 : comparaison de l'espace vocalique des enfants CI, HA et NH dans l'étude de Baudonck et al. (2010) 63
Figure 3.7 : comparaison de l'espace vocalique des enfants CI (cercles et ligne en pointillés), HA (triangles et ligne en pointillés) et NH (carrés et ligne pleine) dans l'étude de Horga et Liker (2006) 63
Figure 3.8 : résultats (proportion de consonnes correctement produites) de l'étude de Geffner (1980), pour les modes (gauche) et lieux d'articulation (droite). 65
Figure 3.9 : résultats (proportion de consonnes correctement produites) de l'étude de Dillon et al. (2004) par lieu d'articulation cible. 66
Figure 3.10 : résultats (proportion de consonnes correctement produites) de l'étude de Peng et al. (2004) par mode d'articulation cible. 68
Figure 3.11 : durée de VOT de /t/ (voiceless) et /d/ (voiced) chez les enfants normo-entendants (HC), porteurs d'implants cochléaires (CI) et sourds profonds (PD), dans l'étude de Horga & Liker (2006) 70
Figure 3.12 : durées des VOT de /t/ et /d/ chez les trois groupes d'enfants (TC : enfants sourds CI, communication signée et orale, Oral : enfants sourds CI, communication orale, NH : groupe contrôle, enfants normo-entendants) dans l'étude de Uchanski & Geers (2003). WNL : « within normal limits » 70
Figure 3.13 : centre de gravité des consonnes /s/ et /ʃ/ du groupe contrôle NH et des quatre sessions d'enregistrement des enfants CI dans l'étude de Mildner & Liker (2008) 72
Figure 3.14 : centre de gravité (haut), skewness (bas, gauche) et kurtosis (bas, droite) des consonnes /s/ et /ʃ/ chez les trois groupes d'enfants (TC : enfants sourds CI, communication signée et orale, Oral : enfants sourds CI, communication orale, NH : groupe contrôle, enfants normo-entendants) dans l'étude de Uchanski & Geers (2003) 73
Figure 3.15 : centre de gravité des consonnes /s/ et /ʃ/ chez les deux groupes d'enfants CI et NH en appariés en âge chronologique (gauche) et en âge auditif dans l'étude de Todd et al. (2011) 74
Figure 4.1 : âge (gauche) et classe (droite) des enfants (groupes NH et CI) ayant participé à l'étude de production 95
Figure 4.2 : annotation du fichier NH21_7_R2_bato.wav avec le logiciel Praat 108
Figure 4.3 : schémas articulatoires des voyelles (orales et nasales) du français, tirés de Léon (2007) 111
Figure 4.4 : valeurs des 1er et 2ème formants (gauche) et représentation F1-F2 du système vocalique du français (droite), tiré de Delattre (1948) 112
Figure 4.5 : représentation spectrale (fréquence-amplitude) d'une voyelle /a/ produite par l'enfant NH26 lors de la tâche de répétition : en haut, son original (fréquence d'échantillonnage de 44 100 Hz), et en bas, son sous-échantillonné (fréquence d'échantillonnage de 11 025 Hz) 114
Figure 4.6 : représentation spatiale F1-F2 des voyelles produites par 76 locuteurs (enfants, femmes et hommes) dans l'étude de Peterson et Barney (1952) 116

Figure 4.7 : représentation spatiale F1-F2 des mêmes voyelles produites par 48 locuteurs (enfants, hommes) dans l'étude de Peterson et Barney (1952) 117
Figure 4.8 : comparaison des valeurs de F1 et F2 (en kHz) de la voyelle /ɑ/, en fonction de l'âge et du genre des participants, dans l'étude de Huber et al. (1999) 118
Figure 4.9 : Représentation spatiale F1-F2 des données de l'étude comparative des six méthodes de normalisation (de haut en bas et de gauche à droite : données en Hertz, en Bark, données normalisées en scores-z de Lobanov, avec notre adaptation du score-z de Lobanov, avec la méthode de Nearey 1 et de Gerstman), ellipses de confiance à 95%, les nombres correspondent aux âges des participants. 125
Figure 4.10 : schémas articulatoires des consonnes occlusives (orales et nasales) du français, tirés de Léon (2007) . 129
Figure 4.11 : schémas articulatoires des consonnes fricatives du français, tirés de Léon (2007) 130
Figure 4.12 : mesures du VOT d'une bilabiale voisée (haut) dans le mot « bicyclette », et d'une bilabiale non-voisée (bas) dans le mot « pistache » produites en répétition par l'enfant NH29 (les échelles temporelles sont différentes pour les deux spectrogrammes) 134
Figure 4.13 : représentation des 1er (haut, gauche), 3ème (haut, droite) et 4ème (bas) moments spectraux, dans l'étude de Nissen & Fox (2005) 136
Figure 4.14 : représentation du 1er moment spectral par lieu d'articulation (gauche) et par lieu et contexte vocalique (droite), dans l'étude de Nittrouer (1995) 137
Figure 4.15 : fenêtre utilisée pour mesurer les quatre moments dans le mot « ficelle » produit en répétition par l'enfant NH24 139
Figure 4.16 : représentation de stimuli synthétiques utilisés pour une expérience de reconnaissance de consonnes (bilabiales, alvéolaires et vélaires) par la transition du F2, tirée de Delattre et al. 1955 140
Figure 4.17 : équations de locus chez des adultes pour /b/ (gauche), /d/ (centre) et /g/ (droite), valeurs moyennes du groupe dans des séquences CV, avec huit voyelles différentes dans l'étude de Fowler (1994) 141
Figure 4.18 : comparaison des pentes (« Slope ») des équations de locus, en fonction de l'antériorité de la voyelle (« front » ou « back »), figure tirée de Fowler (1994) 142
Figure 4.19 : schéma présentant le principe général de l'équation de locus, tiré de Sussman & Shore (1999) 142
Figure 4.20 : évolution de 7 à 40 mois de la pente des équations de locus des bilabiales (en haut à gauche), des alvéolaires (en haut à droite) et vélaires (bas) dans l'étude de Sussman et al. (1999) 145
Figure 4.21 : mesure du F2, dans une séquence /do/ du mot « dauphin », produit par l'enfant NH28 146
Figure 5.1 : représentation des catégories vocaliques dans l'espace F1-F2 (haut) et F2-F3 (bas) par groupe, en scores-z 155
Figure 5.2 : représentation du F1 (en scores-z) par degré de hauteur de la voyelle, pour chaque groupe (enfants CI et NH). 156
Figure 5.3 : représentation du F1 (en scores-z) par voyelle, pour chaque groupe (enfants CI et NH). 157
Figure 5.4 : représentation du F2 (en scores-z) par degré d'antériorité, pour chaque groupe (enfants CI et NH). 158
Figure 5.5 : représentation du F2 (en scores-z) par catégorie vocalique, pour chaque groupe (enfants CI et NH). .. 159
Figure 5.6 : représentation du F3 (en scores-z) par catégorie vocalique, pour chaque groupe (enfants CI et NH). .. 160
Figure 5.7 : représentations par ellipses (haut) et boxplots (bas) de la dispersion de chaque catégorie vocalique sur F1-F2 (en scores-z) par catégorie vocalique, pour chaque groupe (enfants CI et NH) 161
Figure 5.8 : représentations par ellipses (haut) et boxplots (bas) de la distance euclidienne moyenne F2-F3 (en scores-z) par catégorie vocalique, pour chaque groupe (enfants CI et NH) 162
Figure 5.9 : Différences F2-F1 (haut) et F3-F2 (bas) en Bark, par catégorie vocalique, pour chaque groupe (enfants CI et NH), la ligne horizontale étant fixée à 3,5 Bark 163
Figure 5.10 : évolution du F1 (scores-z) en fonction de l'âge chronologique (en mois), pour les enfants NH – un point par voyelle produite et par enfant 166
Figure 5.11 : évolution du F2 (scores-z) en fonction de l'âge chronologique (en mois), pour les enfants NH – un point par voyelle produite et par enfant 166
Figure 5.12 : évolution du F3 (scores-z) en fonction de l'âge chronologique (en mois), pour les enfants NH – un point par voyelle produite et par enfant 167
Figure 5.13 : évolution de la dispersion F1-F2 moyenne (scores-z) en fonction de l'âge chronologique (en mois), pour les enfants NH – un point par catégorie vocalique et par enfant 167
Figure 5.14 : évolution de la dispersion F2-F3 moyenne (scores-z) en fonction de l'âge chronologique (en mois), pour les enfants NH – un point par catégorie vocalique et par enfant 168
Figure 5.15 : F1 en fonction de l'âge à l'implantation (haut) et de l'âge auditif (bas) en mois, par degré de hauteur des voyelles, pour les enfants CI – un point par voyelle produite et par enfant 169
Figure 5.16 : F2 en fonction de l'âge à l'implantation (haut) et de l'âge auditif (bas), en mois, par degré de hauteur des voyelles, pour les enfants CI – un point par voyelle produite et par enfant 170
Figure 5.17 : F3 en fonction de l'âge à l'implantation (haut) et de l'âge auditif (bas), en mois, par degré de hauteur des voyelles, pour les enfants CI – un point par voyelle produite et par enfant 171
Figure 5.18 : Dispersion F1-F2 (en scores-z), en fonction de l'âge à l'implantation (haut) et de l'âge auditif (bas), en mois, par voyelle, pour les enfants CI – un point par catégorie vocalique et par enfant 172

Figure 5.19 : Dispersion F2-F3 par rapport au centre de la catégorie vocalique en fonction de l'âge à l'implantation (haut) et de l'âge auditif (bas), en mois, par voyelle, pour les enfants CI – un point par catégorie vocalique et par enfant 172
Figure 5.20 : F1 (scores-z) pour chaque degré de hauteur de voyelle, par groupe et tâche (répétition ou dénomination) – « modele » : stimuli (locutrice adulte) utilisés lors de la tâche de répétition 174
Figure 5.21 : F2 (scores-z) pour chaque degré d'antériorité de voyelle, par groupe et tâches (répétition ou dénomination) 175
Figure 5.22 : F3 (scores-z) pour chaque degré d'antériorité de voyelle, par groupe et tâche (répétition ou dénomination) 175
Figure 5.23 : Dispersion F1-F2 (en scores-z) pour chaque voyelle, par groupe et tâche (répétition ou dénomination) 176
Figure 5.24 : dispersion F2-F3 (scores-z) pour chaque voyelle, par groupe et tâche (répétition ou dénomination) .. 177
Figure 5.25 : F1 (scores-z) pour chaque degré de hauteur de voyelle, par groupe et degré de complexité du mot (mots simples, codés « 0 » ou mots complexes, codés « 1 ») 178
Figure 5.26 : F2 (scores-z) pour chaque degré d'antériorité de voyelle, par groupe et degré de complexité du mot (mots simples, codés « 0 » ou mots complexes, codés « 1 »)..... 179
Figure 5.27 : F3 (scores-z) pour chaque catégorie vocalique, par groupe et degré de complexité du mot (mots simples, codés « 0 » ou mots complexes, codés « 1 ») 180
Figure 5.28 : Dispersion F1-F2 par rapport au centre de la voyelle (scores-z) pour chaque degré de hauteur de voyelle, par groupe et degré de complexité du mot (mots simples, codés « 0 » ou mots complexes, codés « 1 ») 181
Figure 5.29 : Distance F2-F3 par rapport au centre de la voyelle (scores-z) pour chaque degré de hauteur de voyelle, par groupe et degré de complexité du mot (mots simples, codés « 0 » ou mots complexes, codés « 1 ») 182
Figure 5.30 : distribution du VOT (en sec) par groupe et lieu d'articulation 190
Figure 5.31 : VOT (en sec) par groupe et lieu d'articulation..... 191
Figure 5.32 : distance entre les VOT des consonnes voisées et non-voisées (en sec) par lieu d'articulation, pour chaque groupe 193
Figure 5.33 : VOT (en sec) par groupe, par lieu d'articulation et par contexte vocalique 194
Figure 5.34 : distance entre les VOT des consonnes voisées et non-voisées (en sec) par lieu d'articulation et par contexte vocalique, pour chaque groupe 196
Figure 5.35 : VOT (en sec) en fonction de l'âge chronologique (en mois) par contexte vocalique et lieu d'articulation, pour les enfants NH..... 197
Figure 5.36 : Distance entre VOT des consonnes voisées et non-voisées (en sec) en fonction de l'âge chronologique (en mois) par contexte vocalique et lieu d'articulation, pour les enfants NH..... 198
Figure 5.37 : évolution du VOT (en sec) en fonction de l'âge à l'implantation (haut) et de l'âge auditif (bas), en mois, par contexte vocalique et lieu d'articulation, pour les enfants CI 200
Figure 5.38 : évolution de la distance entre VOT des consonnes voisées et non-voisées (en sec) en fonction de l'âge à l'implantation (haut) et de l'âge auditif (bas), en mois, par contexte vocalique et lieu d'articulation, pour les enfants CI..... 202
Figure 5.39 : VOT (en sec) par lieu d'articulation, par groupe et par tâche (répétition ou dénomination) 204
Figure 5.40 : Distance entre VOT des consonnes voisées et non-voisées (en sec) par lieu d'articulation, par groupe et par tâche (répétition ou dénomination) 205
Figure 5.41 : VOT (en sec) par lieu d'articulation, par groupe et par degré de complexité du mot 207
Figure 5.42 : Distance entre VOT des consonnes voisées et non-voisées (en sec) par lieu d'articulation, par groupe et par degré de complexité (mots simples ou mots complexes) 208
Figure 5.43 : centre de gravité (en Hz) par groupe et lieu d'articulation..... 216
Figure 5.44 : distance entre les centres de gravité de /s/ et /ʃ/ (en Hz) par groupe 217
Figure 5.45 : Ecart-type (en Hz) par groupe et lieu d'articulation..... 218
Figure 5.46 : skewness (coefficient) par groupe et lieu d'articulation 219
Figure 5.47 : kurtosis (coefficient) par groupe et lieu d'articulation 220
Figure 5.48 : centre de gravité (en Hz) par lieu d'articulation, par groupe et par contexte vocalique 224
Figure 5.49 : distance moyenne (en Hz) entre les centres de gravité de /s/ et /ʃ/ par groupe et contexte vocalique225
Figure 5.50 : écart-type (en Hz) par lieu d'articulation, par groupe et par contexte vocalique 226
Figure 5.51 : skewness (coefficient) par lieu d'articulation, par groupe et par contexte vocalique 227
Figure 5.52 : kurtosis (coefficient) par lieu d'articulation, par groupe et par contexte vocalique..... 228
Figure 5.53 : centre de gravité (en Hz) en fonction de l'âge chronologique (en mois) par contexte vocalique et lieu d'articulation, pour les enfants NH 231
Figure 5.54 : distance entre les centres de gravité de /s/ et /ʃ/ (en Hz) en fonction de l'âge chronologique (en mois) par contexte vocalique, pour les enfants NH 232
Figure 5.55 : écart-type (en Hz) en fonction de l'âge chronologique (en mois) par contexte vocalique et lieu d'articulation, pour les enfants NH 233

Figure 5.56 : skewness (coefficient) en fonction de l'âge chronologique (en mois) par contexte vocalique et lieu d'articulation, pour les enfants NH ... 234
Figure 5.57 : kurtosis (coefficient) en fonction de l'âge chronologique (en mois) par contexte vocalique et lieu d'articulation, pour les enfants NH ... 235
Figure 5.58 : Centre de gravité (en Hz) en fonction de l'âge auditif (haut, en mois) et âge à l'implantation (bas, en mois) par contexte vocalique et lieu d'articulation, pour les enfants CI ... 237
Figure 5.59 : distance entre les centres de gravité de /s/ et /ʃ/ (en Hz) en fonction de l'âge auditif (gauche, en mois) et âge à l'implantation (droite, en mois) par contexte vocalique, pour les enfants CI ... 238
Figure 5.60 : Ecart-type (en Hz) en fonction de l'âge auditif (haut, en mois) et âge à l'implantation (bas, en mois) par contexte vocalique et lieu d'articulation, pour les enfants CI ... 239
Figure 5.61 : Skewness (coefficient) en fonction de l'âge auditif (haut, en mois) et âge à l'implantation (bas, en mois) par contexte vocalique et lieu d'articulation, pour les enfants CI ... 241
Figure 5.62 : Kurtosis (coefficient) en fonction de l'âge auditif (gauche, en mois) et âge à l'implantation (droite, en mois) par contexte vocalique et lieu d'articulation, pour les enfants CI ... 242
Figure 5.63 : centre de gravité (en Hz) par lieu d'articulation, par groupe et par tâche (répétition ou dénomination) ... 244
Figure 5.64 : distance moyenne (en Hz) entre les centres de gravité de /s/ et /ʃ/ par groupe et tâche (répétition ou dénomination) ... 245
Figure 5.65 : écart-type (en Hz) par lieu d'articulation, par groupe et par tâche (répétition ou dénomination) ... 246
Figure 5.66 : skewness (coefficient) par lieu d'articulation, par groupe et par tâche (répétition ou dénomination) ... 246
Figure 5.67 : kurtosis (coefficient) par lieu d'articulation, par groupe et par tâche (répétition ou dénomination) ... 247
Figure 5.68 : centre de gravité (en Hz) par lieu d'articulation, par groupe et par degré de complexité du mot ... 248
Figure 5.69 : distance moyenne (en Hz) entre les centres de gravité de /s/ et /ʃ/ par groupe et degré de complexité du mot ... 249
Figure 5.70 : écart-type (en Hz) par lieu d'articulation, par groupe et par degré de complexité du mot ... 250
Figure 5.71 : skewness (coefficient) par lieu d'articulation, par groupe et par degré de complexité du mot ... 250
Figure 5.72 : kurtosis (coefficient) par lieu d'articulation, par groupe et par degré de complexité du mot ... 251
Figure 5.73 : pentes d'équation de locus (coefficient) par groupe et lieu d'articulation ... 259
Figure 5.74 : pente de l'équation de locus (coefficient) en fonction de l'âge chronologique (en mois) par lieu d'articulation, pour les enfants NH ... 261
Figure 5.75 : pente de l'équation de locus (coefficient) en fonction de l'âge à l'implantation (haut) et de l'âge auditif (bas), en mois par lieu d'articulation, pour les enfants CI ... 262
Figure 5.76 : pentes d'équation de locus (coefficient) par groupe (Adultes, enfants CI et enfants NH), par tâche et lieu d'articulation ... 264
Figure 6.1 : écran d'accueil de l'expérience d'intelligibilité, proposée sur Praat (Boersma & Weenink, 2015) ... 288
Figure 6.2 : interface de réponse de la phase d'entrainement et de la phase de test d'intelligibilité ... 289
Figure 6.3 : choix des notes d'intelligibilité données par chaque groupe d'auditeurs (en pourcentage) ... 292
Figure 6.4 : classement des stimuli par niveau d'intelligibilité des groupes CI et NH pour chaque groupe d'auditeurs (1=énoncé le plus intelligible et 130=énoncé le moins intelligible) ... 293
Figure 6.5 : distribution des notes d'intelligibilité pour chaque enfant du groupe CI (haut) et NH (bas) ... 294
Figure 6.6 : distribution des notes d'intelligibilité par groupe (CI et NH) ... 295
Figure 6.7 : pourcentage des notes 5, 6 et 7 données à chaque enfant des deux groupes (CI et NH) ... 295
Figure 6.8 : notes d'intelligibilité obtenues par chaque enfant NH, par ordre d'âge chronologique (de gauche à droite : du plus jeune au plus âgé) ... 296
Figure 6.9 : notes d'intelligibilité obtenues par chaque enfant CI, par ordre d'âge à l'implantation (de gauche à droite : de l'âge à l'implantation le plus précoce au plus tardif) ... 297
Figure 6.10 : notes d'intelligibilité obtenues par chaque enfant CI, par ordre de la durée d'utilisation de l'implant (de gauche à droite : de la durée la plus courte à la plus longue) ... 298
Figure 6.11 : notes d'intelligibilité obtenues par chaque enfant CI, par ordre d'âge chronologique (de gauche à droite : de l'enfant le plus jeune au plus âgé) ... 299
Figure 6.12 : Comparaison des notes d'intelligibilité obtenues par chaque enfant CI, et des caractéristiques acoustiques étudiées dans l'étude de production (de haut en bas : espace vocalique F1-F2, VOT des occlusives voisées et non-voisées, centre de gravité des fricatives) ... 301

Liste des tables

Table 1.1 : chronologie de l'évolution de la perception et de la production de la parole chez l'enfant de 0 à 48 mois (Boysson-Bardies, 1996, cité dans Wauquier-Gravelines, 2005) 24
Table 2.1 : degré de surdité et niveau de perte auditive (d'après Mondain et al. (2005) et Lina-Granade & Truy (2005)) 38
Table 3.1 : caractéristiques des cohortes d'enfants, dans la sélection d'études transversales présentées dans ce chapitre 57
Table 3.2 : caractéristiques des productions enregistrées, dans la sélection d'études transversales présentées dans ce chapitre : voyelles, mesures effectuées, analyses réalisées et méthodologie d'élicitation des voyelles 59
Table 3.3 : effets des caractéristiques individuelles des enfants CI sur l'intelligibilité (classement par ordre alphabétique du premier auteur). Oui : effet sur l'intelligibilité, non : aucun effet sur l'intelligibilité. 82
Table 3.4 : effets des caractéristiques acoustiques des productions de parole sur l'intelligibilité (classement par ordre alphabétique du premier auteur). Oui : effet sur l'intelligibilité, non : aucun effet sur l'intelligibilité. 85
Table 4.1 : âge chronologique, âge à l'implantation, durée d'utilisation de l'implant, sexe, classe, domicile, L1 des enfants, langues parlée à la maison et L1 des parents des enfants 97
Table 4.2 : nombre et types d'implants utilisés par les enfants, degrés de surdité avant implantation, et niveau d'audition après implantation (données récoltées par Corinne Aubert (Aubert, 2016), stagiaire au Gipsa-Lab en 2015-2016 dans le cadre de son stage et mémoire d'orthophonie) – OD : oreille droite, OG : oreille gauche...98
Table 4.3 : Utilisation de LSF (Langue des Signes Française), LPC (Langage Parlé Complété) et Oral par les enfants sourds, porteurs d'implant cochléaire, avant et après implantation (le symbole ✓ signifie oui, et le symbole ✗ signifie non). 98
Table 4.4 : mots utilisés lors de la tâche de répétition pour l'étude des consonnes classés par degré de complexité articulatoire 105
Table 4.5 : mots utilisés lors de la tâche de répétition pour l'étude des voyelles classés par degré de complexité articulatoire 105
Table 4.6 : mots utilisés lors de la tâche de répétition pour l'étude de la coarticulation CV classés par lieu d'articulation de la consonne 106
Table 4.7 : correspondance entre l'Alphabet Phonétique International (IPA) et l'alphabet SAMPA 108
Table 4.8 : système vocalique minimal du français (dix voyelles orales et trois voyelles nasales), d'après Sten (1963), Vallée (1994), Calliope (1989), Walter (1977) et Léon (2007). Les voyelles nasales sont placées en dessous de leur équivalent oral, et les voyelles entre parenthèses sont celles qui ne sont pas communes à toutes les descriptions. 110
Table 4.9 : description de la convergence des formants des voyelles focales, tirée de Vaissière (2007) 113
Table 4.10 : résumé des caractéristiques d'études récentes sur la production de voyelles par des enfants et jeunes adultes porteurs d'implants cochléaires et utilisateurs d'aides auditives traditionnelles (classement par méthode de normalisation puis ordre alphabétique du premier auteur) 119
Table 4.11 : d'après Adank et al. (2004), classement des méthodes de normalisation 122
Table 4.12 : participants de notre étude de normalisation : enregistrements des enfants NH1 à NH réalisés dans le cadre de l'étude de Scarbel et al. (2012), enregistrements des adultes NH9 à NH12 réalisés par nous-mêmes. 123
Table 4.13 : méthodes de normalisation testées (où F_t correspond au formant après transformation et i au numéro du formant) 124
Table 4.14 : indices de distinction inter-catégories et de dispersion intra-catégories 127
Table 4.15 : système consonantique du français d'après Sten (1963), Walter (1977) et Léon (2007). Les cellules grisées correspondent aux consonnes qui ne sont pas présentes dans toutes les descriptions : /R/ est présente uniquement dans Sten (1963), /r/ et /ɣ/ dans Léon (2007) uniquement et /ʁ/ dans Léon (2007) et Walter (1977). 128
Table 5.1 : valeurs moyennes des trois premiers formants (avec écart-type) par groupe, par voyelle, en Hz et en scores-z. 151
Table 5.2 : valeurs moyennes des distances euclidiennes F1-F2 et F2-F3 par rapport au centre de la catégorie vocalique (avec écart-type) par groupe, par voyelle, en scores-z. 152
Table 5.3 : valeurs moyennes des différences F2-F1 et F3-F2 (avec écart-type) par groupe, par voyelle, Bark (en couleur : cases correspondant à des convergences formantiques, F_{i+1}-F_i < 3.5 Bark) 153
Table 5.4 : variables dépendantes, variables indépendantes et variable ayant un effet aléatoire, pris en compte dans l'analyse des catégories vocaliques 154
Table 5.5 : tableau de résultats au test statistique sur toutes les données (effets du groupe et du degré de hauteur sur le F1) 156
Table 5.6 : résumé du test post-hoc sur l'effet d'interaction entre degré de hauteur et groupe sur le F1 des voyelles 156

Table 5.7 : résumé du test post-hoc sur l'effet d'interaction sur les F1 des voyelles par paires de degrés de hauteur 156
Table 5.8 : tableau de résultats au test statistique sur toutes les données (effets du groupe et de la catégorie vocalique sur le F1) ... 157
Table 5.9 : tableau de résultats au test statistique sur toutes les données (effets du groupe et du degré d'antériorité sur le F2) ... 158
Table 5.10 : résumé du test post-hoc sur l'effet d'interaction entre degré de hauteur et groupe sur le F1 des voyelles ... 158
Table 5.11 : résumé du test post-hoc sur l'effet d'interaction sur les F2 des voyelles par paires de degrés d'antériorité ... 159
Table 5.12 : tableau de résultats au test statistique sur toutes les données (effets du groupe et de la catégorie vocalique sur le F2) ... 159
Table 5.13 : tableau de résultats au test statistique sur toutes les données (effets du groupe et de la catégorie vocalique sur le F3) ... 160
Table 5.14 : tableau de résultats au test statistique sur toutes les données (effets du groupe et de la catégorie vocalique sur la dispersion F1-F2) ... 161
Table 5.15 : tableau de résultats au test statistique sur toutes les données (effets du groupe et de la catégorie vocalique sur la dispersion F2-F3) ... 162
Table 5.16 : différences F2-F1 (en Bark), effet du groupe et de la catégorie vocalique ... 164
Table 5.17 : différences F3-F2 (en Bark), effet du groupe et de la catégorie vocalique ... 164
Table 5.18 : résumé des effets des variables groupe, catégorie vocalique et de l'interaction entre groupe et catégorie vocalique sur les formants F1, F2 et F3, ainsi que sur les distances euclidiennes F1-F2 et F2-F4 moyenne au centre de la catégorie ... 164
Table 5.19 : résumé des effets de l'âge chronologique et de l'interaction entre l'âge chronologique et catégorie vocalique sur les formants F1, F2 et F3, ainsi que sur la dispersion sur F1-F2 et F2-F3 ... 168
Table 5.20 : résumé des effets de l'âge (âge à l'implantation et âge auditif), et de l'interaction entre l'âge et catégorie vocalique sur les formants F1, F2 et F3, ainsi que sur les distances euclidiennes F1-F2 et F2-F4 moyenne au centre de la catégorie ... 173
Table 5.21 : tableau de résultats au test statistique sur toutes les données (effets de groupe, degré de hauteur des voyelles, tâche) ... 174
Table 5.22 : tableau de résultats au test statistique sur toutes les données (effets de groupe, catégorie vocalique, tâche sur la dispersion F1-F2 moyenne de chaque catégorie) ... 176
Table 5.23 : tableau de résultats au test statistique sur toutes les données (effets de groupe, catégorie vocalique, tâche sur la dispersion F2-F3 moyenne de chaque catégorie) ... 177
Table 5.24 : résumé des effets de la tâche (tâche de répétition ou dénomination), et de l'interaction entre groupe et tâche, et entre tâche et catégorie vocalique sur les formants F1, F2 et F3, ainsi que sur les dispersions F1-F2 et F2-F3 moyennes de chaque catégorie vocalique ... 178
Table 5.25 : tableau de résultats au test statistique sur toutes les données (effets de groupe, degré de hauteur et degré de complexité du mot sur le F1) ... 179
Table 5.26 : tableau de résultats au test statistique sur toutes les données (effets de groupe, degré d'antériorité et degré de complexité du mot sur le F2) ... 180
Table 5.27 : tableau de résultats au test statistique sur toutes les données (effets de groupe, catégorie vocalique et degré de complexité du mot sur le F3) ... 181
Table 5.28 : tableau de résultats au test statistique sur toutes les données (effets de groupe, catégorie vocalique et degré de complexité du mot sur la distance F1-F2 par rapport au centre de la catégorie vocalique) ... 182
Table 5.29 : tableau de résultats au test statistique sur toutes les données (effets de groupe, catégorie vocalique et degré de complexité du mot sur la distance F2-F3 par rapport au centre de la catégorie vocalique) ... 183
Table 5.30 : résumé des effets de la complexité du mot (« 0 » pour les mots simples, et « 1 » pour les mots complexes), et de l'interaction entre groupe et complexité, et entre complexité et catégorie vocalique sur les formants F1, F2 et F3, ainsi que sur les dispersions F1-F2 et F2-F3 moyennes de chaque catégorie vocalique ... 183
Table 5.31 : résumé des effets du groupe, de la catégorie vocalique de l'âge chronologique (enfants NH) de l'âge à l'implantation (enfants CI), de l'âge auditif (enfants CI), de la tâche et de la complexité du mot, seuls et en interactions sur la réalisation des caractéristiques acoustiques des voyelles. ... 184
Table 5.32 : valeurs moyennes du VOT (avec écart-type) par groupe, par consonne et par contexte vocalique ... 189
Table 5.33 : distance moyenne (en sec) entre les VOT de /p-b/, /t-d/ et /k-g/, par groupe et par contexte vocalique ... 190
Table 5.34 : variables dépendantes, variables indépendantes et variable ayant un effet aléatoire, pris en compte dans l'analyse des occlusives ... 190
Table 5.35 : tableau de résultats au test statistique sur toutes les données (effets de groupe, lieu d'articulation sur le VOT) ... 191
Table 5.36 : résumé du test post-hoc sur l'effet d'interaction entre groupe et lieu d'articulation sur le VOT des consonnes non-voisées ... 192

Table 5.37 : tableau de résultats au test statistique sur toutes les données (effets du groupe sur la distance entre VOT des consonnes voisées et non-voisées) 193
Table 5.38 : résumé des effets des variables groupe, lieu d'articulation et de l'interaction entre groupe et lieu d'articulation sur le VOT et la distance entre VOT des consonnes voisées et non-voisées 193
Table 5.39 : tableau de résultats au test statistique sur toutes les données (effets de groupe, lieu d'articulation et contexte vocalique sur le VOT) 195
Table 5.40 : résumé du test post-hoc sur l'effet d'interaction entre contexte vocalique et lieu d'articulation sur le VOT 195
Table 5.41 : tableau de résultats au test statistique sur toutes les données (effets du contexte vocalique sur la distance entre VOT des consonnes voisées et non-voisées) 196
Table 5.42 : résumé du test post-hoc sur l'effet d'interaction entre contexte vocalique et lieu d'articulation sur la distance entre les VOT des consonnes voisées et non-voisées 196
Table 5.43 : résumé des effets des variables groupe, contexte vocalique et des interactions entre groupe et contexte vocalique et contexte vocalique et lieu d'articulation sur le VOT et la distance entre VOT des consonnes voisées et non-voisées 197
Table 5.44 : tableau de résultats au test statistique sur les données des enfants NH (effets de groupe, lieu d'articulation, contexte vocalique et âge chronologique sur le VOT) 198
Table 5.45 : tableau de résultats au test statistique sur les données des enfants NH (effets de groupe, lieu d'articulation, contexte vocalique et âge chronologique sur la distance entre VOT des consonnes voisées et non-voisées) 199
Table 5.46 : résumé des effets de l'âge chronologique seul et en interactions avec le groupe, le contexte vocalique et le lieu d'articulation sur le VOT et la distance entre VOT des consonnes voisées et non-voisées chez les enfants NH 199
Table 5.47 : tableau de résultats au test statistique sur les données des enfants CI (effets de groupe, lieu d'articulation, contexte vocalique, âge auditif et âge à l'implantation sur le VOT) 201
Table 5.48 : tableau de résultats au test statistique sur les données des enfants CI (effets de groupe, lieu d'articulation, contexte vocalique, âge à l'implantation et âge auditif sur la distance entre VOT des consonnes voisées et non-voisées) 202
Table 5.49 : résumé des effets de l'âge à l'implantation et de l'âge auditif, seuls et en interactions avec le groupe, le contexte vocalique et le lieu d'articulation sur le VOT et la distance entre VOT des consonnes voisées et non-voisées chez les enfants CI 203
Table 5.50 : tableau de résultats au test statistique sur les toutes données (effets de groupe, lieu d'articulation, et tâche sur le VOT) 204
Table 5.51 : résumé du test post-hoc sur l'effet d'interaction entre groupe et tâche sur le VOT des consonnes non-voisées /p/, /t/ et /k/ 204
Table 5.52 : tableau de résultats au test statistique sur toutes les données (effets de groupe, lieu d'articulation et tâche sur la distance entre VOT des consonnes voisées et non-voisées) 205
Table 5.53 : résumé des tests de l'effet de la tâche sur la distance entre VOT des consonnes voisées et non-voisées, par groupe et lieu d'articulation 205
Table 5.54 : résumé de l'effet de la tâche (Dénomination ou Répétition), seule et en interaction avec le groupe et le lieu d'articulation sur le VOT et la distance entre VOT des consonnes voisées et non-voisées 206
Table 5.55 : tableau de résultats au test statistique sur toutes les données (effets de groupe, lieu d'articulation, contexte vocalique et degré de complexité sur le VOT) 207
Table 5.56 : résumé du test post-hoc sur l'effet d'interaction entre degré de complexité du mot et lieu d'articulation sur le VOT 207
Table 5.57 : tableau de résultats au test statistique sur toutes les données (effets de groupe, lieu d'articulation et degré de complexité du mot sur la distance entre VOT des consonnes voisées et non-voisées) 208
Table 5.58 : résumé de l'effet de la complexité du mot, seule et en interaction avec le groupe et le lieu d'articulation sur le VOT et la distance entre VOT des consonnes voisées et non-voisées 209
Table 5.59 : résumé de tous les effets étudiés dans le cadre de notre étude de production d'occlusives par les enfants CI et NH (« n.s. » indique une différence non significative) 210
Table 5.60 : valeurs moyennes (avec écart-type) par groupe, par consonne et par contexte vocalique pour chaque moment spectral étudié (centre de gravité, écart-type, skewness, kurtosis) 214
Table 5.61 : distance moyenne (en Hz) entre les centres de gravité de /s/ et /ʃ/, par groupe et par contexte vocalique 215
Table 5.62 : variables dépendantes, variables indépendantes et variable ayant un effet aléatoire, pris en compte dans l'analyse des fricatives 215
Table 5.63 : tableau de résultats au test statistique sur toutes les données (effets de groupe, lieu d'articulation sur le centre de gravité) 216
Table 5.64 : résumé du test post-hoc sur l'effet d'interaction entre lieu d'articulation et groupe sur le premier moment spectral (centre de gravité) 216

Table 5.65 : résumé du test post-hoc sur l'effet d'interaction entre groupe et lieu d'articulation sur le premier moment spectral (centre de gravité) 217
Table 5.66 : tableau de résultats au test statistique sur toutes les données (effets du groupe sur la distance entre centres de gravité de /s/ et /ʃ/) 217
Table 5.67 : tableau de résultats au test statistique sur toutes les données (effets de groupe, lieu d'articulation, contexte vocalique sur l'écart-type) 218
Table 5.68 : résumé du test post-hoc sur l'effet d'interaction entre lieu d'articulation et groupe sur le deuxième moment spectral (écart-type) 218
Table 5.69 : résumé du test post-hoc sur l'effet d'interaction entre groupe et lieu d'articulation sur le deuxième moment spectral (écart-type) 219
Table 5.70 : tableau de résultats au test statistique sur toutes les données (effets de groupe, lieu d'articulation sur le skewness) 219
Table 5.71 : résumé du test post-hoc sur l'effet d'interaction entre lieu d'articulation et groupe sur le troisième moment spectral (skewness) 220
Table 5.72 : résumé du test post-hoc sur l'effet d'interaction entre groupe et lieu d'articulation sur le troisième moment spectral (skewness) 220
Table 5.73 : tableau de résultats au test statistique sur toutes les données (effets de groupe, lieu d'articulation sur le kurtosis) 221
Table 5.74 : résumé du test post-hoc sur l'effet d'interaction entre lieu d'articulation et groupe sur le kurtosis 221
Table 5.75 : résumé du test post-hoc sur l'effet d'interaction entre groupe et lieu d'articulation sur le kurtosis 221
Table 5.76 : résumé des effets des variables groupe, lieu d'articulation et de l'interaction entre groupe et lieu d'articulation sur chaque moment spectral (le symbole ≃ indique une absence de différence significative) 222
Table 5.77 : tableau de résultats au test statistique sur toutes les données (effets de groupe, lieu d'articulation, contexte vocalique sur le centre de gravité) 224
Table 5.78 : résumé du test post-hoc sur l'effet d'interaction entre contexte vocalique et lieu d'articulation sur le centre de gravité 224
Table 5.79 : résumé du test post-hoc sur l'effet d'interaction entre contexte vocalique et lieu d'articulation sur le centre de gravité, chez les enfants NH 225
Table 5.80 : tableau de résultats au test statistique sur toutes les données (effets de groupe, contexte vocalique sur la distance entre centres de gravité de /s/ et /ʃ/) 226
Table 5.81 : tableau de résultats au test statistique sur toutes les données (effets de groupe, lieu d'articulation, contexte vocalique sur l'écart-type) 226
Table 5.82 : résumé du test post-hoc sur l'effet d'interaction entre contexte vocalique et lieu d'articulation sur l'écart-type 227
Table 5.83 : tableau de résultats au test statistique sur toutes les données (effets de groupe, lieu d'articulation, contexte vocalique sur le skewness) 227
Table 5.84 : résumé du test post-hoc sur l'effet d'interaction entre contexte vocalique et lieu d'articulation sur le deuxième moment spectral (skewness) 228
Table 5.85 : résumé des tests post-hoc sur l'effet d'interaction entre contexte vocalique et lieu d'articulation sur le deuxième moment spectral (skewness) chez les enfants NH (gauche) et les enfants CI (droite) 228
Table 5.86 : tableau de résultats au test statistique sur toutes les données (effets de groupe, lieu d'articulation, contexte vocalique sur le kurtosis) 229
Table 5.87 : résumé du test post-hoc sur l'effet d'interaction entre contexte vocalique et lieu d'articulation sur le troisième moment spectral (kurtosis) 229
Table 5.88 : résumé du test post-hoc sur l'effet d'interaction entre contexte vocalique et lieu d'articulation sur le troisième moment spectral (kurtosis) chez les enfants NH 229
Table 5.89 : résumé des effets de la variable contexte vocalique et de l'interaction entre contexte vocalique et groupe et entre contexte vocalique et lieu d'articulation sur chaque moment spectral (n.s. indique une absence de différence significative) 230
Table 5.90 : tableau de résultats au test statistique sur les données des enfants NH (effets de l'âge chronologique, du lieu d'articulation, du contexte vocalique sur le centre de gravité) 231
Table 5.91 : tableau de résultats au test statistique sur les données des enfants NH (effets de l'âge chronologique et du contexte vocalique sur la distance entre les centres de gravité de /s/ et /ʃ/) 232
Table 5.92 : tableau de résultats au test statistique sur les données des enfants NH (effets de l'âge chronologique, du lieu d'articulation, du contexte vocalique sur l'écart-type) 233
Table 5.93 : tableau de résultats au test statistique sur les données des enfants NH (effets de l'âge chronologique, du lieu d'articulation, du contexte vocalique sur le skewness) 234
Table 5.94 : tableau de résultats au test statistique sur les données des enfants NH (effets de l'âge chronologique, du lieu d'articulation, du contexte vocalique sur le kurtosis) 235

Table 5.95 : résumé des effets chez les enfants NH de la variable « âge chronologique », seule et en interaction avec le groupe, le lieu d'articulation et contexte vocalique sur chaque moment spectral (« n.s. » indique une différence non significative) 236
Table 5.96 : tableau de résultats au test statistique sur les données des enfants CI (effets de l'âge auditif et de l'âge à l'implantation, du lieu d'articulation, du contexte vocalique sur le centre de gravité) 237
Table 5.97 : tableau de résultats au test statistique sur les données des enfants CI (effets de l'âge auditif, de l'âge à l'implantation et du contexte vocalique sur la distance entre les centres de gravité de /s/ et /ʃ/) 239
Table 5.98 : tableau de résultats au test statistique sur les données des enfants CI (effets de l'âge auditif et de l'âge à l'implantation, du lieu d'articulation, du contexte vocalique sur l'écart-type) 240
Table 5.99 : tableau de résultats au test statistique sur les données des enfants CI (effets de l'âge auditif et de l'âge à l'implantation, du lieu d'articulation, du contexte vocalique sur le troisième moment, ou skewness) 241
Table 5.100 : tableau de résultats au test statistique sur les données des enfants CI (effets de l'âge auditif et de l'âge à l'implantation, du lieu d'articulation, du contexte vocalique sur le quatrième moment ou kurtosis) 243
Table 5.101 : résumé des effets chez les enfants CI des variables « âge auditif » et « âge à l'implantation », seules et en interaction avec le groupe, le lieu d'articulation et contexte vocalique sur chaque moment spectral (« n.s. » indique une différence non significative) 243
Table 5.102 : tableau de résultats au test statistique sur toutes les données (effets de groupe, lieu d'articulation, et tâche sur le centre de gravité) 245
Table 5.103 : tableau de résultats au test statistique sur toutes les données (effets de groupe et tâche sur la distance entre centres de gravité de /s/ et /ʃ/) 245
Table 5.104 : résumé des effets de la variable tâche, seule et en interaction avec le groupe, le lieu d'articulation sur chaque moment spectral (« n.s. » indique une différence non significative) 247
Table 5.105 : tableau de résultats au test statistique sur toutes les données (effets de groupe, lieu d'articulation, et degré de complexité sur le centre de gravité) 248
Table 5.106 : résumé du test post-hoc sur l'effet d'interaction entre complexité du mot et lieu d'articulation sur le centre de gravité 248
Table 5.107 : tableau de résultats au test statistique sur toutes les données (effets de groupe et complexité lexicale sur la distance entre centres de gravité de /s/ et /ʃ/) 249
Table 5.108 : tableau de résultats au test statistique sur toutes les données (effets de groupe, lieu d'articulation et degré de complexité sur le troisième moment spectral ou skewness) 250
Table 5.109 : résumé du test post-hoc sur l'effet d'interaction entre complexité du mot et lieu d'articulation sur le troisième moment spectral (skewness) 251
Table 5.110 : résumé des effets de la variable « Complexité lexicale », seule et en interaction avec le groupe et le lieu d'articulation sur chaque moment spectral (« n.s. » indique une différence non significative) 252
Table 5.111 : résumé de tous les effets étudiés dans le cadre de notre étude de production de fricatives par les enfants CI et NH (« n.s. » indique une différence non significative) 253
Table 5.112 : valeurs moyennes des pentes d'équation de locus (avec écart-type) par groupe et par lieu d'articulation 258
Table 5.113 : variables dépendantes, variables indépendantes et variable ayant un effet aléatoire, pris en compte dans l'analyse des équations de locus 259
Table 5.114 : tableau de résultats au test statistique sur toutes les données, adultes et enfants (effets de groupe et lieu d'articulation sur la pente de l'équation de locus) 259
Table 5.115 : résumé du test post-hoc testant l'effet inter-groupe par lieu d'articulation sur la pente d'équation de locus 260
Table 5.116 : résumé du test post-hoc testant l'effet du lieu d'articulation par groupe sur la pente d'équation de locus 260
Table 5.117 : résumé des effets des variables « Groupe » et « Lieu d'articulation », seules et en interaction sur la pente de l'équation de locus (« n.s. » et ≃ indiquent une absence de différence significative) 260
Table 5.118 : tableau de résultats au test statistique sur les données des enfants NH (effets de l'âge chronologique, du lieu d'articulation sur la pente d'équation de locus) 261
Table 5.119 : résumé des effets de l'âge chronologique seul et en interaction avec le lieu d'articulation sur la pente de l'équation de locus 262
Table 5.120 : tableau de résultats au test statistique sur les données des enfants CI (effets de l'âge auditif, de l'âge à l'implantation et du lieu d'articulation sur la pente d'équation de locus) 263
Table 5.121 : résumé des effets de l'âge à l'implantation et de l'âge auditif seuls et en interaction avec le lieu d'articulation sur la pente de l'équation de locus 263
Table 5.122 : tableau de résultats au test statistique sur toutes les données, adultes et enfants (effets de groupe, tâche et lieu d'articulation sur la pente de l'équation de locus) 264
Table 5.123 : résumé des effets de la tâche seule et en interaction avec le groupe et le lieu d'articulation sur la pente de l'équation de locus 264

Table 5.124 : résumé de tous les effets étudiés dans le cadre de notre étude de production de la coarticulation dans des séquences CV par les enfants CI et NH, en comparaison avec un groupe d'adultes 265
Table 6.1 : résumé des procédures d'élicitation de parole, utilisées dans des études récentes sur l'intelligibilité de la parole d'enfants porteurs d'implants cochléaires (classement par ordre alphabétique du premier auteur) 277
Table 6.2 : résumé des méthodes de jugement de l'intelligibilité de quelques études récentes sur l'intelligibilité de la parole d'enfants porteurs d'implants cochléaires (classement par ordre alphabétique du premier auteur) 282
Table 6.3 : comparaison des méthodes de jugement (subjectives et objectives) utilisées par les auteurs de quelques études récentes sur l'intelligibilité de la parole d'enfants porteurs d'implants cochléaires. 283
Table 6.4 : Enfants des deux groupes (CI et NH) dont les productions ont été sélectionnées pour l'étude d'intelligibilité 290
Table 6.5 : âge moyen (avec écart-type) des participants à l'étude d'intelligibilité..... 291
Table 6.6 : distribution des notes attribuées pour chaque enfant, des groupes CI (gauche) et NH (droite), le mode pour chaque enfant est indiqué en couleur 294

Annexes

Annexe 1 : Formulaire de consentement pour l'étude de production – enfants NH

gipsa-lab

DOCUMENT D'INFORMATION et CONSENTEMENT DE PARTICIPATION

Remis aux parents des enfants sollicités pour participer à cette recherche
Investigateurs principaux : A. Vilain, H. Lœvenbruck, S. Schmerber
Intervenante : Bénédicte Grandon

Lieu de l'enregistrement :

GIPSA-lab Département Parole et Cognition
Site Stendhal - Salle E017
Université Stendhal, 1180 avenue Centrale BP 25 - 38040 GRENOBLE CEDEX 9

Cadre général :

Il s'agit d'enregistrer votre enfant en situation de répétition/production de mots et de narration de dessin animé.

Méthodologie :

Tout d'abord, si votre enfant n'a pas passé récemment d'audiogramme, un court test d'acuité auditive lui sera proposé. On lui fera écouter des sons calibrés avec un casque audiophonique et il (ou elle) devra nous dire lorsqu'il (ou elle) entend les sons (en levant la main par exemple). Ce test durera environ cinq minutes.

Ensuite, votre enfant sera assis(e) sur une chaise, la tête et les mains libres de mouvement, devant une table sur laquelle seront posés un ordinateur et un microphone. Sa voix sera enregistrée

Dans une première session, on présentera des photos d'objets quotidiens à votre enfant. Dans un premier temps, chaque photo sera affichée à l'écran en même temps que le son décrivant cet objet. Par exemple, votre enfant verra une photo de carottes et entendra le mot « carotte ». Il sera demandé à votre enfant de répéter le mot entendu. Dans un deuxième temps, il sera demandé à votre enfant de nommer à nouveau certains mots à l'aide des images mais sans le son correspondant. Cette session durera environ 20 minutes.

Dans une deuxième session, votre enfant regardera de courts épisodes d'un dessin animé sur un ordinateur. Après chaque épisode, on lui demandera de raconter à la personne assise à côté de lui (ou d'elle) l'épisode qu'il (ou elle) vient de voir. Il y a aura huit épisodes de dessins animés, de quelques secondes chacun.

Nom de l'enfant :
Prénom de l'enfant :
Date de naissance de l'enfant : Classe : Langue Maternelle :

Nom et prénom du parent présent :

J'ai lu et compris les consignes et j'accepte que mon enfant _____ participe à cette expérience.
Mon enfant peut arrêter de participer à tout moment.
Les données seront anonymes et seuls des résultats de groupe seront pris en compte, les enfants ne seront pas comparés entre eux.
Je donne mon consentement pour autoriser le laboratoire à utiliser les données audio enregistrées lors de cette étude (préalablement rendues anonymes) à des fins de recherche uniquement (exposés scientifiques lors de colloques, publications scientifiques) et ceci pendant toute la durée de la présente étude.

Date **Signature**

Annexe 2 : Formulaire de consentement pour l'étude de production – enfants CI

gipsa-lab

DOCUMENT D'INFORMATION et CONSENTEMENT DE PARTICIPATION

Remis aux parents des enfants sollicités pour participer à cette recherche
Investigateurs principaux : A. Vilain, H. Lœvenbruck, S. Schmerber
Intervenante : Bénédicte Grandon

Lieu de l'enregistrement :
Service ORL
CHU de Grenoble

Cadre général :

Nous recherchons des enfants de 6 à 10 ans pour participer à une étude sur le langage. Il s'agit d'enregistrer votre enfant en situation de répétition, de production de mots et de narration d'un dessin animé.

Méthodologie :

Votre enfant sera assis(e) sur une chaise, la tête et les mains libres de mouvement, devant une table sur laquelle seront posés un ordinateur et un microphone. Sa voix sera enregistrée.

Dans une première session, on présentera des photos d'objets quotidiens à votre enfant. Chaque photo sera affichée à l'écran en même temps que le son décrivant cet objet. Par exemple, votre enfant verra une table et entendra le mot « table ». Il sera demandé à votre enfant de répéter le mot entendu. Dans un deuxième temps, il sera demandé de renommer certains de ces objets avec l'aide des mêmes images, mais sans le son correspondant. Cette première session durera environ 15 minutes.

Dans une deuxième session, votre enfant regardera de courts épisodes d'un dessin animé sur un ordinateur. Après chaque épisode, on lui demandera de raconter à la personne assise à côté de lui (ou d'elle) l'épisode qu'il (ou elle) vient de voir. Il y a aura huit épisodes de dessins animés, de quelques secondes chacun.

Nom de l'enfant :
Prénom de l'enfant :
Date de naissance de l'enfant : Classe : Langue Maternelle :
Date d'implantation :

Nom et prénom du parent présent :

J'ai lu et compris les consignes et j'accepte que mon enfant _____ participe à cette étude.
Mon enfant peut arrêter de participer à tout moment.
Les données seront anonymes et seuls des résultats de groupe seront pris en compte, les enfants ne seront pas comparés entre eux.
Je donne mon consentement pour autoriser le laboratoire à utiliser les données audio enregistrées lors de cette étude (préalablement rendues anonymes) à des fins de recherche uniquement (exposés scientifiques lors de colloques, publications scientifiques) et ceci pendant toute la durée de la présente étude.

Date **Signature**

Annexe 3 : Questionnaire – enfants CI

Questionnaire

Date de naissance de l'enfant :
Lieu de naissance de l'enfant (département) :
Lieux d'habitation successifs (départements) + durée :
Nombre de frères et sœurs (+âge) :
Langue(s) maternelle(s) de la mère :
Langue(s) maternelle(s) du père :
Langue(s) parlée(s) à la maison :

Date de diagnostic de la surdité :
Date d'implantation (1er implant) :
Date d'implantation (2ème implant) :
Type d'implant :

Autres aides auditives avant implantation :
Autres aides auditives après implantation :

Communication

- Avant implantation

	oui/non	à partir de quel âge ?	Avec qui ?
Gestes maison			
LSF			
LPC			
Oral			
Autre			

- Après implantation

	oui/non	à partir de quel âge	Avec qui ?
Gestes maison			
LSF			
LPC			
Oral			
Autre			

Scolarisation

	oui/non	à partir de quel âge	Type d'établissement	Accompagnement scolaire ?
Ecole maternelle				
Ecole primaire				
Collège				

Suivi orthophonique

	oui/non	à partir de quel âge	durée	lieu
Avant implantation				
Après implantation				

Annexe 4 : Fréquence des mots utilisés dans les tâches de répétition et de dénomination

A. Fréquences des mots utilisés pour l'étude des voyelles

Voyelle	Mots simples	Fréquence lexicale	Mots complexes	Fréquence lexicale
a	bateau	64,1	balançoire	50,62
e	bébé	60,92	bétonneuse	X
ɛ	bec	59,67	berceau	52,18
i	bille	54,55	bicyclette	56,95
o	bobo	38,82	bocal	54,52
ɔ	botte	53,67	bol de soupe	55,81 (pour bol)
u	bouton	55,74	bourdon	42,17
y	bureau	58,19	bûcheron	53,55
ø	belette	49,28	besace	33,25
œ	bœuf	57,48	beurre fondu	57,88 (beurre)
ã	banc	57,17	banquise	47,58
ɔ̃	bonbon	50,83	bonbonne	32,75
ɛ̃	bain	58,93	bungalow	X

B. Fréquences des mots utilisés pour l'étude des consonnes

Consonne	Mots simples	Fréquence lexicale	Mots complexes	Fréquence lexicale
p	pipe	54,2	pistache	17,92
p	poupée	58,77	poussière	57,52
t	ticket	48,16	T-shirt	X
t	toupie	49,54	tournesol	44,99
k	kiwi	43,1	kimono	49,55
k	couteau	57,18	couverture	56,7
b	bille	54,55	bicyclette	56,95
b	bouton	55,74	bourdon	42,17
d	dix	63,13	dinosaure	46,1
d	doudou	X	douze	57,77
g	guitare	56,44	guirlande	44,45
g	goutte	55,52	gourmand (adj)	52,89
s	sirop	55,81	cygne	48,7
s	soupe	58,95	sous-marin	51,26
f	filet	56,68	ficelle	57,02
f	four	58,22	fourchette	52,96
ʃ	chiffon	50,49	chinois	52,39
ʃ	chou	54,86	choucroute	41

C. Fréquences des mots utilisés pour l'étude de la coarticulation

/bV/	Fréquence lexicale	/dV/	Fréquence lexicale	/gV/	Fréquence lexicale
bateau	64,1	date	57,54	gâteau	61,04
bébé	60,92	dessin	64,29	guépard	50,24
bouton	55,74	doudou	X	goutte	55,52
bille	54,55	dix	63,13	guitare	56,44
bobo	38,82	dauphin	52,19	gaufre	38,77

Annexe 5 : Correspondance entre mots et codes utilisés pour l'organisation du corpus

mot	code	mot	code	mot	code
bain	bain	bouteille	bute	fourchette	furS
balançoire	bala	bouton	buto	gâteau	gato
banc	banc	bûcheron	bych	gaufre	gofr
banquise	bank	bungalow	bung	gourmand	gurm
bateau	bato	bureau	byro	goutte	gutt
bébé	bebe	chiffon	Sifo	guépard	gepa
belette	beul	chinois	Sinw	guirlande	girl
berceau	bers	chou	Suuu	guitare	gita
besace	beuz	choucroute	Sukr	kimono	kimo
bec	bekk	couteau	kuto	kiwi	kiwi
bétonneuse	beto	couverture	kuve	pipe	pipp
beurre fondu	boer	cygne	sign	pistache	pist
bicyclette	bisi	datte	datt	poupée	pupe
bille	bijj	dauphin	dauf	poussière	pous
bobo	bobo	dessin	dess	sirop	siro
bocal	boka	dinosaure	dino	soupe	supp
bœuf	boef	dix	diss	sous-marin	soum
bol de soupe	bold	doudou	dudu	T-shirt	tiSe
bonbon	bonb	douze	duzz	ticket	tike
bonbonne	bonn	ficelle	fise	toupie	tupi
botte	bott	filet	file	tournesol	turn
bourdon	burd	four	furr		

Annexe 6 : Scénario du test d'intelligibilité réalisé sur Praat (Boersma & Weenink, 2015) – phase d'entraînement

```
"ooTextFile"
"ExperimentMFC 5"
stimuli are sounds? <yes>
"Sounds/" ".wav"
carrier phrase "" ""
initial silence duration 0.5 seconds
medial silence duration 0 seconds
numberOfDifferentStimuli = 8
    "test_CI2_1" ""
    "test_CI2_2" ""
    "test_CI3_1" ""
    "test_CI3_2" ""
    "test_NH3_1" ""
    "test_NH3_2" ""
    "test_NH4_1" ""
    "test_NH4_2" ""


numberOfReplicationsPerStimulus = 1
breakAfterEvery = 0
randomize = <PermuteBalancedNoDoublets>
startText = "Vous allez entendre des phrases, prononcées par des enfants.

Vous devez évaluer l'intelligibilité de ce que vous avez entendu,
c'est à dire la faculté de l'enfant à se faire comprendre.

Merci de donner une note sur une échelle de 1 à 7 :
1 (pas du tout intelligible/compréhensible) - 7 (tout à fait
intelligible/compréhensible).

Votre note doit refléter uniquement votre compréhension
de la prononciation des enfants,
Ne tenez pas compte des éventuelles erreurs de grammaire et de vocabulaire.

Cliquez pour commencer."
runText = "donnez une note de 1 à 7 :
1 (pas du tout intelligible) à 7 (tout à fait intelligible)"
pauseText = "Petite pause. Cliquez pour continuer."
endText = "L'expérience est terminée. Merci !"
2 replays
replay button 0.20 0.40 0.6 0.80 "Réécouter (r)" "r"
okButton = 0.8 0.95 0.05 0.15 "Suivant" " "
oopsButton = 0 0 0 0 "" ""
responsesAreSounds? <no> "" "" "" "" 0 0
1 response categories
   0.6 0.8 0.6 0.8 "noter" 24 "" "suivant"
numberOfGoodnessCategories = 7
    0.05 0.17 0.30 0.40 "1"
    0.17 0.29 0.30 0.40 "2"
    0.29 0.41 0.30 0.40 "3"
    0.41 0.53 0.30 0.40 "4"
    0.53 0.65 0.30 0.40 "5"
    0.65 0.77 0.30 0.40 "6"
    0.77 0.89 0.30 0.40 "7"
```

Annexe 7 : Scénario du test d'intelligibilité réalisé sur Praat (Boersma & Weenink, 2015) – phase de test

```
"ooTextFile"
"ExperimentMFC 5"
stimuli are sounds? <yes>
"Sounds/" ".wav"
carrier phrase "" ""
initial silence duration 0.5 seconds
medial silence duration 0 seconds
numberOfDifferentStimuli = 130

    "CI6_1" ""    "CI6_2" ""    "CI6_3" ""    "CI6_4" ""    "CI6_5" ""
    "CI7_1" ""    "CI7_2" ""    "CI7_3" ""    "CI7_4" ""    "CI7_5" ""
    "CI10_1" ""    "CI10_2" ""    "CI10_3" ""    "CI10_4" ""    "CI10_5" ""
    "CI11_1" ""    "CI11_2" ""    "CI11_3" ""    "CI11_4" ""    "CI11_5" ""
    "CI12_1" ""    "CI12_2" ""    "CI12_3" ""    "CI12_4" ""    "CI12_5" ""
    "CI13_1" ""    "CI13_2" ""    "CI13_3" ""    "CI13_4" ""    "CI13_5" ""
    "CI14_1" ""    "CI14_2" ""    "CI14_3" ""    "CI14_4" ""    "CI14_5" ""
    "CI16_1" ""    "CI16_2" ""    "CI16_3" ""    "CI16_4" ""    "CI16_5" ""
    "CI17_1" ""    "CI17_2" ""    "CI17_3" ""    "CI17_4" ""    "CI17_5" ""
    "CI19_1" ""    "CI19_2" ""    "CI19_3" ""    "CI19_4" ""    "CI19_5" ""
    "CI20_1" ""    "CI20_2" ""    "CI20_3" ""    "CI20_4" ""    "CI20_5" ""
    "CI21_1" ""    "CI21_2" ""    "CI21_3" ""    "CI21_4" ""    "CI21_5" ""
    "CI22_1" ""    "CI22_2" ""    "CI22_3" ""    "CI22_4" ""    "CI22_5" ""
    "NH14_1" ""    "NH14_2" ""    "NH14_3" ""    "NH14_4" ""    "NH14_5" ""
    "NH16_1" ""    "NH16_2" ""    "NH16_3" ""    "NH16_4" ""    "NH16_5" ""
    "NH19_1" ""    "NH19_2" ""    "NH19_3" ""    "NH19_4" ""    "NH19_5" ""
    "NH20_1" ""    "NH20_2" ""    "NH20_3" ""    "NH20_4" ""    "NH20_5" ""
    "NH21_1" ""    "NH21_2" ""    "NH21_3" ""    "NH21_4" ""    "NH21_5" ""
    "NH23_1" ""    "NH23_2" ""    "NH23_3" ""    "NH23_4" ""    "NH23_5" ""
    "NH25_1" ""    "NH25_2" ""    "NH25_3" ""    "NH25_4" ""    "NH25_5" ""
    "NH26_1" ""    "NH26_2" ""    "NH26_3" ""    "NH26_4" ""    "NH26_5" ""
    "NH29_1" ""    "NH29_2" ""    "NH29_3" ""    "NH29_4" ""    "NH29_5" ""
    "NH30_1" ""    "NH30_2" ""    "NH30_3" ""    "NH30_4" ""    "NH30_5" ""
    "NH32_1" ""    "NH32_2" ""    "NH32_3" ""    "NH32_4" ""    "NH32_5" ""
    "NH33_1" ""    "NH33_2" ""    "NH33_3" ""    "NH33_4" ""    "NH33_5" ""
    "NH34_1" ""    "NH34_2" ""    "NH34_3" ""    "NH34_4" ""    "NH34_5" ""

numberOfReplicationsPerStimulus = 1
breakAfterEvery = 26
randomize = <PermuteBalancedNoDoublets>
startText = "Vous allez entendre des phrases, prononcées par des enfants.

Vous devez évaluer l'intelligibilité de ce que vous avez entendu,
c'est à dire la faculté de l'enfant à se faire comprendre.

Merci de donner une note sur une échelle de 1 à 7 :
1 (pas du tout intelligible/compréhensible) - 7 (tout à fait
intelligible/compréhensible).

Votre note doit refléter uniquement votre compréhension
de la prononciation des enfants,
Ne tenez pas compte des éventuelles erreurs de grammaire et de vocabulaire.

Cliquez pour commencer."
runText = "donnez une note de 1 à 7 :
1 (pas du tout intelligible) à 7 (tout à fait intelligible)"
pauseText = "Petite pause. Cliquez pour continuer."
endText = "L'expérience est terminée. Merci !"
2 replays
```

```
replay button 0.20 0.40 0.6 0.80 "Réécouter (r)" "r"
okButton = 0.8 0.95 0.05 0.15 "Suivant" " "
oopsButton = 0 0 0 0 "" ""
responsesAreSounds? <no> "" "" "" "" 0 0
1 response categories
   0.6 0.8 0.6 0.8 "noter" 24 "" "suivant"
numberOfGoodnessCategories = 7
    0.05 0.17 0.30 0.40 "1"
    0.17 0.29 0.30 0.40 "2"
    0.29 0.41 0.30 0.40 "3"
    0.41 0.53 0.30 0.40 "4"
    0.53 0.65 0.30 0.40 "5"
    0.65 0.77 0.30 0.40 "6"
    0.77 0.89 0.30 0.40 "7"
```