HAL
open science

Caractérisation de la voix de l'enfant sourd appareillé et implanté cochléaire: approches acoustique et perceptuelle et proposition de modélisation

Harold Andrés Guerrero López

To cite this version:

Harold Andrés Guerrero López. Caractérisation de la voix de l'enfant sourd appareillé et implanté cochléaire: approches acoustique et perceptuelle et proposition de modélisation. Linguistique. Université Paul Valéry - Montpellier III, 2010. Français. NNT : . tel-00668565

HAL Id: tel-00668565

https://theses.hal.science/tel-00668565

Submitted on 10 Feb 2012

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

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

UNIVERSITÉ PAUL VALÉRY - MONTPELLIER III

Arts et Lettres, Langues et Sciences Humaines et Sociales

E.D. 58 « Langues, Littératures, Cultures, Civilisations »

DOCTORAT DE L'UNIVERSITÉ PAUL VALÉRY - MONTPELLIER III

Discipline : Sciences du Langage

Harold Andrés GUERRERO LÓPEZ

CARACTÉRISATION DE LA VOIX DE L'ENFANT SOURD APPAREILLÉ ET IMPLANTÉ COCHLÉAIRE : APPROCHES ACOUSTIQUE ET PERCEPTUELLE ET PROPOSITION DE MODÉLISATION

Thèse dirigée par

M. le Professeur **Jacques BRES** et M. le Professeur **Michel MONDAIN**

Soutenue publiquement le **19 mars 2010**

Jury :

M. AMY DE LA BRETEQUE Benoit	Praticien Attaché, CHU Gui de Chauliac, AP-Hôpitaux de Marseille
Mme. BARKAT-DEFRADAS Melissa	Chargée de Recherches CNRS, PRAXILING 5267 CNRS
M. BRES Jacques	Professeur de l'Université Montpellier III
Mme. LE NORMAND Marie-Thérèse	Directrice de Recherches INSERM, LPNCog FRE 3292 CNRS
M. MONDAIN Michel	Professeur de l'Université Montpellier I
M. NICOLLAS Richard	Praticien Hospitalier ORL, Hôpital d'enfants La Timone, Marseille
M. SERRAFERO Patrick	Professeur Associé de l'Ecole Centrale de Lyon
Mme. VAISSIERE Jacqueline	Professeur de l'Université Paris III

|__|__|__|__|__|__|__|__|__|__|

N° attribué par la bibliothèque

A mi familia, por el apoyo y el ánimo que me han dado todo
este tiempo sin importar la distancia que nos separa.
Este trabajo de investigación es el resultado de mucho esfuerzo y paciencia…
Gracias por créer en mí. Les expreso todo mi agradecimiento y amor.

(A ma famille, pour le soutien et les encouragements qu'elle a su
me donner pendant tout ce temps, malgré la distance qui nous sépare.
Ce travail est le résultat de beaucoup d'efforts et de patience…
Vous avez cru en moi. Avec ma reconnaissance et tout mon amour).

A la mémoire de Nathalia et de John-Ander.

A mes amis de France pour leur confiance, leur aide et leur compréhension.
J'ai découvert la richesse de la culture française grâce à vous.

A tous mes autres amis qui m'ont donné leur soutien.

Aux enfants qui ont participé à ce travail

« Le savoir que l'on ne complète pas chaque jour diminue tous les jours »

(Proverbe chinois)

Remerciements

Je tiens à remercier Madame **Jacqueline VAISSIERE**, Madame **Marie-Thérèse LE NORMAND** et Monsieur **Richard NICOLLAS**, pour la confiance et l'honneur qu'ils m'ont fait d'accepter de juger ce travail. Leur expérience dans la recherche et leur savoir-faire dans leur discipline me sont très précieux pour avoir un jugement transdisciplinaire du travail effectué.

Je remercie le Professeur **Jacques BRES** pour sa confiance et son accueil en 2004 au sein du laboratoire PRAXILING. Merci de m'avoir écouté et de m'avoir ouvert les portes à ce moment-là pour faire partie de l'équipe de PRAXILING et mener à bien mon projet de thèse.

Je remercie Madame **Melissa BARKAT-DEFRADAS** qui a bien voulu accepter d'encadrer cette thèse « en cours de route », qui a su être présente et a tenu à m'accompagner avec fort dynamisme dans cette période de formation universitaire. Merci pour le temps et la patience accordés.

Je remercie le Professeur **Michel MONDAIN** d'avoir accepté la co-direction de cette thèse. Merci pour les conseils pertinents qui m'ont permis de recadrer mon travail dans certaines périodes charnières. Rassurez-vous, ma réflexion sur l'épistémologie de l'Orthophonie ne fait que commencer…

Je remercie :

Monsieur **Patrick SERRAFERO**, qui m'a permis de découvrir d'autres horizons dans ce travail : l'ingénierie biomédicale, le monde de l'entreprise et la gestion de projets industriels. Merci pour les conseils de « management » de la thèse, la qualité de sa réflexion revêtue d'un esprit pédagogique, les encouragements et le temps dédié. Merci de m'avoir guidé dans l'expérience de conception et développement du logiciel ARTIVox 4.0.

Le Docteur **Benoit AMY DE LA BRETEQUE**, qui a énormément apporté à ce travail. Merci pour les discussions enrichissantes, les conseils avisés et le temps accordé ainsi que pour l'autorisation à l'accès au matériel et à l'équipement de l'Unité de Phoniatrie du CHU Gui de Chauliac utilisés pour le recueil de données.

Madame **Catherine TROTTIER**, qui m'a également accordé beaucoup de temps pour les discussions sur le traitement statistique de ce travail. J'ai éprouvé du plaisir à effectuer ces calculs. Merci pour l'écoute, la patience et l'intérêt porté à ce travail.

Madame **Chantal Charnet**, pour son accueil au sein du Laboratoire PRAXILING.

Je souhaiterais également adresser ma reconnaissance envers :

Joana Revis, Orthophoniste et Directrice Pédagogique du Département d'Orthophonie de l'Université de la Méditerranée, pour ses précieux conseils à propos de l'analyse perceptive.

Monsieur Bernard Teston et Alain Ghio, du Laboratoire Parole et Langage, pour les échanges et les discussions instructifs à propos de l'analyse instrumental avec l'appareillage E.V.A.

Monsieur Olivier Bareille, Monsieur David Lenoir et l'ensemble du personnel de l'Ecole Centrale de Lyon, pour leur support technique dans la conception et le développement du simulateur numérique ARTIVox 4.0.

Adrienne Vieu, Martine Sillon, Géraldine Geffriaud, Anick Bastin, Kattell Blanchard, orthophonistes de l'Institut Saint Pierre, centre d'implantation cochléaire à Palavas-les-Flots (Hérault, France) pour m'avoir permis de contacter les enfants implantés cochléaires pour ce travail.

Jean-Pierre Piron et Marièle Sicard (régleurs); Eliane Gaujoux et Isabelle Retout (orthophonistes), ainsi que Jean-Pierre Arteau et l'ensemble du personnel du service d'O.R.L. au CHU Gui de Chauliac de Montpellier.

Monsieur Jean Marie Jacquot et Madame Laurence Cartier du C.R.O.P. à Saint-Hippolyte-du-Fort (Gard, France) ; Madame Brigitte Coste, Responsable Pédagogique et l'ensemble du personnel du C.E.S.D.A. à Montpellier ; Catherine Silhol, Phoniatre du C.A.M.S.D. du CHU La Timone à Marseille ; ainsi que les orthophonistes Cécile Renault, Rosyne Parpillon, Martine Mir, Francine Jalabert, Geneviève Trouchet et Bernadette Samson, pour leur collaboration dans le recrutement de la population d'enfants appareillés.

Le directeur de l'école primaire Jean Sibélius, le principal du collège Jeu de Mail et le médecin scolaire pour leur collaboration dans le recrutement de la population contrôle.

Christophe Péalat, pour son soutien moral, les multiples lectures, les conseils de rédaction ainsi que pour le support informatique.

J'adresse à tous les enfants et leurs familles un remerciement tout particulier pour leur précieuse participation, leur bonne volonté et leur intérêt à l'enregistrement des voix pour cette étude.

A ma famille, mes amis et toutes les personnes qui ont contribué d'une façon ou d'une autre à l'élaboration de ce travail, veuillez trouver ici l'expression de ma gratitude.

Vous avez tous participé à ce travail de longue haleine. Sans votre aide, mon travail n'aurait pu aboutir. Je garde un bon souvenir de chacun de vous.

Table des matières

INTRODUCTION.....13

CHAPITRE I : RAPPEL THEORIQUE.....19
1. La voix : moyen d'expression du langage humain par excellence 21
1.1 Rôle de la voix dans la communication orale.....21
1.2 La notion de « voix » et son usage au cours du temps.....24
1.3 Anatomie fonctionnelle des mécanismes et structures participant à la production de la voix et de la parole.....26
1.4 Particularités de la voix infantile 47
2. L'audition : mécanisme de perception de la parole humaine.....51
2.1. Anatomie de l'oreille 52
2.2. Physiologie de l'oreille 57
2.3. Transmission de l'influx électrique au cerveau 60
2.4. Perception du son.....61
3. La boucle audio-phonatoire.....67
4. La surdité : déficience de la perception auditive 71
5. L'implant cochléaire.....79

CHAPITRE II : PROBLEMATIQUE 87
1. Etat de l'art de la recherche dans le domaine de l'oralité de l'enfant sourd 89
2. Position de la problématique 93
3. Objectifs de notre contribution.....97
4. Hypothèses et questions scientifiques 98

CHAPITRE III : MÉTHODOLOGIE EXPÉRIMENTALE.....103
1. Population étudiée 105
1.1. Sexe et âge.....107
1.2. Appareillage/implantation. 111
1.3. Surdité 117
2. Protocole expérimental appliqué, matériel et équipe.....120
2.1. Enregistrement de la voix.....120
2.2. Matériau phonétique 121
2.2.1. Echantillon de parole.....121
2.2.2. Échantillon de voyelles tenues.....122
3. Analyse acoustique (Expérience 1) 125
3.1. Objectifs 125
3.2. Méthodologie.....125
3.3. Résultats et discussion 131
3.4. Conclusion.....154
4. Analyse aérodynamique (Expérience 2) 155
4.1. Objectifs 155
4.2. Méthodologie.....155
4.3. Résultats et discussion 158
4.4. Conclusion.....167
5. Analyse perceptive (Expérience 3).....168
5.1. Objectifs 168

5.2. Méthodologie 168
5.3. Résultats et discussion 171
5.3.1. Validité du jury d'écoute 171
5.3.2. Etude de la variabilité des jugements perceptifs 175
5.3.3. Etude du lien entre les résultats de l'analyse instrumentale et les résultats de l'analyse perceptive 177
5.4. Conclusion 181

CHAPITRE IV : PROPOSITION DE MODÉLISATION 183
1. Proposition d'un modèle représentant le rôle de la boucle audio-phonatoire chez l'enfant déficient auditif 185
1.1. Objectifs du modèle 186
1.2. Caractéristiques du modèle proposé 186
1.3. Description du modèle proposé 191
1.3.1. Bloc « Vibrateur » 192
1.3.2. Bloc « Résonateur » 199
1.3.3. Bloc « Récepteur » 202
1.3.4. Bloc « Comparateur » 206
1.3.5. Banc d'essai dans son ensemble dit : « Modèle macro-systémique de la boucle audio-phonatoire » 207
2. Conception d'un simulateur numérique : ARTIVox 4.0 211
2.1. Matériel et méthode 212
2.2. Historique des versions du logiciel ARTIVox 212
2.3. Interface ARTIVox v4.0 213
2.3.1. Interface Homme-Machine ARTIVox v4.0 214
2.3.2. Affichage de résultats de la simulation numérique ARTIVox v4.0 215
2.3.3. Export et import des sons 216
3. Simulation numérique : ARTIVox 4.0 217
3.1. Données d'entrée pour la simulation numérique 217
3.2. Résultats de la simulation numérique 224
3.2.1 Cas de l'enfant entendant 231
3.2.2 Cas de l'enfant déficient auditif 233
4. Conclusions 238

CHAPITRE V : CONCLUSION GENERALE ET PERSPECTIVES 241

BIBLIOGRAPHIE 253

TABLE DES FIGURES 271

LISTE DES TABLEAUX 275

ANNEXES 277

INTRODUCTION

Introduction

Le présent travail est le résultat d'une réflexion orientée vers l'étude de l'interaction intriquée des processus de production et de perception de la voix et de la parole. Cette étude explore plus particulièrement la voix de la population infantile implantée cochléaire dans une perspective transdisciplinaire : l'orthophonie, la phonétique, l'acoustique, la phoniatrie, l'audiologie, la médecine.

La compréhension des processus physiologiques participant au développement de la voix et de la parole chez l'enfant déficient auditif représente un énorme défi. Les recherches dans ce domaine peuvent en effet fournir un éclairage pertinent sur les mécanismes de production de la parole mis en place chez l'enfant mal entendant bénéficiant d'une aide auditive et d'une rééducation orale.

Depuis quelques décennies, plusieurs axes de recherche se sont développés autour de la production orale de l'enfant déficient auditif. Ces axes ont porté sur le mécanisme de la phonation dans l'émergence des vocalisations, les éléments suprasegmentaux ou prosodiques, l'intelligibilité de la parole, le rôle des informations acoustiques provenant de l'environnement et l'apport des appareils de correction auditive. Notre travail s'inscrit à la suite des travaux réalisés dans ces deux derniers axes de recherche. Nous nous proposons de contribuer avec notre étude à la caractérisation acoustique et perceptive de la voix des enfants sourds bénéficiant d'une prothèse classique et/ou d'un implant cochléaire.

La caractérisation acoustique de la voix chez l'enfant appareillé et implanté cochléaire est traitée à travers une étude sur l'analyse des différents paramètres acoustiques et aérodynamiques utilisés en clinique et validés par la communauté scientifique.

Notre étude propose une modélisation numérique qui tient compte de la situation communicative orale manifestée chez l'enfant sourd bénéficiant d'une prothèse auditive. Un modèle *« vibro-acoustique »* fondé sur les principes de la dynamique de systèmes en boucle de rétroaction négative, est envisagé afin de rendre compte des caractéristiques acoustiques vocales observées dans la population étudiée.

Dans le **premier chapitre** de ce travail, nous faisons un rappel général des bases anatomo-physiologiques de la phonation et de l'audition, notions théoriques sur lesquelles se basent

notre étude. Ce chapitre a pour but de rappeler la complexité de ces processus physiologiques ainsi que d'évoquer le phénomène physique qui leur est lié : le son. L'interaction dynamique de ces processus mis en jeu lors de la perception et la production de la parole est accomplie grâce à l'action d'un mécanisme dénommé « boucle audio-phonatoire ». La relation extrêmement étroite entre la perception et la production permet l'acquisition et le développement du langage. Dans le cas de la surdité, la boucle audio-phonatoire est défaillante et ces processus sont altérés entraînant une diminution de l'efficacité dans la communication orale. Grâce aux avancées technologiques, nous disposons aujourd'hui des aides auditives permettant la prise en charge de la surdité et la rééducation du langage oral.

Le **deuxième chapitre** présente le contexte dans lequel s'inscrit notre problématique ainsi que les hypothèses scientifiques régissant notre travail. D'une part, nous supposons que les voix des enfants sourds (appareillés avec une prothèse classique et implantés cochléaires) manifestent des caractéristiques vocales « particulières » associées à des troubles vocaux, du fait de l'absence d'une boucle de régulation auditive « naturelle ». D'autre part, nous supposons que les caractéristiques acoustiques de la voix des enfants sourds implantés cochléaires *sans troubles vocaux* sont sensiblement comparables à celles des enfants entendants. Deux questions sont formulées à partir de ces hypothèses : les caractéristiques vocales de l'enfant sourd appareillé et implanté cochléaire sont-elles pathologiques ? Ces caractéristiques relèvent-elles d'une dysphonie ?

Dans le **troisième chapitre**, nous présentons la méthodologie expérimentale de notre étude. Pour ce faire, deux approches complémentaires ont été abordées afin de caractériser la voix de l'enfant sourd appareillé et implanté cochléaire : une approche perceptive et une approche acoustique. Nous organisons la recherche heuristique des réponses à nos questions par la mise en place d'un protocole expérimental afin de recueillir les données de l'étude à travers trois expériences.

Dans le **quatrième chapitre**, nous proposons un modèle d'asservissement représentant le rôle de la boucle audio-phonatoire dans la production de la voix. Nous avons conçu et développé un *« banc d'essai numérique »* permettant de manipuler un certain nombre de paramètres relatifs aux mécanismes anatomo-physiologiques participant à la régulation de la voix. Ce simulateur permet ainsi de mieux comprendre les relations entre ces mécanismes dans une boucle interactive. Nous verrons que perception et production sont deux notions intimement liées, au point de s'influencer l'une l'autre par le rôle de la boucle audio-phonatoire.

Le **cinquième chapitre** présente les conclusions générales de notre travail en fonction des différents résultats obtenus. Nous traitons les limites de notre méthodologie. Puis, nous envisageons les perspectives de recherche afin de contribuer au domaine de l'exploration de la voix de la population infantile déficiente auditive.

CHAPITRE I : RAPPEL THEORIQUE

Dans ce chapitre, nous présentons un rappel théorique des notions relatives à deux éléments, objet de notre étude, à savoir : la voix et l'audition. Pour aborder sobrement ces vastes et complexes sujets, dans un premier temps, nous situons la place de la voix au sein de la communication orale ; nous présentons de façon succincte l'usage que le terme « voix » a connu au cours du temps ; ensuite, nous évoquons les éléments anatomo-physiologiques qui nous permettrons ainsi de parler plus précisément de la voix chez l'enfant. Dans un deuxième temps, nous exposerons les bases anatomo-physiologiques qui sous-tendent le processus de l'audition. Nous décrirons ce processus dès la réception du son jusqu'à son interprétation au niveau cérébral. Ensuite, nous évoquerons quelques travaux qui ont tenté de décrire le rôle de la boucle audio-phonatoire, mécanisme chargé de la régulation de la voix et de la parole dans la communication orale. Enfin, nous ferons allusion à la surdité comme étant une pathologie auditive qui atteint le fonctionnement du mécanisme auditif et par conséquent l'efficacité de la boucle audio-phonatoire.

1. La voix : moyen d'expression du langage humain par excellence

1.1 Rôle de la voix dans la communication orale

Il nous parait nécessaire de citer brièvement quelles sont les notions co-existant avec le terme « voix » afin de mieux comprendre le rôle de celle-ci dans la communication orale.

Dans les affirmations *« la **voix** est le moyen d'expression sonore de la **parole »** ou « le véhicule de transmission privilégié du **langage oral** dans la **communication orale** humaine »*, plusieurs notions essentielles apparaissent* pour déterminer le processus interactif dont la voix est un des éléments constitutifs. Afin de mieux comprendre la place de la voix au sein de ce vaste processus, nous considérons approprié de représenter cette relation de notions à l'aide d'un diagramme (Figure 1). Dans ce diagramme, chaque cercle représente l'une des notions et l'extension de ses propriétés dans l'ensemble du système déterminant la communication orale : le langage oral, la langue orale, la parole et la voix étant au cœur de l'ensemble ; les dégradés de couleur attribués à chaque cercle du diagramme signalent la densité que la notion représente pour l'ensemble (communication orale) d'un point de vue linguistique.

* Termes soulignés en italique gras.

La ***communication*** est un processus interactif de production de sens par l'interprétation réciproque de signes constituant un message. La communication orale humaine est constituée de plusieurs éléments agissant comme les « fondations » d'un immeuble. C'est l'ensemble des parties sous-jacentes (langage, langue, parole et voix) qui permet de construire une structure aussi complexe que la communication. Shannon (1948)[1] définit la communication comme un processus car les événements et les relations intrinsèques des éléments y participant sont dynamiques et en constante évolution. Ces éléments interagissent, chacun ayant une certaine influence sur l'autre. L'interprétation du message est effectuée par un émetteur et un récepteur, acteurs de la communication. La transmission d'information s'obtient par différents moyens ou canaux de communication. Ces canaux renvoient basiquement aux cinq sens.

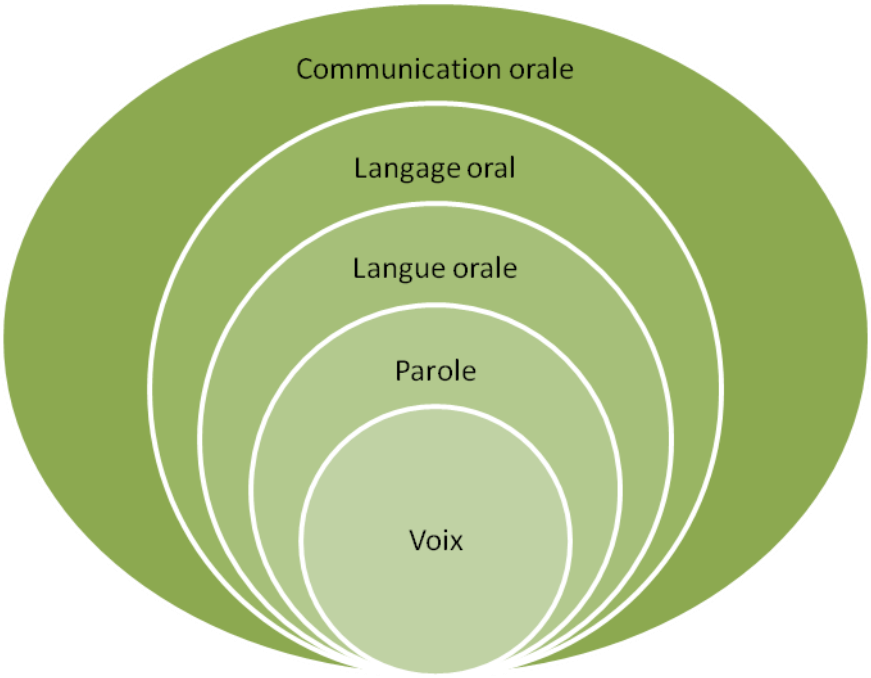


Figure 1 : Place de la voix dans la communication orale humaine.

Dans la ***communication orale***, l'émetteur est le locuteur qui transmet un message sonore par le biais de sa voix entre autres. Le récepteur est l'auditeur qui reçoit le message par le biais de l'oreille. Néanmoins, cette dynamique d'émission et de réception n'est plus la même quand il y a une atteinte d'un ou plusieurs canaux de communication. Par exemple, dans le cas d'une surdité sévère, les individus privilégient la communication gestuelle. En absence de l'oreille, la voix est remplacée par le geste.

Selon notre figure, une deuxième ellipse située à l'intérieur représente le langage. Le ***langage*** est la faculté d'une espèce, en occurrence de l'espèce humaine, à communiquer et à

représenter la réalité du monde dans lequel elle vit au moyen de signes différenciés. Le langage humain peut se représenter et se manifester sous trois formes : verbale (oral), non verbale (mimo-gestuelle,…) et paraverbale (par exemple, les phénomènes prosodiques). On considère le langage comme un outil de la pensée et « éminemment comme un fait social » (Meillet, 1921)[2] . Cette faculté se trouve altérée chez les individus déficients auditifs qui manifestent un retard dans l'acquisition et le développement du langage par rapport aux individus entendant. Ces derniers sont considérés comme la « norme » dans notre société. Ils connaissent des étapes d'acquisition bien déterminées se mettant en place de manière progressive et chronologique.

La ***langue*** est le code d'expression (orale et écrite) employé de manière conventionnelle par une communauté. Un groupe d'individus habitant dans une zone géographique quelconque partage une forme langagière qui l'identifie. La langue sert de marqueur identitaire quant aux caractéristiques de l'individu et de ses appartenances sociales. En effet, la langue est le produit social du langage et de l'ensemble de règles adoptées par le corps social afin de faciliter l'exercice de cette faculté chez les individus (Ferdinand de Saussure, 1916[3])[*]. C'est grâce à la langue que les individus échangent et mettent en commun leurs idées, sentiments, pensées, etc. Humboldt (1822)[4] a déclaré que « *la langue est l'expression de l'âme nationale du peuple qui la parle* ».

Un autre élément à distinguer, depuis Ferdinand de Saussure, est la ***parole.*** C'est un « *acte individuel de volonté et d'intelligence dans lequel il convient de distinguer : les combinaisons par lesquelles le sujet parlant utilise le code de la langue en vue d'exprimer sa pensée personnelle ; le mécanisme psycho-physique qui lui permet d'extérioriser ces combinaisons* » (De Saussure, 1916)[3]. Elle est considérée comme la faculté d'expression orale pour communiquer la pensée dans une langue déterminée à travers un système de sons articulés. C'est l'acte même de production orale du langage par un individu. Chez les sourds, la parole est perturbée car la défaillance de l'appareil auditif entraîne systématiquement et de manière considérable dans certains cas, une altération du comportement vocal.

Enfin, la ***voix*** est située au cœur dans notre représentation de la communication orale humaine, elle est considérée comme le moyen physique d'expression verbale (langage oral) par excellence. Autrement dit, la voix est le support acoustique du message linguistique

[*] Linguiste suisse (1857–1913) fondateur du structuralisme en linguistique. Il définit certains concepts fondamentaux tels que la distinction entre langage, langue et parole.

transmis et donc le vecteur de la parole et du langage dans la modalité orale. La voix est la « matière brute » sur laquelle s'appuie le langage pour véhiculer un message porteur de sens. Ainsi, lorsqu'elle est perturbée, elle peut masquer voire supprimer l'apport du contenu linguistique. D'une part, la voix joue un rôle identitaire d'un individu grâce à son timbre déterminé par les caractéristiques physiques, l'âge et le sexe entre autres. Elle reflète l'humeur, le tempérament, la personnalité, les émotions ainsi que la santé de l'individu. D'autre part, la voix joue un rôle socio-culturel permettant la fonction de lien entre l'individu et le collectif qui partage la même langue (Konopczynski, 2007)[5].

1.2 La notion de « voix » et son usage au cours du temps

Après avoir établit le cadre dans lequel la voix et le rôle que celle-ci accomplit dans la communication, nous considérons nécessaire de consacrer quelques paragraphes du présent chapitre à l'usage donné au terme *« voix »* au cours du temps. Puisque ce terme sera constamment évoqué dans notre ouvrage, nous tenons à en préciser le sens.

Le terme « voix » a généralement été utilisé dans la littérature de deux façons. D'une part, il a été employé de manière vaste pour évoquer la notion de « parole »[*] et l'associer à des références très concrètes : ce qui est produit par un individu qui parle, ce qui sort de sa bouche, ce qui est transmis et qui sert de support à la signification du message reçu. D'autre part, le terme de « voix » a été utilisé de manière plus spécifique, pour évoquer la notion de « phonation »[**]. Les concepts « parole » et « voix » ont émergé et suivi tout un développement historique.

La notion de voix était a été liée à l'identité d'un locuteur particulier et considérée comme une partie du message transmis. Le rôle de la voix a été largement sous-estimé dans la signification du message verbal. Ceci sans doute par un manque de connaissances approfondies sur le fonctionnement laryngé. Malgré des connaissances anatomiques anciennes, il a fallu attendre le développement de techniques de recherche appropriées pour surmonter les difficultés initiales d'observation et progresser dans la connaissance du fonctionnement de la physiologie vocale afin d'établir son importance pour la communication orale. Entre 1782 et 1832, l'édition de l'« Encyclopédie méthodique » de Panckoucke consacre

[*] Message que la voix véhicule.
[**] Processus physique et physiologique de la vibration sonore au niveau des plis vocaux ou cordes vocales. Dans notre travail, nous avons donc adopté le terme « voix » associé à ce sens.

quelques pages à la voix. Elles exposent non seulement le fonctionnement de l'organe vocal humain mais également le processus de production de cris des animaux et des oiseaux avec le rappel d'expériences faites à partir d'organes vocaux excisés d'oiseaux. Grâce aux recherches effectuées dans le domaine des phénomènes prosodiques, et en particulier de l'intonation des langues (Hirst et Di Cristo, 1998[6]), la place de la voix a été valorisée de manière plus juste dans la signification globale du message.

Une grande partie de la réflexion conduite en Occident sur les termes « voix » et « parole » se rattache à deux traditions : l'une plutôt religieuse (judéo-chrétienne) pour laquelle la parole est créatrice et à l'origine du monde. La parole qui crée le monde existe dans de nombreux mythes d'origine. Par exemple, dans la tradition religieuse égyptienne (2.300 à 1.500 ans avant J.C.), à l'époque où la ville de Memphis est capitale de l'Egypte. Pour la cosmogonie memphite, le dieu Ptah agit comme un créateur « en pensant le monde dans son cœur puis en l'appelant à la vie par l'autorité de ses paroles » (Yoyotte, 1956[7]).

Une seconde approche relève de la rhétorique (gréco-latine), la voix et la parole sont considérées comme argumentatives et persuasives. Avec la philosophie grecque naît une réflexion approfondie sur le langage, la parole et la voix. Platon et les Stoïciens distinguaient les termes « parole » et « voix » mais c'est Aristote (384 – 322 ans avant J.C.) qui reviendra souvent sur la différence entre ces termes : « *la voix et la parole sont des espèces du son* » (Durozoi et Roussel, 1997[8]). C'est l'enseignement de la rhétorique qui va assurer la transmission de connaissances acquises sur le rôle de la voix dans le discours par les maîtres de déclamation. A l'âge adulte, le déclamateur allait chez le « phonasque » pour étudier l'art des nuances et des modulations de la voix. Les techniques enseignées leur permettaient ainsi de projeter leur voix dans les théâtres antiques.

Dans son traité « De l'âme », Aristote souligne que : « la voix est un son produit par un animal vivant mais pas avec n'importe quelle partie, elle consiste dans le choc de l'air qui a été inspiré contre ce qui a été appelé la trachée artère, sous l'agencement de l'âme pour ces parties […] Tous les sons produits par des créatures vivantes ne font pas une voix (car quelqu'un peut faire un son avec la langue ou en toussant) mais ce qui cause le choc doit avoir une âme et l'accompagner de quelque phantasme [image mentale] car la voix est un son qui est le signe de quelque chose » (Aristote in Barthélemy-Saint-Hilaire, 1846[9]). Une précision apportée dans « De l'interprétation » : « la voix est le signe des passions de l'esprit » (Aristote in Tricot, 1984[10]). Dans « Politique », Aristote définit la parole comme suit : « l'homme est le

seul que la nature a doté du don de la parole. La voix est une simple indication du plaisir et de la douleur et on la trouve chez d'autres animaux. Le pouvoir de la parole est destiné à exposer ce qui est salutaire ou ce qui est nuisible et donc de la même manière ce qui est vrai et ce qui est faux. Elle est caractéristique de l'homme qui est le seul à avoir le sens du bien et du mal, du vrai ou du faux, et la participation à ces choses fait une famille aussi bien qu'une société […]. C'est dans ce sens que la parole est naturelle à l'homme qui est par nature un animal politique ou civil » (Aristote in Barthélemy-Saint-Hilaire, 1846[11]).

De nos jours, la voix identifie un interlocuteur particulier. Il n'existe pas deux voix identiques, chaque personne possède une voix qui lui est propre, et qui permet de la reconnaître. En effet, elle reflète la personnalité d'un sujet et donne des indices sur ses états d'âme et les nuances de sa pensée (Konopczynski, 2007[5]). Elle exprime les émotions qui caractérisent l'être humain (la joie, la colère, la tristesse, la peur, l'incertitude, le mépris, la surprise, etc.). Plus encore, la voix révèle l'état de santé et la forme du locuteur (Parret, 2002[12]). L'être humain utilise sa voix différemment selon les circonstances sociales. Tel est le cas de l'enfant qui « crie » dans certaines situations pour révéler son existence à son entourage. Cette situation montre que la voix est donc beaucoup plus qu'un simple objet esthétique. Elle devient essentielle dans la vie relationnelle de l'homme (Lafon, 1996[13]). En outre, la voix devient éventuellement un moyen de « subsistance » pour certains individus car elle peut être aussi un outil de travail professionnel dans une société, c'est par exemple le cas de l'enseignant, du chanteur ou de l'avocat.

La parole est la voix articulée qui forme le langage. Ce dernier reste l'attribut essentiel de l'homme qui seul trouve dans la parole le moyen d'exprimer ses sentiments et ses idées. Mais une mise au point est faite : la parole ne se réduit pas à la simple articulation des sons. La parole étant une manifestation du langage, elle ne prend toute sa valeur non pas parce qu'elle est articulée mais parce qu'elle est liée à la pensée.

1.3 Anatomie fonctionnelle des mécanismes et structures participant à la production de la voix et de la parole

Historiquement, l'analyse de faits plus concrets impliqués dans la production de la voix et de la parole a favorisé l'étude des structures et des mécanismes anatomiques y participant. Les premiers dessins du larynx et des plis vocaux réalisés par Léonard de Vinci (1490) et par Bartolomeo Eustachi (1714) en font preuve. De même, les planches anatomiques de Vesale

qui datent de 1543. Hieronymus Fabricius de Aquapendente (1537-1619), professeur d'anatomie à Padoue, donne une bonne description anatomique de l'organe vocal à partir de laquelle il précise bien le rôle des cordes vocales dans la production de la voix (*«De visione, voce et auditu » in* Expertise collective, 2006[14]).

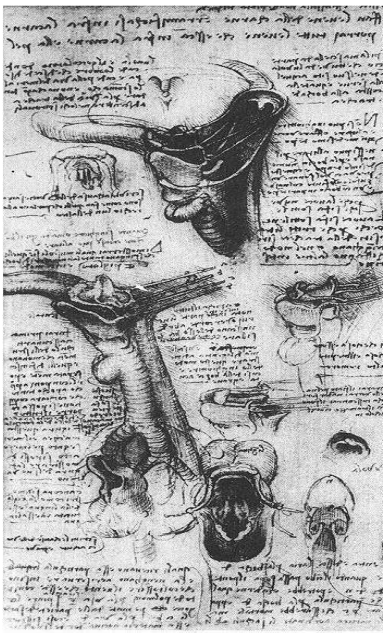


Figure 2 : Le larynx et les plis vocaux dessinés par Léonard De Vinci[*].

La visualisation des cordes vocales est rendue possible par l'utilisation d'un miroir, semblable à celui qu'utilisent les dentistes, d'abord éclairé par la lumière du soleil. Les premières observations datent de 1743 réalisées à l'aide de différents types de speculum, dont le *speculum auris* d'André Levet. En 1854, Manuel Garcia observe les cordes vocales à l'aide du miroir éclairé par la lumière du soleil. En fait, c'est Johan Nepomak Czermak qui, en 1859, publie un mémoire sur la laryngographie et la rhinoscopie et utilise la lumière artificielle. Son ouvrage « Du laryngoscope et de son emploi en physiologie et en médecine » est publié à Paris en 1860 chez Baillère. Czermak étudiera non seulement les cordes vocales mais aussi les mouvements du voile du palais (Expertise collective, 2006[14]).

L'étude de la physiologie de la parole et de la voix en phonétique amène à l'invention d'appareils permettant d'observer et de comparer les mouvements des organes articulateurs grâce à la réalisation et à la conservation de diagrammes (Etienne Jules Marey, 1878 in Teston, 2004[15]). Plusieurs chercheurs tels que Rosapelly (1876) et l'abbé Rousselot (1846 –

[*] Source de l'image : http://www.leonardoda-vinci.org/Anatomical-Studies---Larynx-And-Leg.html

1924) étudieront différents phénomènes phonétiques liés aux lieux d'articulation et aux organes bucco-phonatoires (Expertise collective, 2006[14]).

Tout au long du XIXe siècle, la mise en place d'un contexte scientifique et technique (ces deux aspects interagissant l'un sur l'autre) favorise une analyse scientifique de l'oralité : la voix et la parole. Ces derniers aspects devenant un objet auquel les phonéticiens et cliniciens vont prêter une attention toute particulière. Ils sont assistés, dans leurs analyses de plus en plus fines, par le développement d'une technologie essentiellement mécanique qui relance sans cesse de nouvelles recherches physiologiques, mais aussi, physiques.

L'imagerie anatomique, sous la forme de dessins regroupés en planches, est davantage diffusée. Les représentations en coupe sagittale de la région de la tête et du cou favorisent une compréhension globale des mécanismes de la production de la parole et de la voix. Ainsi, la connaissance du fonctionnement du larynx progresse. Deux grandes conceptions s'opposent et ne cesseront d'être discutées. Celle de Dodart (en 1700), pour qui le larynx est un instrument à vent, les cordes vocales jouant le rôle des lèvres lorsqu'on souffle, par exemple, dans une trompette. Les cavités supérieures agissent alors comme un porte-voix. Mais c'est à Ferrein (1741) que l'on doit l'utilisation du terme de « cordes vocales ». Celles-ci vibrent, d'après cet auteur, comme les cordes d'un violon, la modification du timbre par les régions supérieures du conduit vocal se faisant d'après un phénomène qui évoque celui de la résonance (Expertise collective, 2006[14]). Aujourd'hui, grâce aux avancées technologiques, ces deux grandes conceptions ont été amplement développées et ont évoluées pour proposer de nouvelles hypothèses et théories expliquant le phénomène de la vibration des cordes vocales.

La voix est un phénomène physique particulièrement complexe et riche dont la fonction principale est d'émettre les sons du langage, lesquels résultent de la mise en jeu simultanée et complexe d'un nombre important de structures anatomiques. L'ensemble de ces structures sont disposées de façon à constituer les différents mécanismes responsables de la production de la vibration sonore au niveau du bord libre des cordes vocales. Différents modèles ont été proposés pour expliquer l'ensemble des phénomènes physiologiques permettant la vibration sonore et l'interrelation des mécanismes. Le modèle de Perello (1990[16]) explique le processus de production de la parole résultant de l'interaction de plusieurs systèmes, mécanismes et facteurs externes. Il nous semble important de considérer ce modèle pour deux raisons : 1) il nous rappelle la participation d'autres systèmes aussi nécessaires que celui de la production de la voix proprement dite, 2) nous nous inspirerons de ce modèle pour la conception du

modèle théorique proposé ultérieurement (cf. chapitre IV). En effet, dans ce modèle interactif, la voix est considérée non seulement comme le résultat d'un processus purement physiologique mais également comme le résultat d'un processus « psycho-socio-communicatif ». Perello présente dans son modèle trois systèmes : (i) les systèmes de commande, (ii) de contrôle et (iii) de production.

a) Le système de commande intègre la pensée et l'action d'émettre un son. Perello explique que la voix est ici élaborée et interprétée avec ses caractéristiques acoustiques propres à elle-même (déterminant la forme) et ses caractéristiques linguistiques (déterminant le contenu) puisqu'elle porte un message communicatif. C'est grâce à ce système que nous sommes conscients des informations afférentes provenant de la muqueuse des organes bucco-phonatoires et des sensations proprioceptives des muscles et des articulations. Les informations sont analysées et mises en relation afin de produire une réponse. Ainsi, un ordre moteur volontaire est émis pour que les structures anatomiques du mécanisme de la production de la voix interviennent. Lors de la production de la voix chanté et de la voix parlée, les structures et mécanismes intervenants sont les mêmes. Cependant, c'est au niveau du cortex cérébral que les ordres moteurs d'interrelation seront déterminés en fonction de l'action communicative désirée. Le système de commande est constitué de 3 sous-systèmes : *le système nerveux central (SNC), le système nerveux périphérique (SNP) et le système nerveux autonome ou neurovégétatif (SNA)*. Le système nerveux porteur d'une histoire phylogénétique est le support principal de la grande flexibilité d'adaptation du comportement humain à l'environnement.

- Le S.N.C. comprend les structures de la zone corticale telle que le lobe temporo-parietal. Ce dernier représente une zone corticale importante où le *« schéma corporel vocal »* se développe. Le schéma corporel vocal stocke les « images acoustiques et linguistiques » perçues et les compare afin de produire et coordonner les divers mouvements précis des mécanismes intervenant dans la production de la voix et de la parole. D'autres structures situées au niveau sous-cortical y ont un rôle.

- Le S.N.P. transmet l'influx électrique provenant du S.N.C. par des fibres nerveuses efférentes appelées *« paires crâniennes »* afin de produire les divers mouvements précis des structures anatomiques intervenant dans la production de la voix et de la parole.

- Le S.N.A. régule l'action des muscles participants dans la production de la voix et de la parole soit pour accélérer soit pour ralentir leur activité électrique de réponse à l'ordre sollicité respectivement. Ce système contrôle également les manifestations végétatives telles que l'accélération ou le ralentissement du rythme cardiaque, les modifications respiratoires, digestives, sudorales ou vasomotrices, etc. qui accompagnent les réactions émotionnelles pendant la production vocale.

b) Le système de contrôle régule la production de la voix non seulement de ce qui est dit mais également de comment cela est dit. Par exemple, la régulation du contenu de ce que l'on dit serait effectuée par le « niveau auditif », tandis que la régulation de la forme ou manière de production de la parole et de la voix dépendrait non seulement du « niveau auditif » mais également du « niveau proprioceptif ». Ce monitorage se manifeste de manière simultanée au cours de l'acte de la parole. Pour Mysak (1980[17]), le contrôle et le monitorage de la parole et de la voix sont deux événements différents. Le « contrôle » de la production de la parole et de la voix a lieu durant l'acte de la parole et le « monitoring » poursuit le contrôle juste après l'acte de la production de la parole y compris durant la production de la parole et de la voix de l'interlocuteur. Nous retenons la définition du "système de contrôle" proposé par Perello qui nous semble la définition la plus pertinente pour notre étude. Le système de contrôle surveille l'efficacité de production de la voix à différents niveaux.

- *Niveau endocrinien :* Ce niveau comprend l'activité hormonale qui agit dans les différentes périodes de la vie humaine (0 à 6 ans, 8 à 10 ans, 12-13 ans, sénilité, cycles menstruels et ménopause chez la femme) entraînant des modifications anatomiques, morphologiques et histologiques du larynx. Ces modifications se traduisent par des changements sur les plans acoustique et aérodynamique de la voix. Par exemple, l'augmentation de la portion membraneuse de la corde vocal et l'amincissement des cordes vocales chez l'enfant induisent une diminution du fondamental laryngien (Woisard, 1998[18]; Titze, 1994[19]; Hacki[20], 1999; Nicollas, 2007[21]). Ces aspects seront évoqués dans la section suivante (section 1.4).

- *Niveau émotionnel :* Les émotions exprimées par la voix sont contrôlées par différentes structures anatomiques et aires cérébrales interconnectées. Aucune n'est à elle seule responsable d'un état émotionnel spécifique. Cette implication simultanée et multidirectionnelle ne permet pas d'établir une représentation simple et linéaire des processus émotionnels et rend difficile une conceptualisation globale de la séquence et

des niveaux de traitement. Cependant, les techniques d'imagerie fonctionnelle cérébrale ont permis d'étudier *in vivo* les émotions chez le sujet normal et de déterminer ainsi les différentes régions cérébrales impliquées dans les processus et états émotionnels. Les principales structures classiquement associées sont : le circuit de Papez (Papez, 1937[22]), le cerveau triple (Mac Kean, 1964[23] ; Mesulam, 1985[24]), le système limbique, l'amygdale (contrôle des principaux états affectifs dans l'expression de l'humeur, peur, rage, agressivité) et l'hypothalamus. Le système limbique échange de nombreuses relations avec l'hypothalamus qui coordonne les réactions somatovégétatives sous-jacentes aux états émotionnels. Mais le système limbique est aussi en relation avec les aires associatives du cortex et spécialement les aires préfrontales qui interviennent dans la maîtrise émotionnelle en général, dans le contrôle intentionnel de l'expression des émotions et dans l'évaluation cognitive de la signification émotionnelle des situations sociales complexes. Ces aires corticales sont également le lieu d'élaboration de l'expérience mentale et «privée» de l'émotion, pour reprendre les termes employés par Damasio (1999)[25] lorsqu'il définit ce qu'il considère comme l'étape finale du processus émotionnel: le sentiment.

- *Niveau Postural :* Il est déterminé par la position et l'accommodation des organes bucco-phonatoires les uns par rapport aux autres à un moment donné pendant la phonation. Sonninen (1970)[26] a insisté sur le fait que la position du larynx, et par conséquent, les muscles laryngés extrinsèques, sont impliqués dans le contrôle de la tension des cordes vocales. Iwarsson (1998)[27] conclut dans une étude qu'il semblerait que la position du larynx soit influencée par la posture du corps, elle-même en interaction avec la stratégie respiratoire. En effet, la posture joue un rôle dans la respiration, mais aussi dans la mise en résonance et la formation des sons ainsi que dans la qualité vocale en général. Par exemple, une posture aberrante affecte la mécanique respiratoire en créant des tensions musculaires inadéquates et une dépense d'énergie supplémentaire (Mc Kinney, 1982[28]). Les modifications des courbures physiologiques verrouillent le geste respiratoire alors qu'un meilleur ajustement de la posture vertébrale libère l'expansion pulmonaire (Faure, 1987[29]). L'enjeu du système de contrôle sur le niveau postural est de déterminer quelles sont les tensions musculaires des structures sur lesquelles il va falloir agir, pour se rapprocher le plus

possible de la posture dite « commode »* (Bouisset, 1995[30]) (Figure 3). La figure montre une représentation schématique de la ligne de gravité passant par les différents parties du corps humain.

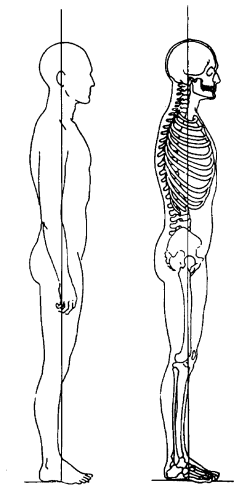


Figure 3 : Vue de profil de la posture debout dite « commode ».

- *Niveau auditif :* Ce niveau contrôle par la perception auditive ce que l'on dit. La perception est l'interprétation des sensations et le résultat d'un fonctionnement bio-mécanique. La sensation auditive est transmise au S.N.C. par l'intermédiaire de l'oreille qui intègre un mécanisme dénommé « boucle audio-phonatoire ». Ce mécanisme facilite le retour auditif de ce qui est exprimé par le locuteur permettant ainsi d'effectuer les modifications nécessaires pour atteindre l'efficacité communicative. Grâce à l'existence d'un tel mécanisme de régulation permanent, l'être humain peut acquérir et développer la parole et le langage oral. Nous ne pouvons émettre vocalement que ce que nous entendons. Nous parlerons de ce niveau de contrôle de manière plus détaillée dans les sections suivantes (section 1.2 et 1.3).

- *Niveau proprioceptif :* La voix chantée nous offre une vision plus pratique de ce qui représente le contrôle proprioceptif pour la production de la parole et de la voix parlée. La voix est le seul instrument de musique qui ne puisse être entièrement contrôlé par manipulation directe. Il est donc nécessaire de représenter mentalement les

* Cette posture correspond à une organisation du corps : alignement selon un axe passant par le sommet du crâne, le lobe de l'oreille, la pointe de l'épaule, la hanche et légèrement en avant de l'articulation de la cheville indiquant que la personne présente un bon équilibre mécanique.

coordinations des mouvements des structures intervenant au moment de l'acte chanté. Cette représentation mentale est remarquablement exécutée grâce à l'intervention du système de contrôle de la « proprioception ». Mais comment peut-on créer intelligemment ces représentations mentales des mouvements et sensations lors de l'acte de la parole? La proprioception est définie comme « la conscience globale du corps » (Dorland, 1951)[31]. Ce terme se réfère aux « terminaisons nerveuses sensorielles qui donnent des informations concernant les mouvements et la position du corps. Elles se trouvent dans les muscles, les tendons, les articulations, la peau et le labyrinthe vestibulaire. Dans son livre *Bel Canto*, James Stark (1999) [32] affirme que Giovanni Lamperti[*] reconnaissait le rôle de la proprioception dans le contrôle de la pression d'air quand il se référait à la sensibilité de « la peau interne » (*the inside skin*), la muqueuse et son réseau de nerfs. Wilbur J. Gould (1971)[33] a mis l'accent sur le fait que les muscles abdominaux et intercostaux contiennent des terminaisons nerveuses proprioceptives ajustant continuellement le système nerveux autonome (ou neurovégétatif) qui contrôle l'interaction entre la posture, la respiration et la voix. Les professeurs de chant demandent aux élèves de sentir les muscles du bas du dos et des côtés pour acquérir une meilleure conscience de ce qui intervient dans l'acte respiratoire. Ils sont mis en garde contre les tensions inutiles dans la mâchoire, la langue, le cou, et les muscles faciaux. Néanmoins, la sensation comme résultante d'une action est tout à fait différente de celle de produire une action en tentant de créer une sensation. Progressivement l'élève devient plus conscient des sensations créées dans tout son corps par les vibrations et les pressions d'air. Ce processus proprioceptif intervient dans l'ensemble du corps, des pieds jusqu'au sommet de la tête ; d'où l'emploi courant des termes « voix de poitrine » et « voix de tête ». Ces sensations sont très personnelles. Attirer l'attention des élèves sur leurs sensations proprioceptives les conduit à conditionner des réponses automatiques. C'est un processus qui demande beaucoup de temps, et qui varie selon les élèves, en fonction de leurs possibilités naturelles, de leur expérience antérieure, et de leur adaptabilité.

c) Le système d'exécution comprend les mécanismes et structures qui accomplissent les ordres provenant du système de commande (SNC, SNP, SNV) et surveillés par le système de contrôle. En clinique, ces structures anatomiques sont décrites classiquement en trois étages : 1) la soufflerie pulmonaire, 2) l'organe vibrateur laryngé, 3) les résonateurs bucco-pharyngés

[*] Giovanni Battista Lamperti (1839-1910) est le maitre italien en chant, auteur de « The Technics of Bel Canto » en 1905.

(Le Huche et Allali, 2001[34]) (Figure 4). Nous présenterons une description générale des processus physiologiques et physiques ainsi que des principaux mécanismes et structures participant à l'exécution proprement dite de l'acte phonatoire et articulatoire.

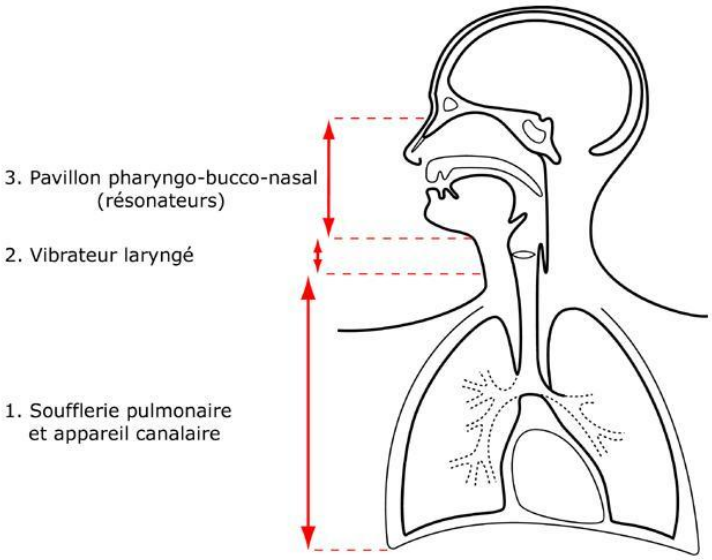


Figure 4 : Schéma des étages de l'appareil phonatoire.

L'énergie aérodynamique est fournie par la pression de l'air trachéal. La vibration sonore qui est la transformation de l'énergie aérodynamique en énergie acoustique apparaît au niveau des cordes vocales lorsqu'elles se rapprochent de la ligne médiane. Le son produit au niveau des cordes est aussitôt modifié par les organes bucco-phonatoires et filtré par les cavités aériennes situées au dessus des cordes vocales jusqu'au niveau des lèvres.

1) La soufflerie pulmonaire

La soufflerie pulmonaire est considérée comme le moteur de la vibration des cordes vocales. Elle fournit le flux d'air qui produit une pression appelée « pression sous-glottique » qui génère une énergie aérodynamique. Le système nerveux central y intervient afin de contrôler l'action des poumons et des muscles antagonistes expiratoires et inspiratoires nécessaires pour la phonation.

Avant d'émettre la voix, il est indispensable d'effectuer une inspiration. Celle-ci produit une dépression qui aspire l'air dans les poumons. Les muscles intercostaux externes soulèvent les côtes vers le haut et vers l'avant en « anse de seau ». Le diaphragme (muscle inspiratoire majeur en forme de dôme qui sépare le thorax de l'abdomen) se contracte en descendant pour

permettre une augmentation de la cage thoracique et, par conséquent, du volume pulmonaire (Figure 5).

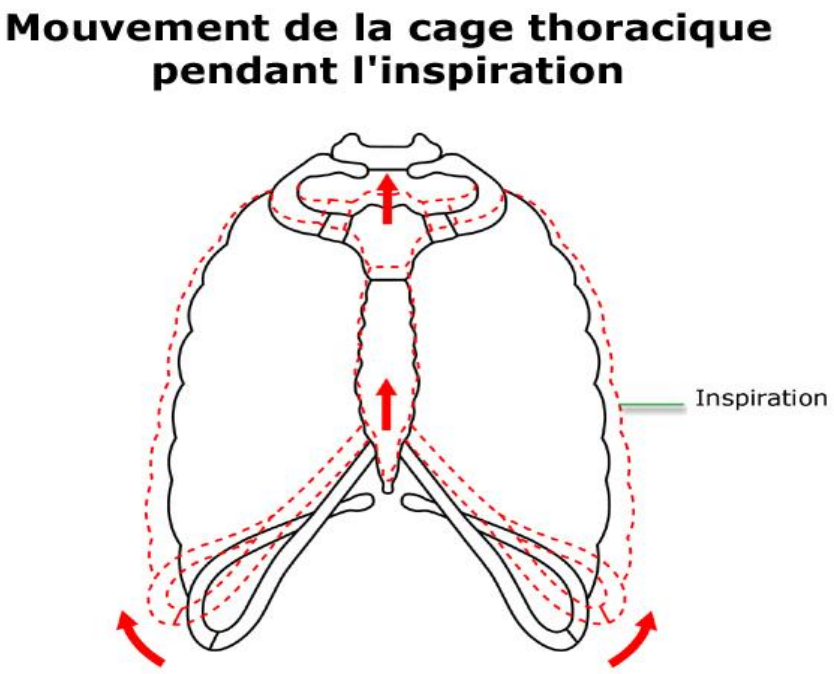


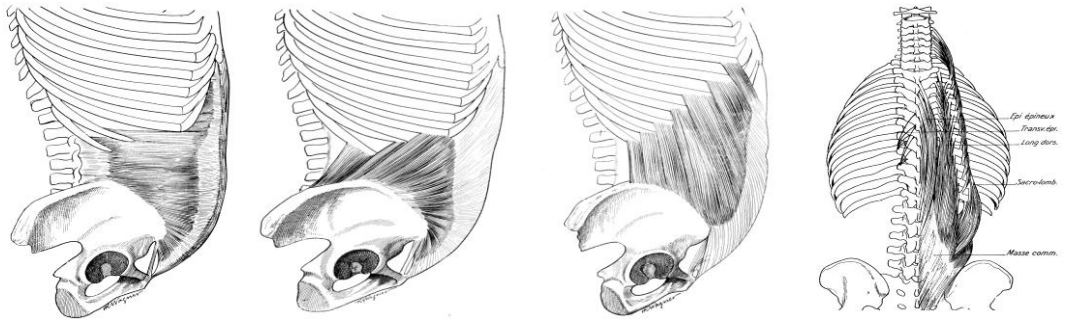
Figure 5 : Mouvement d'inspiration.

Durant l'expiration, au début les forces élastiques peuvent d'abord entrer en jeu seules : les muscles et les tissus qui avaient été mis en tension lors de la phase d'inspiration tendent à revenir à leur état d'équilibre. Ensuite, suivant la tache vocale à effectuer, les muscles expirateurs tels que les muscles intercostaux internes et les muscles de l'abdomen se contractent pour rétrécir la cage thoracique et ainsi augmenter la pression pulmonaire. Les muscles du dos peuvent également intervenir pour rigidifier le thorax. Dans la phase expiratoire, l'objectif du sujet est de maintenir une pression constante.

L'expiration active nécessaire à la production de la voix s'appelle *souffle phonatoire*. Ce souffle n'est pas toujours produit de la même façon. Il peut être produit de trois manières différentes: soit par l'abaissement de la cage thoracique (souffle thoracique supérieur) lors de la voix conversationnelle, soit par l'action des muscles abdominaux (souffle abdominal) lors de la projection vocale, soit par la flexion thoracique (souffle vertébral) lors d'un contexte d'effort relativement important. Ce dernier a lieu dans la voix d'appel, de détresse et dans le comportement de forçage vocal (Le Huche, 2001[34]).

Le larynx accomplirait un rôle de régulateur de débit lors de l'inspiration et surtout lors de l'expiration phonatoire : s'il y a un rétrécissement du tractus vocal, le débit d'air est plus bas. La quantité d'air maximale disponible dans les poumons* chez l'adulte est de 5 à 7 litres et de 3 à 7 litres chez l'enfant. Le volume d'air réellement utilisé** lors d'une respiration calme est de 10 à 20% de la capacité vitale et peut atteindre 50 % lors d'une activité physique intense. Ce volume chez l'adulte est de l'ordre de 4 ou 5 litres alors que chez l'enfant de 5 ans, il est de 1 litre (Thiebauld, 1997)[35]. Par ailleurs, les pressions pulmonaires utilisées pendant la phonation sont de l'ordre de 3 à 12 hPa*** (7 hPa en moyenne) chez l'adulte et de 4 à 8 hPa environ chez l'enfant (Higgins, 2003[82]).

Lors de la phonation, le cycle respiratoire doit être adapté. Le rythme de la respiration perd de la régularité et une augmentation des volumes pulmonaires à chaque cycle se produit. Cela entraîne un allongement du temps expiratoire. Par conséquent, la pression pulmonaire sollicite constamment un réglage. En effet, si les cordes vocales sont en position phonatoire (adduction), la pression pulmonaire doit être augmentée pour vaincre une telle résistance. Par exemple, si le sujet produit une voix aiguë, les cordes vocales deviennent encore plus rigides et demandent encore plus de pression pulmonaire pour maintenir la même amplitude vibratoire. L'ensemble des muscles interviennent suivant les ordres du système de commande afin d'augmenter cette pression. En l'occurrence, les muscles expiratoires accessoires tels que le transverse, le petit oblique, le grand oblique et certains muscles du dos et de la poitrine (muscles para-vertébraux) (Figure 6).



Transverse **Petit oblique** **Grand oblique** **Para-vertébraux**

Figure 6 : Muscles expiratoires accessoires (d'après Le Huche, 2001, p. 38, 47).

* La capacité vitale correspond au volume total disponible pour les échanges d'air.
** Le volume courant étant la quantité d'air qui entre et qui sort des poumons durant un cycle ventilatoire.
*** Par rapport à la pression atmosphérique (une P de 10 hPa = 1/100 P atmosphérique). Il est donc possible de considérer la pression pulmonaire ou « pression sous-glottique » comme négligeable lorsque les lèvres sont ouvertes c'est-à-dire lorsque les voies respiratoires sont en continuité avec le monde extérieur (ce qui se passe lors de la production de voyelles).

La puissance aérodynamique est le produit de la pression pulmonaire par le débit d'air. Cette puissance devient une énergie sonore acoustique par transformation au niveau du vibrateur laryngé. Une telle transformation correspond à un certain degré de perte d'énergie, liée à la transformation d'une partie du flux qui était laminaire sous la glotte en flux turbulent au niveau glottique.

2) Le vibrateur laryngé

Le vibrateur laryngé transforme l'énergie aérodynamique en énergie acoustique. Les cordes vocales se rapprochent de la ligne médiane sur le plan glottique et peuvent ainsi vibrer. Le larynx se situe à l'extrémité supérieure de la trachée.

Le larynx

Le larynx est l'organe central de la production vocale mais est avant tout l'organe clé du carrefour aérodigestif. Il s'agit d'un organe qui assure plusieurs fonctions simultanément : la fonction respiratoire, de déglutition et la phonation. *La fonction respiratoire* est primaire. Cette fonction assure prioritairement la protection des voies aériennes contre l'écrasement par les tissus cervicaux grâce à la rigidité du cartilage thyroïde et du cartilage cricoïde qui sont les cartilages « fixes » du larynx. *La fonction digestive* assure l'occlusion des voies aériennes avec les muscles, les ligaments endolaryngés et les cartilages « mobiles » du larynx tels que l'épiglotte (qui se rabat en arrière pour former un couvercle sur le larynx) et les cartilages aryténoïdes. Cette occlusion facilite le passage des aliments vers l'œsophage et non dans la trachée. *La fonction phonatoire* assure donc la vibration de l'air au niveau des cordes vocales dans certaines circonstances : fermeture contrôlée du larynx pour produire un rétrécissement suffisant et mise en pression de l'air pulmonaire par une expiration active.

Le larynx est constitué de cartilages articulés les uns par rapport aux autres par des ligaments, des tissus aponévrotiques et par des muscles recouverts par une muqueuse. Le larynx comprend cinq cartilages principaux : a) trois cartilages impairs et médians (thyroïde, cricoïde, épiglottique) ; b) deux cartilages pairs latéraux (aryténoïdes). Il comporte également quelques accessoires intervenant d'une manière moins directe que les autres cartilages précédemment cités, les cartilages de Santorini et de Morgagni, les sésamoïdes, les corniculés. On peut estimer que les cartilages thyroïde et cricoïde par leur configuration morphologique rigide sont des éléments de soutien assurant l'ouverture du larynx. Les autres éléments plutôt mobiles assurent la fermeture de la structure laryngée (Figure 7). On note l'articulation des cartilages fixes et mobiles constituant la structure laryngée.

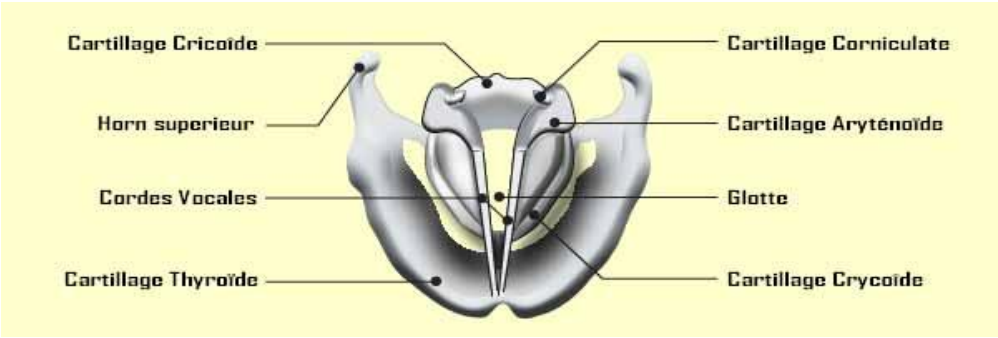


Figure 7 : Vue supérieure antéro-postérieure des cartilages du larynx.

D'autres éléments tels que les muscles intrinsèques et extrinsèques du larynx accomplissent des fonctions importantes pour la production vocale. Les muscles intrinsèques sont responsables de l'abduction (ouverture), de l'adduction (fermeture) et du réglage de la tension des cordes vocales. Ces muscles sont des muscles squelettiques qui contiennent plus de fibres de type IIa* que les autres muscles. Les muscles thyro-aryténoïdien et crico-aryténoïdien sont particulièrement spécialisés dans la contraction rapide. Tous les muscles intrinsèques sont innervés par des branches motrices du X paire crânienne (pneumogastrique). La fonction, l'innervation, l'origine et l'insertion des muscles intrinsèques sont présentées dans le tableau 1.

Les muscles extrinsèques sont responsables du maintien de la position du larynx dans le cou (suspension) afin de faciliter l'action des muscles endolaryngés intrinsèques (Figure 8). La plupart des muscles extrinsèques ont une insertion sur l'os hyoïde. Celui-ci est un anneau osseux concave vers l'arrière, situé juste au-dessus du cartilage thyroïde. Les muscles extrinsèques se divisent en deux groupes : les muscles sous-hyoïdiens et les muscles sus-hyoïdiens.

Les muscles sous-hyoïdiens sont abaisseurs du larynx en attirant vers le bas le cartilage thyroïde et l'os hyoïde. Ils permettent, en outre, d'abaisser la mâchoire inférieure lors de l'ouverture « forcée » de la bouche. La fonction, l'innervation, l'origine et l'insertion des muscles sous-hyoïdiens sont présentées dans le tableau 2.

* Ce type de fibres se contracte rapidement tout en étant résistant à la fatigue.

MUSCLE	FONCTION	INNERVATION	ORIGINE (O) – INSERTION (I)
Thyro-aryténoïdien (TA) - Constricteur de la glotte	Raccourcissement des CV** (épaississement), augmentation de la raideur CV et de l'adduction (portion membraneuse).	Nerf Laryngé Inf. (NLI)	O : Partie inf. fac int. du cartilage thyroïde I : Face lat. cartilage aryténoïde
Crico-aryténoïdien Latéral (CAL) –Constricteur de la glotte	Rotation, adduction (faire pivoter l'aryténoïde par rapport à l'axe vertical, porter en dedans l'apophyse vocale). Abaissement, allongement et affinement de CV	NLI	O: Bord sup. face lat. du cartilage cricoïdien I : Face lat. apophyse musc. aryténoïdien
Inter- aryténoïdien (IA) - Constricteur de la glotte	Adduction partie cartilagineuse des CV (rapprochement des aryténoïdes)	NLI + NL Supérieur	O-I : Flanc int. Aryténoïdes (fibres transverses) O : Base aryténoïdien (fibres obliques) I : Pointe aryténoïdienne contralat. (fibres obliques)
Crico-aryténoïdien Postérieur (CAP) – Dilatateur de la glotte	Abduction (faire pivoter l'aryténoïde par rapport à l'axe vertical, porter en dehors l'apophyse vocale latéralement et vers l'arrière), élévation, allongement de CV, augmentation de tension de CV	NLI	O : Face post. châton Cricoïdien (commisure post.) I : Face post. apophyse musc. aryténoïde
Crico-thyroïdien (CT) – Tenseur des plis vocaux	- Adduction de CV en position para-médiane (rapprochement de l'arc cricoïdien et du bord inf. du thyroïde, décalage vers l'arrière de cet arc) + augmentation de la tension de CV, affinement du bord libre des CV	NLExterne (branche du NLSupérieur)	O : Portion ant. et lat. cartilage cricoïde I : moitié post. face int. de l'aile thyroïdienne (pars obliqua) I : Bord inf. cartilage thyroïde

Tableau 1 : Synopsis de muscles du larynx, leur fonction, innervation et points anatomiques d'origine et d'insertion.

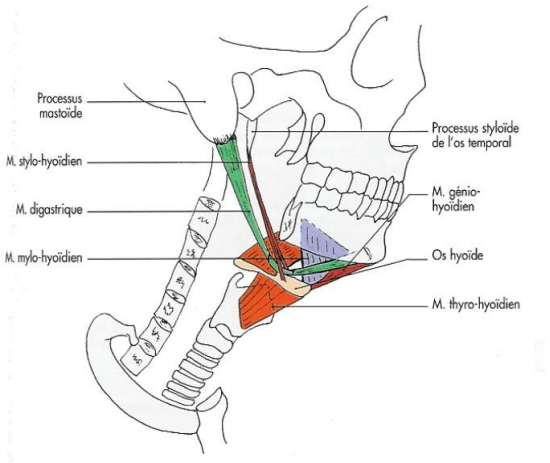


Figure 8 : Vue de profil des muscles extrinsèques responsables du maintien de la position du larynx dans le cou (d'après Bonfils et Chevallier, 1998[40]).

** CV=Corde vocale.

MUSCLES	FONCTION	INNERVATION	ORIGINE (O) – INSERTION (I)
Thyro-hyoïdien	Rapprochement Os hyoïde-cartilage Thyroïde en avant.	XII (Hypoglosse)	O : Crête oblique du cartilage Thyroïde + bord inf de la grande corne de l'os hyoïde I : Deux tubercules de la face ext. de la lame thyroïdienne et sur la ligne oblique
Sterno-thyroïdien	Abaissement du cartilage Thyroïde	XII (Hypoglosse)	O : Bord post. manubrium sternal + partie int. 1er cartilage costal I: Deux tubercules de la face ext. de la lame thyroïdienne et sur la ligne oblique
Sterno-cleïdo-hyoïdien	Abaissement de l'os hyoïde	XII (Hypoglosse)	O: Clavicule et face post manubrium sternal I : Bord inf du corps de l'os hyoïde
Omo- hyoïdien	Abaissement de l'os hyoïde	XII (Hypoglosse)	O : Bord inf du corps de l'os hyoïde I : Bord sup de l'omoplate

Tableau 2 : Synopsis de muscles sous-hyoïdiens, leur fonction, innervation et points anatomiques d'origine et d'insertion.

Les muscles sus-hyoïdiens sont élévateurs du larynx. La fonction, l'innervation, l'origine et l'insertion des muscles sont présentées dans le tableau 3.

MUSCLES	FONCTION	INNERVATION	ORIGINE (O) – INSERTION (I)
Digastrique : ventre post	Elévation de l'os hyoïde vers l'arrière. Le ventre antérieur du digastrique éleve l'os vers l'avant	VII (Facial) + V (trijumeau)	O : Face int. apophyse mastoïde (ventre post) O : Bord inf mandibule près de la symphyse (ventre ant.) I : Tendon intermédiaire qui attache l'os hyoïde (ventre post et ant.)
Mylo-hyoïdien	Elévation de l'os hyoïde en avant	XII (Hypoglosse)	O : Ligne mylo-hyoïdienne, face int de la branche horizontale de la mandibule I : Raphé médian (fibres ant.) + bord inf. du corps de l'os hyoïde,
Génio-hyoïdien	Elévation de l'os hyoïde en avant	XII (Hypoglosse)	O : Epine mentale mandibule I : Face ant du corps de l'os hyoïde
Stylo- hyoïdien	Elévation de l'os hyoïde en arrière	XII (Hypoglosse)	O : Partie postéro- externe de l'apophyse styloïde I : Face ant. Du corps de l'os hyoïde

Tableau 3 : Synopsis de muscles sus-hyoïdiens, leur fonction, innervation et points anatomiques d'origine et d'insertion.

Une coordination extrêmement précise de l'action de tous ces muscles facilite le positionnement correct du larynx dans le cou pendant la production vocale.

Les plis vocaux ou cordes vocales

Ces structures anatomiques impliquées directement dans l'émission du son laryngé correspondent au muscle thyro-aryténoïdien (TA) précédemment cité. Il s'agit d'une structure particulière présentée comme deux lèvres horizontales à l'extrémité supérieure de la trachée. Les structures sous-muqueuses spécifiques de la corde vocale sont appelées « lamina propria ». La partie la plus superficielle de la lamina propria correspond à l'espace de Reinke.

La partie la plus profonde est le ligament vocal qui recouvre directement le muscle vocal. Ce muscle vocal est le compartiment médial du thyro-aryténoïdien (Figure 9).

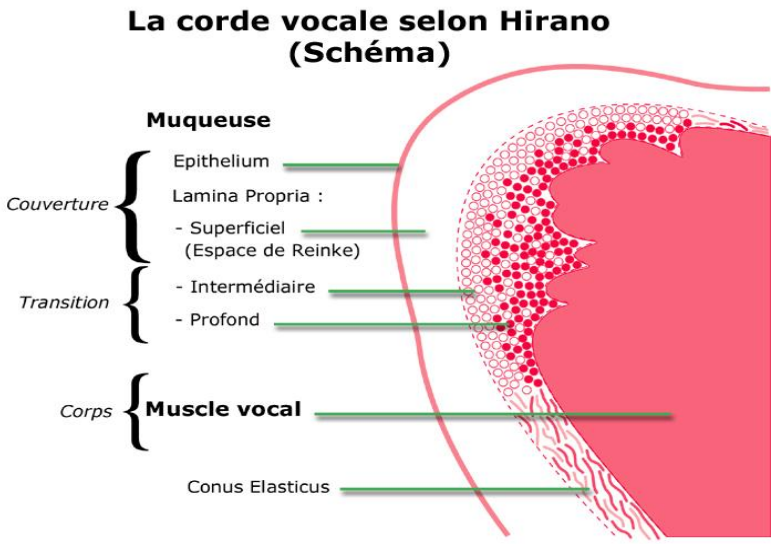


Figure 9 : Schéma de la structure « feuilletée » de la corde vocale proposée par Hirano (Coupe transversale)

La lubrification des cordes vocales est assurée par du mucus sécrété par des glandes acineuses situées en dehors du bord libre de la corde vocale. Dans les circonstances normales, seulement la partie antérieure du larynx est donc réellement responsable de la qualité de la vibration vocale. L'espace triangulaire présent entre les plis vocaux lorsqu'ils sont éloignés l'un de l'autre est nommé « glotte ». Les deux tiers antérieurs de cet espace constituent la glotte ligamenteuse et le tiers postérieur la glotte cartilagineuse. Au-dessus des plis vocaux, existent deux replis un peu semblables, les « plis vestibulaires » (ou bandes ventriculaires, ou fausses cordes vocales). Ces bandes ventriculaires n'ont normalement pas de contact entre elles pendant la parole ou le chant. Elles interviennent dans la résistance du tractus vocal supra-glottique de même que dans la résonance. En pathologie, dans le cadre du forçage vocal en particulier, on peut observer des hypertonies des bandes ventriculaires en phonation, voire un accolement en phonation responsable d'une voix rauque (voix de bandes).

Le phénomène de la vibration glottique a fait l'objet d'une très abondante littérature. La théorie myo-élastique proposée par Ewald (1898 in Le Huche, 2001[34]) a été la théorie la plus adaptée pour expliquer le phénomène de la vibration vocale et c'est celle qui est à l'origine des théories et modèles actuels. Dans un premier temps, les cordes vocales doivent être correctement positionnées sur la ligne médiane (position fermée pré-phonatoire)

accompagnée d'une tension appropriée (Figure 10). Elles opposent ainsi une certaine résistance à l'écoulement de l'air. La vibration commence avec la compression de l'air sous-glottique sous l'action des muscles expiratoires. Lorsque la pression sous-glottique dépasse le seuil de cette résistance, le mécanisme cyclique de la vibration se met en route : l'air peut s'échapper à travers les cordes ce qui fait diminuer la pression sous-glottique. Les cordes vocales se referment et le processus recommence grâce à l'effet « Bernouilli »* et l'élasticité des cordes vocales. Ainsi, le principe de cette théorie est de considérer le phénomène vibratoire comme un phénomène plutôt passif dont le contrôle est obtenu par le réglage des forces mécaniques et aérodynamiques du larynx.

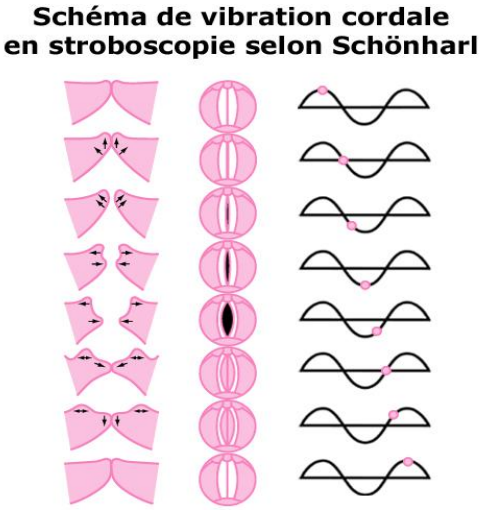


Figure 10 : Configuration glottique d'un cycle vibratoire des cordes vocales.

La vibration glottique (vibration des cordes vocales) a lieu exactement sur le bord libre des lèvres sous l'influence de l'air phonatoire. Les éléments clés en sont d'une part les phénomènes de démarrage tels que le seuil phonatoire pour déclencher l'oscillation vibratoire** et d'autre part les phénomènes d'auto-entretien.

Le seuil de pression phonatoire à partir duquel les phénomènes cycliques de la vibration se produisent dépend de plusieurs facteurs : la raideur de la partie vibrante de la corde vocale, la viscosité de la muqueuse, l'épaisseur du bord libre, la largeur de la fente glottique préphonatoire et la différence de pression transglottique. Ce seuil est donc plus élevé dans un

* Effet de « retro-aspiration » : si le flux d'air est rapide, l'effet Bernouilli est plus important.
** Le seuil phonatoire est le niveau minimum d'énergie potentielle pour que la vibration démarre.

certain nombre de circonstances. Dans les cas de pathologies comme les nodules ou les polypes des cordes vocales[***], l'augmentation de la raideur de la muqueuse est responsable d'une augmentation du seuil de pression phonatoire. Cette pression augmente dans les situations pathologiques et sa valeur devient un indicateur précis du degré de la pathologie, dans ce cas du forçage vocal.

En somme, la production du son correspond à la transformation d'un mouvement continu (l'expiration active et la mise en position phonatoire des cordes vocales) en un phénomène périodique. Le mécanisme de base correspond à la résolution du conflit de forces entre les forces de fermeture des cordes vocales (contraction musculaire, élasticité) et les forces qui tendent à les écarter (pression sous-glottique). Dans le chapitre IV de ce travail, nous développerons les modèles proposés dans la littérature pour expliquer le phénomène de la vibration des cordes vocales.

3) Les résonateurs bucco-pharyngés

Ces résonateurs comprennent le pharynx, la bouche, les fosses nasales et les sinus. Le pharynx est une cavité musculo-membraneuse capable de se rétrécir et de modifier son volume verticalement par l'action des muscles constricteurs du pharynx. Ces variations de forme et de volume dépendent des mouvements d'ascension et de descente du larynx lesquels interviennent de façon très importante dans l'articulation des voyelles. Cette cavité s'étend verticalement en avant de la colonne vertébrale cervicale, en arrière des fosses nasales, de la cavité buccale et du larynx. Les muscles du pharynx sont : les muscles constricteurs supérieur, moyen et inférieur qui rétrécissent les diamètres antéro-postèrieur et transversal du pharynx ; les muscles élévateurs qui élèvent le pharynx et le larynx. Le pharynx est sectionné en trois étages superposés : l'hypophayrnx, l'oropharynx et le rhinopharynx (Figure 11).

L'hypopharynx est la partie du pharynx située au-dessous de l'espace libre de l'épiglotte. Deux conduits débouchent ainsi dans l'hypopharynx : le larynx et l'œsophage. L'oropharynx est la zone du pharynx située derrière les replis de la muqueuse du voile du palais disposés verticalement et qui sont séparés en bas par la base de la langue et qui se rejoignent en haut pour former une ogive. Le rhinopharynx est la zone qui se trouve par dessus le voile du palais lorsque celui-ci est élevé. Pendant la production des voyelles[****] et des consonnes[*****] nasales

[***] Lésions bénignes résultantes d'un forçage vocal.

[****] Les voyelles résultent du passage libre du son laryngé dans la cavité buccale. Ce sont des sons voisés qui se définissent par le lieu d'articulation dans la cavité buccale où se localise la masse de la langue (antérieures,

[ã, ẽ, õ, ɛ, œ m, n, ŋ, ñ], le voile du palais reste abaissé laissant écouler l'air ainsi, en partie ou en totalité, par le nez. Pour toutes les autres voyelles et consonnes (définies comme « orales »), le voile du palais s'élève et laisse écouler l'air par la bouche.

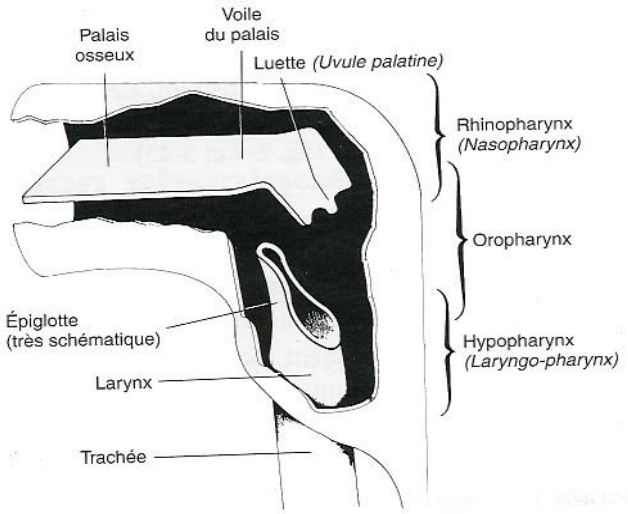


Figure 11 : Etages constituant le pharynx et position de la cavité pharyngée dans le cou (d'après Le Huche, 2001, p.17[34])

La bouche est le résonateur et articulateur principal. Plusieurs organes la composent : la langue, la mandibule, les dents, les lèvres, les joues, le voile du palais. Ces organes interviennent dans la production des bruits et des sons dont ceux du langage. La langue occupe la partie moyenne du plancher de la bouche et est constituée d'une musculature complexe comportant 17 muscles, des vaisseaux et des nerfs propres. Le maxillaire inférieur (ou mandibule) intervient de façon importante dans l'articulation de la parole et le réglage du volume des cavités buccale et pharyngée au cours de la phonation. Le voile du palais est une cloison musculo-membraneuse mobile qui prolonge en arrière et en bas la voûte palatine.

Les fosses nasales sont deux cavités séparées l'une de l'autre constituant une partie des voix aériennes supérieures. Elles sont situées au-dessus de la cavité buccale, au-dessous de la cavité crânienne et à l'intérieur des cavités orbitaires.

---

moyennes ou postérieures) et le degré d'aperture entre la langue et le palais lors de la production (en français, il existe 4 degrés : très fermé, fermé, ouvert et très ouvert).

***** Les consonnes résultent de l'obstruction du son laryngé dans la cavité buccale. Ce sont des sons voisés qui se définissent par le lieu d'articulation dans la la cavité buccale où ils sont produits (consonnes labiales, labio-dentales, dentales, palatales, vélaires).

Les sinus sont des cavités supplémentaires des fosses nasales. Ils contiennent des petits orifices lesquels communiquent avec les fosses et les sinus. Toutes ces cavités sont remplies d'air. On distingue quatre sinus : frontal (au-dessus de l'orbite), maxillaire (entre la cavité orbitaire et l'os maxillaire supérieur), ethmoïdal (constituant la paroi qui sépare l'œil de la fosse nasale) et sphénoïdal (constituant le plafond du rhinopharynx). On dit classiquement que ces cavités accomplissent une fonction de résonateurs pour amplifier le son émis dans la voix projetée et le chant.

Le son produit au niveau des cordes vocales est aussitôt modifié, c'est-à-dire, filtré par les cavités aériennes situées au dessus des cordes vocales jusqu'au niveau des lèvres. Ce filtrage sélectionne un certain nombre de fréquences qui « résonnent » avec le tractus vocal. Comme le corps d'un instrument à vent, le tractus vocal humain résonne pour certaines fréquences de la source selon la forme du tractus, ce qui détermine l'apparition des sons du langage (phonèmes).

Une zone de résonance du tractus vocal correspond à un « formant ». C'est une zone du spectre acoustique présentant une grande accumulation d'énergie sonore. Il existe cinq formants principaux qui interviennent dans la production de sons (production phonique) dont un certain nombre est utilisé dans la parole (production phonétique). Le formant le plus grave est situé dans le pharynx et le formant le plus aigu dans la cavité buccale. Dans le cas des voyelles, deux formants (les plus graves) nommés F1 et F2 suffisent à les caractériser.

Les formants sont modifiés par les changements articulatoires. Premièrement, les formants diminuent ou augmentent conjointement en fonction de la longueur du tractus vocal. Par exemple, si la longueur du tractus vocal s'allonge soit par l'abaissement du larynx soit par une projection des lèvres, les formants s'abaissent. Par ailleurs, l'obturation légère de la bouche comme font les instrumentistes à vent entraîne une modification des formants telle qu'un glissement de fréquences s'effectue vers les graves. La combinaison entre la hauteur du larynx et la position et la forme des lèvres permet effectivement d'éclaircir ou d'assombrir le timbre de la voix. L'abaissement de la mâchoire peut également être utilisé. En particulier, F1 peut être modifié par un abaissement de la mâchoire. Par ailleurs, la réduction du volume buccal abaisse F1 et élève F2, ce qui donne une voyelle au spectre plus large avec un étalement de la puissance acoustique comme dans /i/. Inversement la réduction du volume pharyngé élève F1 et diminue F2 ce qui rend la voyelle plus compacte comme /a/. En résumé,

le volume des cavités de résonance et par conséquent les formants varient constamment au cours de la parole.

Qualités perceptives de la voix

Comme nous l'avons vu, la production de la voix et de la parole est un processus extrêmement complexe qui demande la participation de différents mécanismes et structures anatomiques. De même, divers phénomènes physiques ont lieu pour que la voix et les caractéristiques acoustiques qu'elle comporte soient émises. Les caractéristiques acoustiques de la voix sont : hauteur, intensité et timbre. Elles peuvent être modifiées séparément ou conjointement au cours de la parole.

La hauteur est liée à la fréquence de la composante la plus grave du spectre sonore et se mesure en Hz. Chaque personne possède une hauteur de voix moyenne dénommée « fréquence fondamentale ou Fo » (terme utilisé dans cette étude). Cette fréquence se situe entre :

- 98 Hz (sol 1) et 164,81 Hz (mi 2) chez l'homme
- 196 Hz (sol 2) et 329,63 (mi 3) chez la femme
- 246,96 Hz (si 2) et 392 Hz (sol 3) chez l'enfant (Amy de la Brétèque, 1991[36])

La Fo d'un individu dépend de facteurs anatomo-physiologiques mais aussi familiaux, sociaux, culturels ou psychologiques. La fréquence du son fondamental dépend directement de la périodicité du mouvement vibratoire des plis vocaux. Cette fréquence varie en permanence au cours de la parole constituant ce que l'on appelle « l'intonation » ou la mélodie de la voix parlée.

La hauteur tonale utilisée par un sujet dépend essentiellement de l'anatomie laryngée. On a démontré que l'épaisseur de la corde vocale a un effet sur la hauteur de la voix : si l'épaisseur est faible, le ton de la voix est plus aigu, aussi bien chez l'homme que chez la femme. De même, la taille de toute la structure laryngée et le poids des cordes vocales déterminent la hauteur de la fréquence fondamentale pour chaque sujet. D'une manière générale la fréquence de vibration d'un dispositif quel qu'il soit est inversement proportionnel à sa taille. La Fo est donc plus basse si l'instrument est plus gros. La masse de la corde vocale correspond à la quantité de matériel effectivement en vibration. Plus les cordes vocales sont longues et lourdes, plus la voix est susceptible d'être grave (Le Huche, 2001[34]). C'est pour ces raisons

que la voix de l'enfant est plus aiguë (longueur de plis = 5 à 12 mm) que celle d'une femme (longueur de plis = 14 à 18 mm) ou d'un homme (longueur de plis = 18 à 25 mm).

L'ensemble des fréquences produites par un sujet soit vers les graves soit vers les aigus est appelé *« étendue vocale »*. Pour produire l'ensemble des fréquences de l'étendue vocale, le larynx modifie ses forces musculaires internes ainsi que sa configuration de taille et de forme. La plage fréquentielle produite en voix chantée avec aisance est appelée *« tessiture »*. La hauteur de la voix varie aussi selon le type d'émission vocale et les circonstances : lors d'une conversation « discrète », elle peut être grave; lors de l'expression de colère ou tristesse (excitation), elle peut être aiguë.

L'intensité traduit la puissance de la voix et se mesure en dB. Elle est due à l'amplitude de la vibration par la combinaison de modifications de la pression sous-glottique, de la force d'adduction des cordes vocales et de la géométrie du conduit vocal dans son ensemble. L'intensité est proportionnelle à l'amplitude vibratoire des cordes vocales, et dépend essentiellement de la pression sous-glottique. Elle varie selon les intentions expressives déterminées par les circonstances d'utilisation du locuteur ; ainsi, l'intensité peut varier en permanence pendant la parole et ne pas dépasser 20 dB (voix chuchotée) ou atteindre 90 dB (voix d'appel).

Le timbre représente la couleur (clair, obscur, brillant, etc.) et l'aspect esthétique de la voix. Il dépend de la richesse et de l'intensité relative des harmoniques* de la voix. Ces derniers sont essentiellement produits par la forme de l'onde sonore au sortir de la glotte, et sont modifiés par les résonateurs. Le timbre est donc la résultante des différentes modalités d'accolement des cordes vocales, des caractéristiques anatomiques individuelles des cavités de résonance (pharynx, bouche et cavité nasale) et de leurs modifications.

1.4 Particularités de la voix infantile

Pour mieux comprendre le processus de phonation chez l'enfant, il faut préciser que les caractéristiques de la voix infantile sont déterminées par les particularités topographiques, morphologiques et histologiques du larynx. Le point-clé est l'évolutivité du larynx dans le temps. Ces modifications induisent les différences de forme et de taille du larynx et des structures participants à ce phénomène sonore chez l'enfant.

* Fréquences multiples du son fondamental qui se superposent à celui-ci.

D'après Woisard (1998)[18], les modifications de la forme sont en rapport avec :

- l'augmentation plus importante de la portion membraneuse de la corde vocale par rapport à la portion cartilagineuse,
- un amincissement des cordes vocales lié à l'augmentation relativement plus importante de la longueur des cordes vocales par rapport à leur largeur (Figure 12). On observe l'augmentation de la longueur et l'amincissement des cordes vocales du larynx du nourrisson par rapport au larynx mature.

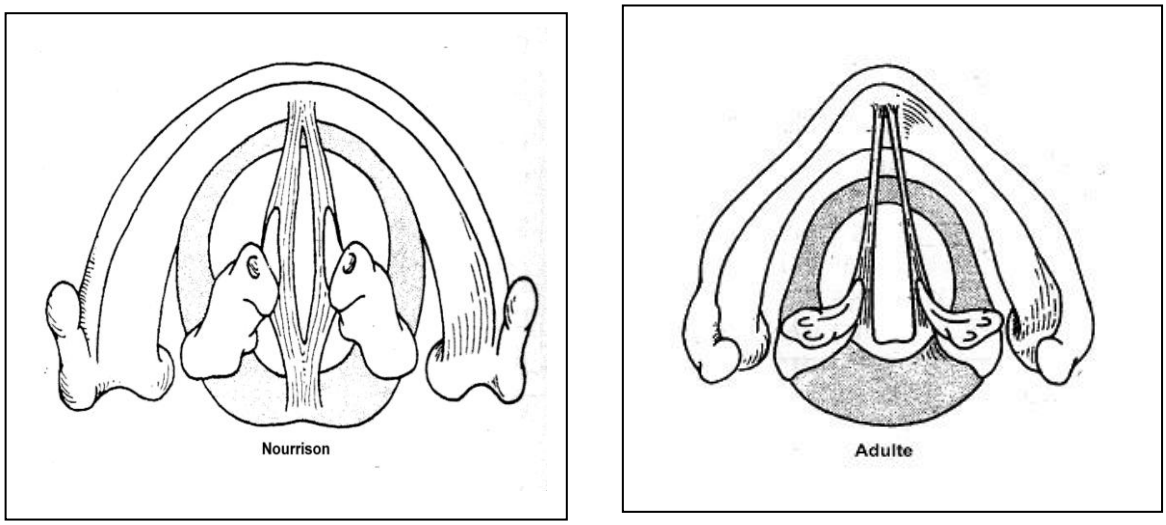


Figure 12 : Différence morphologique du larynx du nouveau-né et de l'adulte. (d'après Fried, 1996[37]).

Les modifications des caractéristiques physiques du larynx (Woisard, 1998[18]) sont avec l'âge :

- l'augmentation de volume de la cavité sous-glottique
- l'augmentation de la masse des cordes vocales
- l'augmentation du débit d'air glottique et
- l'augmentation de l'amplitude de vibration des cordes vocales.
- Par ailleurs, le larynx se situe, à 5 ans, autour des 4ème – 5ème vertèbres cervicales et les cordes vocales passent de 8 mm à environ 12 mm à 11 ans (Figure 12 et Figure 13).

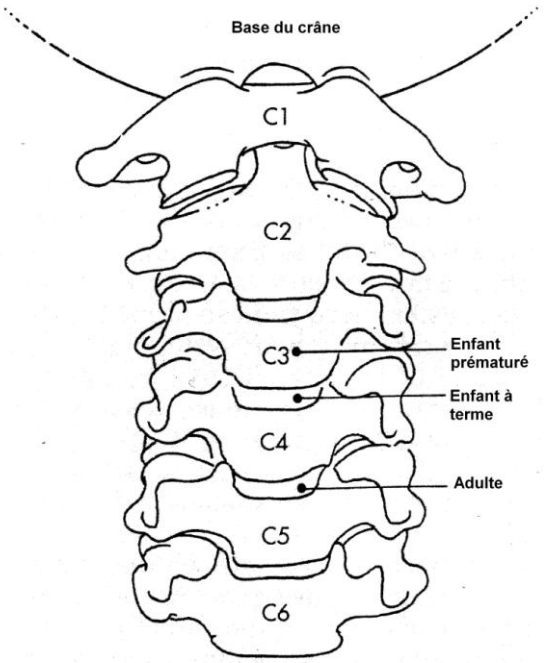


Figure 13 : Variation de position du plan glottique avec l'âge (d'après Fried, 1996[37]).

Ces modifications de forme et des caractéristiques physiques du larynx auront lieu après la période de la mue vocale. Dans cette étude, nous ne les prendrons pas en compte car nous étudierons le larynx chez l'enfant de 5 ans à 13 ans, soit avant cette période.

Pendant la période de croissance de l'enfant, le larynx augmente de forme, taille et volume. De même, sa position sur le plan glottique est modifiée. Ainsi, la principale conséquence acoustique de l'augmentation de la longueur des cordes vocales et de la descente du larynx est la diminution du fondamental laryngien tout au long de la vie (Figure 14).

Celle-ci sera plus lente et linéaire chez la fille (300 à 220 Hz environ) que chez les garçons. La Fo subira une nouvelle diminution rapide chez le garçon à la puberté (220 à 110 Hz). En même temps, l'étendue vocale s'élargit pour atteindre 3 octaves dès 6 ans. La fréquence fondamentale passe de 295 Hz (vers 7 ans) à environ 248 Hz (vers 11 ans) pour les filles et de 268 Hz (vers 7 ans) à 188 Hz (vers 11 ans) pour les garçons (Woisard, 1998[18]). La voix chantée s'épanouit pendant cette période et s'étend sur deux octaves vers l'âge de 8 ans avec des performances compatibles avec certains répertoires lyriques. Plusieurs études montrent que la moyenne de l'intensité de la voix chez l'enfant est 6 dB au dessus de celle de l'adulte quel que soit le niveau sonore demandé (Woisard, 1998[18]).

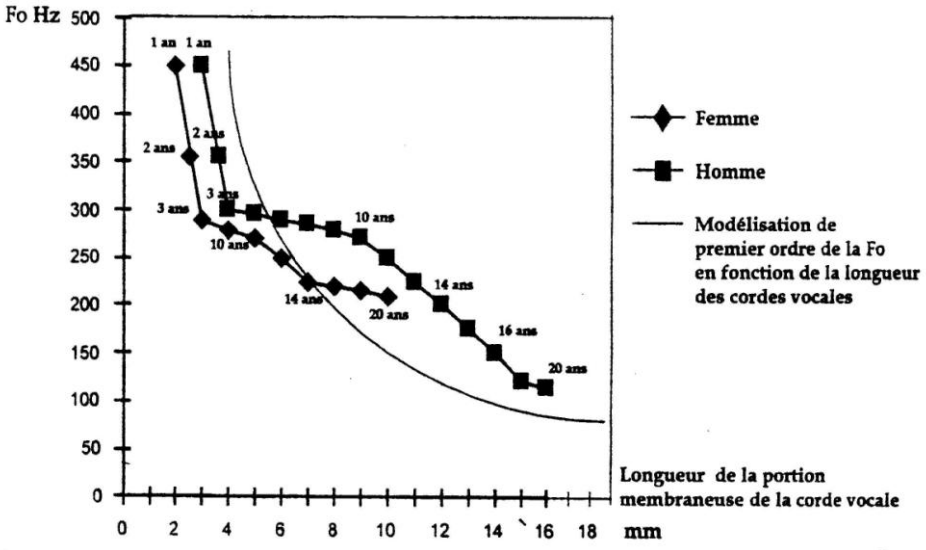


Figure 14 : Fo moyenne de la voix parlée en fonction de la longueur de la portion membraneuse des cordes vocales (d'après Titze, 1994, p.180).

Quant au timbre, le cri du nouveau-né est nasonnée. Ceci est du au larynx qui est situé très en haut dans le cou à la naissance. Progressivement, le larynx grandit et descend dans le cou durant la période de croissance de l'enfant. La variation de la forme et du volume du pavillon pharyngo-buccal donne lieu à des modifications de la résonance laryngée, pharyngé et, par conséquent, du timbre. A partir de 6 ans, il y a une différenciation du timbre entre le garçon et la fille d'environ un demi-ton.

2. L'audition : mécanisme de perception de la parole humaine

La voix et la parole émises *existent* parce qu'elles sont perçues par le locuteur et l'interlocuteur dans la conversation. De même, lorsque le locuteur communique avec soi-même, il écoute ce qu'il dit. Par exemple, le locuteur d'un monologue devient l'interlocuteur de son discours. Cette capacité d'interprétation de sensations sonores fournies par l'environnement s'accomplit grâce à l'action du mécanisme auditif. La perception est pour l'individu l'un des moyens d'accéder au monde et d'en avoir une représentation consciente.

En effet, l'audition est une fonctionnalité du corps humain qui facilite la vie de l'homme dans le monde. Cette fonctionnalité guide la perception et la construction du monde qui l'entoure. D'une part, l'audition se rend nécessaire pour l'apprentissage de la notion d'*espace* par l'appréhension des distances[*], de l'orientation et de l'espace volumétrique, mais aussi pour l'acquisition de la notion du *temps* et de l'organisation temporelle ou de *succession*. D'autre part, l'audition nous offre des informations sur ce que nous ne voyons pas nous permettant ainsi l'anticipation et l'alerte dans les situations de danger[**]. L'audition devient de cette manière une fonction essentielle dans la vie quotidienne de l'homme. En Outre, elle joue un rôle important dans l'acquisition et le développement de la parole et du langage dans la communication humaine.

Dans la communication orale, l'audition est une fonction active. Deux éléments interagissent de façon bidirectionnelle en la rendant dynamique: *le locuteur et l'auditeur*. Le locuteur étant l'émetteur du message sonore par intermédiaire de la voix et l'auditeur étant le récepteur du message par le biais du mécanisme auditif. Dans la conversation, la présence d'un interlocuteur est indispensable pour l'écoute. Si la communication a lieu, l'auditeur répond au stimulus; s'il ne répond pas, la communication n'atteint pas son but.

Parret (Parret, 2002[12]) souligne qu' « *Il n'y aurait pas de voix sans qu'il y ait oreille. La voix, en toute sonorité, n'est que dans la saisie par l'oreille*. ». Dans ce sens, l'audition est indissociable de la voix. Dans le cas de défaillance du mécanisme auditif, la voix est perturbée parfois absente et remplacée spontanément par d'autres moyens de communication

[*] Par exemple, la perception de l'écho produit par l'émission sonore d'un animal, un objet ou une personne dans un espace quelconque.
[**] Par exemple, la perception du bruit produit par le klaxon d'une voiture qui approche derrière nous.

autres que l'expression orale (par exemple, les gestes). La perception auditive est donc la fonction privilégiée de la communication orale.

La sensation auditive est transmise par l'intermédiaire de l'oreille, dont le fonctionnement complexe et astucieux fascine depuis les développementistes jusqu'aux audiologistes. L'impression auditive dépend d'une part des caractéristiques de l'onde, d'autre part des caractéristiques de l'oreille. Concernant les caractéristiques de l'onde, il n'a pas encore été possible de séparer physiquement l'oreille de l'onde qu'elle reçoit. Par conséquent, on est conduit à mentionner les propriétés de l'onde sonore quand on parle de la perception auditive.

Cette section présente successivement les différentes structures anatomiques et les processus physiologiques impliqués dans la perception auditive. Nous donnons ici quelques éléments essentiels et nous renvoyons le lecteur aux traités d'anatomie et de physiologie classique pour une description plus détaillée (Burt, 1993[38]; Lechevalier et Iglesias, 1997[39]). Nous nous intéressons aux notions de base expliquant l'anatomie et la physiologie de l'oreille depuis l'oreille externe jusqu'aux voies nerveuses qui conduisent l'information jusqu'au cerveau où la sensation auditive devient perception des sons. Nous aborderons également les concepts physiques élémentaires du traitement des informations auditives par l'étude de la perception des caractéristiques acoustiques des sons simples et complexes. Ensuite, nous présenterons la façon dont l'interprétation des sons s'organise au niveau cérébral. Nous aborderons les processus intellectuels et les structures cérébrales qui sous-tendent la perception consciente de notre environnement sonore. En effet, rien ne semble plus naturel et simple que « d'entendre ». Néanmoins, l'immense richesse et la quantité colossale de stimulations sonores auxquelles nous sommes soumis, témoignent de la complexité des tâches cognitives qui permettent le traitement de l'information auditive.

2.1. Anatomie de l'oreille

Comme pour la voix, l'audition est le résultat de la mise en jeu complexe et simultanée de plusieurs mécanismes et structures anatomiques. Ces mécanismes et structures constituant *« l'appareil auditif »* sont présentés classiquement en trois compartiments : *l'oreille externe, l'oreille moyenne et l'oreille interne* (Figure 15). Chaque compartiment prend en charge le traitement des signaux acoustiques ou des ondes sonores[*] aériennes[**] pour les transformer en

[*] Les ondes sonores sont dues à des variations de la pression dans un milieu élastique tel que l'air ou l'eau.

influx nerveux (potentiels électriques) lesquels sont par la suite interprétés par les centres cérébraux.

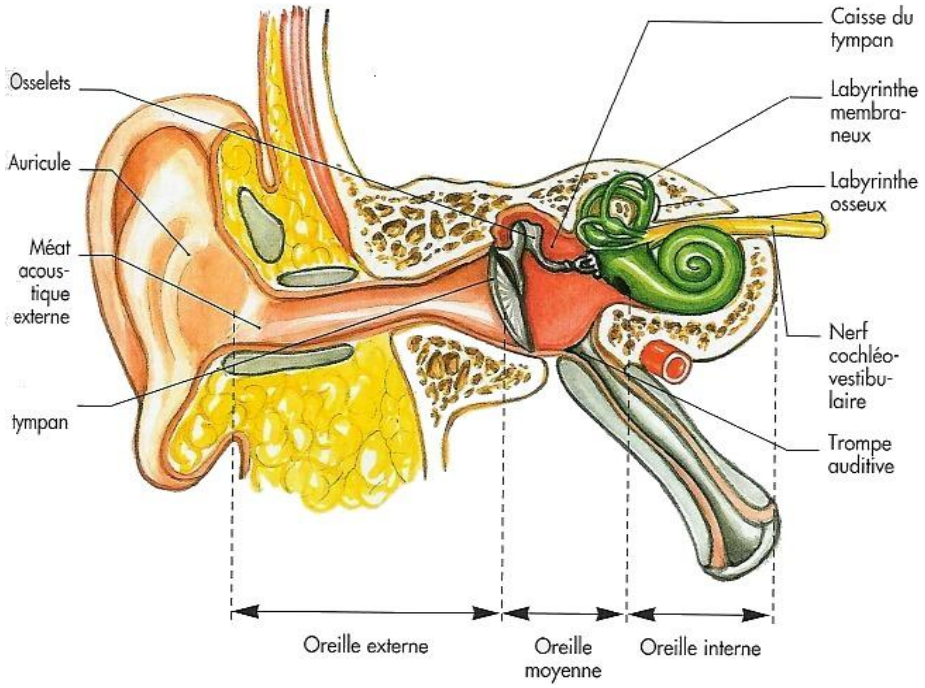


Figure 15 : Schéma des trois compartiments qui constituent l'appareil auditif (d'après Bonfils et Chevallier, 1998[40])

L'oreille externe

L'oreille externe est composée du *pavillon* et *l'auricule.* Le pavillon est une sorte de cornet acoustique qui recherche la direction du son et capte les ondes provenant d'une source sonore cible. Il se prolonge vers l'intérieur en rejoignant le *méat ou conduit auditif externe*, sorte de résonateur de 25 mm de long et de 8 mm de diamètre environ. Ce conduit s'enfonce obliquement dans l'os temporal et se termine en dedans par la *membrane tympanique ou tympan*.

L'oreille moyenne

L'oreille moyenne comprend le *tympan* qui constitue la frontière avec l'oreille externe (Figure 15*)*. Le tympan est sensible aux variations de pression provoquées par l'arrivée d'une onde acoustique. C'est une fine membrane élastique, tendue, très résistante et concave en dehors. Son diamètre est d'environ 10 mm, et son épaisseur de 0,1 mm. L'oreille moyenne est une cavité de forme irrégulière, pleine d'air, enserrée dans les os du crâne, mais susceptible d'être

[**] Les ondes captées par l'oreille sont des vibrations qui se propagent dans le milieu élastique de l'air avec une vitesse de 340 m/s environ. Cette vitesse ne dépend que de la nature, de la température et de la pression du milieu transmetteur.

reliée au pharynx, pendant la déglutition, par la *trompe auditive ou trompe d'Eustache* (Figure 15). Ce passage permet l'égalisation fréquente de la pression de l'air de l'oreille moyenne et de celle de l'air extérieur afin d'éviter des différences de pression pouvant altérer l'audition. On peut constater lors de la déglutition qu'on passe par un court moment de « surdité ». De même, lors d'une descente rapide en avion, des bourdonnements d'oreille se manifestent dus à l'augmentation de la pression atmosphérique. Pour éviter cet effet gênant, il suffit de déglutir plusieurs fois de façon à ramener à sa valeur extérieure la pression de l'oreille moyenne.

L'oreille moyenne comprend également la *caisse du tympan* laquelle contient des *osselets* constituant une chaîne osseuse articulée à partir du tympan jusqu'à la frontière de l'oreille moyenne avec l'oreille interne : *le marteau, l'enclume et l'étrier* (Figure 15). La grande apophyse du marteau est insérée dans la membrane tympanique. La tête du marteau repose sur l'enclume, laquelle se termine par un renflement, l'apophyse lenticulaire, qui s'articule avec l'étrier.

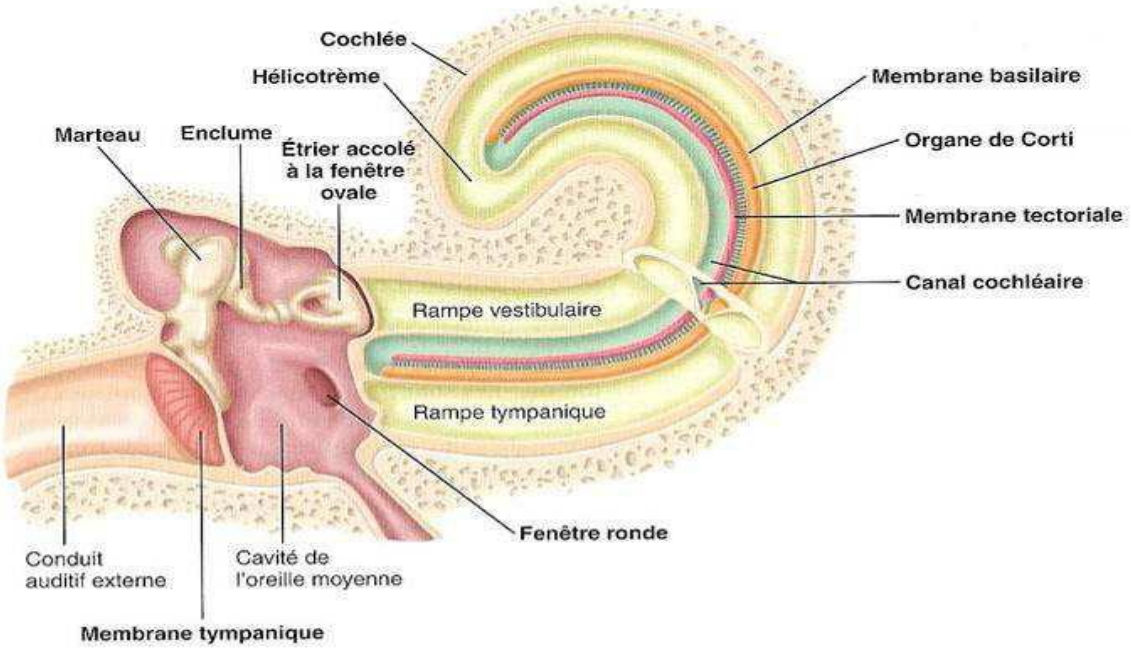


Figure 16 : L'oreille moyenne et l'oreille interne, cochlée dépliée.

La variation de pression exercée sur le tympan entraîne un mouvement synchrone du marteau, répercuté sur l'enclume, puis sur l'étrier. La platine de l'étrier vient s'enfoncer dans la *fenêtre ovale* qui est située sur la paroi interne de la caisse tympanique, et transmet alors la vibration de l'oreille moyenne au vestibule de l'oreille interne. Les osselets sont considérés comme des adaptateurs car ils transmettent les pressions reçues par le tympan après multiplication par

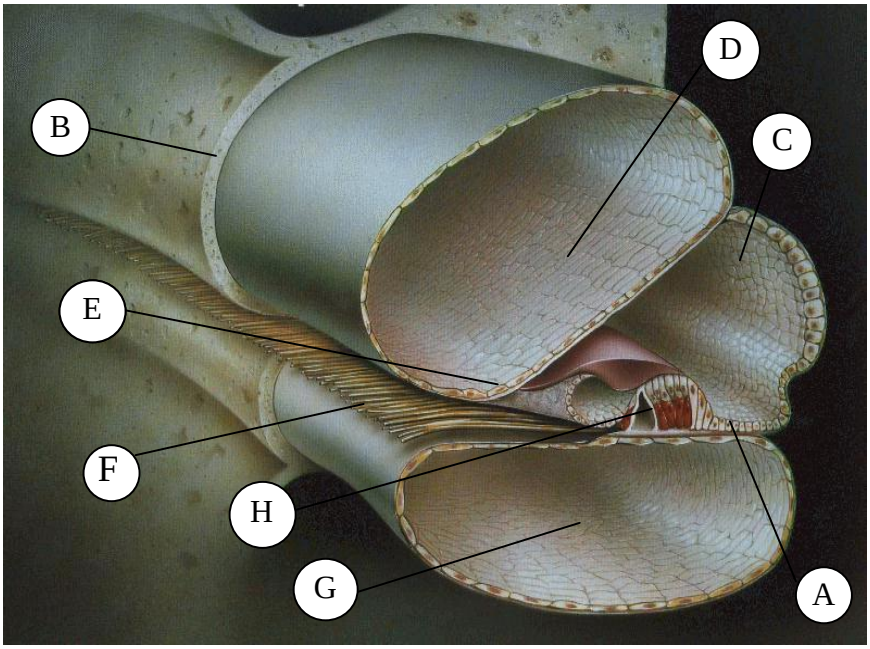
soixante environ (Matras, 1982)[41], les vitesses sont réduites d'autant. De cette manière, l'oreille interne est protégée contre les perturbations trop intenses, en se déformant sous leur action.

L'oreille interne

L'oreille interne est constituée de la *cochlée* en avant, organe proprement dit de l'audition, et de trois *canaux semi-circulaires* en arrière (Figure 15). L'ensemble de ces cavités est rempli de liquide : le *liquide périlabyrinthique*. La cochlée est formée d'un tube osseux enroulé en forme de colimaçon (2 tours et ½) autour du nerf auditif. Elle a une structure extrêmement complexe.

La cochlée est divisée en deux rampes sur toute sa longueur (Figure 16), la *rampe vestibulaire* et la *rampe tympanique*, par une cloison osseuse dans sa partie axiale (*lame spirale*), et membraneuse dans sa partie distale (*membrane basilaire*). Entre ces deux rampes se trouve le canal cochléaire correspondant à la partie externe des tours de colimaçon de la cochlée. Ces rampes sont en contact avec l'oreille moyenne par la présence de deux fenêtres : la *fenêtre ovale* qui reçoit l'impulsion vibratoire par l'intermédiaire de l'étrier, et la *fenêtre ronde* qui, fermée par une membrane élastique appelée tympan secondaire, sert de soupape de décompression. La membrane basilaire a une longueur de 32 mm. Sa largeur est de 0,04 mm à la base et croît progressivement jusqu'à 0,36 mm au sommet alors que la lame spirale osseuse diminue.

Le limaçon, ou petit sac membraneux situé à l'intérieur de la cochlée, présente une forme plutôt triangulaire lors d'une coupe transversale. Il est déterminé par la juxtaposition de trois membranes, dont la plus importante est la membrane basilaire (Figure 17). Le limaçon est séparé de la rampe vestibulaire par la membrane de Reissner et de la rampe tympanique par la membrane basilaire. A la partie centrale de sa base d'implantation sur la membrane basilaire, son épithélium subit une haute différenciation cellulaire qui constitue un renflement : l'organe de Corti. De part et d'autre de celui-ci sont implantées les cellules sensorielles, les cellules ciliées (Figure 18). Il en existe une seule rangée à l'intérieur du tunnel de Corti, les cellules ciliées internes qui sont au nombre de 3 500 et 3 à 4 rangées à l'extérieur du tunnel, les cellules ciliées externes, au nombre de 20 000 environ.



A. Membrane basilaire
B. Lame spirale
C. Limaçon ou canal cochléaire
D. Rampe vestibulaire
E. Membrane de Reissner
F. Nerf cochléaire
G. Rampe tympanique
H. Organe de Corti

Figure 17 : Coupe d'un tour de la cochlée.

Au contact de la base des cellules ciliées se terminent les fibres nerveuses qui se réunissent en faisceau pour constituer le *nerf cochléaire* ou nerf de l'audition lequel est constitué de 30 000 fibres environ. C'est donc à ce niveau que la vibration jusqu'alors mécanique se transforme en influx électrique neuro-sensoriel. L'organe de Corti est recouvert de la *membrane tectoriale* qui s'insère à la face supérieure de la lame spirale en dedans et en contact très étroit avec les cellules ciliées internes et externes qui la pénètrent par leurs cils. L'ensemble est d'une remarquable régularité, toutes les dimensions subissent des altérations régulières de la base au sommet de la cochlée.

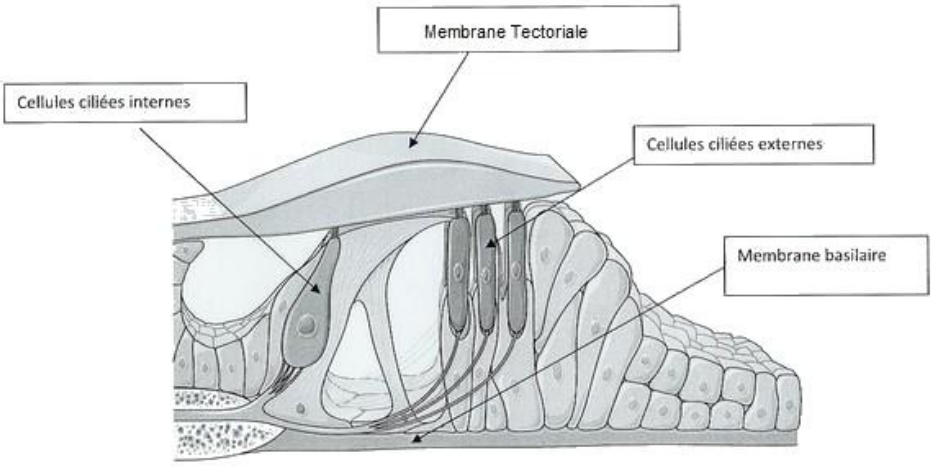


Figure 18 : Vue détaillée de l'organe de Corti.

2.2. Physiologie de l'oreille

Les ondes sonores sont captées par le pavillon de l'oreille. Grâce à la forme du pavillon, la direction de la source sonore peut être relativement déterminée. L'onde sonore captée est ensuite amplifiée* par le conduit auditif externe et transmise jusqu'au tympan lequel se met en vibration entraînant la chaîne d'osselets.

Le mouvement synchrone du marteau, répercuté sur l'enclume, puis sur l'étrier retentit sur la platine de l'étrier qui vient s'enfoncer dans la fenêtre ovale. Les osselets accomplissent un rôle d'adaptateurs de l'impédance** dans la transmission des ondes. Ils participent à l'amplification ou à l'atténuation de l'intensité des ondes perçues par le tympan. Contrairement à l'oreille externe et moyenne, l'oreille interne n'est donc pas un milieu aérien mais liquidien. Le mouvement de la chaîne ossiculaire génère un déplacement de liquide périlymphatique dans la rampe vestibulaire par la fenêtre ovale. L'énergie de l'onde sonore transmise par le déplacement du liquide se dissipe dans les mouvements de la fenêtre ronde accomplissant un rôle de « soupape ». Ses mouvements facilitent la décompression de l'intérieur de l'oreille interne et s'effectuent de manière inverse aux mouvements de la fenêtre ovale. Quand la fenêtre ovale cède vers l'intérieur de l'oreille interne, la fenêtre ronde se déplace vers l'extérieur.

Les mouvements du liquide périlymphatique provenant de la fenêtre ovale stimulent la membrane basilaire. L'onde se propage à partir de la base de la cochlée jusqu'à l'apex. La position maximale atteinte de déplacement de cette onde dépend de la fréquence du son perçu. La réponse à l'onde sonore n'est pas donc la même en tout point : chaque section de la membrane basilaire a une fréquence de résonance précise. La stimulation de la membrane est la plus forte en ce point précis pour la fréquence donnée. C'est ainsi que les cellules de la région basale, près de la fenêtre ovale, sont sensibles aux fréquences aiguës tandis que les cellules de l'apex sont sensibles aux les fréquences graves. On dit que la cochlée a une *« organisation tonotopique »* (Figure 19). Cette organisation couvre une plage fréquentielle

---

* Du fait de la géométrie et de la nature des parois de l'oreille externe, les ondes sonores ne sont pas également transmises. L'ensemble de la conque et du conduit auditif a une fréquence de résonance vers 2,5 kHz et la conque seule vers 5,5 kHz. Il s'ensuit que la transmission des fréquences comprises entre 2 et 7 kHz s'opère avec un gain de l'ordre de 10 à 20 dB par rapport aux fréquences inférieures à 0,5 kHz.

** Diminution du rapport entre l'impédance élevée de la cochlée et celle de la membrane tympanique étant beaucoup plus faible. Ainsi, les ondes faibles sont amplifiées, les ondes fortes sont atténuées et l'oreille interne est protégée. L'intensité de l'onde diminue en fonction de la distance de la source et de la nature de l'onde (si la fréquence et la forme de la vibration ne subissent pas de déformation).

entre 20 et 20 000 Hz. Une fois atteinte la position maximale de déplacement, l'onde s'estompe rapidement. La stimulation mécanique de la membrane basilaire est transmise aux cellules ciliées à travers l'effet de glissement entre l'ensemble formé par le tunnel de Corti et la membrane tectoriale. Il en résulte une inclinaison de cils placés au sommet des cellules sensorielles. Si cette excitation mécanique est suffisante, la cellule délivre un influx électrique aux centres nerveux et au cerveau via le nerf auditif (nerf Cochléaire ou VIII paire crânienne) indiquant l'intensité de la stimulation.

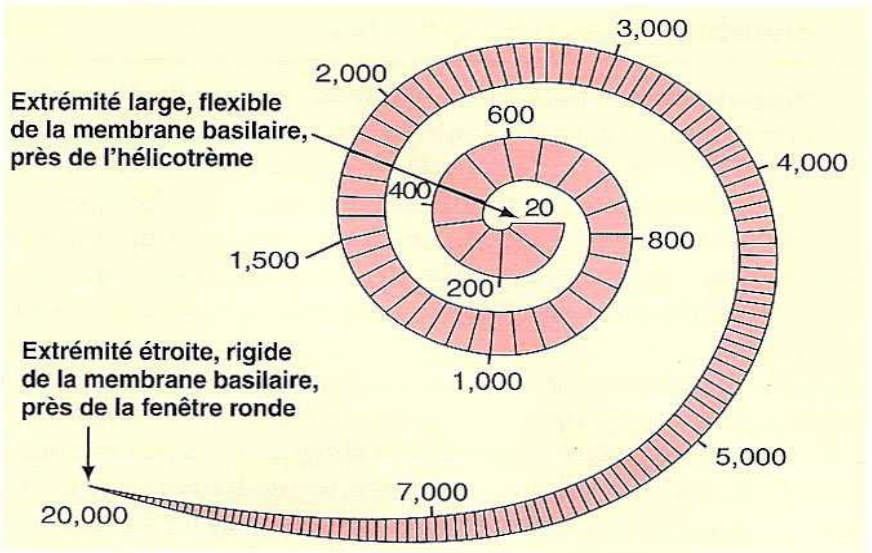


Figure 19 : Schéma de l'organisation tonotopique de la membrane basilaire.

On compte le long de la cochlée humaine environ 3 500 cellules ciliées internes et 20 000 cellules ciliées externes, qui transmettent les ondes et qui sont reliées à environ 40 000 fibres nerveuses. La transformation de l'information mécanique en information électrique au niveau de la cochlée et plus précisément au niveau de l'organe de Corti a été largement étudiée. De nombreux travaux ont été entrepris pour déterminer avec précision le rôle respectif de la membrane basilaire (par exemple les travaux de Békésy) et du nerf auditif dans la discrimination des fréquences.

Helmholtz (1868 in Matras, 1982)[42] avait argumenté l'idée que l'analyse des fréquences perçues par l'oreille se produisait dans le limaçon. Chacune des trente-cinq mille fibres de la membrane basilaire entre en résonance pour une fréquence déterminée et communique la vibration à la fibre nerveuse voisine. Néanmoins, d'autres arguments formulés par Baring (in Matras, 1982)[41] dans la *théorie du téléphone* remettent en cause l'idée de Helmholtz. Pour

Baring, la membrane basilaire jouerait le rôle d'un central téléphonique relié au cerveau par les fibres du nerf cochléaire dont l'ensemble constituerait un véritable fil de transmission. Puis, l'analyse du son se produirait dans le cerveau. La plus grande objection de cette théorie, c'est qu'elle n'explique absolument pas le fonctionnement ni le rôle joué par la structure complexe de l'oreille interne dans l'audition. Les théories de Helmholtz et Baring ainsi que celles de Wever et Bray (in Matras, 1982[41]) semblent à présent dépassées.

Des études plus récentes ont mis en évidence que chaque fibre nerveuse est le siège d'un phénomène de propagation d'impulsions électriques. Pour l'onde d'un « son pur »[*], schématiquement, une seule fibre va être mise en action. Pour l'onde d'un « son complexe »[**], plusieurs fibres. Pour un son faible, les influx sont peu nombreux. Ils augmentent jusqu'à 1 000 par seconde quand l'intensité augmente, mais les fibres nerveuses ne peuvent émettre plus d'influx par seconde. Quand l'intensité augmente encore, les fibres voisines entrent en jeu même si elles ne correspondent pas exactement à la fréquence incitatrice. La spécificité de la perception diminue alors. De manière générale, ce phénomène d'excitation de fibres est soumis à deux lois:

La *loi du tout ou rien* qui déclare que le phénomène électrique n'apparaît que si le stimulus extérieur a une certaine valeur minimale, au-dessus de cette valeur, l'intensité des impulsions est toujours la même. D'une part, l'intensité des impulsions ne serait pas proportionnelle à l'intensité du stimulus. D'autre part, le nombre de fibres excitées donnerait l'impression d'intensité physiologique.

La *loi de la période réfractaire* déclare que les fibres du nerf cochléaire ne répondent à la cadence du stimulus que celle-ci ne dépasse pas une fréquence voisine de 1 000 Hz. Il est donc nécessaire de laisser le nerf au repos pendant 0,001 s (période dite réfractaire) pour qu'il récupère ses possibilités d'action.

A la lumière de ces dernières études, on peut dire que la discrimination des fréquences se produit essentiellement dans la cochlée laquelle est fortement affinée par l'effet très sélectif des fibres nerveuses. Cette discrimination se traduit enfin par une structuration de certaines zones du cortex cérébral.

---

[*] Un son pur est déterminé par une variation sinusoïdale de pression et caractérisé par son intensité et sa fréquence.

[**] Un son complexe est déterminé par plusieurs sons purs élémentaires dont les fréquences respectives sont des multiples de la composante la plus grave, nommée fréquence fondamentale selon la loi de Fourrier. Les autres fréquences sont appelées harmoniques.

2.3. Transmission de l'influx électrique au cerveau

Le nerf cochléaire, constitué par les fibres nerveuses qui rentrent en contact avec les cils des cellules ciliées de l'organe de Corti, achemine les influx nerveux au cerveau. Le nerf quitte le massif pétreux par le conduit auditif interne accolé aux nerfs vestibulaire et facial, pendant 2 ou 3 cm avant de pénétrer dans le bulbe. Toutes les structures du tronc cérébral et du diencéphale impliquées dans l'audition ont en commun des caractéristiques fonctionnelles telles que la tonotopie et la *loi de l'isofréquence*[*]. La fibre nerveuse appartenant au mécanisme auditif, issue d'une cellule sensorielle réceptrice d'une fréquence donnée va faire synapse, dans les étages sus-jacents, avec une autre cellule nerveuse spécialisée dans le traitement de cette fréquence et ainsi de suite jusqu'au cortex cérébral.

Les fibres nerveuses sortant de l'organe sensoriel, et se dirigeant vers les étages supérieurs, vont se regrouper et former différents noyaux ou centres de relais. Le premier centre de relais est constitué par le *noyau cochléaire dorsal et ventral* du tronc cérébral (Figure 20). On observe sur la figure une projection bilatérale à prédominance controlatérale. Ces noyaux contiennent des projections nerveuses controlatérales et homolatérales qui se dirigent vers le *mésencéphale*.

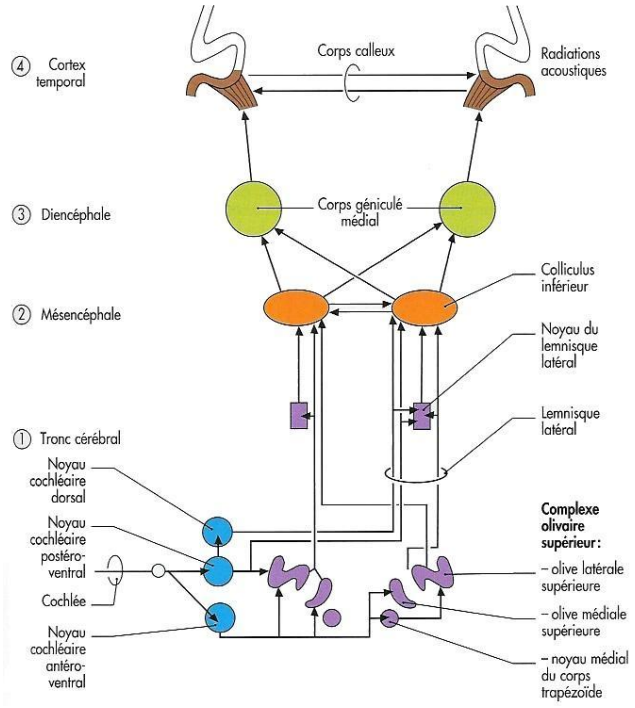


Figure 20 : Schéma simplifié des voies auditives centrales à quatre niveaux. (d'après Bonfils et Chevallier, 1998[40]).

[*] Il existe dans les nombreux systèmes anatomiques des étages de cellules qui traitent chacune une seule fréquence.[39]

La terminaison corticale des fibres est plutôt en « Y », à la fois sur l'aire auditive primaire (A1 ou 41) et sur l'aire auditive secondaire (A2 ou 42) (Figure 21).

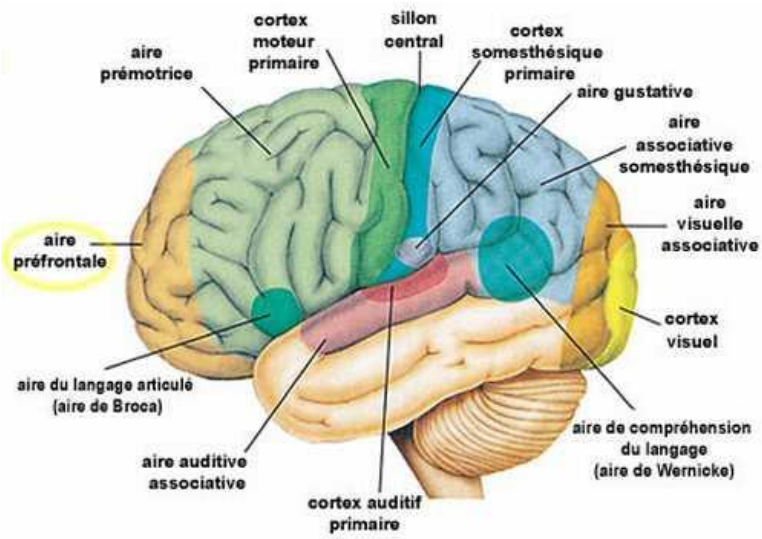


Figure 21 : Vue latérale du cerveau : aires ou zones corticales primaires et secondaires[*].

Parallèlement, il y a l'affluence de fibres efférentes issues du tronc cérébral qui terminent dans l'organe de Corti au niveau des cellules ciliées.

Le *complexe olivaire supérieur* a une organisation tonotopique. Les différents noyaux de ce complexe interviennent dans la localisation du son en fonction de la différence du temps d'arrivée du stimulus à chaque oreille (une différence de 400 ms peut être appréciée). Les noyaux latéraux sont consacrés à l'appréciation de l'intensité sonore et donnent naissance à des fibres descendantes inhibitrices appelées *« corps trapézoïde »*.

2.4. Perception du son

Moles (in Matras, 1982)[41] compare le cerveau avec un ordinateur pour imager le décodage des impulsions électriques provenant du nerf auditif. Il considère que :

« tout se passe comme si, dans le cerveau, se succédaient des cartes perforées, à la vitesse du temps de réaction de l'oreille...; sur ces cartes 's'impriment' les impulsions codées transmises par les nerfs. C'est en fonction de leur allure générale (rapidité d'évolution des perforations)

[*] Source de l'image : http://www.lyc-briand-gap.ac-aix-marseille.fr/mpi_2005/2DE%208%20ET%209/frassatibonnefoy/images/cerveauaires.jpg

que, en se référant à l'ensemble des cartes déjà stockées en mémoire, le cerveau livre telle ou telle sensation » (p.38).

En effet, le processus de perception auditive au niveau du cortex cérébral en quelque sorte pourrait être schématisé comme étant le résultat d'un stockage d'impulsions codées correspondant à des sensations. Néanmoins, ce processus est beaucoup plus complexe que l'on pense et une telle analogie resterait vague pour expliquer les principes sous-jacents de la perception du son et de la parole.

Les impulsions électriques créées remontent au cerveau où elles son traitées par différentes zones. Certaines zones ou aires traitent toutes les modalités de sons perçus (par exemple, le bruit, la musique, le langage), d'autres s'occupent de manière spécifique de sons concernant plutôt une seule catégorie. De manière générale, les fibres du nerf auditif se terminent dans la circonvolution transverse de Heschl située dans l'aire primaire ou aire 41 (*Figure 21).* Celle-ci étant une aire commune à toutes les perceptions auditives.

Quant à la perception de bruits de l'environnement, il a été démontré qu'il n'y a pas de structures anatomiques définies pour le traitement acoustique (Lechevalier, 1997)[39]. En revanche, à l'inverse du traitement de bruits, des études ont conclu que la perception de la musique est réalisée par les deux hémisphères cérébraux. L'hémisphère gauche est réservé à la perception du rythme (Mavlov, 1980)[43], à l'identification d'une œuvre musicale entendue (Lechevalier et al., 1995)[44], à la perception des hauteurs, à l'écriture et à la lecture musicales. L'hémisphère droit analyse sous sa dépendance la perception des timbres et la reconnaissance du contour mélodique (Peretz, 1990)[45]. De cette manière, la perception de la musique serait accomplit grâce à la participation des deux hémisphères cérébraux.

Quant à la perception des sons de la parole, certaines aires telles que 22 (Wernicke), 44 (Broca) et 8 ont été classiquement déterminées comme étant les centres du langage. L'aire 22 était considérée comme un « processeur de sons » qui extrayait les phonèmes des sons verbaux, leur apportant une signification. De nos jours, l'aire 22 apparait désormais comme une vaste zone d'un réseau neuronal diffus reparti en deux zones : une zone antérieure (aires 22, 42) traitant la composante phonologique de l'écoute et une zone postérieure (aires 22, 39) traitant la composante sémantique. Cependant, on ne peut exclure des variations individuelles dans les réseaux neuronaux présidant exclusivement à la compréhension ou à la production des sons verbaux (Wise et col., 1991[46]; Zatorre et col., 1992[47]). Par ailleurs, des travaux ont

démontré la participation d'autres zones dans la perception des sons verbaux (Cambier et col., 1983[48] ; Belin, 2009[49]). Si l'aire motrice secondaire apparaît comme une région primordiale dans l'initiation du langage, son rôle dans la perception est moins démontré. Une des découvertes importantes dans ces dernières décennies porte sur la contribution de l'hémisphère droit dans le langage. Enfin, la perception de la prosodie et de la pragmatique serait assurée par l'hémisphère « mineur » (Hannequin et col., 1987)[50].

En ce qui concerne la voix et la parole, trois approches peuvent être distinguées pour expliquer la perception selon la nature du stimulus : 1) l'approche articulatoire qui attribue aux gestes articulatoires à valeur phonémique un rôle d'unité minimale d'échange dans la communication (Liberman et Mattingly, 1985)[51], 2) l'approche acoustique-auditive qui détermine que le traitement du signal acoustique de la parole doit être exploré selon la perception auditive (Ohala, 1986)[52], 3) l'approche « psycholinguistique » qui tente de caractériser la nature des processus et des représentations mentales permettant à l'auditeur d'associer au signal de parole un élément spécifique de sa langue tels que le phonème, la syllabe, le mot, etc. (Segui in Lambert, 1997)[53]. Dans notre travail, nous adopterons plutôt la deuxième approche (acoustique-auditive) pour expliquer la perception des stimuli sonores.

Les stimuli sonores que nous percevons sont de deux types selon leur forme d'onde : simple et complexe. Les premiers correspondent à un ton pur généré par exemple par la vibration d'un diapason. Il faut savoir qu'un corps vibrant produit toujours un son et que ses propriétés déterminent les caractéristiques du son produit. Ainsi, quand on frappe un diapason, celui-ci comprime les molécules d'air et les disperse. L'onde résultante a une forme sinusoïdale qui peut être caractérisée essentiellement par deux paramètres : fréquence et amplitude (Figure 22). La fréquence correspond au nombre d'oscillations effectuées par le corps vibrant pendant une unité de temps et s'associe à la sensation de « grave » ou « aigu ». Elle est généralement exprimée en cycles par seconde ou Hertz (Hz) dans un phénomène périodique. L'amplitude représente la variation de la pression d'air qui produit l'onde sonore et est liée à la sensation de « fort » ou « faible ». Elle est généralement exprimée en décibels (dB).

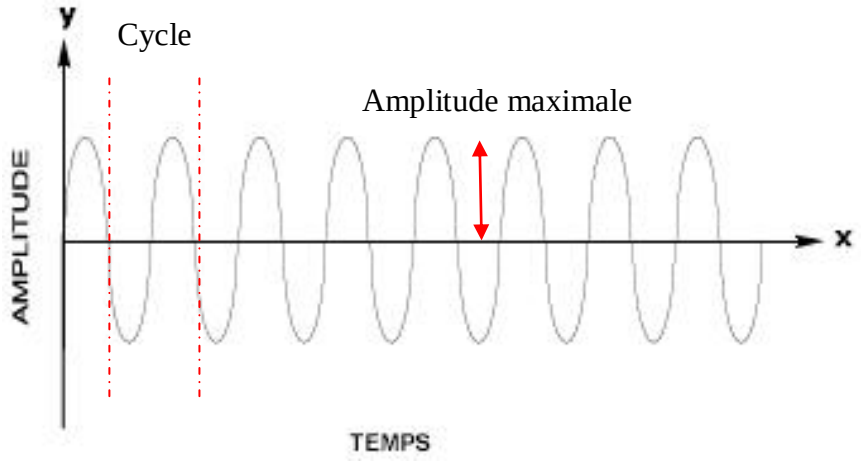


Figure 22 : Exemple d'une onde sonore simple.

Le deuxième type de stimulus sonore est constitué d'une forme d'onde complexe (Figure 23Figure 23). L'onde acoustique est constituée par une multiplicité de fréquences de différentes amplitudes et dont les propriétés varient dans le temps. Ces stimuli correspondent aux sons de la parole générés par les différents composants du mécanisme phonatoire. Toute onde sonore complexe peut être décomposée en une somme d'ondes simples ayant chacune sa propre fréquence et amplitude.

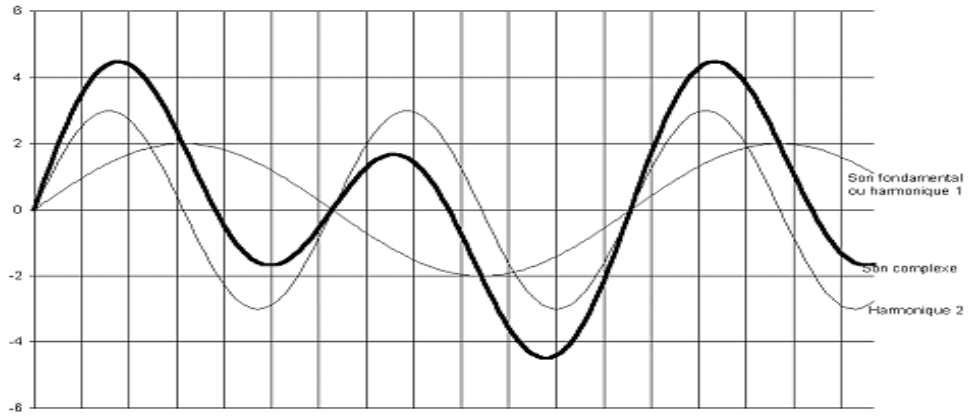


Figure 23 : Exemple d'une onde sonore complexe et des ondes simples la constituant

Schématiquement, on peut dire que la perception des paramètres d'un son s'obtient au niveau de l'oreille interne et du cerveau. Cette perception se traduit par une sensation de hauteur et de force sonore de manière subjective. Concernant la perception de la fréquence, on sait que l'oreille humaine peut percevoir les sons entre 20 Hz et 20 000 Hz. Néanmoins, la sensation de différence entre une fréquence et la fréquence deux fois plus rapide est perçue comme le même intervalle quelque soit la fréquence de départ. Par exemple, notre oreille percevra la

même sensation de hauteur entre 205 Hz et 500 qu'entre 500 Hz et 1 000 Hz. A partir d'une étude réalisée par Stevens et Volkman (1940 in Revis, 2004)[54], il a été démontré qu'en-dessous de 1 000 Hz, toute différence de ±3 Hz est perçue et décodée par l'oreille et le cerveau comme un stimulus différent, quelle que soit la fréquence du son au départ. Au-dessus de 1000 Hz, la sensation de différence varie de 3% en fonction de la hauteur. Par exemple, il est de 3 Hz à 1 000 Hz, 4,5 Hz à 1 500 Hz, etc. La zone la plus sensible de l'oreille humaine se situe entre 2 000 Hz et 3 000 Hz. La fréquence de 3 000 Hz est extrêmement importante car cette fréquence joue un rôle clé dans l'intelligibilité de la voix.

Concernant la perception de l'intensité, il existe la notion de seuil auditif déterminé par l'intensité minimale à laquelle un sujet entend un son et l'intensité maximale provoquant une sensation de douleur, d'intolérance ou d'inconfort. L'intensité maximale se situe aux alentours de 120 dB (HL). L'intensité minimale dépend de la fréquence du son. Par exemple, les fréquences entre 1 000 Hz et 4 000 Hz nécessitent moins d'intensité pour être perçues. Tandis que les fréquences entre 100 et 125 Hz nécessitent une intensité minimale de 40 dB. Le paramètre de l'intensité influence la perception de la fréquence. Par exemple, pour des fréquences inférieures à 1 000 Hz, on a la sensation auditive d'une diminution de la hauteur lorsqu'on augmente l'intensité. Au contraire, pour des fréquences supérieures à 4 000 Hz, on a la sensation d'une augmentation de la hauteur lorsqu'on augmente l'intensité. Par ailleurs, la plus petite variation d'intensité perçue par l'oreille est de 3dB pour une fréquence de 1 000 Hz à 5 dB au-dessus du seuil absolu ; alors que pour le même son à 100 dB au-dessus du seuil absolu, une différence sera perçue dès 0,25dB.

Mais la perception d'un signal sonore simple n'est pas la même que pour un signal complexe. La perception de hauteur pour le son complexe est déterminée par la fréquence fondamentale, celle-ci peut être absente physiquement du signal acoustique. La présence des harmoniques supérieurs et la périodicité de l'onde globale suffisent pour en déterminer la hauteur. Par conséquent, la hauteur est intégrée par une analyse temporelle et non fréquentielle. En outre, la perception de hauteur des sons complexes est moins dépendante de l'intensité que pour les sons simples.

Quant à la perception de l'intensité des sons complexes, il semblerait que l'oreille effectue une analyse spectrale des sons complexes en fonction des caractéristiques de l'oreille, principalement en ce qui concerne la non-linéarité subjective en intensité et fréquence.

En somme, la perception d'un son est un processus physiologique humain extrêmement complexe qui est relié aux propriétés acoustiques du signal sonore. Ce processus est le résultat de la mise en jeu de diverses structures anatomiques suivant différentes lois physiques. Cette section a tenté de rappeler l'anatomie, la physiologie et les principes acoustiques impliqués dans la perception d'un son. Néanmoins, la compréhension des mécanismes participant dans le décodage et dans l'interprétation acoustique au niveau cérébral est restreinte et représente un énorme défi.

3. La boucle audio-phonatoire

Parret (2002)[12] souligne qu'il n'y aurait pas de voix sans qu'il y ait oreille. En effet, la perception et la production de la voix sont deux processus physiologiques intimement liées, au point de s'influencer l'un l'autre par l'intermédiaire d'un mécanisme régulateur dénommé « boucle audio-phonatoire ». La production de la voix et de la parole est régulée par des centres auditifs (section 2). Cette régulation a lieu grâce à l'action d'un mécanisme réflexe involontaire et d'un mécanisme conscient volontaire.

Quant à l'action du mécanisme réflexe, les ajustements articulatoires surviennent dans le processus de phonation grâce à l'action du réflexe cochléo-bulbo-récurrentiel ou acoustico-laryngé. Un certain nombre d'expériences conduites sur la voix parlée ont permis d'explorer les mécanismes utilisés pour la régulation de l'émission vocale. Certaines expériences datent du début du XXe siècle mais leurs conclusions demeurent toujours en vigueur. Une de ces expériences est celle menée par Lombard (1911)[55] lorsqu'il étudiait les phénomènes de régulation des productions vocales en présence de bruit. Lombard démontre l'effet que le bruit produit sur la parole dans des situations d'atténuation du retour auditif de la voix telles que le bruit ambiant ou le port de protections auditives, entre autres. Il s'aperçoit que l'intensité phonatoire augmente sous l'influence de bruit. Cet effet a été constaté par le rapprochement de deux faits d'observation : 1) Les sourds par lésion de l'appareil labyrinthique élèvent anormalement la voix, 2) Un sujet dont l'audition est intacte, élève anormalement la voix au milieu d'un bruit assourdissant. Selon Lombard, le locuteur, qui n'entend pas mieux sa propre voix que les bruits extérieurs, s'efforce de remédier à cette audition défaillante en augmentant l'intensité du son glottique. Tout se passe comme si « le locuteur se crie en quelque sorte à lui-même comme on crie à l'oreille d'un sourd », et ce d'une manière réflexe. D'autres mécanismes réflexes d'origine abdominale thoracique, cervicale, linguale participent probablement aux ajustements constants agissant par feed-back sur le vibrateur laryngé.

A partir de 1920, une autre expérience a été conduite par Fletcher afin d'explorer la régulation de la voix et de la parole (Allen, 1996)[56]. Fletcher introduit la notion de « side-tone » comme étant le niveau perçu de sa propre voix. C'est donc une donnée subjective mais déterminante dans le contrôle de la voix. L'expérience de Fletcher a consisté à renvoyer la propre voix d'un individu dont on peut varier l'intensité. Il a observé que si l'on amplifie le niveau du retour, c'est-dire le *side-tone*, le niveau vocal du sujet diminue ; inversement, si le *side-tone* diminue, le niveau vocal augmente. Les tentatives de quantification aboutissent à un résultat semblable

à celui de l'effet Lombard : si le *side-tone* augmente de 10 dB, la voix diminue de la moitié de cette valeur environ, soit 5 dB.

Ces expériences fournissent des résultats concordants qui ont tenté de démontrer notamment le contrôle de l'intensité de la voix et de la parole par rapport au bruit environnant externe par un mécanisme réflexe. Le locuteur s'efforce de compenser au mieux les changements du rapport signal/bruit qui détermine son intelligibilité. Ces manipulations ont eu lieu dans des situations acoustiques inhabituelles nécessitant d'une régulation spécifique. On peut dire que ces expérimentations ont provoqué des situations de détérioration artificielle de la boucle qui finalement ne tiennent pas compte de la véritable réalité physique subie par ce mécanisme.

L'effet Lombard est interprété dans la littérature tantôt comme une accommodation aux exigences de la communication intelligible (Lane and Tranel, 1971)[57], tantôt comme un réflexe physiologique (Garde, 1965)[58]. Certains auteurs comme Lane et Sisson (1970)[59] et Lane et Tranel (1971)[57], conçoivent la régulation de la voix plus comme une recherche d'intelligibilité envers l'interlocuteur, que comme une motivation de s'entendre soi-même. De ce point de vue, l'effet Lombard n'a de sens que lorsque le locuteur se trouve dans un contexte de communication.

Quant à l'action d'un mécanisme conscient pour la régulation de la voix, on pourrait dire qu'il existe probablement l'envoi des informations efférentes induites par les voies cortico-bulbaires en réponse aux informations acoustiques parvenues jusqu'au cortex auditif. Les centres nerveux participant dans l'action de ce mécanisme sont toujours mal définis (cf. section 2.4). Les aires cérébrales perceptives 22, 44 et 39 envoient des informations aux centres de production de la parole et de la voix de l'hémisphère gauche à partir de la comparaison des « images acoustiques » parvenues et celles de référence stockées en mémoire. Si l'identité des deux images acoustiques n'est pas reconnue, on procède à de rapides ajustements, parfois perçus par l'interlocuteur. Moles (in Matras, 1982)[41] compare le cerveau avec un ordinateur pour imager le décodage des impulsions électriques provenant du nerf auditif. Il considère que :

«… tout se passe comme si, dans le cerveau, se succédaient des cartes perforées, à la vitesse du temps de réaction de l'oreille...; sur ces cartes 's'impriment' les impulsions codées transmises par les nerfs. C'est en fonction de leur allure générale (rapidité d'évolution des

perforations) que, en se référant à l'ensemble des cartes déjà stockées en mémoire, le cerveau livre telle ou telle sensation » (p. 37-38).

L'hémisphère gauche compare les images acoustiques, détermine l'ajustement à effectuer en temps réel, coordonne et donne l'ordre moteur aux structures bucco-phonatoires pour atteindre l'émission désirée. La représentation acoustique interne existant au niveau cortical grâce à une stimulation auditive préalable permettrait d'expliquer la voix perturbée et non modulée des sourds profonds. A ce propos, Husson (1962)[60] a observé que lorsqu'un sujet émet un son à hauteur voulue avec une représentation psycho-auditive préalable de la hauteur, la fréquence de la voix est tenue. Dans le cas contraire, la hauteur vocalique n'est pas tenue.

De manière générale, les ajustements articulatoires surviennent dans le processus de phonation en deux temps : lors de l'ajustement préphonatoire et lors de l'ajustement phonatoire proprement dit. L'ajustement préphonatoire est indépendant du contrôle audio-phonatoire. Un exemple de l'ajustement préphonatoire est le mécanisme utilisé par le chanteur professionnel. Celui-ci est capable d'émettre un son qui est d'emblée à la hauteur et intensité désirées sans attendre le contrôle auditif. Il s'agit d'un contrôle exercé en synergie avec le contrôle proprioceptif permettant l'anticipation préphonatoire. En effet, ce réglage préphonatoire d'origine corticale dépend essentiellement des informations fournies par les mécanorécepteurs laryngés. L'information envoyée des mécanorécepteurs aux centres portent des informations sur l'état de tension et sur la position des différents muscles et articulations. L'ajustement phonatoire ayant lieu au cours de l'émission, il s'accomplit grâce aux informations permettant des ajustements instantanés nécessaires pour le maintien d'une configuration glottique donnée.

La production de la parole est l'exemple parfait de la mise en jeu de ces mécanismes de contrôle réflexes et conscients. La rapidité du flux de la parole et la précision des points d'articulation (sur le plan phonétique) nécessitent des accommodations aussi rapides que fines pour mener à terme l'acte de la parole. Ainsi, les adaptations sont assurées grâce à des mécanismes réflexes extrêmement rapides qui sont supportés par la sensibilité tactile et proprioceptive des lèvres, de la langue, du palais et de la mâchoire.

La boucle audio-phonatoire est un mécanisme de régulation complexe permanent qui permet à l'être humain d'acquérir et de développer la voix, la parole et le langage oral essentiels dans la communication. En effet, nous ne pouvons émettre vocalement que ce que nous entendons.

Les travaux sur l'appauvrissement vocalique des jeunes enfants sourds au cours du temps supposent l'existence et l'action de la boucle audio-phonatoire dans l'acquisition et le développement du langage. Vinter (1994)[61] a observé que l'enfant sourd, comme tous les enfants entendants, babille et s'adonne à des jeux vocaux (jasis) pendant les premiers mois de vie. L'enfant prend du plaisir à sentir les vibrations au niveau du larynx et du pharynx. Les premières phases du babillage sont amorcées par simulation sensorielle et préparent l'enfant à des comportements vocaux plus complexes (Locke et Pearson, in Ferguson,1992)[62]. L'imitation met en jeu les boucles proprioceptives et auditives provenant de la perception des sons environnants. Néanmoins, l'enfant sourd diminue ses productions quantitativement et qualitativement après un certain temps en raison de l'absence de la boucle audio-phonatoire. Parallèlement à cette diminution de productions vocales, l'enfant sourd voit son intérêt pour la parole et le monde sonore décroître car il n'a pas de retour auditif.

Seul un fonctionnement très subtil et coordonné des mécanismes qui sous-tendent ces performances peut permettre cette mise en relation permanente entre ce qui est produit et ce qui est perçu. Dans cette section nous avons présenté les processus physiologiques sous-jacents de la production et de la perception des sons et des mots mettant en évidence l'étroite interdépendance des ces deux processus.

4. La surdité : déficience de la perception auditive

La surdité est une déficience de la perception auditive qui peut entrainer un handicap. Cette déficience altère la qualité de détection et d'identification des signaux sonores notamment ceux de la parole, et entrave la communication de soi avec le monde environnant et ses interlocuteurs. La surdité se caractérise par une perte partielle ou complète de la capacité d'entendre d'une oreille (surdité unilatérale) ou des deux oreilles (surdité bilatérale). La perte auditive peut rester stable tout au long de la vie ou évoluer. Quand la surdité se manifeste tardivement, il y a une baisse progressive des facultés auditives souvent à l'insu de la personne concernée. La plupart du temps, ce sont les proches des personnes malentendantes qui décèlent les premiers le problème.

D'une manière générale, la surdité affecte 17% de la population mondiale (Deggouj et Devresse, 2006)[63]. La prévalence de la surdité sévère et profonde chez l'enfant est de l'ordre de 1/1 000 à 1/2 000 à la naissance et de 3/1 000 à 3 ans (Lina-Grenade, 2000)[64]. Cette prévalence augmente chez les nouveau-nés hospitalisés en réanimation néonatale de 1 à 2/100 (François, 2001)[65].

Types de surdité

Il existe différents types de surdité se manifestant en fonction de la partie du mécanisme auditif atteint : surdité de transmission, de perception ou mixte.

La *surdité de transmission* est due à une atteinte au niveau de l'oreille externe et moyenne avec une cochlée normale. La conduction aérienne du signal sonore est compromise. Cette surdité ne se manifeste pas de manière totale et l'audition peut être récupérée. Le plus souvent c'est la perception des sons graves qui est altérée. Le sujet s'entend parler par résonance crânienne et de ce fait, même dans le cas d'une surdité profonde de transmission, la voix et la parole sont peu perturbées.

La *surdité de perception* correspond à une atteinte de l'oreille interne où la cochlée est généralement touchée mais le nerf auditif et les centres nerveux peuvent parfois être aussi endommagés. La conduction électrique du signal auditif est compromise. Dans la majorité des cas, la perception des fréquences aigues est altérée. Le sujet ne s'entend pas parler, et donc cela aura un retentissement dans sa production de la voix et de la parole. Ici, la surdité est traitée à l'aide d'une rééducation orthophonique et prothétique et, pour certains cas, par une intervention chirurgicale. La surdité de perception est le déficit sensoriel le plus fréquent.

La *surdité mixte* est due à une atteinte plus ou moins importante de l'oreille interne qui altère la transmission et la perception de sons. Le traitement chirurgical est le plus souvent adopté pour débuter la rééducation.

Degrés de surdité

On peut également caractériser la surdité en fonction du degré ou de la perte auditive en décibels (dB HL) par rapport à l'oreille normale et en référence aux normes ISO. En clinique, il est possible de déterminer le degré de surdité grâce à l'audiométrie tonale HTL (Hearing Tone Level). L'audiométrie est une épreuve audiologique qui permet de déterminer de manière quantitative et subjective le seuil auditif d'un individu. Le calcul de la perte auditive est assez simple selon la recommandation du Bureau International d'Audio-phonologie (Recommandation BIAP 02/1 bis, 1997)[*] :

« Une perte totale moyenne est calculée à partir de la perte en dB aux fréquences 500 Hz, 1000 Hz, 2000 Hz et 4000 Hz. Toute fréquence non perçue est notée à 120 dB de perte. Leur somme est divisée par quatre arrondie à l'unité supérieure. En cas de surdité asymétrique, le niveau moyen de perte en dB est multiplié par 7 pour la meilleure oreille et par 3 pour la plus mauvaise oreille. La somme est divisée par 10 ».

Nous présentons dans le tableau ci-dessous, la classification du BIAP que nous adopterons dans notre étude (Tableau 4).

Suivant le degré de perte auditive, la perception de la voix et de la parole peut être fortement altérée ou non. Par exemple, dans la surdité légère, l'enfant ne perçoit pas la voix faible et ne reconnait pas les contrastes phonétiques de la parole. Le langage peut se développer mais avec un retard. Dans la surdité moyenne, le jeune enfant ne développera que difficilement le langage. Alors que, dans la surdité sévère, l'enfant ne pourra pas développer de langage ou très partiellement. L'enfant identifie difficilement les consonnes et peut entendre sa voix mais à forte intensité. Dans la surdité profonde, la parole articulée n'est pas perçue mais des éléments prosodiques (mélodie et rythme) peuvent être conservés (Dumont, 1988)[66]. Si l'enfant n'est pas pris en charge tôt, un mutisme prendra place chez l'enfant.

[*] http://www.biap.org/recom02-1.htm

PERTE (dB)	DEGRE	INTERPRETATION
0 à 20	Audition normale ou subnormale	Il peut s'agir éventuellement d'une atteinte légère sans incidence sociale.
21 à 40	Surdité légère	La parole est perçue à intensité normale mais difficilement à voix basse ou lointaine. Les bruits familiers sont perçus.
41 à 55	Surdité moyenne 1er degré	La parole est perçue si on augmente l'intensité de la voix. Le contact visuel facilite la compréhension de la parole. Quelques bruits familiers peuvent être perçus.
56 à 70	Surdité moyenne 2e degré	
71 à 80	Surdité sévère 1er degré	La voix est perçue forte près de l'oreille. Les bruits forts sont perçus.
81 à 90	Surdité sévère 2e degré	
91 à 100	Surdité profonde 1er degré	Aucune perception de la parole. Seuls les bruits très puissants sont perçus.
101 et 110	Surdité profonde 2e degré	
111 à 119	Surdité profonde 3e degré	
120	Surdité totale ou cophose	Rien n'est perçu.

Tableau 4 : Classification de la surdité selon le degré (d'après le BIAP).

Age d'apparition de la surdité

La manifestation de la surdité peut avoir lieu tout au long de la vie. Selon la période de manifestation, la surdité peut être :

- *Congénitale* : lorsque la surdité se manifeste dès la naissance. L'enfant n'a aucune expérience auditive.

- *Pré-linguale* : lorsque la surdité survient avant la mise en place du langage chez l'enfant, c'est-à-dire, avant deux ou trois ans d'âge environ. Elle est souvent consécutive à une maladie (par exemple la méningite).

- *Péri-linguale* : lorsque la surdité survient pendant la mise en place du langage chez l'enfant, c'est-à-dire entre trois et quatre ans d'âge environ. L'enfant peut connaître quelques mots, il n'a donc qu'une courte expérience auditive.

- *Post-linguale* : lorsque la surdité se manifeste après la mise en place du langage chez l'enfant, c'est-à-dire, après cinq ans d'âge. L'enfant a déjà une expérience auditive qui lui permet de parler.

Dans le cas de l'enfant sourd, de nombreuses informations provenant de l'environnement sonore ne sont pas perçues, et son accès à la communication en est de ce fait rétréci (Van Lem et Peault, 1984)[67]. L'objectif crucial de la rééducation, en ce qui concerne l'enfant sourd, est donc de lui faciliter la communication le plus tôt possible. Le résultat de la prise en charge sur le développement du langage est d'autant meilleur qu'il a été entrepris précocement (Downs et Yoshinaga-Itano, 1999)[68]. Plus on attend et plus il sera difficile de revenir à un niveau satisfaisant d'audition. Idéalement, il est souhaitable de commencer la rééducation orthophonique et prothétique avant l'âge de 6 mois pour les surdités sévères et profondes congénitales (Yoshinaga-Itano, 1998)[69]. Depuis quelque temps, la prise en charge des surdités est de plus en plus précoce. Un dépistage et une prise en charge précoce des troubles auditifs chez les nourrissons permettent d'améliorer l'acquisition du langage oral par rapport aux enfants dépistés plus tardivement (Dauman et col., 2009)[70]. Néanmoins, les délais s'accumulent et le temps s'écoule entre le diagnostic et la prescription de la prothèse, puis entre la prescription et l'adaptation prothétique. Le délai entre le début supposé de la déficience auditive (naissance pour les affections congénitales) et le diagnostic peut être raccourci par une politique active de dépistage, sans attendre les signes cliniques tels que le désintérêt du monde sonore, la régression des productions vocales et le retard du langage qui sont toujours très retardés.

Bien qu'un dépistage s'effectue auprès des enfants à risque, quelque 50% des surdités ne sont pas détectés chez ces enfants. Par ailleurs, 20% à 30% des surdités ne sont pas détectables au cours des premiers mois de vie (Deggouj et Devresse, 2006)[63]. D'où la nécessité de réaliser des dépistages en maternité de manière systématique et de les réitérer entre 9 et 12 mois, voire plus tard lorsque l'enfant intègre l'école pour la première fois (Parving, 1999[71] ; François, 2003[72]).

Etiologie de la surdité

Les causes de la surdité sont multiples. Dans le cas des surdités de perception, on considérait jusqu'à ces dernières années qu'un tiers des surdités étaient liées à des causes environnementales (infections et toxiques en période pré ou postnatale, anoxie, prématurité), qu'un tiers seulement avait une origine génétique et un tiers était de cause indéterminée

(Denoyelle, 2001)[73]. La fréquence de ces groupes étiologiques était mieux étudiée dans la population pédiatrique. Actuellement, on sait que la majorité des cas sporadiques de surdité congénitale sont de formes génétiques récessives de surdité, et près de trois quarts des surdités congénitales sont donc d'origine génétique.

Les surdités génétiques sont des maladies monogéniques. L'altération d'un seul gène est en cause dans chaque forme de surdité et le déficit auditif résulte d'une atteinte cochléaire en dehors de rares syndromes avec surdité d'origine centrale. Dans un cas sur trois, la surdité est liée à d'autres pathologies ou malformations, ou elle est syndromique. Dans les deux autres tiers, la surdité est non syndromique. Pour ces dernières, il a été identifié et localisé 60 gènes entrainant une surdité isolée (VanCamp et Smith, 2001)[*]. Par ailleurs, il a été mis en évidence la large prédominance du gène « Connexine 26 » responsable de la moitié des surdités non syndromiques congénitales (Denoyelle, 2001[73] ; Morzaria et col. 2004[74]).

Rééducation Orthophonique et prothétique de la surdité

En France, 6 millions des personnes environ connaissent des difficultés d'audition. Parmi ces personnes, seulement 15% portent des aides auditives alors que dans d'autres pays comme les Etats-Unis (45%) et le Danemark (50%), les pourcentages de la population appareillée sont plus élevés. Comment expliquer ce retard? Peut-être suffit-il de rappeler qu'en France, 1 Français sur 2 ne fait jamais contrôler son audition. Une situation paradoxale alors qu'il existe désormais de nombreuses aides auditives très performantes.

La rééducation orthophonique et prothétique dépend de multiples facteurs notamment de la nature de la surdité et de sa sévérité précédemment évoquées. Le profil auditif varie d'un sujet à l'autre. Chaque défaillance auditive nécessite une démarche personnalisée et une attention particulière. En effet, le choix d'une aide auditive dépend des capacités à entendre mais également du mode de vie du patient malentendant. D'une manière générale, le but de la rééducation orthophonique d'un enfant sourd est de lui donner la possibilité de communiquer, de développer une parole intelligible et de trouver « sa voix ». Cette dernière lui permettra d'affirmer son identité et d'occuper une place au sein de la société dans laquelle il existe (Van Lem et Peault, 1984[67]). Le port de la prothèse et la rééducation auditive précoce facilitent à l'enfant son intégration dans la communauté entendante dont il fait partie.

[*] http://www.eurohear.org/pdf/EuroHear_Brochure_fr.pdf, http://webhost.ua.ac.be/hhh/

L'appareillage d'un enfant déficient auditif tente de pallier aux problèmes de déficit auditif et de rétablir le « feed-back ». Selon les résultats obtenus des divers examens audiologiques, une prothèse auditive lui est proposée afin de « récupérer » la perception du modèle sonore entendant. Le bilan audiologique qui, par la suite, aboutit à un diagnostic, doit être réalisé de façon précoce. Il peut être établi dans les premiers mois de la vie de l'enfant. L'appareillage doit être proposé sans délai par une équipe multidisciplinaire (médecins ORL, audiologistes, audioprothésistes, orthophonistes ou logopèdes, pédagogues et psychologues), laquelle devra avoir acquis par la formation et l'expérience une compétence particulière dans ce domaine. Cette équipe aide à choisir la solution la mieux adaptée aux besoins de l'enfant.
L'adaptation prothétique est en principe binaurale et stéréophonique, sauf contre-indication clinique. Ceci car très fréquemment la surdité concerne les deux oreilles et la correction d'une seule oreille ne permet pas une localisation sonore spatiale correcte. De plus, ce qui est plus grave encore, la compréhension de la parole ne sera pas satisfaisante dans les conversations en groupe et dans le bruit. L'intérêt de l'appareillage stéréophonique est qu'il restitue à l'enfant malentendant le relief sonore et lui procure un confort d'audition de meilleure qualité.

Quant à la rééducation prothétique des surdités de transmission, elles sont principalement rééduquées avec des aides auditives amplificatrices. Celles-ci fourniront le son amplifié soit par voie aérienne dans le conduit auditif externe (appareils intra CAE ou contour placé derrière et sur le pavillon), soit par voie osseuse, par intermédiaire d'un vibrateur placé sur ou à travers la peau (BAHA) dans la région mastoïdienne ou sur l'écaille temporale.

Actuellement, il existe plusieurs types d'aides auditives amplificatrices disponibles visant à faciliter la communication des enfants et des adultes atteints de déficience auditive. Parmi ces aides, on trouve les « appareils de correction auditive » qui sont des dispositifs actifs[*]. Ces derniers sont différents des systèmes ou dispositifs amplificateurs usuels de commerce. Les appareils de correction auditive sont adaptés aux caractéristiques individuelles anatomiques, physiologiques, psycho-acoustiques, pathologiques de l'oreille et de l'audition résiduelle du patient. Chez l'enfant, l'appareillage devient difficile car une évolution des caractéristiques anatomiques s'accomplit pendant cette période. L'adaptation prothétique doit tenir compte de l'âge et des capacités perceptives de l'enfant. Un type de prothèse auditive classiquement utilisé par l'enfant est un appareil électronique qui fonctionne avec des piles miniaturisées. L'amplification du son n'est pas linéaire. Un réglage précis par l'audioprothésiste permet

[*] Dispositif dépendant pour son fonctionnement d'une source d'énergie électrique ou de toute autre source d'énergie que celle générée directement par le corps humain ou la pesanteur (définition extraite de la directive européenne n° 90/385/CEE, J.O.C.E. 20/07/90 In Recommandations du BIAP).

d’ajuster l’amplification à la courbe auditive de l’enfant. Ainsi, certaines fréquences sont privilégiées. D’autres ont une amplification limitée par un système de compression du son, pour éviter les traumatismes auditifs. Cette prothèse permet une audition stéréophonique et comporte :

- un microphone,
- un amplificateur,
- un écouteur/un vibrateur.

Le principe de fonctionnement est le suivant : le signal acoustique est capté par un microphone qui convertit le signal acoustique en signal électrique. Ce dernier est modifié par un amplificateur et adapté selon le réglage précédemment établi par l’audioprothésiste. Le signal électrique est à nouveau converti en signal acoustique et transmis par un écouteur. Celui-ci amplifie les variations de pressions du signal acoustique en les répercutant sur le tympan grâce à un embout auriculaire placé dans le conduit auditif externe. Dans le cas d’une prothèse type vibrateur, le signal acoustique est transmis directement à l’oreille interne par voie osseuse. Il est utilisé dans les cas où l’embout serait contre-indiqué ou si la voie aérienne serait défaillante. Ce genre d’aide auditive est placé sur la mastoïde.

Aujourd'hui l'apport de l'informatique permet d'effectuer des réglages très fins et de personnaliser réellement la correction auditive. La technologie numérique offre désormais des possibilités aussi nombreuses que performantes. En effet, les progrès réalisés ces dernières années dans le domaine de la correction auditive ont été considérables. Les aides auditives actuelles conjuguent haute qualité sonore, extrême discrétion et facilité d'utilisation. Tous les éléments peuvent être placés dans un seul compartiment. Grâce à l'intégration d'une minuscule puce électronique, les aides auditives numériques actuelles traitent les sons de façon beaucoup plus fine et rapide que les appareils analogiques. La pureté des sons est comparable à celle du disque laser. Programmés par ordinateur, ces appareils assurent une qualité d'écoute et un bien-être auditif qui s’approchent de ceux de l’audition humaine.

Il existe plusieurs modèles de prothèses auditives classiques. Certains modèles sont portés derrière chaque oreille et reliées à un embout (Figure 24). Ces modèles de prothèse sont appelées « prothèses de contour d’oreille » et sont les mieux adaptées à l’enfant. Au fil du temps, les contours d'oreilles ont gagné en discrétion notamment avec l’arrivée récente des appareils à adaptation ouverte. De plus, les modèles actuels intègrent des processeurs de dernière génération extrêmement puissants.

Un autre modèle de prothèse est celle qui est placée directement dans le conduit auditif externe appelé « prothèse intra-auriculaire » (Figure 24). Cette prothèse est contre-indiquée lorsque le conduit est malformé ou trop petit (par exemple chez l'enfant) ou la surdité est sévère ou profonde car la prothèse n'est pas assez puissante. En revanche, les intra auriculaires sont aujourd'hui des modèles de prothèses miniaturisées qui offrent une très grande discrétion. Ils existent sous différentes formes : l'intra-conque qui se place dans la coque de l'oreille, l'intra-conduit (Figure 25) qui se glisse à l'entrée du conduit de l'oreille et le mini intra-conduit qui se loge dans le canal auditif. La prothèse intra-conduit contient le microphone, le bouton programme et le compartiment de piles sur la face externe.

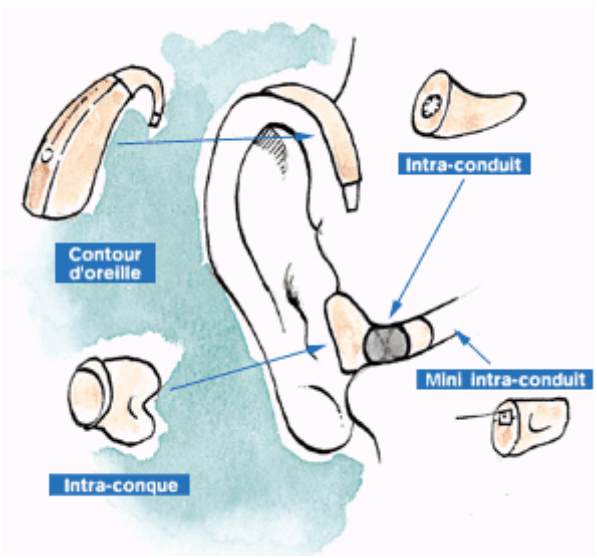


Figure 24 : Prothèses auditives classiques.
(d'après Audition Conseil)[*].

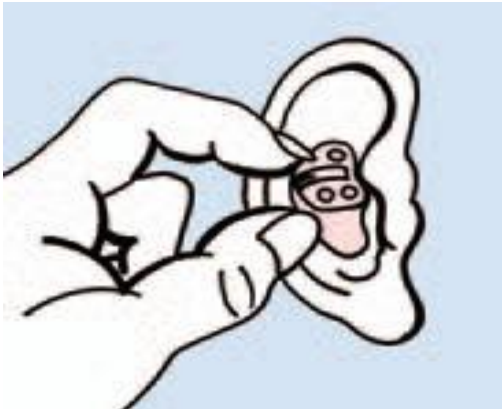


Figure 25 : Prothèse auditive classique intra-conduit.
(d'après Siemens)[**].

[*] Source de l'image : http://auditionconseil.fr/page_appareillage.htm
[**] Source de l'image : http://www.siemens-audiologie.fr

Un dernier modèle de prothèse classique, qui tend à disparaître, est celui composé d'un *boîtier portatif* où sont placés le microphone et l'amplificateur (ressemblant à un walkman). Un fil relie le boîtier à l'écouteur fixé à l'embout.

Quant à la rééducation prothétique des surdités de perception légères et moyennes, elles peuvent être effectuées grâce aux aides auditives conventionnelles précédemment citées. Dans ce cas, les aides auditives conventionnelles compenseront seulement la perte en quantité de l'audition. Il sera nécessaire à ce moment là d'avoir un quota minimum en cellules ciliées pour garder l'analyse en fréquence, intensité et durée des sons. De cette manière, la qualité de l'audition pourra être préservée avec ce type de prothèses. En effet, la perte auditive des surdités de perception est corrélée à la perte de cellules ciliées et des fibres nerveuses cochléaires. Cela entraine une perte de la quantité (seuil auditif) mais également de la qualité de traitement de l'information acoustique.

Lorsqu'il s'agit de surdités profondes, la perte de cellules sensitives est très importante parfois complète. Dans ce cas là, les aides auditives classiques ne suffiront pas pour fournir une information acoustique complète en quantité mais aussi en qualité au malentendant. Le patient va au mieux percevoir des sons mais ne va pas les discriminer : il entend qu'on parle mais il ne comprend rien. Tout est perçu comme du bruit. C'est ainsi que les sourds profonds qui n'ont plus de discrimination efficace de la parole avec des prothèses auditives classiques, sont des excellents candidats pour bénéficier d'un autre type de prothèse auditive : l'implant cochléaire.

5. L'implant cochléaire

L'implant cochléaire est une aide auditive électrophysiologique sophistiquée qui a été assez développée depuis 1980. Cette aide auditive comporte une partie externe mobile constituée d'un microphone, un microprocesseur et une antenne ; et une partie interne fixe constituée d'un récepteur stimulateur, d'un porte-électrodes et de plusieurs électrodes* (Figure 26). La partie externe est contenue dans un seul compartiment semblable au contour d'oreille d'une prothèse conventionnelle. La partie interne est placée chirurgicalement dans la cochlée.

* Le nombre d'électrodes varie selon le type d'implant.

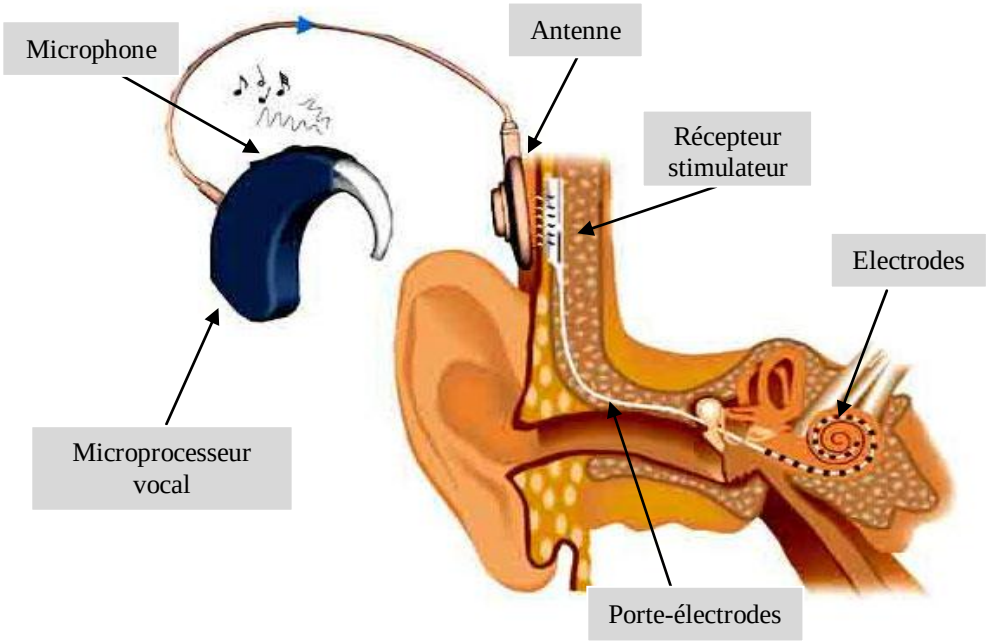


Figure 26 : Composants de l'implant cochléaire[*]

Le principe de fonctionnement de cette prothèse auditive est basé sur la transduction des signaux acoustiques en signaux électriques. Le microphone contenu dans la partie externe de l'implant capte et recueille les signaux acoustiques et les transmet au microprocesseur lequel les convertit en signaux électriques. Le microprocesseur vocal extrait des informations du signal acoustique entrant, les codifie en paramètres de stimulation électrique. Les informations codées sont transmises à la partie intérieure de l'implant par l'antenne. Cette dernière a la forme d'un petit anneau comprenant un aimant dans sa partie centrale. Elle reçoit l'information électrique et la transmet par voie transcutanée à la partie interne de l'implant. L'antenne est maintenue sur le cuir chevelu par attraction avec un aimant localisé dans le récepteur stimulateur au niveau de l'os temporal. Le récepteur stimulateur est constitué d'un ensemble hermétique en titane qui contient une bobine réceptrice et l'aimant. Dans un premier temps, la bobine interne reçoit les signaux électriques à travers la peau. Cette transmission est de type électromagnétique. Dans un deuxième temps, le récepteur stimulateur distribue ces impulsions électriques aux minuscules électrodes acheminés jusqu'à la cochlée par un porte-électrodes. Les électrodes sont connectées au récepteur stimulateur de façon indépendante et placées suivant le tour de spiral de la cochlée (2 tours et demi, cf. Figure 27). C'est ainsi qu'il est possible de délivrer une stimulation aux cellules ciliées et créer ainsi des sensations

[*] Source de l'image : http://pagesperso-orange.fr/anso/images/implant.gif

auditives. Chaque électrode intra-cochléaire possède une bande fréquentielle spécifique, par exemple de 350 Hz à 500 Hz. Les électrodes sont placées en respectant la tonotopie de la cochlée pour se rapprocher du fonctionnement normal de l'oreille humaine. De cette manière, l'électrode basale transmettant les sons aigus sera appelée « électrode 1 » et ainsi de suite jusqu'à l'électrode apicale qui transmettra les sons graves.

Figure 27 : Le porte-électrodes de l'implant dans la cochlée[*].

Les divers systèmes d'implants cochléaires se différencient par : le traitement du signal, le nombre d'électrodes et la position des électrodes.

Le traitement du signal

Selon la façon dont le signal acoustique entrant est transformé par le microprocesseur en impulsions électriques, le traitement peut être :

a. Analogique, lorsque la totalité du message sonore est transmise. C'est le cortex auditif qui effectue un tri de l'information. Les implants cochléaires existant dans le marché utilisant ce type de traitement sont : Clarion (premières générations), Ineraid, Frachet, Monosonic, Medel.

[*] Source de l'image : http://www.bcm.edu/oto/jsolab/cochlear_implants/cochlear_implant.htm

b. Digital, lorsque la compréhension de la parole est privilégiée. Le processeur extrait l'information des bandes de fréquences de 80 à 8000 Hz. Chaque fréquence correspond à une électrode. Les implants cochléaires utilisant ce type de traitement du signal sont : Clarion, Nucleus, Digisonic, Combi-40, Laura, Minimed.

Le nombre d'électrodes

Selon le nombre d'électrodes placées dans la cochlée, l'implant peut être :

a. Mono-électrode, lorsque l'électrode posée stimule le nerf auditif d'une manière globale, comme par exemple l'implant cochléaire Monosonic

b. Multi-électrodes, lorsque plusieurs électrodes stimulent les différentes fibres du nerf auditif. La stimulation des zones respecte au maximum le fonctionnement naturel de la cochlée donnant une sensation de hauteur tonale plus fine et précise. Ainsi, pour une fréquence aigüe, les électrodes stimulent les fibres basales et pour une fréquence grave, les électrodes stimulent les fibres apicales. Selon le modèle, le nombre d'électrodes varie, par exemple Nucleus en possède 22 ou 24, Clarion en a 16.

La position des électrodes

Les nouvelles générations d'électrodes surpassent la première génération en terme de profondeur d'insertion et de proximité du modiolus comme les électrodes périmodiolaires HIFocus (Advanced Bionics) et Contour (Cochlear). En effet, ces nouvelles électrodes permettent une stimulation plus précise fournissant des excellents scores de reconnaissance vocale. L'insertion de ces électrodes peut être profonde (plus de deux tours cochléaires) comme pour l'implant MED-EL, partielle (plus de 1,5 tours cochléaires) comme pour les l'implant Cochlear et Advanced Bionics, ou électroacoustique.

Bilan pré-implantation

Comme nous l'avons déjà expliqué, l'implantation cochléaire est envisagée pour les patients manifestant une surdité neurosensorielle profonde. Dans le cas de l'enfant, cette prothèse s'adresse aux enfants manifestant de surdités congénitales et acquises profondes et aux surdités sévères dont l'intelligibilité de la parole est inférieure à 50% avec les prothèses auditives classiques comme proposé par l'HAS[*]. L'implantation doit survenir dans les délais les plus rapides possibles par rapport à la date de survenue de la surdité. Une équipe pluridisciplinaire

[*] Haute Autorité de Santé : http://www.has-sante.fr

doit procéder à la réalisation de différents bilans préopératoires afin de déterminer si l'enfant répond aux critères de sélection. Ce bilan pré-implantation consiste à effectuer :

Un bilan audiologique et médical afin de déterminer les antécédents de la surdité, d'effectuer un diagnostic précis de la surdité et de l'état fonctionnel de la cochlée et de vérifier le gain prothétique de la prothèse portée à ce moment-là. Ce bilan comprend une anamnèse de la surdité de l'enfant, un examen otologique, des tests audiologiques subjectifs (audiométrie vocale, tonale), des tests objectifs (électro-cochléographie, potentiels évoqués auditifs).

Un bilan radiologique afin de vérifier l'intégrité des structures de l'oreille interne et éliminer ainsi des contre-indications pour l'implantation (Tomodensitométrie, I.R.M.).

Un bilan orthophonique afin de déterminer le niveau de compréhension et de production orale du langage ainsi que les aptitudes non verbales de l'enfant et les moyens de communication utilisés spontanément par lui et sa famille. L'enfant doit montrer avoir envie de communiquer oralement.

Un bilan psychologique afin de repérer les difficultés psychologiques de l'enfant et sa famille concernant les attentes de l'implantation et la motivation car le travail de rééducation après l'implantation est long.

Après tous ces bilans, la décision d'implantation cochléaire est prononcée par l'ensemble de l'équipe du centre d'implantation, l'équipe qui prend en charge habituellement l'enfant et les parents. Ces derniers sont avertis des avantages de l'implant mais également de ses limites et des complications éventuelles (rejet, allergie, panne, processus d'adaptation, entre autres).

Suivi post-implantation

Une fois l'implantation réalisée, l'enfant et sa famille sont accompagnés par l'équipe de rééducation pendant plusieurs années. Le suivi post-implantation diffère d'un centre à l'autre. A l'Institut Saint-Pierre de Palavas-Les-Flots (Hérault) où nous avons recruté les enfants de notre étude, le suivi post-implantation se déroule comme suit :

Un mois après l'intervention chirurgicale, la mise en fonction de l'implant s'effectue en trois jours dans le centre d'implantation. Ces séances de réglages sont appelées « séjour ». La première année, le séjour s'effectue trois jours par mois. La deuxième année, le séjour est de

trois jours tous les deux ou trois mois, selon l'évolution. Puis le séjour est de trois jours tous les six mois.

Lors du premier séjour, le régleur programme le microprocesseur qui est encore vierge. Il charge la mémoire à l'aide d'un ordinateur. Pour ce faire, il recherche le seuil minimum de stimulation électrique entraînant une sensation auditive pour chaque électrode. L'opération est répétée pour établir le seuil maximum (ou de « confort »). Le régleur procède ensuite à la recherche de l'équilibre des électrodes entre elles car toutes doivent provoquer les mêmes niveaux de sensations. Cette recherche établit un écart entre les seuils minimum et maximum qui correspond à la « dynamique », laquelle est variable selon les individus.

Pendant ces séjours, le régleur, en collaboration avec l'orthophoniste, fait deux réglages afin de faciliter la meilleure perception possible car les seuils auditifs évoluent. L'enfant suit une rééducation orthophonique intensive concentrée sur l'éducation auditive, afin de développer l'écoute en donnant du sens aux nouvelles sensations sonores, développer la perception de la parole et développer la communication à travers des échanges *orthophoniste/parents/enfants.* L'enfant participe également à des séances collectives avec une éducatrice spécialisée dans le but d'analyser de manière plus fine l'utilisation de l'implant en contexte communicatif avec d'autres enfants. Les parents et l'enfant rencontrent d'autres enfants implantés, leurs parents et l'équipe rééducatrice au cours de laquelle sont abordées les questions relatives à l'implant.

Le suivi post-implant permet de mesurer les progrès des perceptions sonores de l'enfant et de modifier les réglages afin d'améliorer la discrimination de la parole et par-la même, une meilleure communication orale.

L'objectif de l'implant cochléaire est de restaurer un canal auditif utile permettant l'entrée d'informations acoustiques suffisamment pour la perception, la production et la compréhension du langage oral. Grâce à l'implant cochléaire, l'enfant sourd découvre l'environnement sonore qui l'entoure et s'y situe. Il peut percevoir les bruits et donc anticiper ce qui va arriver, cela apporte un apaisement dans le comportement de l'enfant. Néanmoins, il faut rappeler que l'implant cochléaire reste avant tout une prothèse et bien qu'elle restitue en grande partie la fonction auditive, si l'on débranche la partie extérieure, le porteur redevient sourd.

Dans des conditions optimales, les résultats de l'implantation cochléaires sont encourageants car les productions orales des enfants qui en bénéficient semblent évoluer vers celles des entendants, tout en gardant un décalage. Bien que l'implant cochléaire constitue la meilleure solution technique disponible pour améliorer l'audition en fournissant une entrée qui augmente considérablement l'accès auditif au langage oral, celui-ci n'offre pas un signal aussi riche et complet qu'une audition normale (Le Normand, Ouellet et Cohen, 2001[75] ; Geers, 2006). Ce type de prothèse comme les autres rencontre des limites dans la transmission des caractéristiques acoustiques de la parole et de la voix en présence de bruit environnemental. Une solution actuellement proposée à ces limites est l'implantation bilatérale (Litovsky et col., 2004)[76].

Le port d'une prothèse auditive a pour but de récupérer le « feed-back auditif » indispensable pour la communication orale. Les prothèses auditives ont pour fonction de tirer le meilleur parti possible des restes auditifs, et la transformation des signaux acoustiques ne s'opère jamais sans distorsion (Lepot-Froment et Clerebaut, 1996, p. 29). Lorsque leur usage est soutenu par une « éducation auditive », la sensibilisation à l'environnement acoustique paraît avoir une influence favorable sur les possibilités ultérieures d'utilisation d'une audition résiduelle précédemment non fonctionnelle (Ito, 2004)[77]. Selon le degré de déficience auditive de l'enfant et les fréquences que la prothèse lui permet de récupérer, la perception du modèle sonore sera différente de celle de l'enfant entendant et, par là même, conséquemment, la production de la parole.

CHAPITRE II : PROBLEMATIQUE

1. Etat de l'art de la recherche dans le domaine de l'oralité de l'enfant sourd

La surdité est le handicap sensoriel le plus fréquent. Un enfant sur mille est atteint à la naissance d'une surdité sévère ou profonde (Lina-Granade et col., 2000[64] ; Garabedian et col., 2003[78]). La surdité ou déficience auditive[*] entraîne de multiples difficultés au niveau linguistique et communicatif. Des études ont montré un décalage dans le temps et dans les délais d'acquisition des systèmes linguistiques chez l'enfant sourd même étant implanté précocement, par rapport aux systèmes linguistiques décrits chez l'entendant (Cochard et col. 1998)[79]. Il en va de même pour l'acquisition et le développement des aspects suprasegmentaux de la parole qui contribue à une diminution de l'intelligibilité (Boothroyd et Decker, 1972 *In Gilbert et Campbell, 1980*)[80]. L'enfant sourd n'a pas accès aux éléments prosodiques dans la période normale d'acquisition du langage, contrairement à l'enfant entendant qui les perçoit et les utilise dès les premiers mois de sa vie. L'enfant déficient auditif ne perçoit pas les traits distinctifs propres à chaque phonème, il ne peut donc pas les reproduire. En raison de l'absence d'un retour auditif, l'enfant diminue ses productions tant qualitativement que quantitativement. Parallèlement à cette diminution d'émissions, son intérêt pour la parole et le monde sonore décroit car il n'a pas de plaisir à interagir avec son interlocuteur.

Lorsque l'apparition de la surdité se manifeste à la naissance ou de manière précoce, avant trois ans, l'acquisition et le développement du langage et des fonctions auditives se trouvent fortement altérés. Il y a un nombre réduit de phonèmes, des erreurs articulatoires et un manque de maîtrise de l'intensité, du timbre et des qualités acoustiques des émissions vocales.
Le mécanisme auditif joue donc un rôle capital sur le contrôle et la production de la parole et de la voix (Yates, 1963 *In Hocevar-Boltezar et col., 2006*)[81]. Certains auteurs affirment plus précisément que le canal auditif est le canal sensitif privilégié pour le traitement et le contrôle de la fréquence fondamentale de la voix et de la parole (Mallard et col., 1987 ; Angelocci et col., 1964 ; Nickerson, 1975 cités en *Hocevar-Boltezar et col., 2006*[81]*;* Higgins et col., 2003[82]).

[*] Le BIAP (Bureau International d'Audiophonologie) recommande d'utiliser le terme de « *déficience auditive* » plutôt que de surdité. Ceci en raison du caractère trop absolu du mot « surdité ».

La littérature concernant la caractérisation des productions vocales des enfants déficients auditifs est abondante et s'intègre à différents axes de recherche largement développés depuis les années soixante.

Les premiers travaux ont été réalisés en langue anglaise. Ces études ont notamment mis l'accent sur le rôle de la *physiologie et de la maturation du mécanisme de la phonation dans l'émergence des vocalisations*, et sur le rôle important des *informations acoustiques provenant de l'environnement* (Lenneberg et col., 1965, 1967 ; Mavilya, 1972 ; Oller, 1980 ; Maskarinec et col., 1981 ; Stark, 1980, 1983 ; Oller et col. 1985 ; Lynch et col., 1989 ; Oller et Eilers, 1988, 1992 cités en *Lepot-Froment et Clerebaut, 1996*, p. 37-43)[83]. Des divergences conceptuelles et méthodologiques à propos des paramètres retenus pour caractériser les productions sonores ont abouti et donné lieu à des résultats contradictoires, partiels, dans certains cas, contestables. Dans les pays francophones, *« des conclusions abusives ont été tirées de l'un ou l'autre de ces travaux »* (Lepot-Froment et Clerebaut, 1996, p. 38[83]). Toutefois, le consensus partagé par l'ensemble de la communauté scientifique tend à poser que l'enfant sourd vocalise indépendamment de l'importance de sa déficience auditive.

D'autres travaux ont porté sur l'aspect *phonologique,* et plus précisément sur *l'articulation et l'intelligibilité* de la parole de l'enfant sourd. Ces recherches ont démontré, par exemple, les différences quantitatives et qualitatives ainsi que les retards dans l'acquisition du répertoire consonantique des enfants à déficience auditive moyenne, sévère et profonde de 4 à 18 mois et de 4 à 30 mois (Stoel-Gammon, 1983[84] in *Hochberg et col. 1988*, 1988[85], 1991[86] ; Stoel-Gammon et Otomo, 1986[87] ; Higgins, 2001[88] ; Seifert et col., 2002[89] ; Higgins et col., 2003[82]). Les travaux de Stoel-Gammon et Otomo (1986) révèlent que les caractéristiques déviantes du babillage et de la production consonantique chez les enfants sourds ne doivent pas être attribuées uniquement à la perte d'informations acoustiques en provenance de l'environnement mais, également, à la perte de « feed-back » auditif. Le feedback auditif est un mécanisme de rétroaction permettant le retour auditif de l'information acoustique produite lors de l'émission vocale. Cette perte de feed-back entraîne une défaillance (dans une mesure inconnue, selon l'auteur) du contrôle des mouvements articulatoires et affecte considérablement la motivation à produire des sons « pour le plaisir » (Lepot-Froment et Clerebaut, 1996[83]).

D'autres études se sont penchées sur l'analyse des *éléments suprasegmentaux ou prosodiques* des émissions vocales de l'enfant sourd. Ces études ont permis d'évoquer et d'introduire

quelques paramètres acoustiques par exemple, l'intensité et la hauteur, intervenant dans la phonation et qui sont en rapport avec les signaux sonores perçus de la langue parlée dans une communauté (traits prosodiques). Certains chercheurs tels que Mounier-Kuhn et col. (1967)[90] comparent la fréquence fondamentale moyenne des enfants entendants et des enfants sourds de 2 à 4 ans. Cette fréquence apparait comparable dans les 2 groupes mais avec des contours mélodiques perturbés dans le groupe d'enfants sourds. Manolson (1972)[91] met en évidence les altérations du contour intonatif représentées par des variations brusques de la hauteur et de l'intensité des enfants sourds de 12 à 24 mois. La privation momentanée d'information acoustique induit des perturbations de l'espace tonal utilisé (Vinter, 1989)[92]. Or, chez des enfants sourds appareillés plus âgés, la privation momentanée de la prothèse n'affecte plus la qualité des émissions vocales chez ceux d'entre eux qui manifestent une parole intelligible (Harmegnies, 1985 ; cités en *Lepot-Froment et Clerebaut, 1996*[83]). Toutes ces études ont analysé les variations de quelques paramètres acoustiques et de leurs effets sur la perception de la parole de l'enfant sourd dans un contexte linguistique. Mis à part le langage avec tout ce qu'il comporte de syntaxe, de prosodie et d'implicite, les paramètres physiques et physiologiques caractérisant la phonation de l'enfant sourd n'ont pas été exploités en recherche de façon exhaustive. En effet, un nombre moins important de travaux ont tenté de caractériser physiquement et/ou physiologiquement la voix de l'enfant sourd sans circonscrire les paramètres comme des éléments prosodiques (Higgins et col., 2003[82] ; Seifert et col., 2002[89] ; Campisi, 2005[93]).

Un autre axe de recherche assez développé relève de l'étude de *l'apport des appareils de correction auditive*. Les avancées technologiques en ce domaine sont considérables proposant différents types de prothèses auditives à caractéristiques particulières. Parmi les premiers travaux entrepris en France pour étudier les effets bénéfiques de l'appareillage prothétique et de l'entraînement auditif sur la parole et la voix, se trouvent les travaux de Périer (1987)[94], Vinter (1987[95], 1989) et Bizaguet (1992)[96]. L'étude de Kent et col. (1987)[97] a mis en évidence les particularités du comportement vocal de l'enfant déficient auditif profond du groupe I, diagnostiqué à l'âge de trois mois, appareillé et pris en charge au moment du diagnostic. Ce travail révèle le déplacement de l'espace tonal dans les hautes fréquences avec des variations brusques de la fréquence fondamentale moyenne.

Les différents types de prothèses proposés pour rétablir la fonction auditive - acoustiques, vibro-tactiles ou électrophysiologiques - ont beaucoup évolué ces dernières décennies. Plus

particulièrement, l'implant cochléaire*, cette prothèse électrophysiologique a été développée depuis les années 70 (House, 1976 in *Campisi, 2005*[93]). Grâce aux premiers résultats satisfaisants obtenus par l'implantation cochléaire chez l'enfant sourd (Portmann, 1992[98] ; Dejean et col., 1992[99]), les avancées technologiques continuent à proposer des prothèses plus performantes et plus adaptées aux besoins communicatifs de l'enfant. Les résultats positifs à court, moyen et long termes ont encouragé de nouvelles recherches afin d'envisager l'implantation le plus précocement possible (Heath, 2001[100] ; James et col., 2004[101] ; Spencer et col., 2004[102] ; Manrique et col., 2004[103] ; Lee et col., 2005[104] ; Lesinski-Schiedat et col., 2004[105], 2005[106] ; Haensel et col., 2005[107] ; Colletti et col., 2005[108] ; Harrison et col., 2005[109]).

Initialement, le recours aux implants cochléaires était envisagé dans les cas de surdité totale chez les adultes dans le but de rétablir une fonction auditive d'alerte. Actuellement, la prescription thérapeutique de cette technique s'élargit aux enfants. On en fait bénéficier les enfants atteints de façon congénitale ou précoce de déficience auditive profonde et depuis peu, d'une surdité sévère. De plus en plus, un grand nombre d'enfants sourds bénéficient de l'implant cochléaire. Le port de l'implant permet aux enfants d'avoir, au bout de quatre ans et du point de vue phonologique, des productions comparables à celles d'enfants entendants (même si quelques phonèmes demeurent imprécis)**. D'importants progrès se mettent en place vers trois ans d'implantation (Flipsen et Colvard, 2006)[110]. Quant à la voix, des améliorations ont été démontrées après six mois d'implantation chez l'enfant implanté avant l'âge de quatre ans (Hocevar-Boltezar et col., 2005[81]).

D'autres axes de recherche développés dans le domaine de l'implantation cochléaire concernent :

- l'amélioration de la biocompatibilité des matériaux de l'implant et les effets d'une stimulation électrique efficace (Heid et col., 1998[111]),
- la miniaturisation des systèmes, la numérisation et le traitement du signal (Summerfield, 1996[112]),
- les aspects éthiques (Lane et Grodin, 1997[113]) et économiques de l'implantation (Barton et col., 2004[114] ; Sach et col., 2005[115]),
- les effets de l'implant sur la fonction mentale de l'enfant (Dumont, 1994[116] ; Yang et col., 2004[117]).

---

* Nous présentons une description de cette aide auditive et les principes de fonctionnement dans le chapitre I.
** Nous ne rentrerons pas dans les détails de la description des différences phonologiques et les facteurs associés à ces différences car ce n'est pas l'objet d'étude dans ce travail.

2. Position de la problématique

A propos de l'étude de la boucle audio-phonatoire

Les recherches menées sur le contrôle audio-phonatoire débutant de l'époque de Lombard (1911)[55] jusqu'à nos jours (Hanley et Steer, 1949[118] ; Draegert, 1951,[119] ; Dreher et O'Neill, 1957[120] ; Ladefoged, 1967,[121] ; Summers et col., 1988,[122]), manifestent un grand intérêt de la plupart des auteurs à l'étude de l'action de la rétroaction auditive dans la production de parole et ce, dans un environnement bruyant. Ces études se basent sur le postulat que « l'effet Lombard » est un réflexe physiologique fortement présent dans la régulation de la voix et de la parole (Hage et col., 2006[123])[*]. Il semblerait que l'effet Lombard implique l'action des aires corticales supérieures (Lane et Tranel, 1971[124] ; Jürgens, 2009[125]). Cet effet met en évidence le contrôle de l'émission vocale par le mécanisme auditif et se manifeste en altérant principalement l'intensité et la hauteur du son pour compenser la présence de bruits environnants. Le locuteur adapte automatiquement son intensité de voix quand le niveau de bruit environnemental augmente et quand le niveau de sa propre voix diminue (le « sidetone », selon les anglo-saxons) pour que son ou ses interlocuteurs puissent mieux le comprendre, malgré le bruit. L'hypothèse de l'effet lombard suppose ainsi que le locuteur tente d'entretenir un rapport parole-bruit favorable à la communication.

Dans cette ligne de recherche expérimentale, une étude menée par Levelt (1989)[126] propose un modèle de production de la parole comportant deux boucles de contrôle audio-phonatoire, une boucle interne et une boucle externe. Ces boucles contrôleraient la performance du langage parlé et traduiraient l'existence de connexions entre le mécanisme de production et le mécanisme de perception de la parole. La boucle interne permettrait au locuteur de « vérifier » que ce qu'il produit correspond à ses intentions linguistiques. La boucle externe lui permettrait de déterminer que ce qu'il a déjà produit l'a été correctement. Si une erreur avait lieu (par exemple, une faute de prononciation), le locuteur pourrait rectifier sa propre production en corrigeant son énoncé. Cette fonction d'autocorrection est indispensable pour assurer une communication efficace. Lackner et Tuller (1979)[127] avaient déjà envisagé un modèle de production de la parole similaire sur l'existence de ces deux boucles. Leurs travaux consistaient à mettre en évidence la distinction entre la boucle interne et la boucle externe. Ils ont testé la prédiction que la détection d'une erreur serait plus rapide si elle est effectuée via la

[*] Le réflexe est essentiellement cochléo-bulbo-récurrentiel. Ce réflexe serait conditionné par la diffusion d'un influx auditif dans le bulbe jusqu'aux noyaux moteurs de la Xème paire crânienne. En effet, Il a été démontré une participation de la zone péri-olivaire du complexe Olivaire Supérieur et la formation réticulaire du pont cérébelleux selon Hage et col. (2006).

boucle interne que via la boucle externe. Ils concluent que le circuit de détection correspondant à la boucle interne est plus direct que celui correspondant à la boucle externe. Tandis que le processus de contrôle par la boucle interne se fait sur la base des seules informations provenant du système de production vocale, celui assuré par la boucle externe demande les informations articulatoires et auditives. Les points intéressants à retenir du modèle de parole de Levelt et de l'expérience de Lackner et Tuller, sont le postulat de l'existence de deux boucles qui régulent l'émission vocale, la mise en évidence d'une différence de rôle pour chacune et l'intervention précoce de la boucle interne lors du processus de production.

L'action de contrôle par la boucle interne pendant la production peut se produire bien avant l'action d'émission du signal sonore émis. Dans le cas de sujets déficients auditifs, le dysfonctionnement des boucles de contrôle se traduirait par une incapacité à régler la qualité de leur émission vocale, parfois en étant conscients des erreurs produites, parfois en ne l'étant pas.

A propos de l'évaluation de la voix

Actuellement, il n'y a pas d'unanimité concernant les mesures exactes à réaliser pour évaluer les patients dysphoniques à l'aide d'un protocole précis et complet. Néanmoins, en Europe, la Société Européenne de Laryngologie a publié des recommandations ainsi qu'un protocole d'évaluation pour atteindre un accord et établir une méthode standard d'évaluation fonctionnelle. Ce protocole prend en compte les données et les résultats actuels présentés dans la littérature sur la rééducation de la voix, les techniques phono-chirurgicales mises au point, les avancées technologiques ou les nouvelles procédures proposées sur la recherche de la voix pathologique. Le protocole propose d'évaluer dans une perspective multidimensionnelle, un nombre minimal de mesures pouvant être considérées comme les plus pertinentes pour l'analyse des dysphonies les plus courantes. Ce protocole comprend cinq approches d'appréciation différentes des troubles vocaux : 1) l'analyse perceptive de la dysphonie (grade, raucité et souffle) ; 2) l'évaluation vidéostroboscopique (fermeture, régularité, symétrie et ondulation muqueuse) ; 3) l'évaluation des paramètres acoustiques (Jitter, Shimmer, Fo, intensité la plus faible) ; 4) l'évaluation des paramètres aérodynamiques (quotient de phonation) ; 5) l'estimation subjective (Dejonckere et col., 2001 *in Hartl et col., 2005*[140]).

Concernant la voix infantile, aucun protocole standardisé n'a été mis au point jusqu'à présent. Un certain nombre d'auteurs ont envisagé d'étudier la voix infantile pathologique et la voix normale et son évolution avant la période de maturité laryngée (Hacki et Heitmuller, 1999[20] ; Campisi et col, 2005[93] ; Nicollas, 2007[21]). Néanmoins, ces études sont basées sur une méthodologie concentrée notamment sur l'analyse des paramètres acoustiques de la voix. L'évaluation multiparamétrique incluant des mesures aérodynamiques a été très rarement envisagée (Higgins, 2003[82]).

A propos de la voix de l'enfant sourd

Nous constatons que la voix de l'enfant sourd, plus particulièrement, la voix de l'enfant implanté cochléaire a été essentiellement étudiée et caractérisée à l'aide des paramètres acoustiques (Tableau 5) (Szyfter et col, 1996[152] ; Monini et col, 1997[156] ; Perrin et col, 1999[133] ; Seifert et col, 2002[89] ; Campisi et col, 2005[93] ; Hocevar-Boltezar et col, 2005[128], 2006[81] ; Lenden, 2007[148]). Dans le tableau 5, on notera en gris les études utilisant des paramètres aérodynamiques. En effet, les mesures physiologiques n'ont pas encore été exploitées de façon exhaustive. Très peu d'études ont envisagé l'évaluation multiparamétrique de la voix incluant à la fois des paramètres acoustiques et aérodynamiques (Higgins et col. 1996[129], 2003[82] ; Uchanski et Geers, 2003[229]). Ces derniers s'avèrent pertinent pour identifier des anomalies de la production orale qui ne seraient pas identifiées par des mesures perceptives ou acoustiques (Higgins et col., 2003[82]).

Une approche récemment envisagée pour aborder cette problématique est l'analyse des mesures physiologiques instrumentales. Elles peuvent en effet révéler des troubles vocaux chroniques qui ne seraient pas identifiés par des mesures perceptives ou acoustiques chez les enfants implantés. Les mesures physiologiques instrumentales ont rarement été exploitées pour caractériser la voix chez l'enfant implanté cochléaire (Higgins et col., 1996[129], 2000[130], 2001, 2003; Svirsky et col.,1998[131]).

L'étude longitudinale de Higgins et col. (2003) est jusqu'à présent une des références en termes de base de données sur les paramètres acoustiques et aérodynamiques de la voix des enfants implantés. Toutefois, son étude présente quelques limites : le nombre d'enfants implantés n'est pas statistiquement fiable (n=7) ; en outre, la population étudiée a été implantée après l'âge de cinq ans, après l'acquisition et le développement du langage, ce qui ne représente pas des populations implantés précocement, notamment avant la période d'acquisition et de développement du langage. L'étude de Higgins et col. (2003[82]) ainsi que

les travaux d'autres auteurs (cf. Tableau 5) sont descriptifs du au nombre faible d'effectifs. L'ensemble de ces études révèlent la carence d'une méthodologie standardisée pour l'enregistrement et l'évaluation de la voix. De plus, certaines études manquent d'une base de données normative sur la voix infantile permettant d'établir des comparaisons fiables. Par conséquent, les conclusions de ces travaux sont contradictoires, partielles, voire contestables dans certains cas. Elles laissent toujours un nombre important de questions en suspens.

Etude	n	Age (ans)	Apparition de la surdité	Équipement	Échantillon de voix	Paramètres évalués
Szyfter et col, 1996	2	10-12	Prélingual	MDVP	Texte standard	Fo, Jitter, Shimmer.
Monini et col, 1997	3	12-14	Pré (2), Péri (1),	IDEM	Phrases	Fo, intensité.
Perrin et col, 1999	4	9-14	Pré (3), Post (1)	ILS	Texte standard	Intensité, formants.
Hamzavi et col, 2000	3	10-14	Postlingual	XTools	Texte standard	Fo
Higgins et col, 2001	2	6	Prélingual	CSL+ Pneumotachomètre	Syllabes	Voice Onset Time (VOT), F1, F2, débit d'air nasal (DAN), pression intra-orale (PIO).
Seifert et col, 2002	20	3-10	Prélingual	CSL	Phrase	Fo, F1, F2, F3.
Uchanski et Geers, 2003	181	8-9	Prélingual	DAT	Phrases	VOT, F2, analyse spectrale, mesure de la nasalité d'une syllabe ; durée de voyelles, mots et phrases.
Higgins et col, 2003	7	5-10	Prélingual	CSL+ Pneumotachomètre	/pa/, /pi/, /ma/, /pop/, /puppy/, /baby/, /bye-bye baby/	Fo, VOT, DAN, debit d'air oral, PIO, (DAO).
Campisi et col, 2005	21	3-17	Pré (15), Post (6)	CSL	/a/ tenu	Fo, vFo (contrôle à long terme de la fréquence), vAm (contrôle à long terme de l'amplitude), F1, F2, F3, étendue vocale.
Van Lierde et col, 2005	9	5-11	Prélingual	Analyse perceptive + CSL, MDVP	/a/ tenu	GRBAS (Hirano), DSI (Dysphonia Severity Index), Fo, F-high, I-low, Jitter, Shimmer, TMP (Temps Maximal de Phonation).
Hocevar-Boltezar et col, 2006	29	2-13	Prélingual	MDVP	Phrase, /a/ tenu	Fo, Jitter, Shimmer, vFo, vAm, Rapport H/B.
Lenden et Flipsen, 2007	6	4-6	Prélingual	DAT + PVSP	Parole conversationnelle	**Prosodie** (fluidité, debit, accent tonique) + **voix** (intensité, hauteur, qualité laryngée, qualité de resonnace)

ILS = Interactive Laboratory System
IDEM = Sofware package
DAT = Digital Audio Tape
MDVP = Multidimensional Voice Program
CSL = Computerized Speech Lab
PVSP=Prosody-Voice Screening Profile

Tableau 5 : Synopsis de quelques études réalisées sur la voix des enfants implantés cochléaires.

A notre connaissance, en France, la caractérisation de la voix de l'enfant sourd appareillé et implanté cochléaire de 5 à 13 ans n'a pas encore donné lieu à l'établissement d'une base de données expérimentale des principaux paramètres acoustiques et aérodynamiques. Nous constatons donc une carence de données sur cette population ainsi que l'inexistence d'un protocole d'évaluation vocale multiparamétrique adapté aux particularités de la voix de l'enfant à déficience auditive.

Notre travail s'inscrit dans la continuité de ces recherches et, pour pallier ce manque, nous avons entrepris une étude expérimentale comparative de la voix d'enfants de 5 à 13 ans déficients auditifs appareillés avec une prothèse classique (ESA) et implantés cochléaires (ESIC). De plus, nous avons abordé la réflexion autour de la boucle audio-phonatoire par le biais d'une modélisation théorique afin d'éclairer la complexité du contrôle de la voix par le retour auditif. Cela nous permettra ainsi de mieux expliquer la qualité vocale altérée observée de notre population dans notre étude expérimentale.

3. Objectifs de notre contribution

Ce travail a comme principal objectif de caractériser sur le plan acoustique, aérodynamique et perceptif la voix de l'enfant sourd sévère et profond âgé de 5 à 13 ans, ayant été appareillé/implanté avant l'âge de 3 ans et portant une aide auditive depuis 3 ans ou plus. Nous cherchons également à identifier et décrire des caractéristiques vocales anormales chez la population d'enfants sourds (appareillés et implantés). Nous espérons ensuite déterminer si ces caractéristiques vocales anormales relèvent d'une dysphonie* et, le cas échéant, décrire quantitativement cette éventuelle dysphonie. D'autres objectifs spécifiques sont également assignés à notre étude :

- Concevoir un protocole d'évaluation multiparamétrique de la voix, dénommé par la suite « Protocole PEMVox », recueillant le plus fidèlement possible les particularités de la production vocale des enfants sourds (ESA et ESIC) de 5 à 13 ans.

* Selon la définition présentée par Hartl (2005[140]) : « Les dysphonies sont caractérisées par une anomalie de la voix d'origine laryngée, secondaire à une anomalie de la vibration des plis vocaux. La dysphonie se caractérise par une altération de ce que l'on appelle la qualité vocale, c'est-à-dire une altération de l'esthétique subjective de la voix ».

- Comparer sur le plan acoustique et aérodynamique la voix de l'enfant ESA et ESIC avec celle d'un enfant EE.

- Se faire une idée de l'incidence du traitement du signal sonore des aides auditives portées par la population sur la qualité de la voix.

- Identifier les paramètres acoustiques les plus pertinents pour évaluer la voix des enfants sourds.

Cette étude permettra ainsi d'établir les bases d'une évaluation vocale multiparamétrique de l'enfant ESA ou de l'enfant ESIC utile pour une prise en charge orthophonique plus adaptée aux particularités de la voix de ces populations. Ainsi, l'intervention thérapeutique, en charge de la rééducation de la voix, sera plus efficace car elle se concentrera sur les paramètres de la voix les plus altérés.

4. Hypothèses et questions scientifiques

D'une part, nous supposons que les voix des enfants sourds (appareillés avec une prothèse classique et implantés cochléaires) manifestent des caractéristiques vocales « particulières » associées à des troubles vocaux, du fait de l'absence d'une boucle de régulation auditive « naturelle ». D'autre part, nous supposons que les caractéristiques acoustiques de la voix des enfants sourds implantés cochléaires *sans troubles vocaux* sont sensiblement comparables à celles des enfants entendants. Trois questions sont formulées à partir de ces hypothèses :

i) les caractéristiques vocales de l'enfant sourd appareillé et implanté cochléaire sont-elles pathologiques ?

ii) Si oui, ces caractéristiques relèvent-elles d'une dysphonie ?

iii) Si non, de quelle nature physique relèvent ces caractéristiques ?

De ces hypothèses affleurent d'autres questionnements que nous tenterons d'élucider au cours de ce travail et qui sont énoncées ici sous forme d'hypothèses.

H1 : *Les paramètres acoustiques de la voix de l'ESA sont significativement différents de ceux de l'EE.*

H2 : *Les paramètres acoustiques de la voix de l'ESIC sont sensiblement comparables à ceux de l'EE.*

- Le protocole d'évaluation multiparamétrique de la voix PEMVox permet-il d'établir ces différences ?

- Les paramètres acoustiques retenus dans notre protocole d'évaluation sont-ils pertinents pour caractériser la voix de l'ESA et de l'ESIC ?

H3 : *Il existe des caractéristiques vocales anormales aussi bien chez l'enfant sourd appareillé que chez l'enfant implanté cochléaire, qui sont mises en évidence par la mesure de paramètres acoustiques*[*].

- Quels sont les paramètres acoustiques les plus sensibles signalant une tendance à la manifestation de caractéristiques vocales anormales chez l'ESA et l'ESIC ?

H4 : *Les paramètres acoustiques de la voix sont notamment plus perturbés chez l'ESA que chez l'ESIC.*

- Les caractéristiques vocales anormales éventuellement manifestés chez l'ESIC sont-elles différentes de celles manifestées chez l'ESA? Qu'est-ce qui les caractérise au niveau acoustique ?

H5 : *Les caractéristiques vocales anormales observées chez l'ESA et/ou chez l'ESIC relèvent de la dysphonie.*

H6 : *La durée d'appareillage/implantation a une incidence positive (amélioration) sur la pression sous-glottique estimée (PSGE) et le débit d'air (DAB).*

---

[*] Cela du fait que ces enfants sont déficients auditifs et que leur boucle audio-phonatoire est naturellement défaillante.

H7 : *La différence des traitements du signal sonore (stratégies de codage) entre la prothèse classique et l'implant cochléaire peut être mise en évidence par un nombre important de différences significatives existant entre l'ESA et l'EE.*

- Quel est l'effet du traitement du signal de la prothèse auditive classique et de l'implant cochléaire sur les paramètres acoustiques de la voix de l'ESA et de l'ESIC respectivement ?

- Peut-on établir une relation entre la qualité et quantité d'informations acoustiques du signal sonore entrant et sortant dans un modèle systémique en « *boucle de rétroaction négative* » ?

CHAPITRE III : MÉTHODOLOGIE EXPÉRIMENTALE

Pour tenter de répondre aux questions scientifiques préalablement posées (cf. chap. II), nous avons abordé deux approches : une approche expérimentale acoustique et une approche perceptive. L'approche expérimentale formalise les procédés appliquées pour le recueil de données à l'aide d'un protocole d'expérimentation que nous allons présenter ici.

1. Population étudiée

Nous avons étudié la production vocale de 78 enfants (33 filles et 45 garçons) âgés de 5 à 13 ans, soit : 38 enfants entendants ou groupe contrôle (EE), 20 enfants sourds porteurs d'une prothèse classique (ESA) et 20 enfants implantés cochléaires (ESIC).

Nous avons limité l'étude à l'analyse des voix dans cette tranche d'âge pour des raisons méthodologiques car : l'étude s'adresse à des larynx infantiles (moins de 13 ans) et nécessite une certaine coopération de l'enfant (plus de 5 ans). Les difficultés de recrutement ont limité la taille des populations ESA et ESIC à 20 enfants pour chaque groupe. D'une part, certains parents n'autorisent pas leurs enfants à participer à ce genre d'étude car ils sont régulièrement sollicités par d'autres études. D'autre part, il a été constaté au cours des dernières années une diminution considérable du nombre d'enfants sourds porteurs d'une prothèse classique. Une notice d'information et une demande d'autorisation ont été distribuées aux parents préalablement à l'évaluation de la voix de l'enfant (Annexe 1 et 2).

Les enfants du groupe EE ont été sélectionnés parmi la population d'élèves du groupe scolaire *« Jeu de Mail »* de la ville de Montpellier. Les enfants du groupe ESA ont été sélectionnés dans le « Centre d'Education Spécialisée pour les Déficients Auditifs - C.E.S.D.A. », dans le « Centre de Rééducation de l'Ouïe et de la Parole – C.R.O.P Paul Bouvier » et dans le C.A.M.P.S. - Hôpital La Timone de Marseille. Enfin, les enfants du groupe ESIC ont été sélectionnés parmi la population suivie à l'Institut Saint Pierre de Palavas-Les-Flots et au CHU Gui de Chauliac de Montpellier.

Afin de rendre le plus homogène possible les groupes de population de notre étude, nous avons défini des critères de sélection restreints. Concernant le groupe EE, nous avons retenu les enfants ne manifestant pas :

- de signe traduisant des troubles du langage et/ou de la parole et/ou de la déglutition,
- de trouble moteur et/ou intellectuel,
- de pathologie associée d'origine sensorielle et/ou neurologique,
- d'antécédent auditif et/ou d'infection intercurrente*,
- d'antécédent chirurgical vocal et/ou de rééducation orthophonique,
- d'antécédent chirurgical auditif lourd (type tympanoplastie ou otite après l'âge de 5 ans).

Cette sélection s'est effectuée grâce à la collaboration du médecin scolaire et à l'aide d'un questionnaire de dépistage (voir Annexe 3). Concernant le groupe ESA et le groupe ESIC, nous avons retenu les enfants ne manifestant pas :

- de trouble spécifique du langage (par exemple, dysphasie, apraxies bucco-phonatoires) et/ou de la déglutition associés,
- de trouble moteur (par exemple, paraplégie, hémiplégie,) et/ou intellectuel spécifique (par exemple, troubles de la personnalité, trisomie 21),
- de pathologie neurologique (type paralysie de corde vocale) et/ou d'origine sensorielle extra-auditive associée (type handicap visuel associé),
- d'antécédent chirurgical vocal (type nodule, kyste),
- de malformation et/ou d'ossification cochléaire.

Quant aux caractéristiques de la surdité et du port de l'appareillage ou de l'implant :

- surdité bilatérale sévère ou profonde : seuil auditif supérieur à 70 dB sans aide auditive**,
- surdité congénitale ou prélinguale***,
- âge d'appareillage ou d'implantation : moins de 3 ans****,
- durée d'appareillage ou d'implantation : plus de 2 ans*****,
- stratégie privilégiée de communication : oralisme.

---

* Manifestation d'infections des voies respiratoires supérieures, états grippaux à répétition.

** Selon la classification audiométrique des déficiences auditives par le BIAP (Bureau International d'Audiophonologie). Dans le cadre de nos travaux, nous adoptons ces grades de surdité afin d'établir un critère de sélection plus proche de la population d'implantés cochléaires et ainsi homogénéiser les populations concernant le gain prothétique.

*** Congénitale : apparition de la surdité dès la naissance; prélinguale : apparition de la surdité avant l'acquisition du langage (3 ans environ).

**** Afin d'étudier l'effet d'une prise en charge avant l'acquisition du langage.

***** Afin d'étudier l'effet du port d'une aide prothétique à moyen et long terme.

La procédure de sélection a été réalisée avec la collaboration des orthophonistes qui suivaient les enfants dans les différents centres de rééducation précédemment cités. Nous avons élaboré une grille pour cet effet (Annexe 6).

Par ailleurs, nous avons conçu un questionnaire qui a été rempli par les parents pendant l'enregistrement vocal (Annexe 4 et 5). Ceci afin d'appréhender les caractéristiques de l'environnement familial, les activités de loisir, les habitudes vocales des enfants des trois groupes de population ; ainsi que l'intégration scolaire, le mode de communication privilégié et les habitudes de port de la prothèse auditive par les enfants des groupes ESA et ESIC. Les informations concernant les caractéristiques de la surdité et de l'appareillage et implant cochléaire porté par l'enfant ont été recueillies auprès des orthophonistes impliqués dans le suivi des enfants.

1.1. Sexe et âge.

L'ensemble de l'échantillon étudié (n=78) était constitué de trois groupes de profil différent : 38 enfants normo-entendant (EE), 20 enfants sourds appareillés (ESA) et 20 enfants implantés cochléaires (ESIC).

Concernant le groupe EE, l'échantillon était constitué de 18 filles (47%) et 20 garçons (53%) (Figure 28). On observe une répartition plutôt équilibrée du nombre de filles et de garçons dans la population témoin. L'âge des enfants se situe entre 6 et 12 ans. La moyenne d'âge du groupe est 9,9 ans (±1,86). La répartition des âges en fonction du sexe est représentée sur la figure 29.

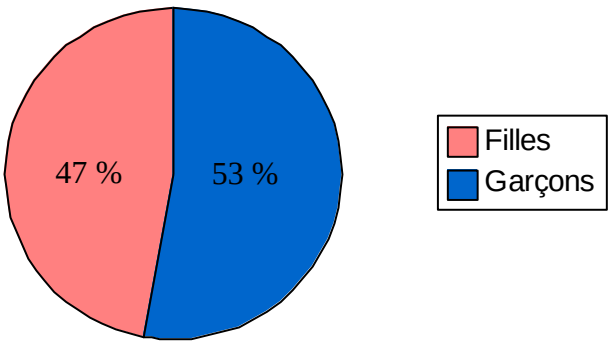


Figure 28 : Répartition du groupe EE en fonction du sexe.

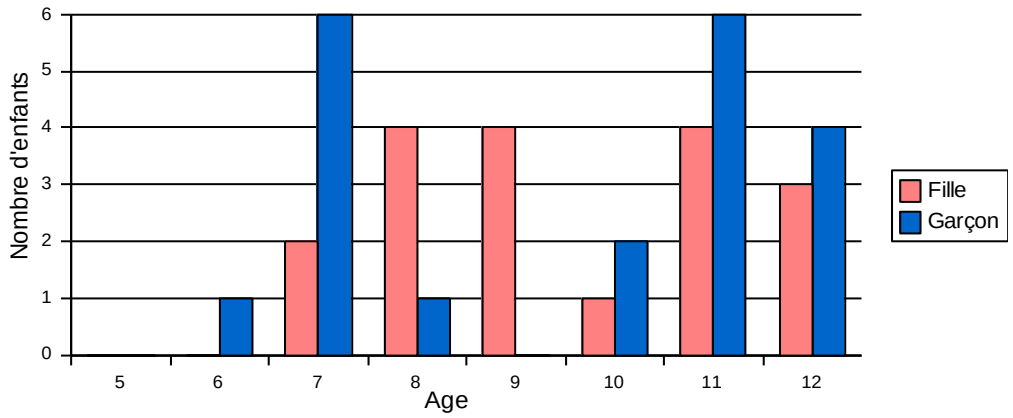


Figure 29 : Répartition de l'âge du groupe EE en fonction du sexe.

Concernant le groupe ESA, la population était représentée par 9 filles (45%) et 11 garçons (55%) (Figure 30). Comme pour le groupe EE, on observe une répartition plutôt équilibrée du nombre de filles et de garçons. L'âge des enfants se situe entre 5 et 13 ans couvrant une tranche d'âge plus étendue que celle du groupe EE. La moyenne d'âge du groupe est 9,5 ans (±2,99). La répartition des âges en fonction du sexe est représentée sur la figure 31.

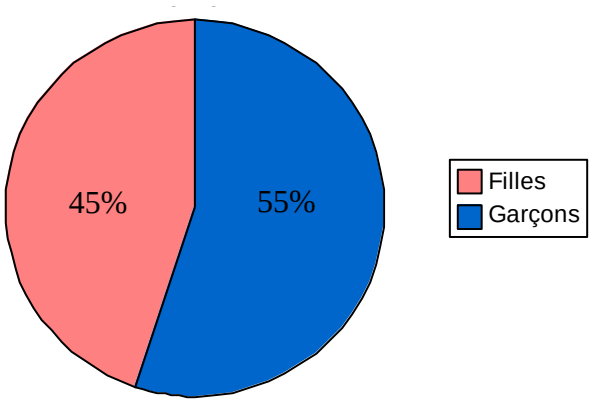


Figure 30 : Répartition du groupe ESA en fonction du sexe.

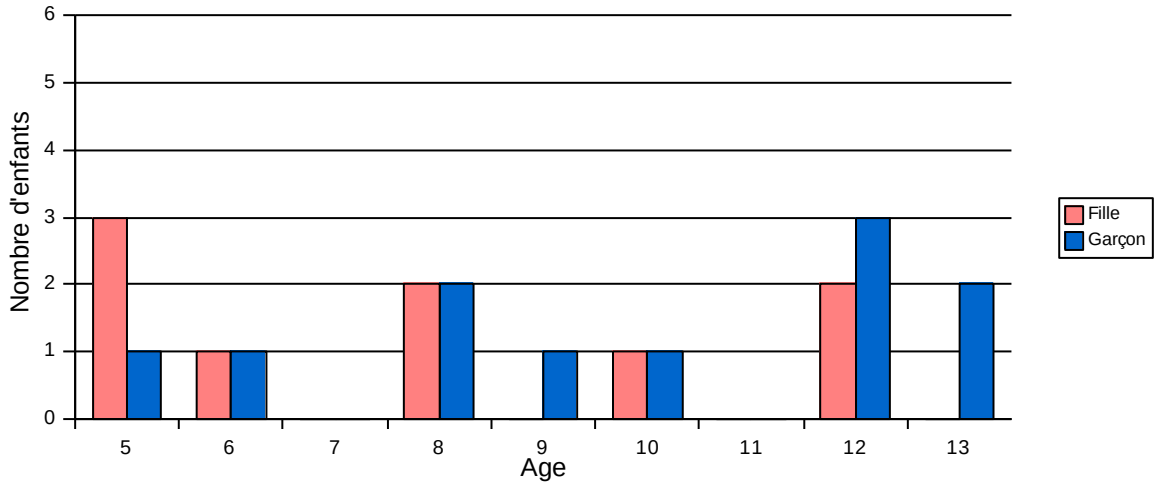


Figure 31 : Répartition de l'âge du groupe ESA en fonction du sexe

Concernant le groupe ESIC, l'échantillon était constitué de 6 filles (30%) et 14 garçons (70%) (Figure 32). L'échantillon ici est majoritairement masculin. L'âge des enfants se situe entre 6 et 13 ans. La moyenne d'âge du groupe est 10,4 ans (±2,46). La répartition des âges en fonction du sexe est représentée sur la figure 33.

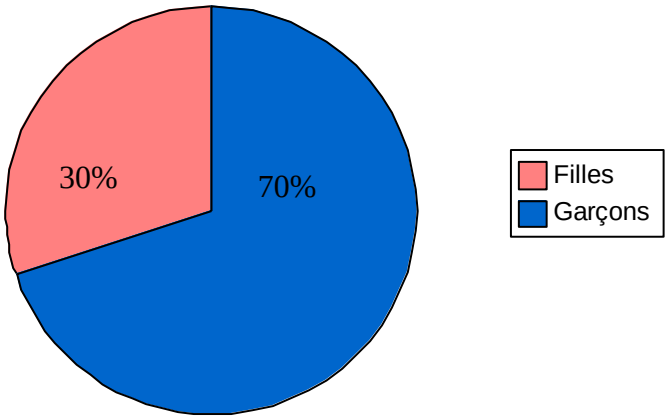


Figure 32 : Répartition du groupe ESIC en fonction du sexe.

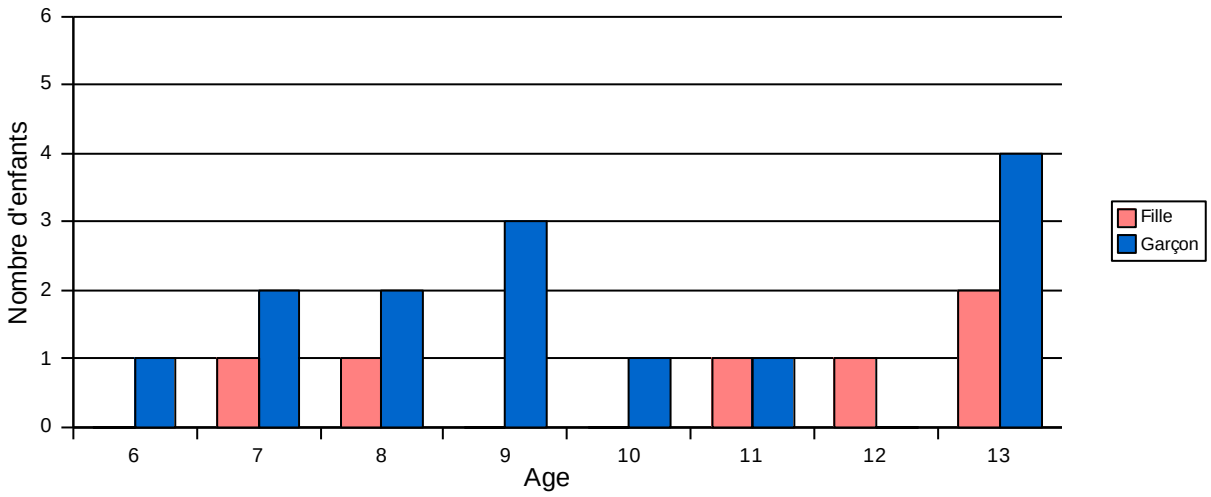


Figure 33 : Répartition de l'âge du groupe ESIC en fonction du sexe.

Globalement, l'échantillon de notre étude était constitué de trois groupes plutôt homogènes et équilibrés concernant le sexe et l'âge. Nous avons voulu constituer des groupes de population les plus homogènes possibles afin de réduire l'effet intergroupe.

Par ailleurs, nous avons recueilli des informations sur quelques facteurs environnementaux, socio-familiaux et sur les habitudes vocales des enfants de l'étude auprès de leurs parents. Ce recueil a été effectué lors de l'enregistrement vocal à l'aide du questionnaire QPEMVox décrit précédemment (Annexe 4 et 5). Les résultats du questionnaire concernant les items 1 à 7 sont présentés dans l'annexe 10. Nous présentons ci-dessous un résumé de ces items.

- Il existe deux facteurs environnementaux communs aux trois groupes de population : les habitudes tabagiques parentales et la cohabitation avec un déficient auditif dans le foyer.
- L'appréciation des parents sur les performances vocales de leurs enfants pour communiquer quotidiennement rejoint la tendance que nous avons trouvée dans l'évaluation acoustique instrumentale : la performance vocale de l'ESIC est plutôt semblable à celle de l'EE.
- Les enfants ESA et ESIC pratiquent des activités de loisir majoritairement collectives.
- Un certain nombre de parents d'enfants ESA et ESIC estiment que la voix de leurs enfants est altérée. Des signes d'anomalie vocale sont référés. Ces signes surviennent de façon spontanée le soir et sont notamment perçus comme une raucité et/ou un accompagnement de souffle de la voix.

- La réussite scolaire est plus avantageuse chez l'ESA que chez l'ESIC.
- Certains enfants appareillés et implantés utilisent d'autres modes de communication que l'oral : LPC* et LSF**.
- Les parents d'ESA ont perçu une amélioration de la voix de leurs enfants avant 3 ans d'appareillage alors que les parents d'ESIC ont perçu une amélioration après 3 ans.
- Les parents des enfants ESA et ESIC sont satisfaits des résultats obtenus avec les aides auditives notamment avec l'implant cochléaire.

Concernant les items 8 et 9 du questionnaire QPEMVox (Annexe 4 et 5), nous présentons ici les résultats obtenus.

1.2. Appareillage/implantation.

Concernant les caractéristiques de l'aide auditive portée par les enfants de notre étude, nous avons observé que :

Chez l'ESA :

- 18 enfants ESA (90%) portent une prothèse auditive classique bilatérale de type « contour d'oreille »; 1 enfant porte une prothèse bilatérale de type « intra-auriculaire »; 1 enfant porte une prothèse unilatérale (côté gauche) de type « Vibreur BAHA »***.
- Les modèles des prothèses auditives portées par l'échantillon de l'étude sont résumées dans le tableau 6.
- L'âge d'appareillage des enfants se situe entre 10 mois et 3 ans et 2 mois. La moyenne d'âge d'appareillage est 2 ans et 4 mois. Nous avons observé que la population de l'étude avait majoritairement été appareillée entre 2 et 3 ans d'âge réel (Figure 34).
- La durée de port de l'appareillage des enfants oscille entre 3 ans et 4 mois, et 11 ans et 7 mois. La moyenne de durée de port de la prothèse de la population est de 7 ans et 2 mois. L'échantillon de l'étude est majoritairement appareillé depuis plusieurs années, ce qui permet d'apprécier les résultats de l'apport de l'appareillage à moyen et à long terme (Figure 35).

---

* Langage Parlé Complété
** Langue des Signes Française
*** La conduction de cette prothèse s'accomplit par voie osseuse. L'enfant manifeste une hypoplasie du pavillon de l'oreille associée au syndrome de Franceschetti-Klein. Bien que des difficultés respiratoires se manifestent précocement perturbant l'émission vocale, nous avons constaté chez cet enfant une activité respiratoire plutôt normale (Temps Maximale Phonatoire = 9 secondes). De ce fait, nous l'avons inclus dans notre étude.

Fabricant	Modèle	Nombre d'enfants
Siemens	ARTIS 2P	3
	CENTRA C HP	2
	SIGNIA 8 DF	2
	ARTIS P	2
	CENTRA CIC	1
	ARTIS POWER	1
	CENTRA CM	1
	ARTIS 2SP	1
	CIELO DP	1
	CIELO DIR	1
	ACURIS AP	1
Phonak	SUPERO 411	1
	CLARO 311	1
	EXTRA 411 AZ POWER	1
Nobel-Pharma	BAHA CLASSIC 300	1

Tableau 6 : Modèles des prothèses auditives classiques portées par les enfants de l'étude.

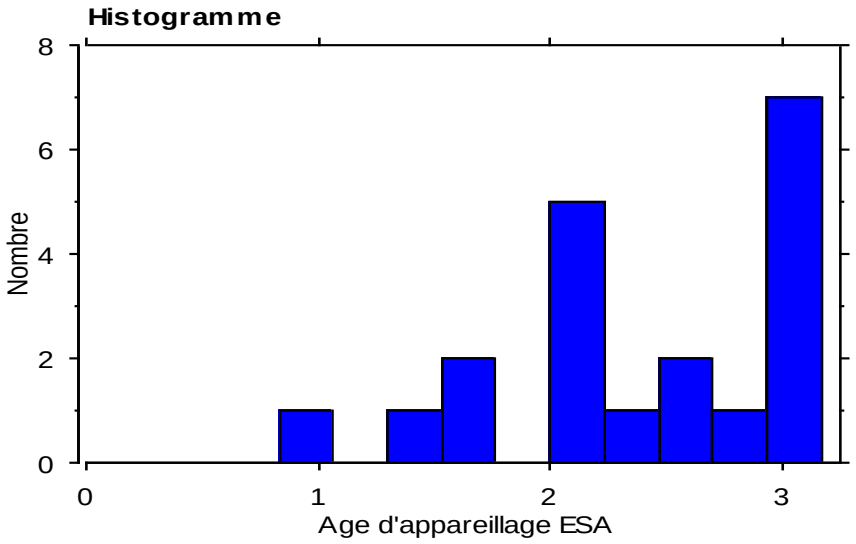


Figure 34 : Répartition de la population selon l'âge d'appareillage.
Le graphe montre une distribution dissymétrique, décalée à droite.

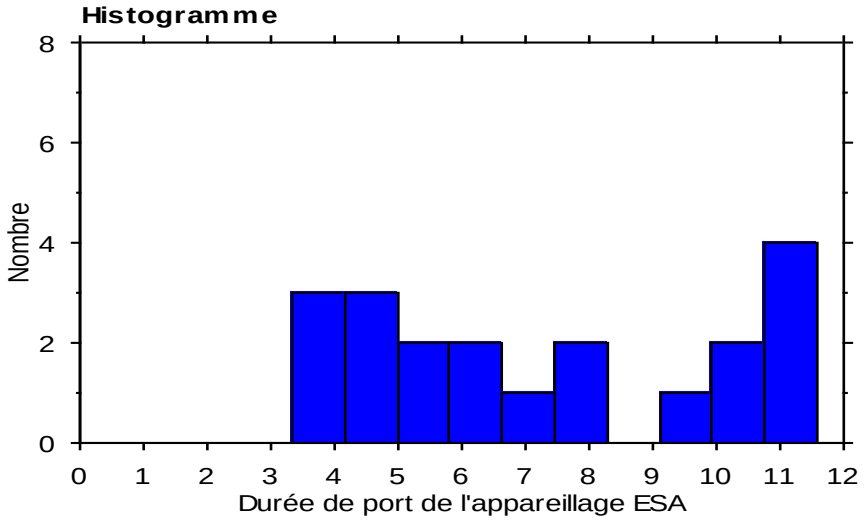


Figure 35 : Répartition de la population selon la durée du port de l'appareillage. Le graphe montre une distribution symétrique régulière.

Résumé des caractéristiques de l'appareillage de l'ESA :

- La majorité des enfants utilisent une prothèse bilatérale de type « contour d'oreille » (Siemens).
- Age moyen d'appareillage : 2 ans et 4 mois.
- Durée moyenne du port de la prothèse : 7 ans et 2 mois.

Chez l'ESIC :

- 19 enfants représentant 95% du groupe portent un implant cochléaire *Nucleus*®; 1 enfant (5%) porte un implant cochléaire *Clarion*® (Tableau 7).
- Les processeurs vocaux des implants utilisés par les enfants de l'étude sont présentés dans le Tableau 7.

Fabricant	Implant cochléaire	Processeur vocal	Nombre d'enfants (n=20)
Cochlear™	Nucleus®	SPrint	9
		ESPrit 3G	5
		Freedom	3
		ESPrit 22	2
Advanced Bionics	Clarion®	S-Series	1

Tableau 7 : Processeurs vocaux des implants cochléaires portés par les enfants.

- La répartition du groupe en fonction des différents modèles des implants cochléaires est présentée par la figure 36.

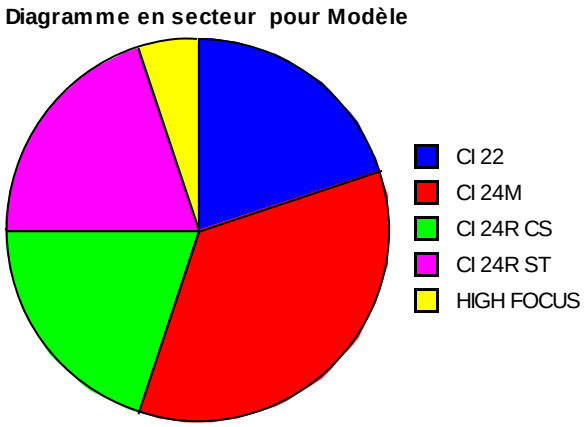


Figure 36 : Répartition du groupe selon le modèle de l'implant cochléaire utilisé.

- La stratégie de codage la plus utilisée est ACE (70%) et SPEAK (25%) pour l'implant cochléaire *Nucleus*®, et SAS (5%) pour l'implant cochléaire *Clarion*® (Figure 37).

Figure 37 : Distribution du groupe selon la stratégie de codage de l'implant cochléaire.

- L'implantation cochléaire a été effectuée sur une oreille : implantation unilatérale. Le côté implanté est manifestement le côté droit, 85% du groupe. Seulement 60% du groupe porte une prothèse contre-latérale au moment de l'étude (Figure 38).

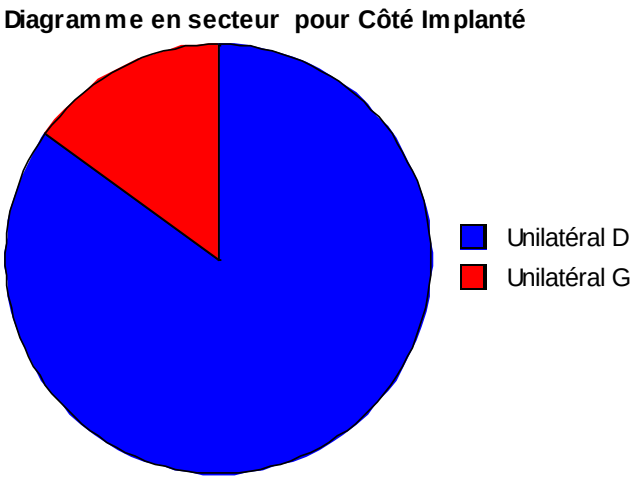


Figure 38 : Répartition du groupe en fonction du côté implanté.

- Le nombre d'électrodes actives, après la dernière séance de réglage, est présenté sur la figure 39. Nous observons que l'activation de 20 électrodes est celle qui prédomine dans la population.

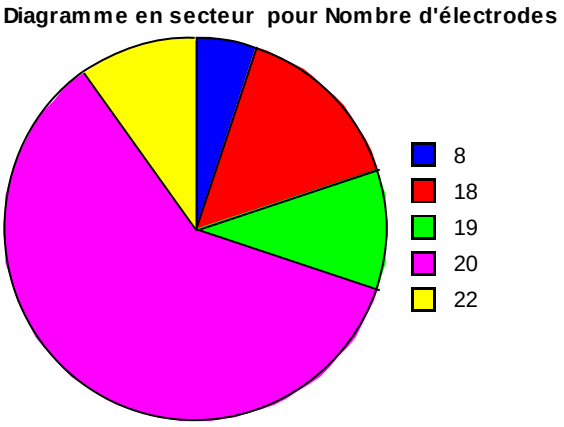


Figure 39 : Répartition du groupe en fonction du nombre d'électrodes actives de l'implant cochléaire.

- L'âge à l'implantation des enfants se situe entre 1 an et 9 mois, et 3 ans et 4 mois. La moyenne d'âge à l'implantation est de 2 ans et 5 mois (Figure 40).

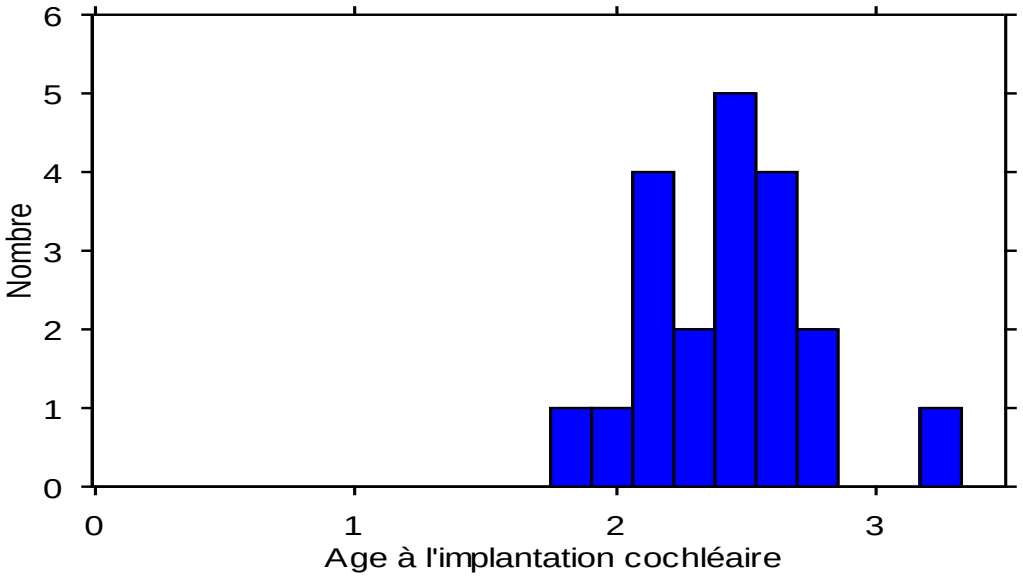


Figure 40 : Répartition du groupe selon l'âge à l'implantation.
Le graphe montre une distribution plutôt symétrique.

- La durée d'implantation cochléaire des enfants oscille entre 3 ans et 8 mois, et 11 ans et 11 mois. La moyenne de la durée d'implantation de la population est 8 ans. La population de l'étude est essentiellement implantée depuis plusieurs années, ce qui permet d'apprécier les résultats de l'apport de l'implant cochléaire à moyen et à long terme (Figure 41).

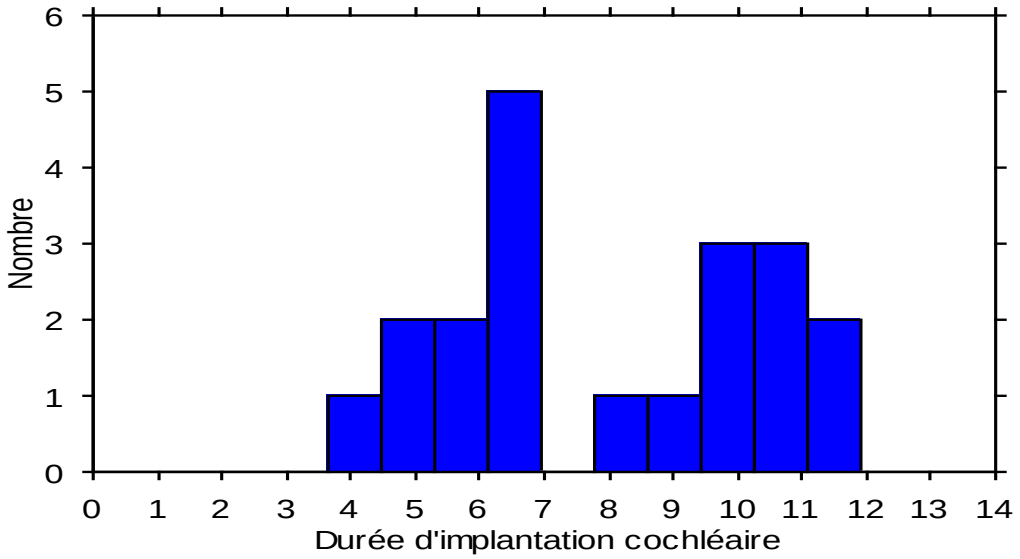


Figure 41 : Répartition du groupe selon la durée d'implantation.
Le graphe montre un faible nombre d'individus au centre de la distribution.

Résumé des caractéristiques de l'implantation de l'ESIC :

- La majorité des enfants utilisent un implant cochléaire Nucleus® (CochlearTH).
- Modèle d'implant : CI24M.
- Stratégie de codage : ACE et SPEAK.
- Implantation unilatérale (majoritairement du côté droit). La prothèse contre-latérale n'est pas portée par tous les enfants au moment de l'étude.
- Nombre d'électrodes actives : 20.
- Age moyen à l'implantation : 2 ans et 5 mois.
- Durée moyenne d'implantation : 8 ans.

1.3. Surdité

Quant aux caractéristiques de la surdité des enfants de notre étude, nous avons observé que :

– Tous les enfants manifestaient une surdité bilatérale. Le degré de surdité prédominant dans le groupe ESA était la surdité sévère Type I (55%) et dans le groupe ESIC était la surdité profonde Type I (65%), (Figure 42).

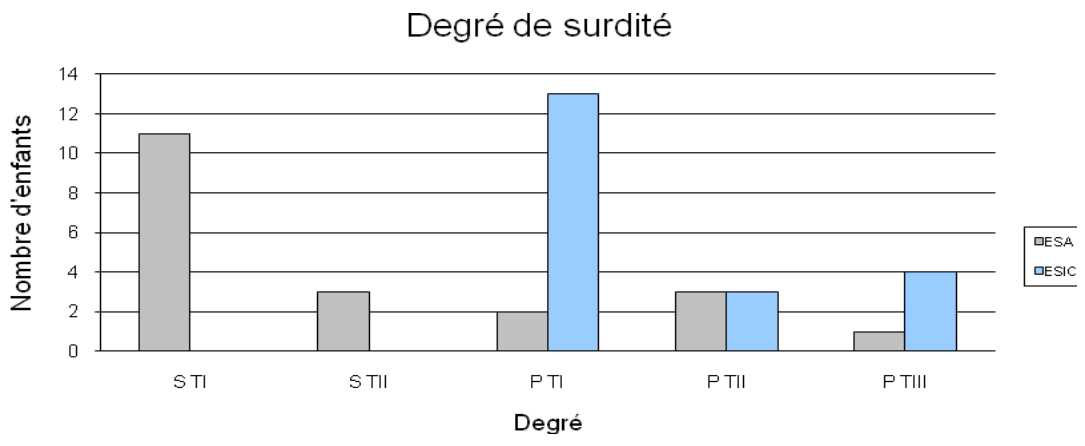


Figure 42 : Distribution des groupes ESA et ESIC en fonction du degré de surdité[*].

- De manière générale, la manifestation des « signes » de la surdité dans le groupe ESA (55%) a débuté à la période de la naissance (Figure 43). Cependant le diagnostic de la surdité s'est effectué tardivement vers l'âge de 2 ans pendant la période d'acquisition et de développement du langage (Figure 44). Par ailleurs, la manifestation de la surdité dans le

[*] ST I, II : Surdité Sévère Type I, II ; PT I, II, III : Surdité Profonde Type I, II, III.

groupe ESIC (85%) a débuté à la naissance (Figure 43). Le diagnostic de la surdité dans ce dernier groupe s'est plutôt effectué avant l'âge de 1 an avant l'acquisition du langage (Figure 44). Le diagramme montre une tendance inverse des deux groupes.

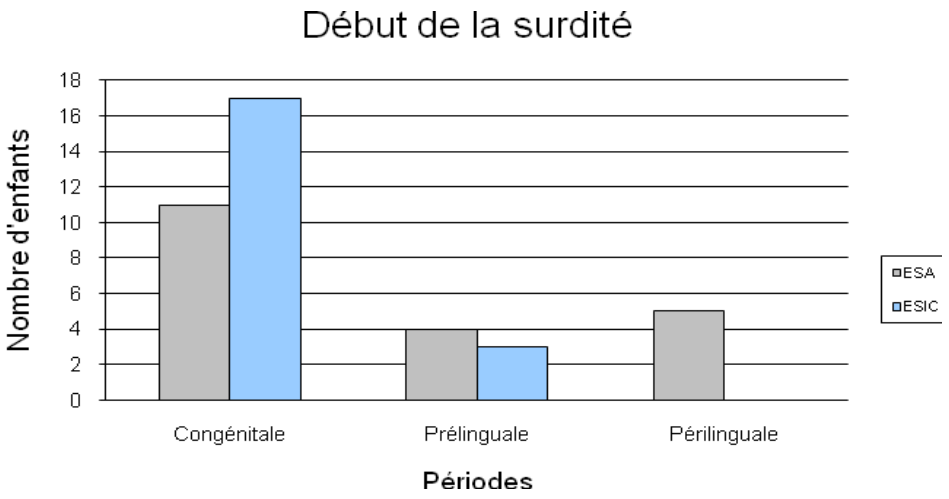


Figure 43 : Distribution des groupes ESA et ESIC en fonction de la période de début de la surdité.

- L'étiologie de la surdité reste inconnue aussi bien chez l'ESA que chez l'ESIC. La deuxième cause de surdité est d'origine génétique chez l'ESA et d'origine environnementale chez l'ESIC (Figure 45).

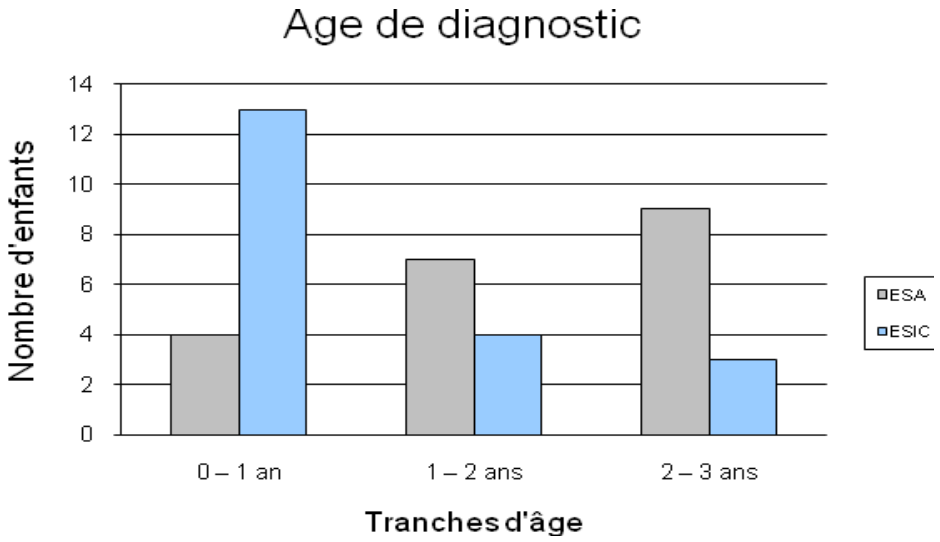


Figure 44 : Distribution des groupes ESA et ESIC en fonction de l'âge de diagnostic de la surdité.

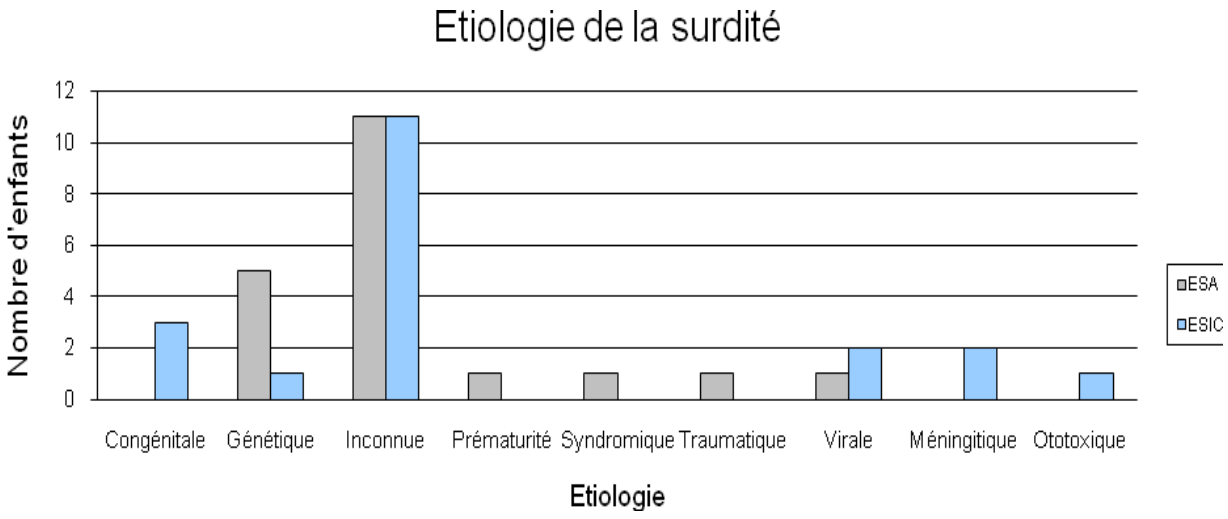


Figure 45 : Distribution des groupes ESA et ESIC en fonction de l'étiologie de la surdité.

Caractéristiques de la surdité chez l'ESA :

- Surdité bilatérale, prédominant une surdité sévère Type I.
- Période d'apparition : plutôt congénitale et périlinguale.
- Etiologie : principalement inconnue et génétique.

Caractéristiques de la surdité chez l'ESIC :

- Surdité bilatérale, prédominant une surdité profonde Type I.
- Période d'apparition : plutôt congénitale et prélinguale.
- Etiologie : principalement inconnue et environnementale.

En somme, nous pouvons dire que les caractéristiques d'appareillage/implantation et de surdité des enfants constituant les groupes ESA et ESIC de notre étude sont plutôt similaires. Nous avons sélectionné un échantillon de population pour notre étude à la fois le plus représentatif et le plus homogène possible malgré les difficultés rencontrées pour le recrutement.

2. Protocole expérimental appliqué, matériel et équipe

Le recueil de données expérimentales a été effectué à l'aide d'un protocole d'évaluation multiparamétrique de la voix que nous avons conçu. Ce protocole dénommé *« PEMVox »* (Annexe 8 et 9) comporte la procédure d'évaluation et d'enregistrement des voix de notre étude. Le protocole PEMVox tente de recueillir des données qualitatives et quantitatives notamment sur les mesures des paramètres acoustiques et aérodynamiques de la voix de l'enfant à déficience auditive. Le protocole a été conçu pour évaluer la voix dans des circonstances d'examen clinique. Nous présentons ici la procédure d'évaluation appliquée et le matériel utilisé pour constituer notre base de données expérimentales.

2.1. Enregistrement de la voix

L'enregistrement des échantillons de voix a été réalisé dans une salle insonorisée à l'aide d'un magnétophone DAT (SONY Digital Audio Recorder PCM–2 300) et d'un microphone (SENNHEISER® e815s). Le microphone était placé à 20 cm de la bouche de l'enfant et légèrement décalé sur le côté afin d'assurer un enregistrement stable et éliminer ainsi les bruits de souffle (Figure 46). L'examinateur surveillait le gain afin d'éviter toute saturation du système d'enregistrement.

Figure 46 : Contexte d'enregistrement de la voix d'un enfant.

L'enregistrement et le recueil de mesures quantitatives des paramètres acoustiques et aérodynamiques des voix ont été effectués à l'aide de l'appareillage EVA™2 (SQ-Lab, Marseille)[*]. Ce système permet de recueillir simultanément des mesures acoustiques et aérodynamiques.

2.2. Matériau phonétique

Dans un premier temps, les enfants avaient pour tâche de produire des émissions de parole. Les techniques d'évaluation pour solliciter ces émissions sont ensuite décrites. Dans un deuxième temps, les enfants ont produit une série de vocalises sur des voyelles tenues et pendant plusieurs secondes. Les consignes d'évaluation étaient données à chaque enfant de manière identique. De même, un exemple était auparavant donné à l'enfant par l'examinateur. Plusieurs essais étaient effectués jusqu'à ce que la qualité des enregistrements et des émissions soit jugée satisfaisante par l'examinateur. Nous avons demandé à l'enfant de rester débout et d'adopter une posture confortable pendant les épreuves d'enregistrement. Par ailleurs, le comportement postural accompagnant les émissions vocales a été observé et évalué (Annexe 15). L'examinateur veillait toujours à ce que la posture n'interfère pas avec l'enregistrement. Un guide comportant les consignes de la passation a été conçu (Annexe 7).

2.2.1. Echantillon de parole

Nous avons enregistré divers types d'échantillon de parole. Ceci en raison des différents types de voix que nous envisagions d'analyser, à savoir : la voix conversationnelle, la voix forte et la voix d'appel. En clinique, examiner les différents types de voix d'un patient permet à l'examinateur d'avoir une idée perceptive du comportement vocal de la voix dans les différentes circonstances communicatives.

Pour évaluer la voix conversationnelle le plus naturellement possible, nous avons demandé à l'enfant de raconter des événements personnels. Certains auteurs préfèrent la lecture des extraits d'un texte[**] ou la lecture d'une phrase conçue spécifiquement pour l'étude de la voix conversationnelle. Cela dans le but de réduire la variabilité phonétique de la production orale et d'homogénéiser les échantillons à traiter (Revis, 2004[132]; Perrin et cols, 1999[133]). Néanmoins, quand il s'agit d'une population d'étude infantile, comme dans notre étude, le

[*] Pour plus de détails sur la conception du dispositif EVA2, consulter : http://www.sqlab.fr/evaRootFR.htm

[**] « La petite chèvre de Monsieur Seguin » d'Alphonse Daudet, utilisé par l'équipe du laboratoire Parole et Langage d'Aix-en-Provence, ou « La bise et le soleil » de l'Association Phonétique Internationale.

problème se pose. Etant donné qu'une partie de notre population n'était pas lectrice, nous avons envisagé plutôt d'enregistrer des échantillons de parole spontanée permettant d'analyser davantage, et de façon plus écologique, les paramètres acoustiques. En effet, nous avons choisi ce corpus utilisé en pratique orthophonique pour évaluer le langage oral. La parole spontanée produite dans une situation d'observation dite « naturelle » permet de relever des informations plus riches pour en faire des analyses verbales (Le Normand, 2007[134]). Par ailleurs, nous avons écarté la procédure de répétition d'une phrase standard par l'enfant car cela demandait un temps supplémentaire à l'examinateur pour l'entraînement de l'enfant avant la passation, cette tâche ne représentant pas vraiment une situation communicative. L'échantillon de parole sélectionné pour l'analyse de la voix conversationnelle présentait une durée de 15 secondes environ.

Concernant l'évaluation de la voix forte, nous avons effectué l'épreuve du « comptage projeté » (Le Huche, 2001[34]). Nous avons demandé à l'enfant de « compter de zéro jusqu'à vingt » à haute voix comme s'il s'adressait à des gens placés à une dizaine de mètres. L'échantillon de parole sélectionné pour l'analyse comprenait toute la durée de l'émission du comptage.

Enfin, pour évaluer la voix d'appel, nous avons demandé à l'enfant d'interpeller quelqu'un se trouvant à une distance d'une cinquantaine de mètres environ.

L'intégralité de l'échantillon de parole enregistré était numérisée par ordinateur (fréquence d'échantillonnage à 22050 Hz, résolution en amplitude à 16 bits, mono) grâce au logiciel Cool Edit Pro® version 2.0. au format *.*wav*.

2.2.2. Échantillon de voyelles tenues

Le choix d'évaluer le profil vocal des enfants de notre étude à partir d'une voyelle tenue est basé sur des arguments plutôt pragmatiques présentés dans la littérature. Il existe un très grand nombre d'études employant davantage les voyelles tenues lors de l'évaluation instrumentale. Ceci car l'on considère que les appareillages de mesure, y compris celui que nous avons utilisé pour notre étude, ne sont performants que sur des voyelles tenues (Giovanni, 1995[135], 1996[136]; Revis, 2004[132]).

Des auteurs argumentent que les voyelles représentent davantage la voix chantée et la parole représentent mieux la voix conversationnelle. Certains effets survenant au cours de la parole (transitoires d'attaque et de sortie, interruptions de la phonation, etc.) n'apparaissent pas lors de l'émission d'une voyelle tenue. De ce fait, la validité de l'évaluation de la qualité de la voix obtenue à partir de voyelles tenues est remise en cause (De Krom, 1994[137]; Hammarberg, 1980[138]).

Cependant, la plupart des études de nos jours fondent leur protocole d'évaluation sur le /a/ tenu. Le /a/ tenu est facilement reproductible et interprétable par les logiciels de mesure et par conséquent est un matériau phonétique représentatif (Wolfe, 1995[139]; Hartl et cols., 2005[140]). Aussi, nous avons opté pour l'utilisation d'un /a/ tenu comme matériau phonétique pour une partie de l'analyse de notre étude.

Trois types d'échantillon à partir des voyelles tenues ont été recueillis pour chaque enfant. En effet, nous souhaitions analyser le Profil Vocal, le Temps Maximal de Phonation et la Tessiture et l'Etendue Vocale. Ces paramètres acoustiques seront décrits en détail ultérieurement.

Pour le profil vocal, l'enfant produisait une série de cinq /a/ tenus sur plusieurs secondes à une intensité et hauteur confortables, le plus naturellement possible comme pour la voix conversationnelle habituelle. Seulement, trois productions médianes étaient retenues pour l'étude : le premier /a/ était supprimé afin d'éviter l'effet d'inhibition et de stress dû à la situation d'examen; le dernier était supprimé afin d'éviter les risques de fatigabilité de la voix. Nous considérons que les trois productions médianes sont les plus représentatives d'une situation « normale ».

Pour le Temps Maximal de Phonation, l'enfant émettait un /a/ tenu le plus longtemps possible à hauteur et intensité confortables, le plus naturellement possible comme pour la voix conversationnelle habituelle. Nous avons demandé à l'enfant d'effectuer trois essais. La mesure la plus élevée était rapportée dans les résultats car il s'agit de mesurer la performance de phonation maximale de l'enfant.

Enfin, l'évaluation de la tessiture et de l'étendue vocale ont été effectuées sur un /a/ et /i/ tenus respectivement. Pour la tessiture, l'enfant émettait un /a/ tenu de manière ascendante et descendante en hauteur comme « la sirène d'une ambulance » mais toujours à une intensité

confortable sans faire d'effort. Nous avons demandé à l'enfant d'effectuer trois essais. Nous avons retenu la production de meilleure qualité. Pour l'étendue vocale, nous avons demandé à l'enfant d'imiter « une fusée » mais cette fois-ci avec un /i/ tenu. L'enfant devait démarrer d'un son le plus grave possible à émettre et aller jusqu'au son le plus aigu avec un /i/ tenu sans interruption. L'intensité augmentait selon l'ascension du /i/ tenu. Les voyelles tenues étaient ensuite numérisées de la même façon que pour les échantillons de parole (*section 2.2.1*).

******

En résumé, nous avons inclus dans notre protocole d'évaluation deux types de corpus : une voyelle tenue et un extrait de parole spontanée. Le premier permet de minimiser les effets acoustiques du conduit vocal et fournit un son voisé relativement stable et de durée adéquate. Le deuxième permet d'étudier les effets des troubles vocaux sur la parole. Au total, ces deux corpus sont complémentaires pour donner ainsi des informations sur la qualité vocale de la parole.

Après avoir décrit le matériau phonétique choisi et la procédure d'enregistrement des échantillons vocaux employée, nous présenterons les expériences menées pour le recueil de données ainsi que les résultats et leur interprétation. Nous commençons par une expérience sur la mesure des paramètres acoustiques et aérodynamiques puis sur l'évaluation perceptive des extraits vocaux.

3. Analyse acoustique (Expérience 1)

3.1. Objectifs

Le but de cette expérience était de :

- ✓ décrire les caractéristiques vocales des enfants sourds sévère et profond sur le plan acoustique et de manière quantitative.

- ✓ comparer les caractéristiques acoustiques de la voix des groupes EE, ESA et ESIC afin de déterminer si les caractéristiques acoustiques de la voix du groupe ESIC étaient plus semblables à celles du groupe EE. Cela nous donnera une appréciation sur l'effet de la prothèse auditive portée par l'ESA et l'ESIC sur les caractéristiques acoustiques de la voix.

- ✓ identifier et décrire d'éventuelles caractéristiques vocales pathologiques chez les enfants sourds.

3.2. Méthodologie

L'analyse des échantillons de parole et des voyelles tenues précédemment décrits ont été effectués grâce au dispositif EVA[*™]2 (SQ-Lab, Marseille). Ce système permet de recueillir simultanément des mesures quantitatives acoustiques et aérodynamiques de la voix. Avant tout, l'examinateur vérifiait l'état et le calibrage de l'appareillage (raccordement des divers fils), ainsi que le paramétrage applicable à l'enregistrement des voix infantiles selon les réglages préconçus par le fabricant pour cet effet.

Les échantillons de parole ont été traités par l'application « PROSODIE » du dispositif EVA™2 (Figure 47). Le premier graphe (de haut en bas) représente le spectre d'un extrait d'échantillon de voix conversationnelle sélectionné et analysé. Le deuxième graphe représente la courbe mélodique de l'extrait; le troisième représente la courbe de variation d'intensité; et enfin, le dernier représente les pauses effectuées au cours de la parole spontanée. Nous n'avons pas pris en compte cette dernière donnée.

[*] EVA= Evaluation Vocale Assistée par ordinateur

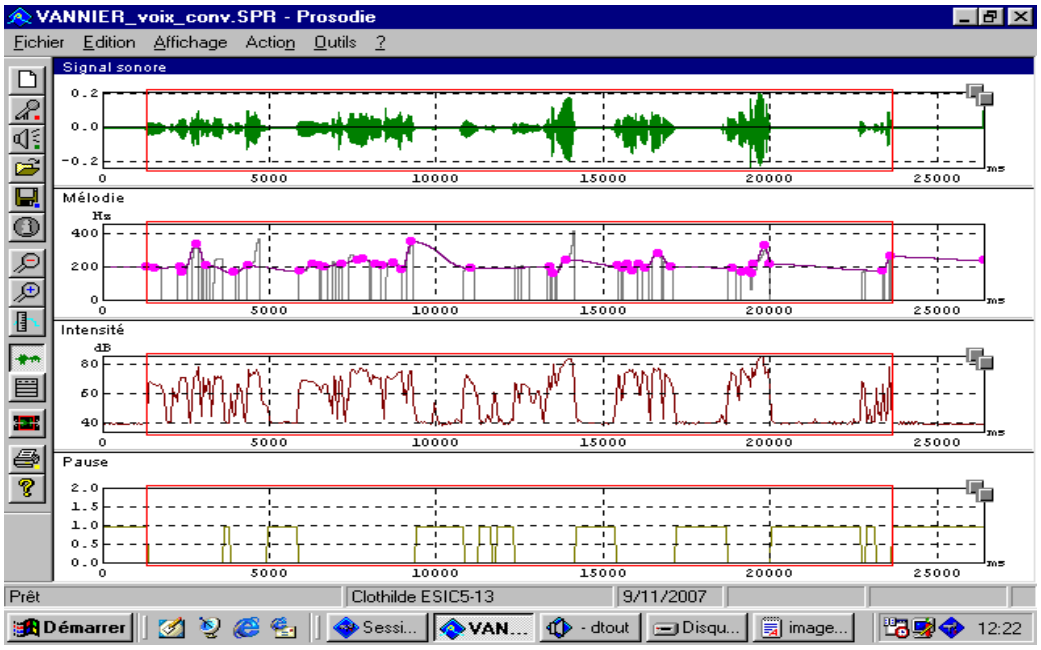


Figure 47 : Fenêtre d'affichage de l'application « PROSODIE » du dispositif EVA™2

Après l'enregistrement de l'échantillon (voix conversationnelle, forte et d'appel), ce logiciel lance automatiquement une fenêtre présentant différentes courbes dont celle de la courbe mélodique de la parole. Les valeurs étaient ensuite recueillies et présentées dans un tableau. Les valeurs des mesures suivantes ont été relevées :

1) Le fondamental usuel moyen (FoUm), ou la fréquence moyenne la plus souvent utilisée par l'enfant au cours de la parole. La valeur était donnée en Hertz (Hz) et sa représentation homologue selon la notation musicale[**] en demi-tons.

2) La dynamique de la parole, ou l'écart entre le son le plus grave (fréquence minimale) et le son le plus aigu (fréquence maximale) atteints au cours de l'émission. La valeur était donnée en Hertz (Hz) et en demi-tons.

Les échantillons des voyelles tenues ont été traités par différentes applications. En ce qui concerne l'analyse des coefficients de stabilité de la fréquence et de l'intensité sur un /a/ tenu, ces paramètres ont été traités par l'application « PROFIL VOCAL » du dispositif EVA™2 (Figure 48). Le premier graphe (de haut en bas) représente le spectre d'un /a/ tenu analysé. Le deuxième graphe représente la courbe de stabilité de la fréquence fondamentale de l'extrait; le

[**] Selon la notation musicale de la gamme diatonique tempéré (LA_3 = 440 Hz).

troisième graphe représente la courbe de variation d'intensité; et enfin, le dernier graphe représente la courbe du débit d'air oral qui n'a pas été prise en compte.

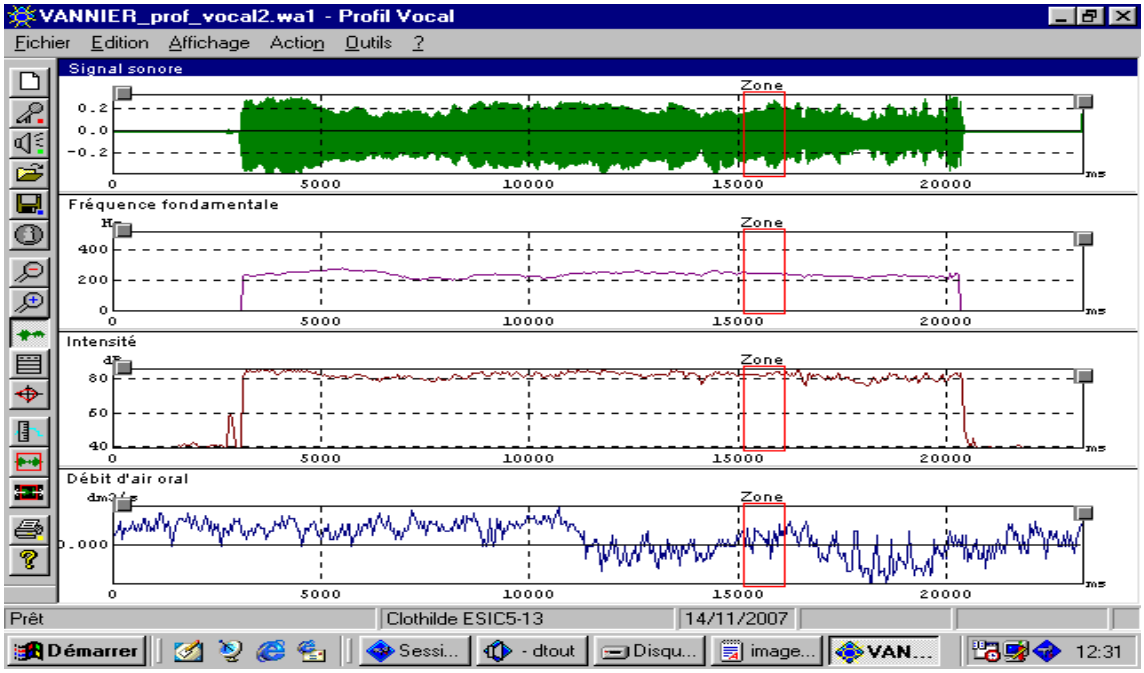


Figure 48 : Fenêtre d'affichage de l'application « PROFIL VOCAL » du dispositif EVA™2.

Suite à l'enregistrement de l'échantillon, le logiciel lance automatiquement une fenêtre présentant différentes courbes dont celle de la fréquence de l'émission en fonction du temps. Nous avons placé une fenêtre normalisée de durée de 1000 ms sur la partie la plus stable du signal et placée après l'attaque vocalique afin d'éliminer les phénomènes liés à celle-ci selon la méthode de Yu (2001)[141]. Ensuite, les valeurs étaient recueillies et présentées dans un tableau de résultats.

Les valeurs sont également présentées dans cette application sous forme de diagrammes, de représentation en radar où chaque axe explore un paramètre de la voix. Cela permet de visualiser de façon synthétique l'ensemble des paramètres de la voix par rapport à des données établies sur la normalité. Nous avons relevé les mesures suivantes pour notre étude :

1) La Fréquence Fondamentale (Fo), ou la fréquence moyenne de vibration des cordes vocales permet d'évaluer la hauteur de la voix. La valeur est donnée en Hertz (Hz).

2) Le Coefficient de variation Fo permet de relativiser l'écart-type obtenu en le comparant à la Fo moyenne lors des calculs statistiques présentés dans le tableau de résultats. Il correspond au pourcentage de l'ampleur des variations de la Fo par rapport à la Fo moyenne (Coef. Var Fo = 100 x E-type F0/moyenne F0). C'est un bon indice pour explorer la stabilité de la fréquence fondamentale à moyen terme (Ghio, 2007)[142]. Dans la littérature anglosaxonne, ce paramètre correspond à « Fundamental frequency variation ou vFo ».

3) Le Jitter factor permet de relativiser le Jitter moyen obtenu en le comparant à la Fo moyenne lors des calculs statistiques présentés dans le tableau de résultats de l'appareillage EVA. Le Jitter correspond aux perturbations de la fréquence fondamentale. Il permet de déterminer le degré d'apériodicité entre chaque cycle d'oscillation des cordes vocales et d'apprécier le caractère de raucité. Lorsque ces variations sont assez importantes, elles donnent l'impression de raucité ou d'éraillement de la voix. C'est un bon indice de la régularité de la fréquence fondamentale à « court terme » (Ghio, 2007[142]; Yan et col., 2005[143]). Puisque l'expression « court terme » n'a pas été établie, aucune définition précise mathématique ne peut être donnée (Yan et col., 2005).

4) L'intensité, est un flux d'énergie par unité de surface et par seconde qui se mesure en décibels (dB). La valeur moyenne mesurée que nous avons reportée ici dans notre étude signale la puissance de l'émission.

5) Le coefficient de variabilité de l'intensité permet de relativiser l'écart-type de l'intensité en le comparant à l'intensité moyenne. Il correspond à l'ampleur en pourcentage des variations d'intensité par rapport à l'intensité moyenne.

6) Le Shimmer factor permet de relativiser le Shimmer moyen en le comparant à l'amplitude moyenne. Le Shimmer correspond aux perturbations d'amplitude entre chaque cycle d'oscillation. C'est un indicateur de la stabilité de l'intensité à court terme. Le Shimmer est calculé comme le pourcentage de la valeur de l'écart-type en le comparant à la valeur moyenne de variation de chaque cycle de l'intensité (Ghio, 2007[142]; Yan et col., 2005[143])

7) Le rapport Harmoniques/Bruit (*Harmonic to Noise Ratio ou HNR* selon la littérature anglaise), correspond à une méthode de calcul temporel de 25 cycles consécutifs d'un signal sonore voisé afin d'établir une forme ondulatoire moyenne. Si l'on soustrait la valeur de la forme ondulatoire moyenne de la valeur du signal d'origine, on obtient comme résidu le *bruit t*

ajouté au cours de la phonation. Ainsi, la valeur du rapport Harmoniques/Bruit résultera du calcul régulier de l'énergie de *l'onde moyennée périodique (H)* et de *l'énergie du bruit résiduel (B)* (Ghio, 2007[142]; Yumoto et Gould, 1982[144]). Cette valeur est présentée dans la fenêtre de résultats en dB. Ce paramètre est un indicateur de l'instabilité de la vibration cordale.

En ce qui concerne l'analyse vocale sur un /a/ (vocalise imitant une « sirène ») et /i/ tenus (vocalise imitant une « fusée »), l'application employée du dispositif EVA™2 était « PROSODIE » (Figure 49 et Figure 50 respectivement).

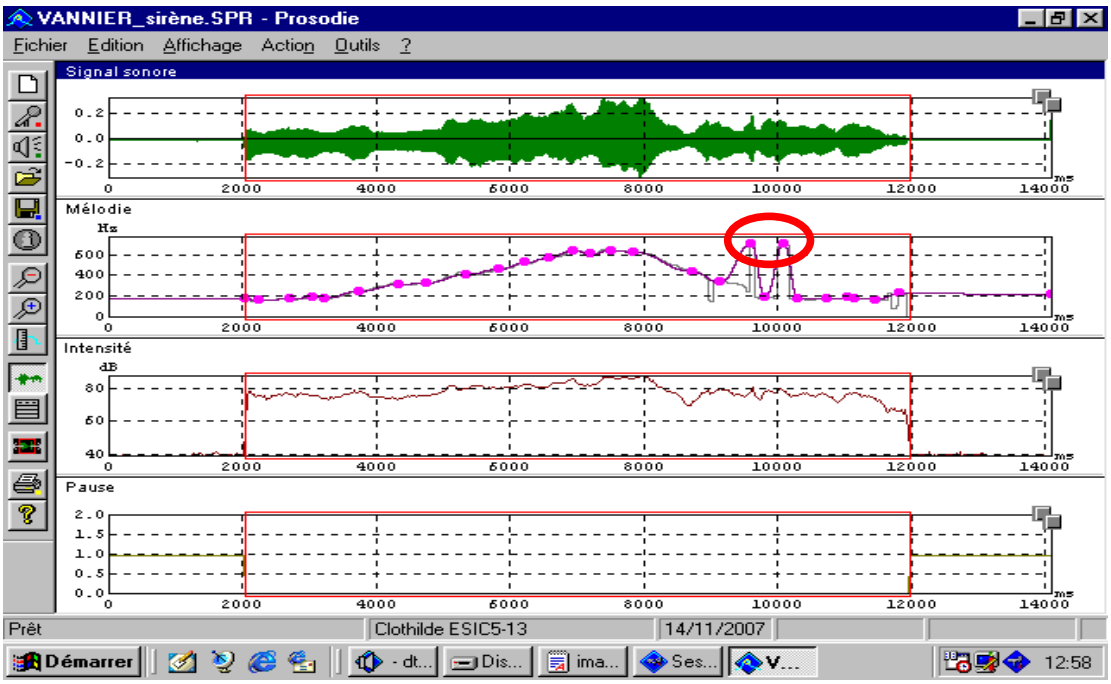


Figure 49 : Fenêtre d'affichage de l'application « PROSODIE » du dispositif EVA™2 pour l'analyse de la tessiture.

Le premier graphe de la figure 49 (de haut en bas) représente le spectre d'un /a/ tenu en imitant une « sirène ». Le deuxième graphe représente la courbe mélodique ascendante et descendante de l'extrait; le troisième graphe représente la courbe de variation d'intensité; et enfin, le dernier représente les pauses effectuées au cours de l'émission. Nous n'avons pas pris en compte ce dernier paramètre. La présence « d'artefacts » (entouré en rouge) pendant l'émission vocale était à surveiller.

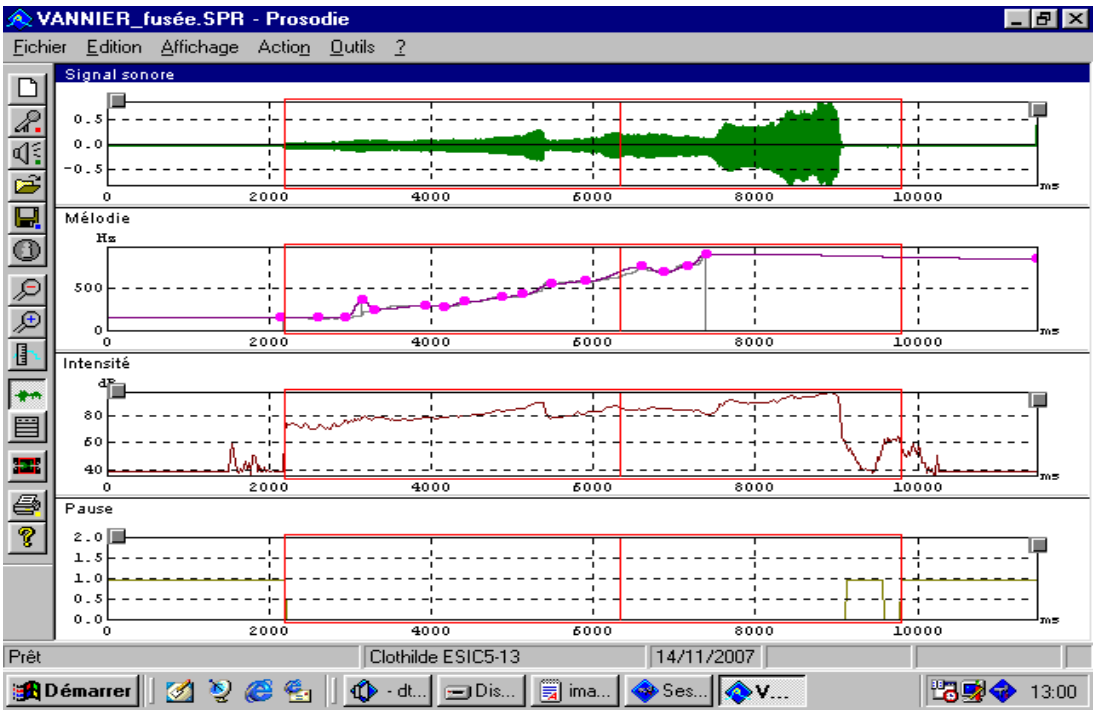


Figure 50 : Fenêtre d'affichage de l'application « PROSODIE » du dispositif EVA™2 pour l'analyse de l'étendue vocale.

Le premier graphe de la figure 50 (de haut en bas) représente le spectre d'un /i/ tenu en imitant une « fusée ». Le deuxième graphe représente la courbe mélodique ascendante de l'extrait; le troisième graphe représente la courbe de variation d'intensité; et enfin, le dernier représente les pauses effectuées au cours de l'émission.

8) La tessiture, ou la gamme de sons émis avec aisance par l'enfant, était tirée de la courbe mélodique de l'émission. Les valeurs étaient recueillies à partir d'un tableau de résultats y apparaissant. Nous avons relevé la mesure de la dynamique (fréquence minimale soustraite de la fréquence maximale) en Hertz et en demi-tons.

9) L'étendue vocale physiologique[*], ou la gamme de sons existant entre le son le plus aigu et le son le plus grave que l'enfant était capable d'émettre, a été obtenue à partir d'un tableau de résultats. De la même manière que pour la tessiture, nous avons relevé la mesure de la dynamique (fréquence minimale soustraite de la fréquence maximale) en Hertz et en demi-tons. C'est un bon indicateur de la fonction vocale (Piccirillo, 1998[145]; Giovanni, 1998[146]).

[*] On distingue deux possibilités d'étendue vocale : soit on teste l'étendue sans tenir compte de la qualité de la voix (étendue vocale physiologique), soit on ne prend en compte que les notes qui possèdent une qualité musicale minimale acceptable (étendue vocale musicale). Cette dernière est bien évidemment plus restreinte que la première.

10) Le Temps Maximal de Phonation (TMP), c'est le temps d'émission vocale sur une voyelle tenue et sur une seule expiration succédant à une inspiration profonde. Il dépend de la quantité d'air disponible et du rendement glottique (rapport d'une énergie produite sur une énergie consommée, représentée par l'intensité du son produit). Ce paramètre est considérablement raccourci chez le sujet dysphonique qui pour finir ses phrases, augmente encore les tensions afin de prolonger la vibration des cordes vocales. Puisqu'il représente un reflet tout à fait fidèle de la qualité de la fermeture glottique, il est considéré comme un paramètre indicateur du bon fonctionnement vocal (Gordon, 1978[147]; Piccirillo, 1998). L'analyse de ce paramètre a été effectuée avec l'application qui porte son nom « TEMPS MAXIMAL DE PHONATION » du dispositif EVA™2. La durée de l'émission était mesurée par l'intermédiaire des curseurs placés en début et fin du signal. Les valeurs étaient rapportées en secondes.

3.3. Résultats et discussion

Le traitement statistique des données expérimentales a été effectué en différentes étapes. Dans cette section, nous présenterons la première étape qui a consisté à analyser les résultats obtenus pour les paramètres acoustiques. Les calculs et traitements statistiques ont été effectués à l'aide des logiciels *StatView version 5.0, Statistica version 7.0* et *« R » version 2.7.2* pour *Windows.*

Dans un premier temps, nous avons effectué une analyse descriptive de 16 paramètres acoustiques retenus dans notre étude pour chaque groupe. Trois groupes ont été constitués : 38 enfants normo-entendant (EE), 20 enfants sourds appareillés (ESA) et 20 enfants implantés cochléaires (ESIC). Au total, nous avions 78 enfants dans l'échantillon.

Puisque nous étions intéressés à analyser la hauteur et la dynamique au cours de la parole, nous avons évalué le fondamental usuel moyen (FoUm[238])* et la dynamique de l'émission de différents types de voix : voix conversationnelle, voix forte et voix d'appel. Ces paramètres sont présentés dans les item 1, 2, 3, 4, 5 et 6 avec les résultats de l'analyse.

1) FoUm de la voix conversationnelle. Les indices statistiques de l'analyse descriptive de la voix conversationnelle pour les différents groupes sont présentés dans le tableau 8 et la figure 51. Le diagramme représente les moyennes FoUm et le graphe en boites représente les

* Fréquence moyenne autour de laquelle évoluent différentes variations de tonalité constituant la mélodie ou l'intonation de la parole (Dalleas et Dalleas, 1987).

médianes, les quartiles ainsi que les « valeurs extrêmes » pour chaque groupe. Pour la voix conversationnelle, les valeurs moyennes de FoUm des groupes ESA et ESIC apparaissent semblables. Par ailleurs, on observe une différence du groupe EE aux groupes ESA et ESIC. Le graphe en boites montre une forte dispersion et une étendue importante chez le groupe ESIC. Une dispersion est également à noter pour les valeurs inférieures à la médiane chez le groupe ESA.

	FoUm EE	FoUm ESA	FoUm ESIC
Moyenne (Hz)	232	248	245
Ecart-type (Hz)	33,15	55,74	52,4
Médiane (Hz)	229	254	260
Minimum (Hz)	158	101	129
Maximum (Hz)	314	368	310

Tableau 8 : Fondamental Usuel Moyen (FoUm en Hz) de la voix conversationnelle pour les différents groupes.

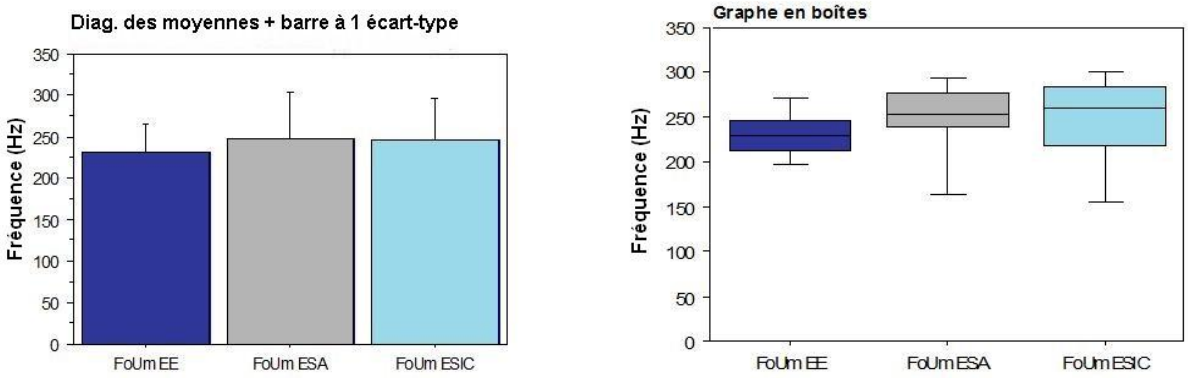


Figure 51 : Diagramme des moyennes et graphe en boites des médianes de FoUm (Hz) de la voix conversationnelle pour les groupes EE, ESA et ESIC.

2) Dynamique de la voix conversationnelle. Quant à la dynamique de la voix conversationnelle, les résultats sont résumés dans le Tableau 9. On note que les moyennes des groupes ESA et ESIC sont plutôt proches (Figure 52). Cependant, si l'on observe les médianes, les valeurs des groupes ESA et ESIC marquent une nette différence. De même, une différence est à noter entre la moyenne des groupes EE et ESA. La dispersion est plus accentuée chez le groupe EE. Il existe une amplitude importante entre les valeurs des fréquences les plus basses et les fréquences les plus hautes émises au cours de la parole.

	Dyn EE	Dyn ESA	Dyn ESIC
Moyenne (Hz)	179	194	189
Ecart-type (Hz)	82,63	80,41	78,87
Médiane (Hz)	166	163	181
Minimum (Hz)	66	78	64
Maximum (Hz)	405	343	357

Tableau 9 : Dynamique (Hz) de la voix conversationnelle pour les différents groupes.

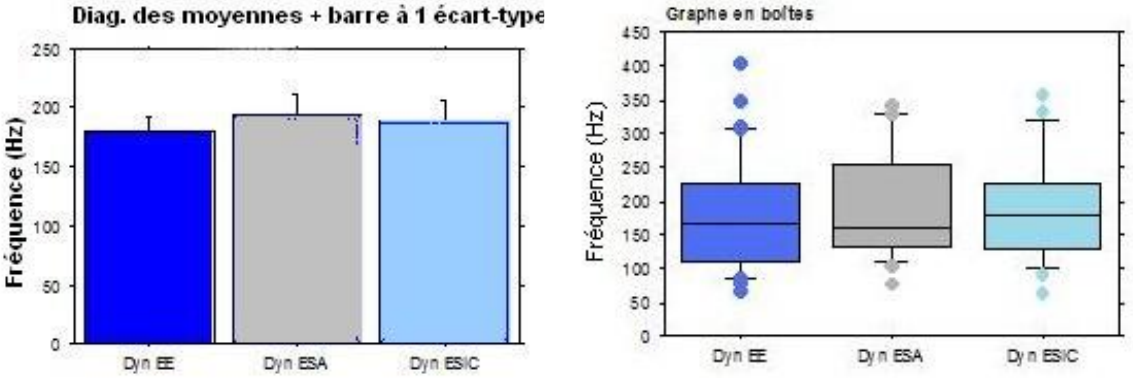


Figure 52 : Diagramme des moyennes et graphe en boites des médianes de la dynamique (Hz) de la voix conversationnelle pour les populations EE, ESA et ESIC.

Concernant la voix conversationnelle, nous observons à simple vue que :

- Les valeurs de FoUm sont plus faibles pour EE que pour les groupes ESA et ESIC, ce qui indique une différence entre le groupe EE et les groupes ESA et ESIC.
- Les valeurs de dynamique de la parole sont différentes entre tous les groupes.
- La dispersion est plus forte chez le groupe EE pour la dynamique.

3) FoUm de la voix forte. Les valeurs des moyennes FoUm sont pratiquement semblables pour les trois groupes. En revanche, les valeurs des médianes signalent une différence entre les groupes ESA et ESIC. On constate une dispersion très importante ainsi qu'une étendue assez grande chez la population ESIC (Tableau 10, Figure 53).

	FoUm EE	FoUm ESA	FoUm ESIC
Moyenne (Hz)	301	306	310
Ecart-type (Hz)	57,86	67,82	92,33
Médiane (Hz)	290	304	286
Minimum (Hz)	181	171	141
Maximum (Hz)	494	459	471

Tableau 10 : Fondamental Usuel Moyen (FoUm en Hz) de la voix forte pour les différents groupes.

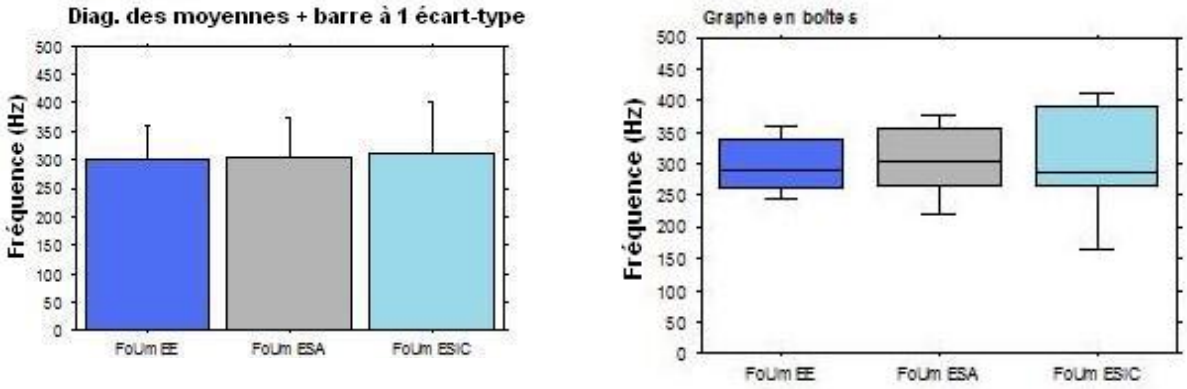


Figure 53 : Diagramme des moyennes et graphe en boites des médianes de FoUm (Hz) de la voix forte pour les groupes EE, ESA et ESIC.

4) Dynamique de la voix forte. La dynamique de la voix forte s'avère semblable pour les groupes EE et ESA. On constate une différence importante entre ces derniers et le groupe ESIC. De même, les valeurs de la médiane indiquent une différence entre les groupes EE et ESIC. La dispersion se révèle très importante chez la population ESIC (Tableau 11, Figure 54).

	Dyn EE	Dyn ESA	Dyn ESIC
Moyenne (Hz)	149	149	197
Ecart-type (Hz)	73,33	69,39	149,23
Médiane (Hz)	135	145	158
Minimum (Hz)	54	52	36
Maximum (Hz)	404	276	659

Tableau 11 : Dynamique (Hz) de la voix forte pour les différents groupes.

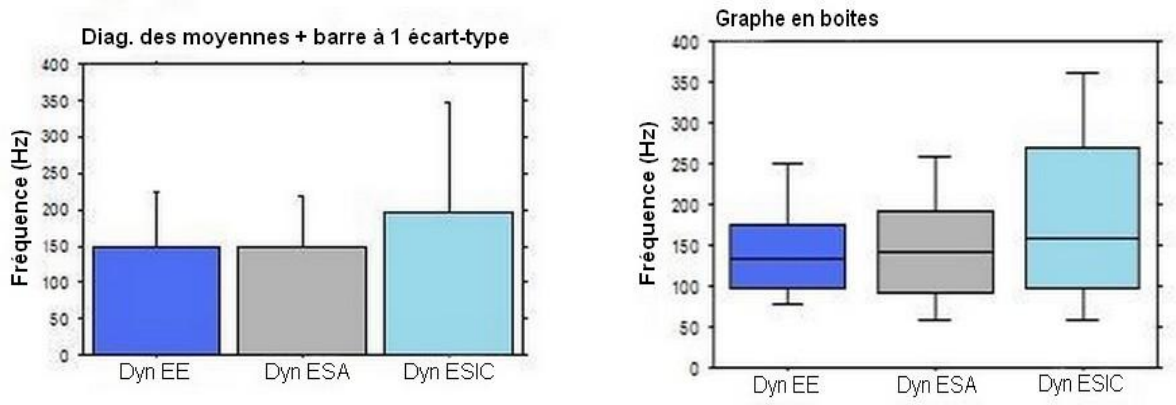


Figure 54 : Diagramme des moyennes et graphe en boites des médianes de la dynamique (Hz) de la voix forte pour les groupes EE, ESA et ESIC.

Concernant la voix forte, nous remarquons que :

- Les différences sont légères entre les groupes pour FoUm.
- Les valeurs de dynamique de la parole sont différentes entre les groupes ESA et ESIC, et EE et ESIC.
- La dispersion est plus forte chez le groupe ESIC pour FoUm et la dynamique.
- La valeur de la dynamique de la parole est plus élevée pour le groupe ESIC.

5) FoUm de la voix d'appel. Les valeurs des moyennes FoUm sont manifestement différentes entre les groupes EE et ESIC, et légèrement semblables entre les groupes EE et ESA. En revanche, les valeurs de la médiane n'indiquent pas de différences importantes entre les groupes. On observe une dispersion non négligeable chez la population ESIC (Tableau 12, Figure 55).

	FoUm EE	FoUm ESA	FoUm ESIC
Moyenne (Hz)	390	387	366
Ecart-type (Hz)	63,04	84,9	98,2
Médiane (Hz)	390	387	380
Minimum (Hz)	276	251	177
Maximum (Hz)	556	527	494

Tableau 12 : Fondamental Usuel Moyen (FoUm en Hz) de la voix d'appel pour les différents groupes.

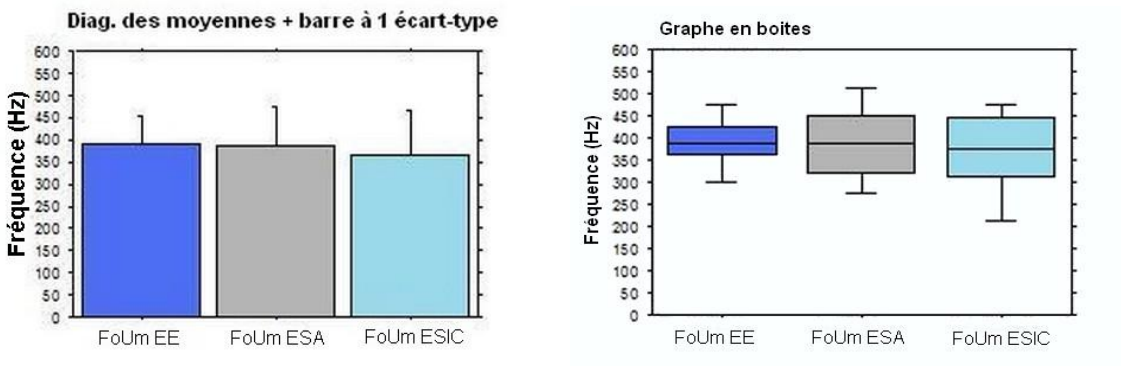


Figure 55 : Diagramme des moyennes et graphe en boites des médianes de FoUm (Hz) de la voix d'appel pour les groupes EE, ESA et ESIC.

6) Dynamique de la voix d'appel. Concernant la dynamique de la voix d'appel, on peut noter que les valeurs moyennes et de la médiane sont manifestement différentes entre les groupes EE et ESA. De même, on observe une différence entre les groupes ESA et ESIC. La dispersion est très importante chez la population EE (Tableau 13, Figure 56).

	Dyn EE	Dyn ESA	Dyn ESIC
Moyenne (Hz)	172	185	179
Ecart-type (Hz)	100,61	88,56	84,2
Médiane (Hz)	162	174	167
Minimum (Hz)	36	80	40
Maximum (Hz)	383	377	333

Tableau 13 : Dynamique (Hz) de la voix d'appel pour les différents groupes.

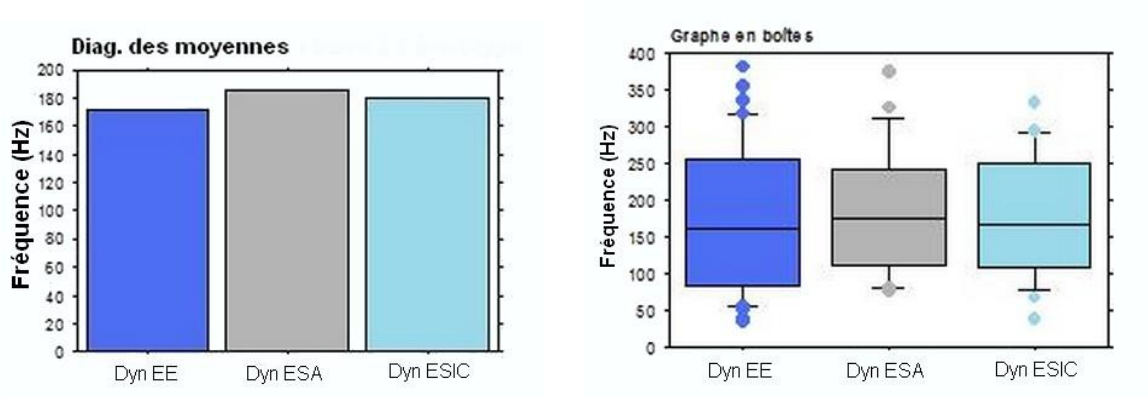


Figure 56 : Diagramme des moyennes et graphe en boites des médianes de la dynamique (Hz) de la voix d'appel pour les groupes EE, ESA et ESIC.

Concernant la voix d'appel, nous observons que :

- Les valeurs de FoUm sont différentes entre les groupes ESA et ESIC ainsi qu'entre les groupes EE et ESIC.
- Les valeurs de dynamique de la parole sont différentes entre les groupes EE et ESA.
- La dispersion est plus forte chez le groupe ESIC pour FoUm.
- La dispersion est plus forte chez le groupe EE pour la dynamique.
- La valeur de FoUm est nettement moins élevée pour le groupe ESIC.
- La valeur de dynamique de la parole est plus élevée pour le groupe ESA.

Concernant l'intensité de la voix conversationnelle, la voix forte et la voix d'appel, nous ne présenterons pas les résultats obtenus car nous avons constaté certaines incohérences dans les données. Ces incohérences ont été associées à un problème de calibrage de l'intensité mesurée. Par conséquent, ce paramètre acoustique ne sera pas pris en compte ici dans la présentation et la discussion des résultats.

Les résultats obtenus de l'analyse de la FoUm et de la dynamique de la parole émise pour les différents types de voix (conversationnelle, forte et d'appel) signalent diverses tendances. Les indicateurs statistiques descriptifs observés tels que la moyenne et la médiane ne se comportent pas de la même façon à l'intérieur de chaque paramètre. Pour certains paramètres, la valeur de la moyenne entre certains groupes est proche alors que si l'on observe la médiane, ce n'est pas du tout le cas. Par exemple, c'est le cas pour le paramètre de la dynamique de la voix conversationnelle. La valeur moyenne des groupes ESA et ESIC est proche tandis que la valeur de la médiane des mêmes groupes est différente à simple vue. Nous avons décidé de surveiller le comportement des valeurs moyennes. Cela nous a conduit par la suite à effectuer un test ANOVA et vérifier si les différences observées étaient significatives.

La littérature sur l'analyse des paramètres acoustiques de la voix conversationnelle infantile n'est pas abondante (Lenden et Flipsen, 2007)[148]. On trouve des données de la Fo analysée pour la production de syllabes comme /pa/ et /pi/ (Higgins et col. 2003)[82] ou pour la parole répétée comme étudié par Peterson et Barney (1952)[149]. Ces auteurs présentent une base de données qui portent sur 76 locuteurs américains (33 hommes, 28 femmes, 15 enfants) ayant

prononcé deux fois une liste de 10 mots contenant une voyelle sur laquelle ont été mesurés la Fo et les formants. Dans ce travail, ils reportent une Fo de 264 Hz (±35 Hz). Par rapport à ces données, le FoUm de notre population contrôle (232 Hz, ±33,15) est moins élevé. Néanmoins, à notre sens, ces valeurs restent quand même proches.

Par rapport à notre base de données normative, nous remarquons que le FoUm de la voix conversationnelle est plus élevé pour la population ESA. Nous ne sommes pas étonnés de trouver que le FoUm soit élevé chez l'ESA quand on sait que le comportement vocal de l'enfant sourd est caractérisé par le déplacement de l'espace tonal vers la zone aiguë. Par ailleurs, la dynamique de la parole chez l'ESA est observée comme étant plus étendue. Généralement, les courbes mélodiques des enfants sourds ne sont pas variées, elles sont plutôt plates (Vinter, 1994)[61]. Nous n'attendions pas que la dynamique de la parole de l'ESA soit plus étendue que chez l'EE. Cependant, une des explications que nous pouvons donner à ces résultats réside dans les caractéristiques de la surdité et de l'appareillage de la population ESA de notre étude. Ce sont des enfants qui ont été appareillés tôt, tout au début de l'acquisition du langage, et qui manifestent une surdité sévère de Type I.

Concernant la voix forte, nous trouvons que c'est le groupe ESIC qui manifeste une dynamique plus étendue que les autres populations. Nous en déduisons que ceci est également du aux caractéristiques de la surdité et de l'implantation du groupe. D'une part, le groupe ESIC est constitué notamment d'enfants manifestant une surdité profonde Type I. Vinter (1994) a observé que les courbes mélodiques complexes sont plus présentes chez l'enfant appareillé manifestant une surdité profonde de Type I que chez celui manifestant une surdité profonde Type de III. D'autre part, la plupart d'enfants du groupe ESIC a bénéficié d'une implantation précoce, c'est-à-dire au début de l'acquisition du langage. La particularité du traitement du signal de l'implant conjugué à une implantation précoce permet une excellente acquisition acoustique chez le jeune enfant sourd facilitant les jeux de variations mélodiques entrainant des contours mélodiques plus variés (Schauwers et col., 2004)[150].

Concernant la voix d'appel, le FoUm est moins élevé pour le groupe ESIC et la dynamique de la parole aiguë est plus élevée pour le groupe ESA.

Dans l'ensemble, nous avons observé que le FoUm varie pour tous les groupes en fonction du type de voix émise, augmentant progressivement de la voix conversationnelle à la voix d'appel. Cette adaptation de hauteur selon l'intention communicative est tout à fait normale et

semble être développée chez les groupes ESA et ESIC. Enfin, nous avons noté des différences entre les groupes du FoUm et de la dynamique de la parole quasiment pour tous les types de voix.

Etant donné que nous étions intéressés à analyser la stabilité de la fréquence et de l'amplitude du signal sonore, nous avons analysé la fréquence fondamentale et ses coefficients de stabilité à moyen et à court terme à partir d'un /a/ tenu. Ces paramètres sont présentés dans les items 7, 8, 9. De même, nous avons analysé l'intensité et ses coefficients de stabilité à moyen et à court terme ainsi que le bruit ajouté au cours de la phonation indiqué par le rapport harmoniques sur bruit. Ces paramètres sont présentés dans les items 10, 11, 12 et 13. Les résultats de l'analyse de ces paramètres sont présentés ensuite.

7) Fréquence fondamentale ou Fo d'un /a/ tenu. Les valeurs moyennes et de la médiane de la Fo s'avèrent semblables entre les groupes ESA et ESIC. On observe une différence entre ces derniers et le groupe EE. On observe une dispersion importante chez le groupe ESIC (Tableau 14, Figure 57).

	Fo EE	Fo ESA	Fo ESIC
Moyenne (Hz)	233	246	246
Ecart-type (Hz)	27,45	50,9	69,42
Médiane (Hz)	231	251	252
Minimum (Hz)	166	98	119
Maximum (Hz)	304	333	362

Tableau 14 : Fréquence fondamentale ou Fo (Hz) d'un /a/ tenu pour les groupes.

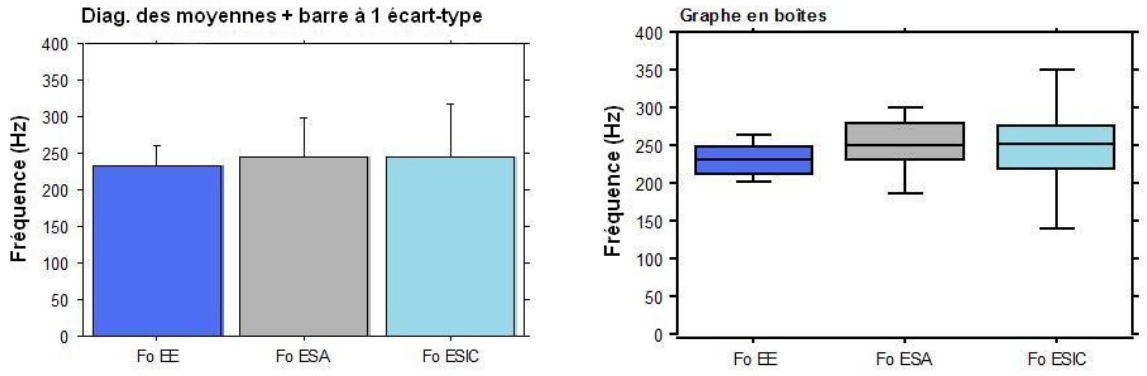


Figure 57 : Diagramme des moyennes et graphe en boites des médianes de Fo (Hz) d'un /a/ tenu pour les groupes EE, ESA et ESIC.

8) Coefficient de variation de Fo d'un /a/ tenu. Nous avons observé une valeur moyenne du coefficient de variation ESA plus élevée que celui du groupe ESIC pour la même Fo (246 Hz). On note une différence entre les groupes EE et ESA, ainsi qu'entre les groupes ESA et ESIC. On observe également une dispersion très importante chez le groupe ESA (Tableau 15, Figure 58).

	Coeff var Fo EE	Coeff var Fo ESA	Coeff var Fo ESIC
Moyenne (%)	0,79	1,62	1,25
Ecart-type (%)	0,26	0,76	0,61
Médiane (%)	0,77	1,47	1,2
Minimum (%)	0,45	0,37	0,52
Maximum (%)	1,72	3,66	2,97

Tableau 15 : Coefficient de variation de Fo (%) d'un /a/ tenu pour les trois populations.

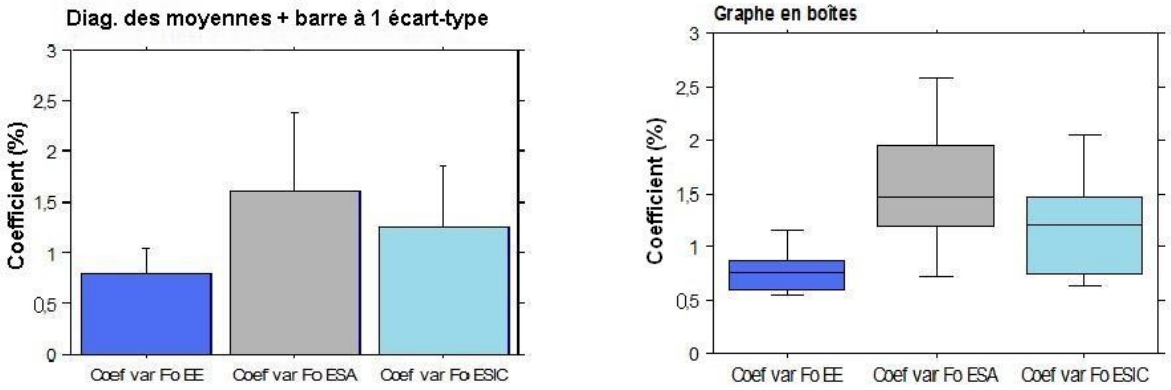


Figure 58 : Diagramme des moyennes et graphe en boites des médianes du Coefficient de variation de Fo (%) d'un /a/ tenu pour les groupes EE, ESA et ESIC.

9) Jitter factor d'un /a/ tenu. Les valeurs moyennes de ce paramètre sont différentes pour tous les groupes. La valeur la plus faible est observée pour EE puis ESIC puis ESA. La population EE se repartit de manière symétrique autour de la valeur 0,56. On note une dispersion très importante dans les groupes ESA et ESIC par rapport au groupe EE (Tableau 16, Figure 59).

	Jitter EE	Jitter ESA	Jitter ESIC
Moyenne (%)	0,56	1,67	0,92
Ecart-type (%)	0,16	0,97	0,83
Médiane (%)	0,56	1,63	0,57
Minimum (%)	0,24	0,27	0,2
Maximum (%)	0,94	3,72	2,53

Tableau 16 : Jitter factor (%) d'un /a/ tenu pour les trois groupes de population.

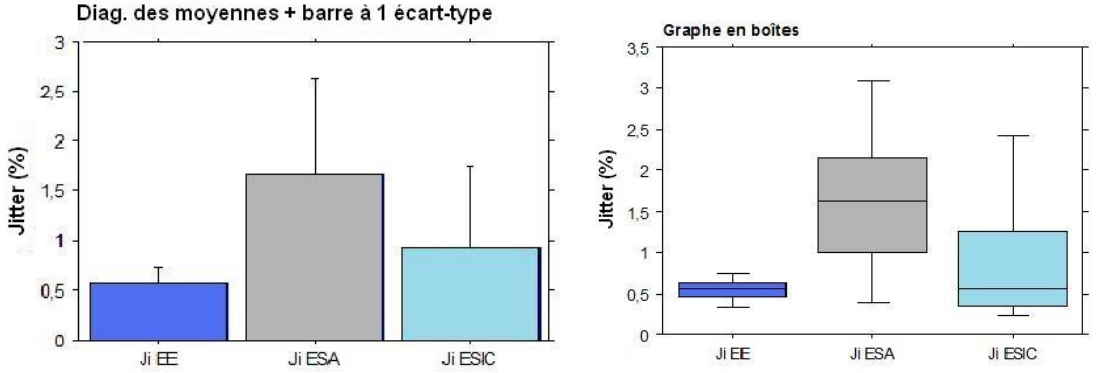


Figure 59 : Diagramme des moyennes et graphe en boites des médianes de Jitter factor (%) d'un /a/ tenu pour les populations EE, ESA et ESIC.

En résumé, nous observons concernant la Fo que :

- Il y a une différence du groupe EE aux groupes ESA et ESIC.
- Il y a des différences entre tous les groupes pour le coefficient de variation.
- Il y a des différences entre tous les groupes pour le Jitter.
- La dispersion est plus élevée chez le groupe ESIC pour Fo.
- La dispersion est plus forte chez le groupe ESA pour le coefficient de variation et le Jitter.
- Les valeurs de Fo, du coefficient de variation et du Jitter sont plus élevés chez les groupes ESA puis ESIC.

10) Intensité d'un /a/ tenu. L'intensité est légèrement différente pour les trois groupes. Cette différence est plus marquée entre les groupes EE et ESA où l'intensité est la plus faible. Le groupe ESA montre une dispersion plus accentuée par rapport aux autres groupes (Tableau 17, Figure 60).

	Int EE	Int ESA	Int ESIC
Moyenne (dB)	64	60	63
Ecart-type (dB)	2,27	4,82	3,4
Médiane (dB)	64	61	63
Minimum (dB)	59	52	55
Maximum (dB)	67	68	69

Tableau 17 : Intensité moyenne (dB) d'un /a/ tenu pour les trois populations.

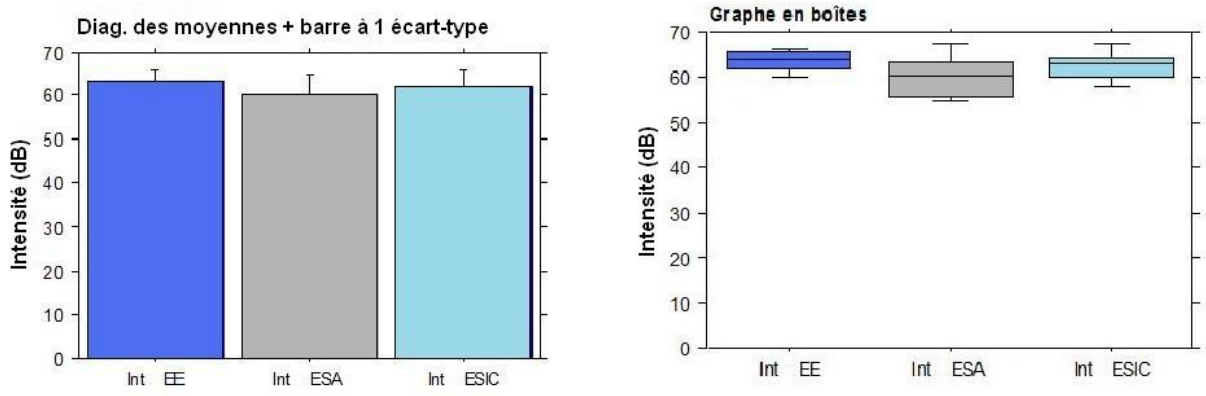


Figure 60 : Diagramme des moyennes et graphe en boites des médianes de l'intensité moyenne (dB) d'un /a/ tenu pour les populations EE, ESA et ESIC.

11) Coefficient de variation de l'Intensité d'un /a/ tenu. Les valeurs moyennes de ce paramètre se montrent élevées particulièrement dans le groupe ESA puis ESIC. On observe une différence entre les moyennes des groupes EE et ESA, ainsi qu'entre les groupes ESA et ESIC. La moyenne et la médiane du groupe ESA se manifestent comme étant les valeurs les plus élevées. On note une forte dispersion des groupes ESA et ESIC dans les 3ème et 4ème quartiles (Tableau 18, Figure 61).

	Coeff var Int EE	Coeff var Int ESA	Coeff var Int ESIC
Moyenne (%)	0,74	1,38	1,08
Ecart-type (%)	0,25	0,72	0,52
Médiane (%)	0,7	1,25	0,99
Minimum (%)	0,36	0,52	0,35
Maximum (%)	1,48	3,21	2,06

Tableau 18 : Coefficient de variation de l'intensité moyenne (%) d'un /a/ tenu pour les trois populations.

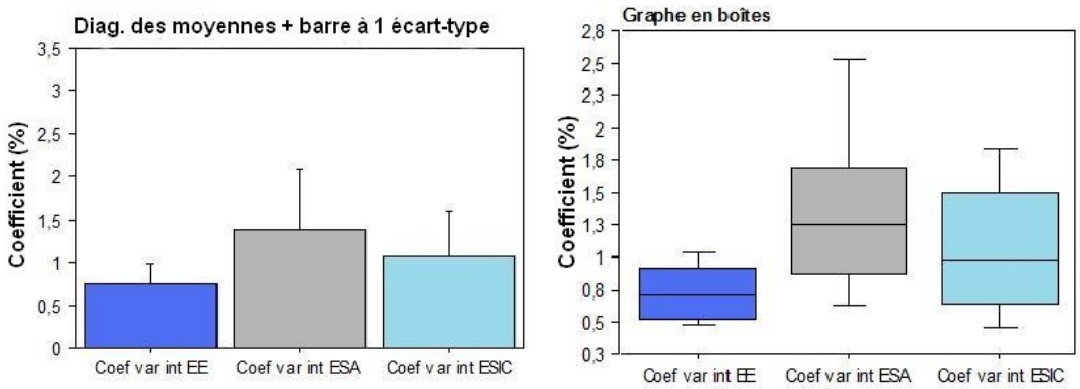


Figure 61 : Diagramme des moyennes et graphe en boites des médianes de Coefficient de variation de l'intensité moyenne (%) d'un /a/ tenu pour les groupes EE, ESA et ESIC.

12) Shimmer factor d'un /a/ tenu. Les valeurs moyennes du paramètre *Shimmer* sont proches pour les groupes EE et ESA. En revanche, une différence importante existe entre les groupes ESA et ESIC (Tableau 19, Figure 62).

	Shimmer EE	Shimmer ESA	Shimmer ESIC
Moyenne (%)	0,53	0,58	0,4
Ecart-type (%)	0,23	0,25	0,18
Médiane (%)	0,52	0,57	0,42
Minimum (%)	0,15	0,21	0,13
Maximum (%)	0,96	1,35	0,81

Tableau 19 : Shimmer factor (%) d'un /a/ tenu pour les trois populations.

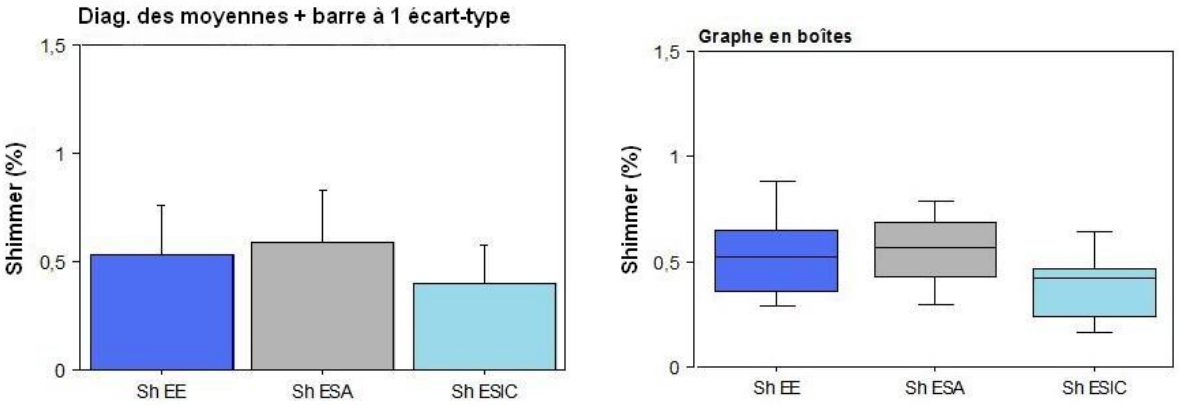


Figure 62 : Diagramme des moyennes et graphe en boites des médianes de Shimmer factor (%) d'un /a/ tenu pour les groupes EE, ESA et ESIC.

13) Rapport Harmoniques/bruit d'un /a/ tenu. Les valeurs du rapport harmoniques/bruit sont comparables entre les groupes EE et ESIC. En revanche, il existe une différence plus marquée entre ces derniers et le groupe ESA. On observe une dispersion importante pour le groupe ESIC (Tableau 20, Figure 63).

	R H/B EE	R H/B ESA	R H/B ESIC
Moyenne (dB)	17,16	11,21	16,84
Ecart-type (dB)	2,41	5,11	6,14
Médiane (dB)	16,72	10,27	17,96
Minimum (dB)	12,37	5,43	5,4
Maximum (dB)	20,71	23,74	26,71

Tableau 20 : Rapport Harmoniques/Bruit (dB) d'un /a/ tenu pour les trois populations.

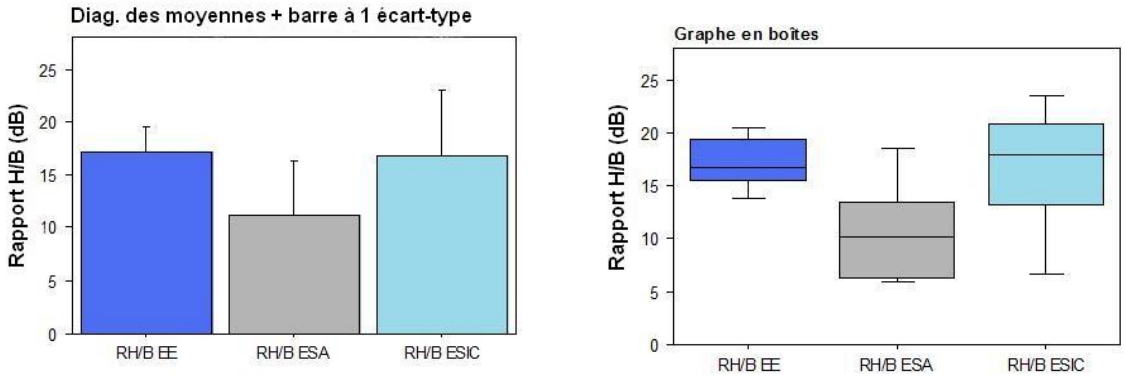


Figure 63 : Diagramme des moyennes et graphe en boites des médianes de Rapport Harmoniques/Bruit (dB) d'un /a/ tenu pour les groupes EE, ESA et ESIC.

En résumé, nous observons concernant l'intensité que :

- Les différences sont légères entre les groupes pour l'intensité.
- Il y a des différences entre tous les groupes pour le coefficient de variation.
- Il y a des différences entre les groupes ESA et ESIC pour le Shimmer.
- Il y a des différences entre les groupes EE et ESA ainsi qu'entre ESA et ESIC pour RHB (rapport Harmoniques/Bruit).
- La dispersion est plus élevée chez le groupe ESA pour l'intensité, le coefficient de variation et le Shimmer.
- La dispersion est plus forte chez le groupe ESIC pour RHB.
- Les valeurs d'intensité et de RHB sont plus faibles chez le groupe ESA.
- Les valeurs du coefficient de variation et Shimmer sont plus élevées chez le groupe ESA.

Afin d'explorer les possibilités vocales de l'enfant, nous avons évalué l'étendue vocale physiologique* et la tessiture. Ces paramètres sont présentés dans les items 14 et 15. Nous avons également analysé la qualité de la fermeture glottique par le paramètre T.M.P. (Temps Maximale de Phonation). Ce paramètre est présenté dans l'item 16. Les résultats de l'analyse de ces paramètres sont présentés plus loin.

* On distingue deux façons d'étudier l'étendue vocale : soit on ne tient pas compte de la qualité vocale, il s'agit de *l'étendue vocale physiologique* ; soit on ne prend en compte que les notes qui possèdent une qualité musicale minimale acceptable, il s'agit alors de *l'étendue vocale musicale*.

14) Dynamique de l'étendue vocale. L'étendue vocale est différente entre le groupe EE et ESA ; différence encore plus marquée entre les groupes ESA et ESIC. En effet, on observe que la dynamique de l'étendue vocale du groupe ESA se trouve étendue vers les sons très aigus et les sons graves, alors que la dynamique des groupes EE et ESIC reste dans un registre plutôt médium et aigu. L'écart est très important entre le 3eme quartile et la valeur maximale du groupe ESIC (Tableau 21, Figure 64). Les lignes (graphe à droite) représentent l'écart fréquentiel entre les sons le plus aigu et le plus grave que l'enfant est capable d'émettre.

	Dyn EE	Dyn ESA	Dyn ESIC
Moyenne (Hz)	431	475	393
Ecart-type (Hz)	146,64	193,6	206,83
Médiane (Hz)	426	451	334
Minimum (Hz)	134	49	88
Maximum (Hz)	769	860	762

Tableau 21 : Dynamique (Hz) de l'étendue vocale pour les trois populations.

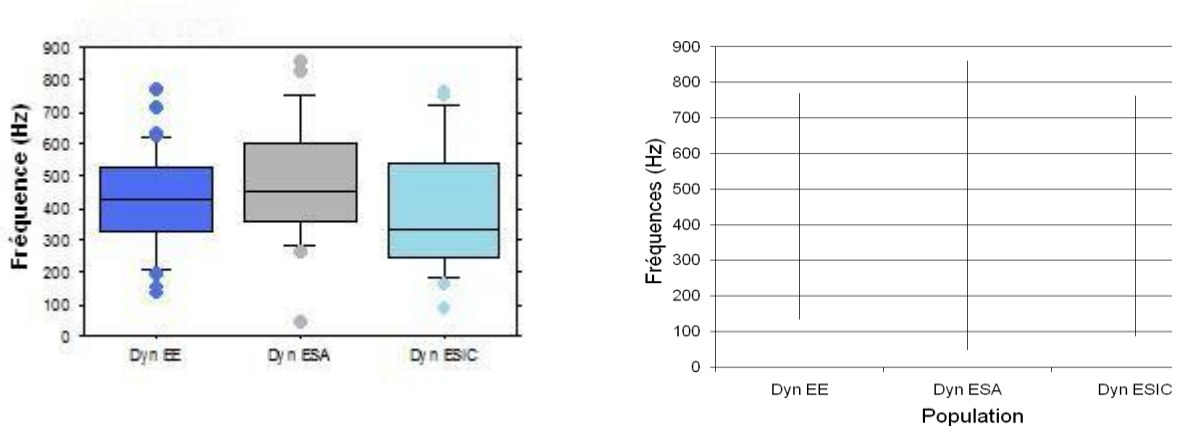


Figure 64 : Graphe en boite des médianes et graphe en ligne de l'étendue vocale (Hz) pour les groupes EE, ESA et ESIC.

15) Dynamique de la tessiture. Les valeurs moyennes de l'écart fréquentiel entre le son grave et le son aigu émis confortablement sont semblables entre les groupes EE et ESIC. Il existe une différence plus importante entre ces groupes et ESA. En effet, on observe que la dynamique de la tessiture du groupe ESA se trouve étendue vers les sons très aigus, alors que la dynamique des groupes EE et ESIC reste dans un registre plutôt médium et aigu. L'écart est très important entre le 3eme quartile et les valeurs les plus grandes du groupe ESA (Tableau 22, Figure 65). Les lignes (graphe à droite) représentent l'écart fréquentiel entre les sons le plus aigu et le plus grave émis avec aisance par l'enfant.

	Dyn EE	Dyn ESA	Dyn ESIC
Moyenne (Hz)	284	403	287
Ecart-type (Hz)	140,88	251,53	148,56
Médiane (Hz)	261	328	283
Minimum (Hz)	81	64	61
Maximum (Hz)	669	994	594

Tableau 22 : Dynamique (Hz) de la tessiture pour les trois populations.

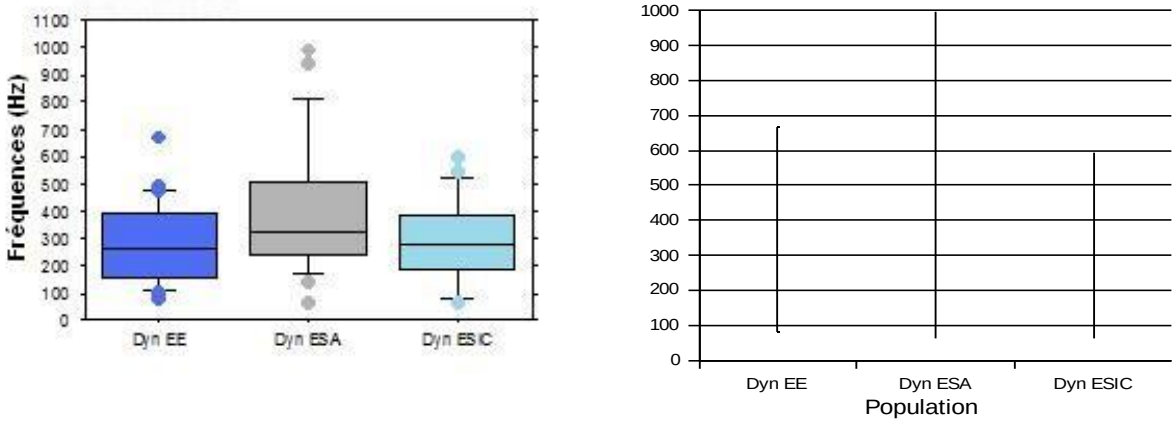


Figure 65 : Graphes en boites des médianes et graphe en ligne de la tessiture vocale (Hz) pour les groupes EE, ESA et ESIC.

16) Temps Maximal de Phonation ou TMP. Le groupe ESIC montre un TMP légèrement plus élevé avec une médiane plus importante. La dispersion est comparable pour les groupes EE et ESIC et est plus accentuée que pour le groupe ESA (Tableau 23, Figure 66).

	TMP EE	TMP ESA	TMP ESIC
Moyenne (sec)	13	12	14
Ecart-type (sec)	4,68	3,82	4,64
Médiane (sec)	12	11	15
Minimum (sec)	5	6	6
Maximum (sec)	23	19	25

Tableau 23 : Temps Maximal de Phonation (TMP en sec) pour les trois groupes de population.

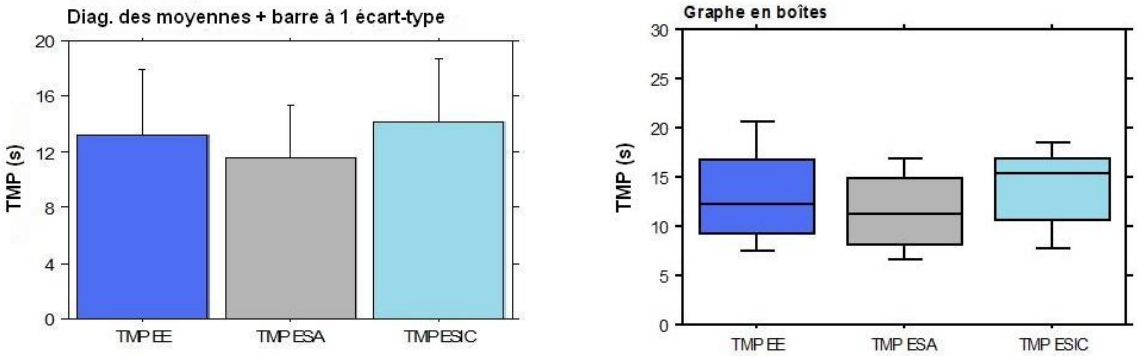


Figure 66 : Diagramme des moyennes et graphe en boites des médianes de Temps Maximal de Phonation (TMP en sec) pour les groupes EE, ESA et ESIC.

En résumé, nous observons que :

- Il y a des différences entre les groupes EE et ESA ainsi qu'entre ESA et ESIC pour l'étendue vocale et la tessiture.
- La dispersion est plus élevée chez le groupe ESIC pour l'étendue vocale.
- La dispersion est plus élevée chez le groupe ESA pour la tessiture.
- La dispersion est plus élevée chez le groupe EE pour le TMP.
- Les valeurs de l'étendue vocale et de la tessiture sont plus élevées chez le groupe ESA.
- Les valeurs du TMP sont proches pour tous les groupes.

Discussion :

Les résultats de notre étude montrent une augmentation de la Fo chez les groupes ESA et ESIC par rapport au groupe contrôle. Concernant ce paramètre, la littérature signale des résultats divergents. D'une part, certains auteurs trouvent une tendance à la diminution de la Fo chez l'enfant implanté cochléaire (Hamzavi et col., 2000[151] ; Szyfer et col., 1996[152] ; Leder et col., 1987[153] ; Hocevar-Boltezar et col., 2006)[81]. D'autres auteurs reportent une Fo plus élevée que la normale après plusieurs années d'implantation (Kirk et col., 1983[154] ; Perrin et col., 1995[155] ; Monini et col., 1997[156] ; Higgins et col. 2003[82]). Concernant le groupe d'enfants implantés, nous estimons que les valeurs moyennes que nous avons obtenues sont élevées à cause d'une forte dispersion de la population (Figure 57). On constate que les valeurs extrêmes sont très écartées pour les groupes ESA et ESIC notamment pour ce dernier.

Nous avons trouvé que les valeurs du coefficient de variation de la Fo et du Jitter sont à la frontière de la normalité pour le groupe d'enfants sourds (ESA et ESIC). Ces valeurs sont plus élevées chez le groupe ESA que chez le groupe ESIC. L'augmentation de ces paramètres chez le groupe ESA peut expliquer la présence d'une légère raucité correspondant bien à la perception de dysphonie. Par ailleurs, Campisi et col. (2005)[93] ont observé une augmentation du coefficient de variation de la Fo et du Jitter dans les profils vocaux des enfants sourds avant et après l'implantation. Campisi souligne que le coefficient de variation de la Fo peut rester perturbé jusqu'à six mois après l'activation de l'implant cochléaire. Néanmoins, Hartl et col. (2005[140]) préconisent que le Jitter et le Shimmer peuvent être également élevés chez les sujets normaux selon l'émission vocale effectuée. Ils déclarent que ces mesures ne sont pas spécifiques à la dysphonie et ne sont pas un caractère perceptif particulier. Elles ne seraient utiles que pour les échantillons vocaux comportant peu d'irrégularités comme pour les voix normales ou peu dysphoniques (Titze, 1994[157]). Compte-tenu de ces arguments, nous remettons donc en question l'interprétation des résultats que nous avons formulée tout au début pour nos groupes de population. De ce fait, nous envisageons deux hypothèses que nous confirmerons par l'analyse perceptive présentée ultérieurement : soit les résultats que nous avons obtenus s'avèrent fiables du fait que les échantillons vocaux des groupes ESA et ESIC sont normaux ou peu dysphoniques, soit les résultats sont biaisés à cause de la nature des échantillons comportant beaucoup d'irrégularités typiquement des voix dysphoniques. Toutefois, nos résultats montrent que les valeurs du coefficient de variation de F0 et de Jitter sont plus élevées chez le groupe ESA que chez l'ESIC. Ces résultats suggèrent une différence inter-groupe qui réside dans la différence de prothèse auditive portée par chaque groupe. Enfin, nous avons identifié des différences entre les groupes pour tous les paramètres.

Concernant l'intensité, les valeurs se trouvent plutôt normales pour les groupes ESA et ESIC. Van Lierde et col. (2005)[158] reportent une valeur moyenne de l'ordre de 55,33 dB pour une intensité minimale et de 97 dB pour une intensité maximale. Ces valeurs sont observées pour une population d'enfants implantés âgés entre 5 et 13 ans (âge moyen= 8 ans) ayant été implantés à l'âge de 2 ans jusqu'à 11 ans. Les valeurs normales de l'intensité présentées dans la littérature oscillent entre 60 et 70 dB pour la voix conversationnelle. Ces valeurs données à titre indicatif ont été mesurées avec un sonomètre placé à 1 mètre de distance sur le plan horizontal de la bouche d'un locuteur (Dalleas, 1987)[159].

Quant au coefficient de variation d'intensité, les valeurs sont à la limite de la norme pour les groupes ESA et ESIC. D'ailleurs, la valeur de ce coefficient est plus élevée chez le groupe

ESA que chez le groupe ESIC. Cette élévation du coefficient indique une instabilité de l'amplitude du signal sonore à moyen terme chez l'ESA. Par ailleurs, les résultats obtenus de l'analyse du Shimmer signalent des valeurs normales pour tous les groupes. Ces résultats suggèrent que chez l'ESA, il y a d'une part, une amélioration de la variation d'amplitude entre les cycles d'oscillation de l'émission d'un /a/ tenu (Shimmer) et d'autre part, il y a une persistance de la variabilité de l'amplitude du signal par rapport à l'intensité moyenne. Le même phénomène a lieu chez l'ESIC mais dans une proportion moins importante.

Quant au rapport Harmoniques/Bruit (RHB), nous avons obtenu une valeur moins élevée pour le groupe ESA. Nous en déduisons que l'énergie de l'onde sonore du /a/ tenu de cette population est caractérisée par une importante énergie de bruit résiduel qui se traduit perceptivement par une impression de « souffle ». Toutefois, les valeurs obtenues pour le groupe ESA est proche de la norme.

Concernant l'étendue et la tessiture vocale, le groupe ESA est capable d'émettre de manière générale une gamme de fréquences plus large que les autres groupes. La tessiture est caractérisée par l'émission notamment des fréquences aiguës allant jusqu'à 1 000 Hz environ. A première vue, ces résultats paraissent étonnants voire inconsistants. Néanmoins, de la même façon que pour la dynamique de la parole, la population ESA manifeste d'excellentes possibilités vocales. D'une part, on peut expliquer la cohérence de ces résultats du fait que la population ESA est constituée d'enfants manifestant notamment une surdité sévère Type I et que l'appareillage s'est effectué relativement tôt pour ces enfants. Les travaux de Vinter (1989[92], 1994[61]) ont démontré que l'appareillage précoce, avant l'âge d'acquisition du langage permettait d'améliorer les performances sur les schémas mélodiques et d'élargir la dynamique vocale. D'autre part, nous expliquons ces résultats du fait de la facilité de l'exécution de la consigne d'évaluation.

Enfin, concernant le TMP, les valeurs apparaissent normales pour tous les groupes (Finnegan, 1985)[160]. Les valeurs sont proches entre les groupes.

Après avoir effectué une analyse statistique descriptive des paramètres acoustiques de notre étude pour l'ensemble de l'échantillon (n = 78), nous avons identifié des différences et des ressemblances entre les valeurs moyennes des groupes pour tous les paramètres (16 au total). Les paramètres pour lesquels nous avons eu des doutes pour déterminer une différence des moyennes entre les groupes ont été classés dans la colonne « incertains ». Nous les avons

listées et classées dans le tableau ci-dessous (Tableau 24). On notera entre parenthèses les groupes pour lesquels ces paramètres se manifestent semblables, différents ou incertains.

Corpus	Semblables	Différents	Incertains
Voix conversationnelle	FoUm *(ESA≈ESIC)*	FoUm *(EE vs. ESA/ESIC)* Dyn *(EE vs. ESA)*	Dyn *(ESIC vs.EE/ESA)*
Voix forte	Dyn *(EE≈ESA)*	Dyn *(ESIC vs.EE/ESA)*	FoUm *(EE vs. ESA vs. ESIC)*
Voix d'appel	FoUm *(EE≈ESA)*	FoUm *(EE/ESA vs. ESIC)* Dyn *(EE vs. ESA)*	Dyn *(ESIC vs.EE/ESA)*
/a/ tenu + Vocalise	Fo *(ESA≈ESIC)*	Fo *(EE vs. ESA/ESIC)* Coeff var Fo *(EE vs. ESA/ESIC)* Jitter *(EE vs. ESA/ESIC)*	Coef. var Fo *(ESA vs. ESIC)* Jitter *(ESA vs. ESIC)*
	Intensité *(EE≈ESIC)* R H/B *(EE≈ESIC)*	Intensité *(EE vs. ESA)* Coeff var Int *(EE vs. ESA)* Shimmer *(ESIC vs. EE/ESA)* R H/B *(EE/ESIC vs. ESA)*	Intensité *(ESA vs. ESIC)* Coef. var Int *(ESIC vs. EE/ESA)* *Shimmer* *(EE vs. ESA)*
	Tessiture *(EE≈ESIC)*	Dyn Tessiture *(EE vs. ESA ; ESA vs. ESIC)* Dyn Etendue *(EE/ESA vs. ESIC)*	Dyn Etendue *(EE vs. ESA)*
		TMP *(ESA vs. ESIC)*	TMP *(EE vs. ESA/ESIC)*

Tableau 24 : Classement des paramètres acoustiques avant l'application des tests ANOVA.

Afin de confirmer nos hypothèses concernant ce classement, nous avons testé la significativité de ces différences par un test ANOVA. Le choix de ce test réside essentiellement dans le fait que nos données sont de nature « quantitative-continue ». L'ANOVA est un outil statistique classique qui permet de comparer les moyennes des différentes populations. C'est un outil robuste quant à l'hypothèse de normalité de la distribution des données.

Nous avons effectué deux ANOVA. Le premier test ANOVA (ANOVA 1). Notre hypothèse nulle (Ho) était : « il n'y a pas de différence entre les moyennes des populations EE et celles des populations ESA et ESIC ». Donc : EE = ESA; EE = ESIC. Notre hypothèse alternative (H1) était : « il y a au moins une différence entre les moyennes des populations EE et ESA, et/ou EE et ESIC. Donc au moins l'une des différences EE < > ESA; EE < > ESIC est vérifiée. L'établissement de la significativité de ces différences permettait d'inférer s'il y avait un effet de groupe pour les paramètres en question, notamment pour ceux qui se trouvaient dans la colonne « différents » et « incertains » du Tableau 24.

Les résultats du test ANOVA 1 indiquent que ce sont plutôt les paramètres acoustiques analysés sur la fréquence et l'intensité d'un /a/ tenu qui se montrent plus sensibles à l'effet de

population (Tableau 25). Dans le tableau 25, on note en bleu, les paramètres acoustiques étant significatifs et, en orange, les paramètres signalant une tendance à la significativité (le seuil de significativité étant défini à 5%). Lorsqu'on compare la population contrôle EE et la population appareillée ESA, il apparait une différence significative pour 6 sur 16 paramètres; 5 parmi elles étant très significatives. Lorsqu'on compare la population contrôle EE et la population implantée ESIC, on note une différence significative pour 4 sur 16 paramètres ; 1 étant très significatif, et, 1 se trouvant à la limite du seuil établi. De manière globale, on note que les différences *EE vs. ESIC* sont moins nettes que celles *EE vs. ESA*. Ces premiers résultats nous ont permis de constater désormais une erreur dans le classement initial (cf. Tableau 24).

	ANOVA 1			
	EE vs ESA		EE vs ESIC	
Paramètre	**Signif**	**p.c.**	**Signif**	**p.c.**
Voix conv FoUm	NON	0.209	NON	0.297
Voix conv Dyn	NON	0.497	NON	0.668
Voix forte FoUm	NON	0.814	NON	0.648
Voix forte Dyn	NON	0.997	NON	0.0803
Voix d'appel FoUm	NON	0.865	NON	0.272
Voix d'appel Dyn	NON	0.614	NON	0.770
Fo /a/	NON	0.325	NON	0.328
Co Var Fo	OUI	2.06e-07	OUI	0.00203
Jitter	OUI	3.10e-08	OUI	0.0457
Intensité	OUI	0.000551	NON	0.220955
Co Var Int	OUI	7.25e-06	OUI	0.0121
Shimmer	NON	0.4098	OUI	0.0287
R H/B	OUI	4.6e-06	NON	0.791
TMP	NON	0.172	NON	0.477
Tessiture	OUI	0.0171	NON	0.9572
Etendue	NON	0.362	NON	0.447

Tableau 25 : Résultats du test ANOVA 1.

Ensuite, nous avons effectué un deuxième test ANOVA (ANOVA 2). Notre hypothèse nulle (Ho) était : « il n'y a pas de différence entre les moyennes des populations ESA et ESIC ». Donc : ESA = ESIC. Notre hypothèse alternative (H1) était : « il y a une différence entre les moyennes des groupes ESA et ESIC ». Donc : ESA < > ESIC. L'établissement de la significativité de cette différence permettait d'inférer, comme pour le test ANOVA 1, s'il y avait un effet de groupe pour les paramètres en question, notamment pour ceux qui se trouvaient dans la colonne « différents » et « incertains » du Tableau 24.

Il ressort de ce test que le nombre de paramètres acoustiques sensibles à l'effet de groupe est plus faible. Seulement 3 paramètres montrent des différences significatives (Tableau 26) dont 2 très significatives; 1 significative et 2 paramètres à la limite du seuil de 5%.

	ESA vs ESIC	
Paramètre	**Signif**	**p.c.**
Voix conv FoUm	NON	0.875
Voix conv Dyn	NON	0.823
Voix forte FoUm	NON	0.867
Voix forte Dyn	NON	0.201
Voix d'appel FoUm	NON	0.487
Voix d'appel Dyn	NON	0.841
Fo /a/	NON	0.997
Co Var Fo	NON	0.103
Jitter	OUI	0.0127
Intensité	NON	0.102
Co Var Int	NON	0.141
Shimmer	OUI	0.0093
R H/B	OUI	0.00314
TMP	NON	0.0619
Tessiture	NON	0.0819
Etendue	NON	0.205

Tableau 26 : Résultats du test ANOVA 2.

Un tableau récapitulatif des résultats obtenus des tests ANOVA est présenté ci-dessous (Tableau 27).

	ANOVA 1				ANOVA 2	
	EE vs ESA		**EE vs ESIC**		**ESA vs ESIC**	
Paramètre	**Signif**	**p.c.**	**Signif**	**p.c.**	**Signif**	**p.c.**
Voix conv FoUm	NON	0.209	NON	0.297	NON	0.875
Voix conv Dyn	NON	0.497	NON	0.668	NON	0.823
Voix forte FoUm	NON	0.814	NON	0.648	NON	0.867
Voix forte Dyn	NON	0.997	NON	0.0803	NON	0.201
Voix d'appel FoUm	NON	0.865	NON	0.272	NON	0.487
Voix d'appel Dyn	NON	0.614	NON	0.770	NON	0.841
Fo /a/	NON	0.325	NON	0.328	NON	0.997
Co Var Fo	OUI	2.06e-07	OUI	0.00203	NON	0.103
Jitter	OUI	3.10e-08	OUI	0.0457	OUI	0.0127
Intensité	OUI	0.000551	NON	0.220955	NON	0.102
Co Var Int	OUI	7.25e-06	OUI	0.0121	NON	0.141
Shimmer	NON	0.4098	OUI	0.0287	OUI	0.0093
R H/B	OUI	4.6e-06	NON	0.791	OUI	0.00314
TMP	NON	0.172	NON	0.477	NON	0.0619
Tessiture	OUI	0.0171	NON	0.9572	NON	0.0819
Etendue	NON	0.362	NON	0.447	NON	0.205

Tableau 27 : ANOVA 1 et 2.

De ces tests statistiques, il ressort que :

- Il existe des différences significatives entre le groupe contrôle EE et les groupes d'enfants sourds (ESA et ESIC) pour les paramètres acoustiques signalant une instabilité de la fréquence et de l'intensité à court et moyen terme du signal acoustique: Coefficient de variation de F0, Jitter et Coefficient de variation d'intensité.

- Il existe des différences significatives entre la population d'enfants sourds ESA et ESIC pour les paramètres acoustiques signalant une instabilité de la fréquence et de l'amplitude du signal acoustique : Jitter, Shimmer et Rapport H/B. Les valeurs sont plus élevées chez le groupe ESA que chez le groupe ESIC.

- Les paramètres évalués sur un corpus de parole ne signalent aucune différence significative inter-groupes. Nous en induisons que le corpus utilisé pour l'analyse acoustique instrumentale n'est pas adéquat ni sensible à la méthode de prise de mesures employée dans notre étude.

- Nous avons pu constater des erreurs dans le classement initial des paramètres acoustiques qui paraissaient semblables ou différents inter-groupes d'après une observation naïve purement visuelle à partir des graphes obtenus du traitement statistique descriptif (cf. Tableau 24). Nous proposons un nouveau tableau récapitulatif ci-dessous (Tableau 28).

Corpus	Paramètres Semblables	Paramètres Différents
Voix conversationnelle	FoUm Dyn	*Aucun*
Voix forte	FoUm Dyn	*Aucun*
Voix d'appel	FoUm Dyn	*Aucun*
/a/ tenu + Vocalise	Fo Coeff var Fo *(EE ≈ ESIC)*	Coeff var Fo *(EE vs. ESA; EE vs. ESIC)* Jitter *(EE vs. ESA; EE vs. ESIC; ESA vs.ESIC)*
	Int *(EE ≈ ESIC; ESA ≈ ESIC)* Coeff var Int *(ESA ≈ ESIC)* Shimmer *(EE ≈ ESA)* R H/B *(EE ≈ ESIC)*	Intensité *(EE vs. ESA)* Coeff var Int *(EE vs. ESA; EE vs. ESIC)* Shimmer *(EE vs. ESIC; ESA vs. ESIC)* R H/B *(EE vs. ESA; ESA vs. ESIC)*
	Tessiture *(EE ≈ ESIC)* Etendue	Tessiture *(EE vs. ESA; ESA vs. ESIC))*
	TMP	*Aucun*

Tableau 28 : Classement des paramètres acoustiques semblables ou différents après l'application des tests ANOVA 1 et 2.

3.4. Conclusion

L'analyse acoustique nous a permis de décrire – sur le plan acoustique - les voix des enfants sourds de notre étude et ce, de manière objective et quantitative. De cette analyse, il ressort que :

- ✓ Trois paramètres acoustiques sont discriminants entre le groupe d'enfants sourds et le groupe d'enfants entendants, à savoir : coefficient de variation de F0, Jitter et Coefficient de variation d'intensité. Ces paramètres témoignent d'une instabilité de la fréquence du signal à court et moyen terme mais également une instabilité de l'intensité à moyen terme.

- ✓ Trois paramètres acoustiques sont discriminants entre le groupe d'enfants appareillés et le groupe d'enfants implantés cochléaires, à savoir : Jitter, Shimmer et Rapport H/B. Ces paramètres signalent une instabilité de la fréquence et de l'intensité à court terme ainsi qu'une présence de bruit accompagnant le signal plus accentuées chez le groupe d'appareillés que chez le groupe d'implantés.

- ✓ Il existe un effet inter-groupe de la prothèse auditive portée par les enfants sur les caractéristiques acoustiques de la voix. Les voix du groupe qui porte un implant cochléaire sont sensiblement comparables à celles du groupe contrôle.

- ✓ Enfin, l'ensemble de résultats de l'analyse acoustique suggère que la voix des enfants sourds ne présente pas les caractéristiques traditionnellement retenues pour déterminer la voix dysphonique. En effet, nous n'avons pas identifié un nombre de paramètres acoustiques altérés suffisamment « robustes » (Jitter, Shimmer, Rapport H/B, Etendue vocale, TMP et F0) nous permettant de caractériser les voix comme étant pathologiques.

4. Analyse aérodynamique (Expérience 2)

4.1. Objectifs

Puisque les résultats de l'analyse acoustique n'ont pas confirmé notre hypothèse que les voix des enfants sourds manifestaient des caractéristiques pathologiques, nous avons tenté d'explorer la fonction laryngée de la population d'étude et d'obtenir par là d'autres indices de la qualité vocale de manière indirecte. Cela nous permettrait d'évaluer l'utilisation laryngée de l'air pulmonaire pour la fonction du voisement et de la parole. A ce propos, Higgins et col. (2003[82]) signalent que les mesures physiologiques peuvent révéler des anomalies qui ne seraient pas mises en évidence à partir des mesures acoustiques et perceptives.

Dans ce sens, nous avons mené cette deuxième expérience afin de :

- ✓ caractériser – au plan aérodynamique – la voix de l'enfant sourd sévère et profond ayant été implanté/appareillé avant l'âge de 3 ans et portant l'implant/l'appareillage pendant 3 ans ou plus,
- ✓ comparer les caractéristiques de la fonction laryngée des groupes EE, ESA et ESIC afin de déterminer un effet inter-groupe de la prothèse auditive portée par l'ESA et l'ESIC sur les caractéristiques acoustiques de la voix,
- ✓ identifier et décrire d'éventuels comportements de la fonction laryngée pathologiques chez les enfants sourds.
- ✓ déterminer en termes de faisabilité la pertinence d'un protocole d'évaluation clinique incluant des mesures aérodynamiques afin d'évaluer et caractériser la voix des enfants sourds appareillés et implantés cochléaires.

4.2. Méthodologie

Le recueil et le traitement de mesures des paramètres aérodynamiques des voix ont été effectués à l'aide de l'application « EFFICACITE GLOTTIQUE » de l'appareillage EVA™2 (Figure 67). Les mesures aérodynamiques ont été seulement exploitées pour la population d'enfants implantés cochléaires (ESIC). Cela car les enfants entendants (EE) et sourds appareillés (ESA) n'ont pas été autorisés à effectuer l'évaluation clinique de ces paramètres au sein du CHU de Montpellier.

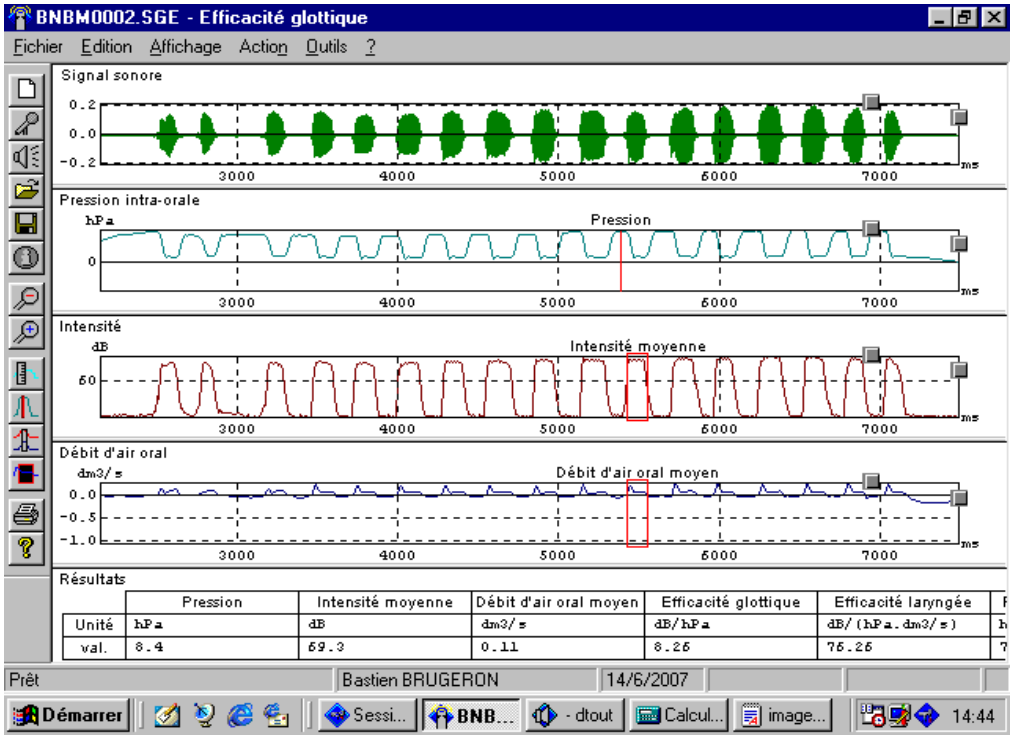


Figure 67 : Fenêtre d'affichage de l'application « EFFICACITE GLOTTIQUE » du dispositif EVA™2 pour l'analyse des paramètres aérodynamiques.

Le premier graphe (de haut en bas) représente le spectre d'une série de plusieurs /pa/. Le deuxième graphe représente les courbes de la pression intra-orale ou PSGE; le troisième graphe représente la courbe des variations d'intensité pour chaque syllabe /pa/; et enfin, le dernier graphe représente les courbes du débit d'air oral.

La technique d'évaluation était indirecte non invasive (Smitheran et Hixon, 1981[161]). Nous avons utilisé la même technique décrite par Yu (2001[141,162]). En effet, différents paramètres aérodynamiques de la voix (tels que la pression sous-glottique estimée, le débit d'air buccal) peuvent être mesurés grâce à la mise au point de cette technique et à l'aide d'un capteur de pression placé dans la bouche de l'enfant par l'examinateur. Le dispositif EVA™2 a un ensemble de capteurs regroupés au niveau d'une embouchure spécifique placée sur le visage de l'enfant. La capture des signaux acoustiques a été réalisée par un microphone coaxial contenu dans l'embouchure, placé à cinq cm de la bouche environ, en aval de la grille de tranquillisation du pneumotachographe thermostaté.

L'examinateur demandait à l'enfant de maintenir la sonde entre les molaires supérieures et inférieures du côté droit sans l'écraser. Il s'agit d'une sonde d'aspiration pour canule de

trachéotomie (Vygon, N°12) dont la relative rigidité permet de ne pas être écrasée. A l'intérieur de la bouche, 10 mm de l'extrémité de la sonde étaient dégagés pour faciliter l'entrée de l'air intra-buccal pendant l'enregistrement. Le masque de l'appareillage EVA™2 a été placé de manière hermétique contre le visage. L'examinateur veillait lors de l'examen à ce que la sonde ne soit pas obstruée par de la salive et que le visage de l'enfant soit plaqué contre le masque (Figure 68). Le cas échéant, l'examinateur changeait la sonde.

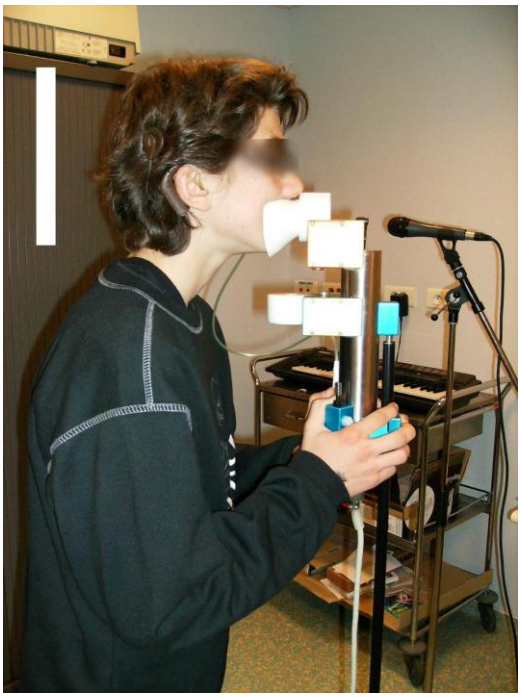


Figure 68 : L'enfant pendant l'enregistrement.

Plusieurs essais ont été effectués avant l'enregistrement afin de vérifier le calibrage du système. Les mesures aérodynamiques ont été traitées à partir de l'émission d'une série de dix syllabes /pa/ à hauteur et intensité usuelles à raison de trois par seconde environ. Après l'enregistrement de l'échantillon (syllabes), le logiciel lançait automatiquement une fenêtre qui affichait simultanément les courbes et le tableau de résultats des différents paramètres (Figure 67). Les mesures étaient relevées à partir des huit syllabes « centrales » consécutives en enlevant systématiquement la première et la dernière. La valeur moyenne des huit résultats a été retenue.

1) La pression sous-glottique estimée (ou PSGE), ou la pression à laquelle l'air arrive des poumons afin de mettre en vibration les cordes vocales, a été mesurée à l'aide d'un curseur placé précisément sur le pic de pression de la courbe correspondant à la phase de fermeture

des lèvres et d'explosion du /p/ lors de l'émission de la syllabe /pa/. Dans notre étude, la PSGE a été mesurée en hPa.

2) Le débit d'air oral (ou débit d'air buccal, DAB), ou volume d'air expiré pendant la phonation, a été calculé sur l'émission du /a/ après l'explosion du /p/ lors de l'émission de la syllabe /pa/. Le curseur a été placé juste après le pic de pression ou la chute de la courbe correspondant à la phase d'ouverture des lèvres et d'accolement des cordes vocales. Dans notre étude, le DAB a été mesuré en dm^3/s.

3) L'intensité, a également été calculée sur l'émission du /a/ après l'explosion du /p/ comme pour le débit d'air. Dans notre étude, l'intensité a été mesurée en dB.

4) Quant à *l'efficacité glottique*, les valeurs de ces paramètres ont été relevées simultanément lorsque le curseur était placé sur le pic de la courbe de pression de chaque courbe pour mesurer la pression sous-glottique estimée.

5) Pour *l'efficacité laryngée*, les valeurs ont été relevées simultanément lorsque le curseur était placé sur la chute du pic de pression de chaque courbe pour mesurer le débit d'air buccal.

L'efficacité glottique et l'efficacité laryngée sont des valeurs calculées instantanément par l'appareillage EVA à partir de la PSGE et du DAB (composantes aérodynamiques) et de l'intensité (composante acoustique). Les valeurs moyennes pour tous les paramètres présentés ci-avant ont été obtenues moyennant les valeurs mesurées des huit courbes centrales.

4.3. Résultats et discussion

Les résultats des paramètres aérodynamiques présentés ici concernent uniquement les enfants ESIC (n=20). En effet, en l'absence de l'autorisation des parents d'enfants EE et ESA, nous n'avons pu obtenir ces données pour ces populations.

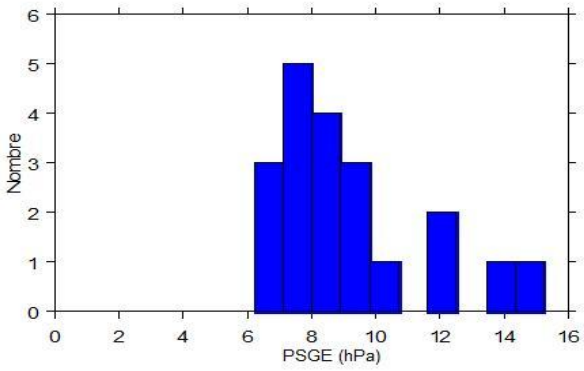
Dans un premier temps, nous avons effectué une analyse statistique descriptive des paramètres aérodynamiques à partir des indices statistiques observés (valeurs de la moyenne, écart-type, médiane et valeurs extrêmes). Au total, nous avons analysé les cinq paramètres aérodynamiques (pression sous-glottique, débit d'air oral, intensité, efficacité glottique et efficacité laryngé). Le résumé des résultats de l'analyse statistique descriptive des paramètres

aérodynamiques pour l'ensemble du groupe (n=20) sont présentés dans le tableau ci-dessous (Tableau 29, Annexe 11).

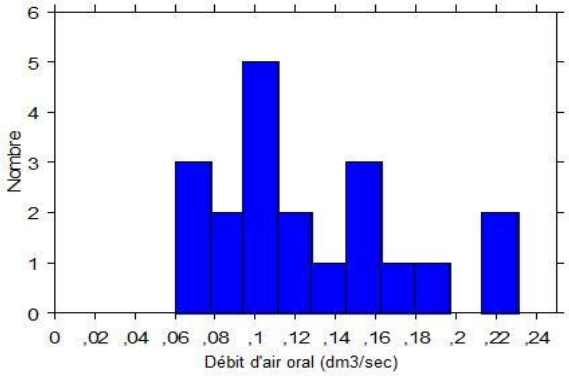
Paramètre (unités)	*Moyenne*	*Ecart-type*	*Médiane*	*Minimum*	*Maximum*
PSGE (hPa)	9,14	2,49	8,26	6,2	15,23
Débit d'air (dm^3/s)	0,13	0,05	0,12	0,06	0,23
Intensité (dB)	72,51	4,68	71,52	65,68	81,6
Eff Glottique (dB/hPa)	8,44	1,81	8,46	5,29	11,74
Eff Laryngée (dB/hPa*dm^3/s)	78,22	32,75	76,14	33,98	155,55

Tableau 29 : Valeur des principaux indices statistiques descriptifs pour les paramètres aérodynamiques du groupe ESIC (n=20).

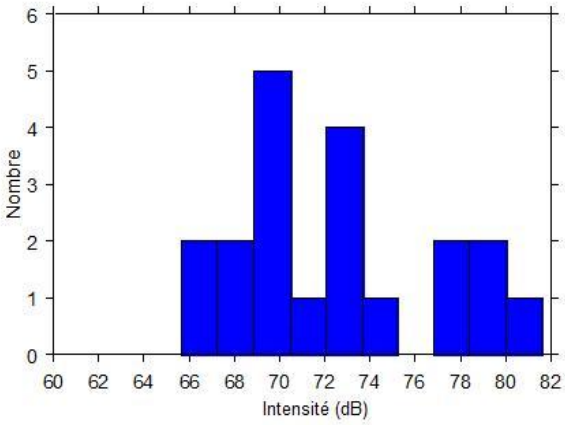
De cette analyse, il ressort que les valeurs de la médiane se situent de manière générale légèrement en-dessous des valeurs moyennes indiquant une légère dissymétrie à gauche de la distribution (sauf pour l'efficacité glottique). Les valeurs extrêmes signalent un écart assez important témoignant d'une forte étendue des observations pour tous les paramètres. Les histogrammes ci-dessous (Figure 69) montrent la distribution d'effectifs pour chaque paramètre.



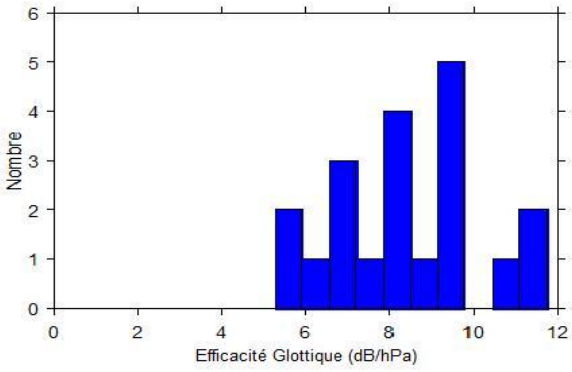
a) Pression sous-glottique estimée (PSGE)

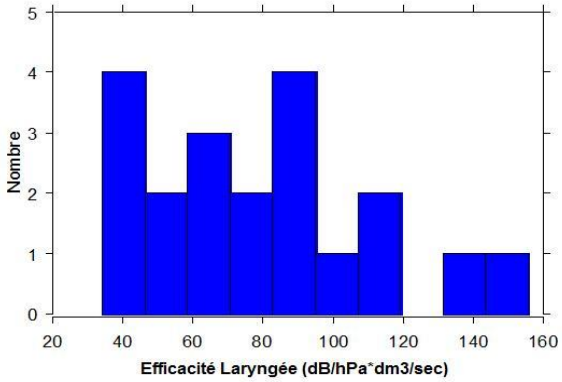


b) Débit d'air oral (DAB)



c) Intensité (Int)





d) Efficacité glottique (Eff glott) *e) Efficacité laryngée (Eff laryng)*

Figure 69 : Histogrammes de la distribution d'effectifs du groupe ESIC pour les différents paramètres aérodynamiques.

L'histogramme de la PSGE montre une distribution des données concentrée sur les valeurs situées entre 6 et 11 hPa et un mode situé à 8 hPa environ. La distribution des données pour le débit d'air oral est plutôt concentrée entre 0,060 dm^3/s et 0,160 dm^3/s avec un mode situé à 0,100 dm^3/s. L'histogramme de l'intensité signale des modes situés à 70 dB et à 72 dB. La distribution des données de l'efficacité glottique est décalée vers la droite avec des valeurs des fréquences qui augmentent progressivement de 6 à 10 dB/hPa. Enfin, l'histogramme de l'efficacité laryngée montre des fréquences situées sur les valeurs de 40 et 90 dB/hPa*dm^3/s.

Etant donné que nous n'avions pas de base de données normative sur la mesure des paramètres aérodynamiques, nous n'avons pas pu effectuer de comparaison inter-groupes.

La revue bibliographique nous a permis de constater que l'analyse des mesures physiologiques chez l'enfant implanté cochléaire est récente et n'a pas fait l'objet d'une littérature abondante. En effet, les mesures physiologiques ont été rarement exploitées pour caractériser la voix de l'enfant sourd et, plus précisément, la voix de l'enfant implanté cochléaire. Une raison qui pourrait expliquer ce manque de données est que les enfants sourds manifestent souvent des productions orales très limitées, altérées, variables et difficiles pour l'application des procédures ou des méthodes d'évaluation physiologique standard. Très peu d'études ont tenté d'explorer le comportement physiologique de la parole et de la voix des enfants sourds. Ces travaux ont mis l'accent souvent sur un seul aspect ou paramètre physiologique de la parole, par conséquent la connaissance de la physiologie de la parole et de la voix reste toujours limitée (Arends et col., 1990[163] ; Dagenais et col., 1991[164], 1992[165] ;

Higgins et col., 1996[129] ; Leeper et col., 1980[166] ; Svirsky et col., 1998[131], 2000[167]). La plupart des études réalisées sont plutôt descriptif car l'effectif est assez réduit. Par ailleurs, ces études ne s'appuient pas sur une méthodologie standardisée. Elles conduisent, par conséquent, à des conclusions inconsistantes voire contradictoires (Higgins et col., 2003[82]).

Nous avons comparé nos données avec celles rapportées par Higgins et col. (2003) pour un groupe contrôle et un groupe d'enfants implantés. Concernant la pression sous-glottique estimée (PSGE), les auteurs reportent une valeur moyenne de 9,96 hPa pour un groupe contrôle de 6,5 à 12,5 ans (n=28) et une valeur moyenne de 10,5 hPa pour un groupe d'enfants implantés du même âge (n=7). D'une part, nous remarquons que la valeur moyenne de 9,14 hPa que nous avons obtenue pour la PSGE est proche de celle du groupe contrôle de Higgins. D'autre part, nous constatons que le groupe d'implantés de Higgins manifeste une valeur de PSGE plus élevée que la notre. L'élévation de la PSGE pour le groupe d'implantés de Higgins et col. s'explique du fait que ce sont des enfants ayant été implantés après l'âge de 5 ans ou plus. D'ailleurs, ce sont des enfants bénéficiant de l'implantation au début des années 1990. Depuis, les avancées technologiques ont amélioré la qualité de traitement du signal proposé par l'implant cochléaire.

Concernant le débit d'air oral (DAB), les auteurs reportent une valeur moyenne de 0,139 dm^3/s pour un groupe contrôle de 6,5 à 12,5 ans (n=28) et une valeur moyenne de 0,111 dm^3/s pour un groupe d'enfants implantés du même âge (n=7). A notre sens, ces données sont étonnantes car le DAB pour le groupe d'implantés de Higgins se situe en-dessous des limites normales. On espère en principe que les enfants implantés manifestent des valeurs au-dessus de la norme. A ce propos, Higgins suggère que ces valeurs illustrent la performance relative obtenue par un paramètre physiologique de la production orale même si d'autres aspects ou paramètres acoustiques apparaissent déviants. Higgins signale dans son étude que les enfants manifestant des valeurs de DAB en-dessous de la « norme » avaient une Fo élevée pour certains sons vocaliques et un VOT long pour certains sons consonantiques occlusifs. Néanmoins, ces enfants faisaient partie du groupe d'enfants ayant une parole intelligible et de très bonne qualité. En ce qui concerne nos données, nous avons obtenu comme Higgins, une valeur moyenne en-dessous de la norme. Cependant, la valeur du DAB de notre groupe est relativement proche de celle du groupe contrôle de Higgins. Nous restons prudents dans l'interprétation de cette valeur pour plusieurs raisons. Premièrement, on sait bien que la différence méthodologique utilisée pour le recueil de données ne permet pas de comparer si facilement des données d'une étude à l'autre. Deuxièmement, il est important de rappeler que

la mesure des paramètres aérodynamiques est covariante. Il s'avère nécessaire de corréler les résultats de la mesure des paramètres aérodynamiques et acoustiques à l'évaluation perceptive de la production orale d'une population afin d'établir des conclusions plus consistantes (Giovanni et col., 1996[136] ; Yu, 2001[162]).

Concernant l'intensité, nous remarquons que la valeur moyenne est légèrement au-dessus de la norme. Nous avons comparé ces valeurs à celles présentées dans la littérature pour /a/ tenu (Van Lierde et col., 2005[158] ; Dalleas, 1987[159]). Le paramètre de l'intensité n'a pas été analysé simultanément avec d'autres paramètres aérodynamiques. L'intensité a été mesurée sur l'émission du /a/ après l'explosion du /p/ lors de l'émission de la syllabe /pa/ (voir Chapitre III). La consigne donnée à l'enfant était d'émettre une série de syllabes /pa/ à *hauteur et intensité usuelles* afin d'obtenir une émission vocale semblable à celle de la voix conversationnelle.

Concernant l'efficacité glottique, la valeur moyenne se situe en-dessous de la frontière de normalité d'après les valeurs présentées par Guerrier et col. (2004)[168] pour un adulte. En revanche, la valeur moyenne de l'efficacité laryngée (ou rendement glottique)[*] apparaissent normale. Les indices d'efficacité glottique et d'efficacité laryngée sont fournis et déduits instantanément des mesures de la PSGE, du débit d'air oral et de l'intensité par l'appareillage EVA™2. De manière générale, ces indices signalent d'une part un « rendement dans le débit (ou glottique) »[**] et d'autre part, un « rendement dans la pression ». Nous interprétons cette donnée comme un indice de la performance du vibrateur laryngé chez l'ESIC comme étant « rentable » dans la transformation de l'énergie aérodynamique mise à disposition par les poumons en énergie acoustique émise. Ces paramètres seraient très intéressants à explorer. Ils permettraient de caractériser davantage les phénomènes liés à la source de phonation tels que la fuite glottique ou l'estimation d'un forçage vocal éventuellement existant et qui ne seraient pas perçus par les mesures acoustiques ou l'évaluation perceptive. Pour ces raisons, il nous a paru pertinent d'inclure dans cette étude les coefficients d'efficacité glottique et laryngée afin de les comparer ultérieurement dans d'autres travaux. Etant donné qu'il n'existe pas actuellement de base de données sur ces indices chez une population infantile, nous préférons rester prudents dans l'interprétation de ces résultats préliminaires.

---

[*] Terme figurant dans l'ouvrage de Guerrier et col. (2004).

[**] Terme figurant dans la thèse de : PILLOT Claire, Sur l'efficacité vocale dans le chant lyrique. Aspects physiologique, cognitif, acoustique et perceptif, Thèse de doctorat, Université Paris III, 2004.

Dans un deuxième temps, nous avons observé la corrélation linéaire entre les paramètres aérodynamiques et l'âge de l'enfant, la durée d'implantation et l'âge à l'implantation. De même, nous avons observé la corrélation linéaire des paramètres aérodynamiques entre eux. Nous étions intéressés à observer si la durée d'implantation avait une relation inversement proportionnelle avec la magnitude de la PSGE. Egalement, nous voulions observer la tendance d'évolution du fonctionnement physiologique laryngé chez l'enfant implanté cochléaire. Cela afin de mieux comprendre les caractéristiques des comportements physiologiques mis en place pour le développement de la parole et de la voix chez l'enfant implanté. La connaissance de la physiologie de la parole et de la voix des enfants sourds, comprenant les enfants implantés cochléaires, est toujours limitée (Higgins et col., 2003).

Pour ce faire, nous avons appliqué le test-Z de corrélation et le test sur la significativité du *« coefficient de corrélation linéaire »* (Tableau 24)**.** Le coefficient de corrélation linéaire mesure le lien linéaire entre deux variables pour un même groupe et établit leur « proportionnalité ». Le coefficient de corrélation oscille entre -1 et 1. Le signe de ce coefficient indique le sens de variation respective des deux variables : signe « + », elles évoluent dans le même sens ; signe « - », elles évoluent en sens opposé.

Dans le tableau 30, la deuxième colonne présente les coefficients de corrélation. La quatrième colonne présente la probabilité critique (p.c.), ou le seuil de significativité du coefficient de corrélation linéaire (le seuil de référence dans notre travail étant fixé à $\alpha = 0.05$). On observe en bleu les corrélations significatives et en orange les corrélations étant à la limite de la significativité.

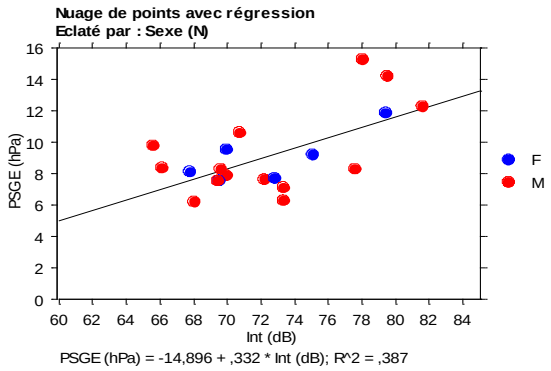
Les graphes correspondant aux corrélations significatives sont présentés ci-dessous (Figure 70). On peut apprécier que pour certains couples de variables comme par exemple *PSGE vs. Durée d'implantation* (Figure 70c) et *Efficacité glottique vs. Durée d'implantation* (Figure 70e), la corrélation est plus importante : plus les points sont proches de la droite, plus la corrélation linéaire est forte.

Nous avons constaté l'existence d'une forte corrélation significative entre la PSGE et l'âge de l'enfant ainsi qu'entre la PSGE et la durée d'implantation : plus l'âge de l'enfant et la durée d'implantation augmentent plus la PSGE diminue (Figure 70 b, c). On observe une diminution de la PSGE et, plus particulièrement, une tendance à la stabilisation de la PSGE après 8 ans d'implantation.

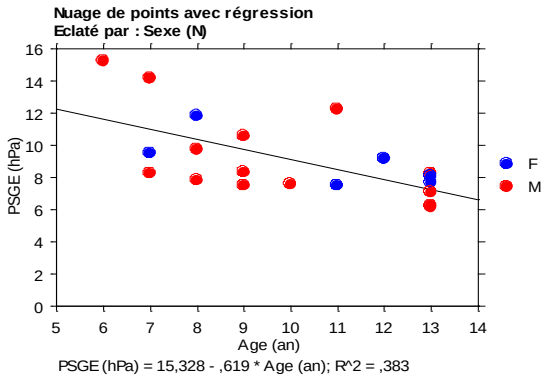
Paramètres corrélés	Coefficient	Significatif	p.c.
PSGE / Intensité	0.622	OUI	0.0027
PSGE / Débit d'air	-0.087	NON	0.7180
PSGE / Eff glottique	-0.916	OUI	<0.0001
PSGE / Eff laryngée	-0.350	NON	0.1317
PSGE / Age	-0.619	OUI	0.0029
PSGE / Age à l'implant	-0.297	NON	0.2059
PSGE / Durée d'implant	-0.652	OUI	0.0013
Débit d'air / Intensité	0.096	NON	0.6903
Débit d'air / Eff glott	0.042	NON	0.8631
Débit d'air / Eff Laryngée	-0.778	OUI	<0.0001
Débit d'air / Durée	0. 092	NON	0.7044
Intensité / Eff glott	-0.343	NON	0.1407
Intensité / Eff laryngée	-0.186	NON	0.4373
Intensité / Age	-0.241	NON	0.3105
Intensité / Age à l'implant	-0.304	NON	0.1957
Intensité / Durée d'implant	-0.235	NON	0.3237
Eff glott / Age	0.645	OUI	0.0016
Eff glott / Age à l'implant	0.279	NON	0.2367
Eff glott / Durée d'implant	0.688	OUI	0.0005
Eff glott / Eff laryngée	0.404	NON	0.0770
Eff laryngée / Age	0.474	OUI	0.0337
Eff laryngée / Age à l'implant	0.347	NON	0.1353
Eff laryngée / Durée d'implant	0.433	NON	0.0561

Tableau 30 : Résumé des résultats du test de significativité du Coefficient de corrélation.

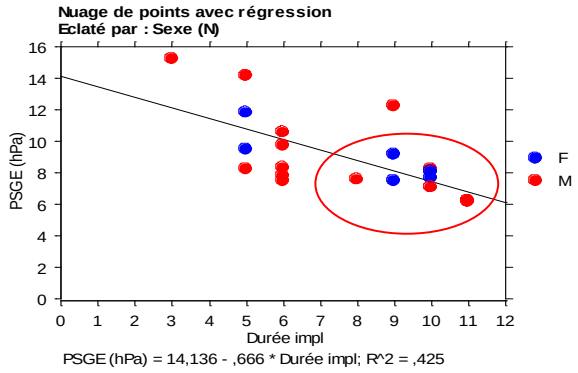
Par ailleurs, la corrélation de la PSGE et l'intensité est directement proportionnelle témoignant d'un fonctionnement laryngé normal : plus la pression augmente plus l'intensité s'accroit. Bien que la corrélation entre le DAB et l'âge à l'implantation ne soit pas significative ($p<0,05$), nous observons qu'il y a une tendance à l'amélioration du DAB quand l'implantation a lieu avant l'âge de deux ans. Quant aux coefficients d'efficacité glottique et laryngée, nous observons une corrélation linéaire « positive » avec l'âge et la durée d'implantation. Le coefficient de l'efficacité glottique accroit notamment en fonction d'une augmentation de l'âge et de la durée de port de l'implant notamment (Figure 70 d, e).



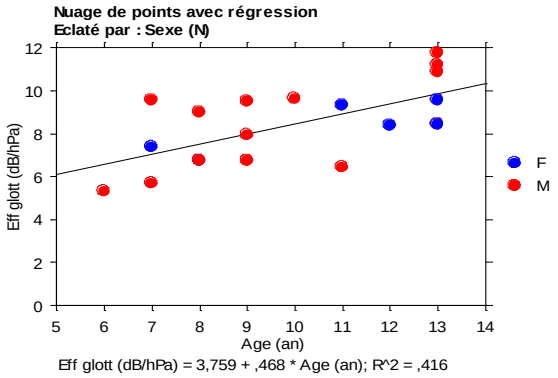
a) PSGE vs. Intensité



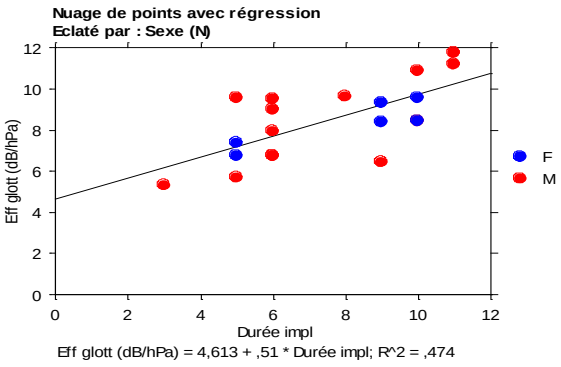
b) PSGE vs. Age



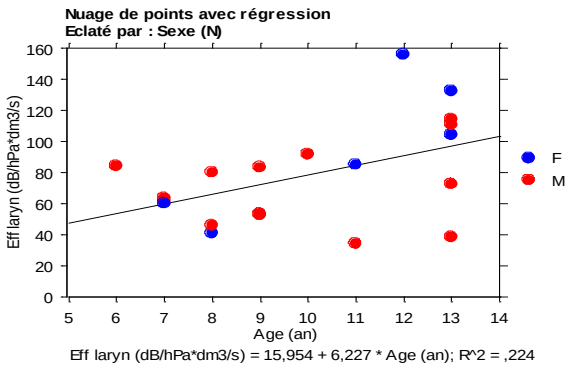
c) PSGE vs. Durée de l'implantation



d) Efficacité glottique vs. Age



e) Eff. glottique vs. Durée d'implantation



f) Eff. Laryngée vs. age

Figure 70 : Nuages de points et droite de régression dans les cas où le coefficient de corrélation linéaire est significatif dans le groupe ESIC.

Les résultats de la corrélation linéaire entre les paramètres aérodynamiques et l'âge de l'enfant, la durée d'implantation et l'âge à l'implantation mettent en évidence un développement physiologique « non déviant » chez l'enfant implanté cochléaire. La dynamique du flux aérien nécessaire dans la production de la parole et de la voix s'inscrit dans un comportement tout à fait normal que l'on pourrait s'attendre à observer chez un enfant normo-entendant. D'après l'analyse de corrélation entre la PSGE et la durée d'implantation, on pourrait supposer que la maitrise de la régulation de la pression sous-glottique nécessaire à l'émission vocale se stabilise à partir de huit ans d'implantation (Figure 70b), moment où l'enfant semblerait avoir développé une certaine expérience auditive et avoir acquis une capacité de régulation de l'intensité de la voix.

Sur le plan aérodynamique, nous n'avons pas identifié de troubles de la fonction laryngée ou de comportements phonatoires déviants. Nous estimons que les valeurs obtenues de la PSGE et du DAB chez notre population s'avèrent cohérentes et non déviantes. La technique de mesure des paramètres aérodynamiques que nous avons appliquée, mise au point par Smitheran et Hixon (1981)[169] a été initialement conçue pour l'étude de la voix chez l'adulte. En termes de faisabilité, cette technique a pu être appliquée à notre population avec certaines dispositions supplémentaires. Il sera nécessaire de développer cette technique et de l'adapter aux particularités de la population infantile. Nous constatons que l'analyse de la voix de l'enfant implanté cochléaire ne peut se faire qu'à l'aide d'un protocole d'évaluation multiparamétrique incluant des mesures aérodynamiques. Ces dernières facilitent la compréhension des processus physiologiques impliqués dans le développement de la parole et dans la production de la voix.

Une base de données sur les paramètres aérodynamiques les plus utilisés en clinique de la voix de l'enfant implanté cochléaire de 5 à 13 ans, a pu être constituée. Elle s'inscrit dans le cadre d'une étude préliminaire qui fournit désormais des éléments pour l'élargissement des données présentées. Par ailleurs, il est nécessaire d'explorer les mesures aérodynamiques sur un groupe contrôle avec des caractéristiques les plus proches à celles du groupe d'enfants implantés afin d'effectuer une comparaison plus détaillée et plus fiable.

4.4. Conclusion

Etant donné que nous n'avons pu évaluer les paramètres aérodynamiques pour les groupes EE et ESA, cette étude expérimentale nous a permis d'accomplir partiellement les objectifs proposés. De l'ensemble de résultats obtenus pour le groupe ESIC, nous concluons que :

- ✓ la fonction laryngée chez l'enfant implanté cochléaire s'accomplit manifestement sans comportements pathologiques,

- ✓ la durée de l'implantation a une incidence significative sur la PSGE. Nous constatons que la PSGE a tendance à une diminution notamment après huit ans ou plus d'implantation,

- ✓ en termes de faisabilité, la technique de recueil de mesures physiologiques de Smitheran & Hixon (1981[169]) initialement proposée aux adultes a pu être adaptée à une population infantile en prenant en compte certaines contraintes (matériel : taille du masque, position de l'enfant pendant l'enregistrement, utilisation d'une sonde, entre autres). Des moyens techniques existants peuvent être adaptés afin d'exploiter ces mesures auprès de cette population,

- ✓ nous avons pu caractériser – au plan aérodynamique – la voix de l'enfant implanté cochléaire à partir des mesures physiologiques. Une base de données sur les paramètres aérodynamiques en fonction de l'âge et du sexe a été effectuée (Annexe 11).

5. Analyse perceptive (Expérience 3)

L'analyse perceptive est la méthode la plus utilisée en pratique clinique pour évaluer la voix. Puisqu'elle est considérée comme la méthode de référence permettant de valider les autres moyens d'évaluation de la qualité vocale (Kreiman, 1996)[170], nous l'avons inclus dans notre étude.

5.1. Objectifs

L'objectif principal de cette analyse était de :

- ✓ déterminer si les voix des enfants sourds (ESA et ESIC) étaient dysphoniques ou non grâce à l'évaluation d'un jury d'écoute,
- ✓ valider les résultats de l'évaluation instrumentale des paramètres acoustiques et aérodynamiques,
- ✓ déterminer si l'échelle GRBAS de Hirano était pertinente pour l'analyse et l'évaluation des caractéristiques vocales des enfants sourds.

5.2. Méthodologie

Initialement, les échantillons de voix ont été analysés par un jury d'écoute dont la composition était de cinq membres. Étant donné que l'un des membres du jury manifestait une forte variabilité intra-individuelle de jugement perceptif, nous nous sommes vus contraints de réduire le jury au nombre de quatre. Les membres du jury étaient des experts dans l'évaluation de la voix, de la dysphonie et des troubles de la parole (1 phonéticien, 1 phoniatre et 2 orthophonistes).

Afin de faciliter les écoutes par le jury et de standardiser la procédure d'évaluation, nous avons procédé à la conception d'une notice d'information contenant les consignes détaillées et les explications rigoureuses de la passation. Les consignes étaient données par écrit et demandaient aux auditeurs de dérouler chaque séance de manière identique (Annexe 12). Nous avons organisé trois séances d'écoute distantes d'une semaine. Le jury exécutait la même tâche d'évaluation pour chaque séance comme proposé par certains auteurs (De Bodt, 1997[171]; Giovanni, 1999[172]; Yu, 2001[162], 2002[173]; Revis, 2004[132]). Ceci dans le but d'analyser

et vérifier notamment la fiabilité de cotation pour chaque auditeur au cours des trois séances. La passation s'est déroulée en mode de « session individuelle » afin d'éviter le risque de concertation et d'influence des auditeurs entre eux ainsi que pour des raisons pratiques de mise en place.

Nous avons créé un fichier informatique contenant les 78 échantillons* de voix numérisés (au format **.wav)* et nous les avons rendu anonymes. Un dossier pour chaque session d'écoute a été créé. Les dossiers contenaient les échantillons de voix présentés dans un ordre aléatoire différent pour chaque séance. Ceci afin d'éviter tout biais lié à une possibilité de lien logique entre les voix (Kreiman, 1992[174]; De Bodt, 1997[171]) et un effet de contexte pouvant influencer les auditeurs (Goldstone, 1998[175]).

Quant au matériau phonétique, nous avons opté pour l'évaluation perceptive des échantillons de parole. Notre choix s'est porté sur le discours spontané tel qu'Anders (1988[176]) le suggère car ce support vocal est le plus représentatif de la voix conversationnelle d'un patient. En effet, le clinicien se fait une idée de la voix d'un patient par sa parole. Pour le jury d'écoute, c'est le matériau le plus intuitif et facile à évaluer. Les échantillons de parole utilisés ont été exploités à partir de ceux utilisés pour l'analyse acoustique de la voix conversationnelle (cf. *section 2.2.1*).

Concernant l'instrument d'évaluation, d'une part, nous avons choisi d'effectuer la cotation de l'analyse perceptive à l'aide d'une échelle visuelle analogique. Cette échelle permettait d'appréhender visuellement la quantité attribuée à la présence d'un paramètre de la voix. Elle facilitait une discrimination plus fine et plus précise des voix. L'auditeur devait porter son évaluation sur un axe de 10 cm par une marque. Le chiffre attribué à la voix correspondait à la distance en cm entre le 0 et la croix (Figure 71).

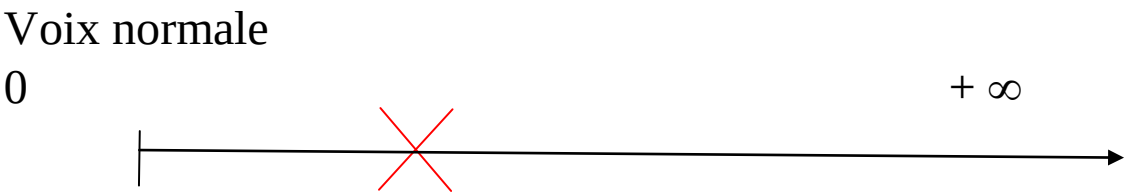


Figure 71 : Échelle visuelle analogique.

* Les voix des trois populations ont été mélangées.

Le « 0 » étant la valeur qui représentait la normalité et l'extrême opposé étant la valeur qui représentait la pathologie. D'autre part, nous avons choisi l'échelle GRBAS de Hirano (Hirano, 1981**Erreur ! Signet non défini.**) comme la méthode d'évaluation perceptive pour notre étude. Cette méthode est actuellement la plus utilisée en clinique pour évaluer notamment les dysphonies sur la base de cinq paramètres (Tableau 31).

INITIALE	TERME	SIGNIFICATION	DÉFINITION
G	*Grade*	*Grade global de dysphonie*	*Impression globale du degré d'anormalité de la voix*
R	*Rough*	*Raucité*	*Impression d'irrégularité des vibrations de la corde vocale qui correspond aux fluctuations irrégulières de la Fo et/ou à l'amplitude du son glottique*
B	*Breathy*	*Caractère soufflé*	*Impression d'une fuite d'air assez importante à travers les cordes vocales, relative à des turbulences*
A	*Asthenic*	*Asthénie*	*Manque de puissance de la voix relatif à une intensité faible du son et/ou un manque des harmoniques élevés*
S	*Strained*	*Forçage*	*Impression d'un état hyperfonctionnel de phonation relatif à une Fo anormalement haute (bruit dans les hautes fréquences et/ou richesse en harmoniques dans les hautes fréquences)*

Tableau 31 : Tableau contenant la définition des paramètres de l'échelle GRBAS. Nous présentons le nom des paramètres en anglais (version originale) et sa respective traduction en français.

Notre choix s'est porté sur cette méthode car l'un des objectifs préalablement formulées dans notre travail était de répondre à la question si les troubles vocaux manifestés chez la population d'enfants sourds relevaient d'une dysphonie ou avaient une tendance à la dysphonie.

De la même façon que Dejonckere (1993)[177], nous avons modifié l'échelle GRBAS en fonction des objectifs de notre étude. Pour cela, nous avons ajouté deux paramètres (I : Intelligibilité; N : Nasalité) visant à élaborer un protocole d'évaluation vocale plus adapté aux caractéristiques vocales des enfants sourds. L'idée d'intégrer à l'échelle GRBAS deux paramètres supplémentaires associés à la parole et de les évaluer en premier, nous assurait que le jury se concentrerait bien à noter les paramètres de l'échelle GRBAS notamment le paramètre G (Grade Global de la dysphonie).

Chaque auditeur a donc évalué au total sept paramètres pour chaque voix sur l'échelle visuelle analogique (Annexe 13). Les valeurs analogiques ont été *à posteriori* converties en valeurs ordinales (échelle de classe) selon la méthode proposée par Yu et Revis (2002)[173]. Nous

expliquerons ultérieurement l'objectif et la procédure de cette manipulation de données (section 5.3.1).

5.3. Résultats et discussion

Le traitement des données de l'évaluation perceptive a été celui proposé par Yu et Revis (2002)[173]. Tout d'abord, nous avons validé le jury d'écoute (cohérence intra et inter-juge); ensuite, nous avons étudié la variabilité des jugements des auditeurs; enfin, nous avons analysé le lien entre les données instrumentales et les résultats de l'évaluation perceptive.

5.3.1. Validité du jury d'écoute

L'évaluation perceptive de notre protocole a fait l'objet de trois sessions d'écoute à intervalle d'une semaine. Les 78 enregistrements de voix (EE n=38, ESA n=20, ESIC n=20) ont été présentés dans un ordre aléatoire lors de chaque session individuelle afin de limiter l'effet de contexte (Revis, 2004[132]). Au total, 12 notes étaient attribuées à chaque voix (4 auditeursx3 sessions = 12) et pour chaque paramètre de l'échelle I (Intelligibilité), N (Nasalité), G (Grade), R (Raucité), B (Souffle), A (Asthénie), S (Forçage) proposée au jury d'écoute. Étant donné que nous avions utilisé une cotation sur une échelle visuelle analogique, les résultats obtenus étaient des valeurs continues. En effet, la marque apposée par l'auditeur sur l'échelle était mesurée sur une droite de 100 mm (10 cm). La note attribuée correspondait à la distance mesurée (en millimètres) entre le « 0 » et la marque enregistrée.

Nous avons ensuite utilisé une conversion de ces valeurs numériques en valeurs ordinales. Pour cela, nous avons exploité l'échelle mixte ou « mVAS » proposée par Yu et Revis (2002)[173] et leur tableau de conversion (Tableau 32).

Echelle visuelle analogique (cm)	Echelle mixte
0.0 – 0.9	Grade 0
1.0 – 5.0	Grade 1
5.1 – 9.0	Grade 2
9.1 – 10	Grade 3

Tableau 32 : Conversion des valeurs continues d'une échelle visuelle analogique en valeurs ordinales d'une échelle mixte à 4 niveaux.

Nous avons opté pour cette méthode car l'échelle mixte proposée par Revis nous permettait de combiner les avantages de l'échelle visuelle analogique (finesse de discrimination) avec ceux de l'échelle de classes (résultats compacts). Yu et Revis réfèrent que cette échelle semble améliorer à la fois la discrimination spécifique des dysphonies intermédiaires de l'évaluation perceptive et l'adéquation entre analyse perceptive et analyse instrumentale.

Après transformation de chaque note sur l'échelle mixte, nous avons effectué une étude de la concordance intra/inter-auditeur de notre jury. L'outil statistique utilisé était le « coefficient Kappa de Cohen ». Cet indice statistique permet une mesure pour des données ordinales du degré d'accord entre les juges (Droesbeke et col., 2005)[178]. Le coefficient « *K* » est compris entre 0 et 1. En cas d'accord parfait entre les juges, $K = 1$. S'ils jugent d'une façon aléatoire, $K = 0$. Afin de faciliter l'interprétation des résultats, nous proposons la table des valeurs des coefficients Kappa présentée par De Bodt (1997)[171] (Tableau 33). Un test peut alors être mis en place à partir de ce coefficient *K* pour tester l'hypothèse nulle des jugements aléatoires.

Valeur du coefficient Kappa	Degré de fiabilité
<0.20 Faible	Faible
0.21 – 0.40	Raisonnable
0.41 – 0.60	Assez bon
0.61 – 0.80	Bon
0.81 – 1.00	Très bon

Tableau 33 : Interprétation des valeurs du coefficient Kappa (d'après De Bodt, 1997).

Premièrement, nous avons effectué le calcul du coefficient pour établir les concordances des jugements d'un même auditeur au cours des trois séances d'évaluation et pour chacun des paramètres de l'échelle I N G R B A S : concordance intra-auditeur. Les résultats du calcul du coefficient Kappa sont présentés dans les tableaux ci-dessous (Figure 34). Dans ces tableaux, on observe les valeurs du coefficient Kappa pour chaque juge, leur probabilité critique « p » et leur interprétation. On note également en jaune, un score étant non significatif (le seuil ayant été fixé à 0,05). Les résultats révèlent que :

- Au cours des trois séances d'évaluation, le « juge a » et le « juge d » ont présenté un score les qualifiant comme les auditeurs les plus fiables du jury. Par ailleurs, les juges « b » et « c » ont obtenu des scores plus bas, néanmoins ces scores les qualifient comme des auditeurs fiables selon le tableau d'interprétation (Tableau 33).

Intelligibilité	Kappa	P	Interprétation
Juge a	0,93	0,000	Très bon
Juge b	0,50	0,000	Assez bon
Juge c	0,68	0,000	Bon
Juge d	0,78	0,000	Bon

a) Paramètre I

Nasalité	Kappa	P	Interprétation
Juge a	0,97	0,000	Très bon
Juge b	0,34	0,073	Raisonnable
Juge c	0,55	0,000	Assez bon
Juge d	0,79	0,000	Bon

b) Paramètre N

Grade	Kappa	P	Interprétation
Juge a	0,97	0,000	Très bon
Juge b	0,47	0,000	Assez bon
Juge c	0,70	0,000	Bon
Juge d	0,77	0,000	Bon

c) Paramètre G

Raucité	Kappa	P	Interprétation
Juge a	0,96	0,000	Très bon
Juge b	0,36	0,003	Raisonnable
Juge c	0,71	0,000	Bon
Juge d	0,84	0,000	Très bon

d) Paramètre R

Souffle	Kappa	P	Interprétation
Juge a	0,97	0,000	Très bon
Juge b	0,43	0,001	Assez bon
Juge c	0,67	0,000	Bon
Juge d	0,74	0,000	Bon

e) Paramètre B

Asthénie	Kappa	P	Interprétation
Juge a	1,00	0,000	Très bon
Juge b	0,43	0,000	Assez bon
Juge c	0,67	0,000	Bon
Juge d	0,75	0,000	Bon

f) Paramètre A

Forçage	Kappa	P	Interprétation
Juge a	0,98	0,000	Très bon
Juge b	0,35	0,000	Raisonnable
Juge c	0,58	0,000	Assez bon
Juge d	0,73	0,000	Bon

g) Paramètre S

Tableau 34 : Concordance intra-auditeur pour chacun des paramètres de l'échelle INGRBAS.

- Tous les juges ont obtenu des scores très significatifs sauf le juge « b » qui obtient un score « raisonnable » pour le paramètre N (Nasalité) non significatif (en jaune, Tableau 34b).

- De manière générale, nous constatons une légère variabilité intra-auditeur des paramètres N (Nasalité) et S (Forçage). Ce qui nous laisse supposer qu'il existe des difficultés d'évaluation pour les paramètres N (Nasalité) et S (Forçage) par le jury d'écoute.

Deuxièmement, nous avons effectué le calcul du coefficient Kappa pour établir les différences de jugements existant entre les auditeurs sur chaque voix. Le Tableau 35 montre les résultats de la concordance inter-auditeur au cours de chacune des trois séances et pour chacun des paramètres de l'échelle I N G R B A S.

Ces résultats révèlent une concordance statistique largement significative. La concordance inter-individuelle est manifestement élevée au cours des trois séances. A partir des résultats obtenus concernant la variabilité intra/inter du jury d'écoute, nous validons donc la fiabilité de la cotation et des jugements attribués par le jury aux voix des enfants de l'étude.

Session 1	Kappa	P	Interprétation
I	0,649	0	Bon
N	0,508	0	Assez bon
G	0,476	0	Assez bon
R	0,500	0	Assez bon
B	0,478	0	Assez bon
A	0,472	0	Assez bon
S	0,45	0	Assez bon

Session 2	Kappa	P	Interprétation
I	0,660	0	Bon
N	0,459	0	Assez bon
G	0,553	0	Assez bon
R	0,546	0	Assez bon
B	0,538	0	Assez bon
A	0,560	0	Assez bon
S	0,55	0	Assez bon

Session 3	Kappa	P	Interprétation
I	0,614	0	Bon
N	0,410	0	Assez bon
G	0,515	0	Assez bon
R	0,543	0	Assez bon
B	0,501	0	Assez bon
A	0,515	0	Assez bon
S	0,58	0	Assez bon

Tableau 35 : Concordance inter-auditeur pour chacun des paramètres de l'échelle I N G R B A S au cours de trois sessions.

Le paramètre I est le mieux noté par les juges, c'est-à-dire que les juges sont en accord mutuel sur la note attribuée à chaque voix concernant ce paramètre. Il semblerait que le paramètre I soit le paramètre le plus facile à évaluer. Or, il s'agit d'un paramètre évaluant *l'Intelligibilité de la parole* mais pas la voix. Nous l'avons inclus dans l'évaluation perceptive conjointement avec le paramètre N évaluant la *nasalité de la parole* afin d'éviter que le jury d'écoute évalue de manière faussée le paramètre G (Grade Global de la dysphonie) en appréciant d'autres paramètres « anormaux » de la « voix caractéristique des enfants à déficience auditive » (nasalité, prosodie, aspects articulatoires,...). De cette manière, nous étions sûrs que le jury noterait bien les voix sans fausser les résultats de l'évaluation de la voix. En effet, si notre jury d'écoute avait comme consigne d'évaluer d'autres paramètres plutôt associés à la parole tels que « I » (Intelligibilité) et « N » (Nasalité), notre jury ferait plus attention à la façon d'évaluer les paramètres de l'échelle GRBAS. Ainsi, il n'entrerait pas « de force » des appréciations perceptives de la parole dans l'évaluation perceptive des paramètres de la voix.

5.3.2. Etude de la variabilité des jugements perceptifs

Nous reprenons ici les notes des juges sur l'échelle analogique où les données sont plus précises et permettent une analyse plus fine de la variabilité. Pour cela, nous avons d'abord calculé la moyenne et l'écart-type des notes obtenues réalisées pour chaque voix. Afin d'éviter l'influence de l'émergence de valeurs extrêmes sur la moyenne, nous avons enlevé la note la plus basse et la plus élevée[*]. A partir de ces moyennes et du tableau de conversion de l'échelle mixte (cf. Tableau 32), nous avons regroupé les voix en fonction du niveau de dysphonie 0, 1, 2 et 3 de l'échelle mixte. Puis, pour chaque groupe de voix, nous avons enfin calculé la moyenne des écart-types. Cette étude de variabilité des jugements perceptifs a été effectuée sur la cotation du paramètre G qui évalue le grade global de la dysphonie (Tableau 36).

Echelle Mixte	Moyenne Ecart-type
Grade 0	0,398
Grade 1	1,069

Tableau 36 : Valeurs moyennes des écart-types en fonction du grade global de la dysphonie.

[*] Cette méthode est largement utilisée en statistique dans le cas d'une variable quantitative car la moyenne est très sensible à un résultat de mesure déraisonnable.

Les résultats montrent que les valeurs recueillies se regroupent sur les grades 0 (voix normale) et 1 (voix légèrement dysphonique) de l'échelle mixte. On note une valeur de la moyenne des écart-types plus élevée pour le grade 1 que pour le grade 0 indiquant une forte variabilité de jugements pour noter les voix classées dans un grade intermédiaire que celles recueillies sur le grade extrême.

Le graphique ci-dessous (Figure 72) montre la répartition des groupes EE, ESA et ESIC en fonction du niveau de dysphonie d'après le classement par le jury d'écoute. 60% (n=47) de l'échantillon a été classé dans le grade 0 et 40% dans le grade 1 (n=31). Concernant le grade 0, le groupe EE (n=33) était le plus nombreux, ce qui est normal par rapport à ce que nous attendions car il s'agit du groupe « contrôle ». Ce dernier est suivi en nombre du groupe ESIC (n=9) puis ESA (n=5). Concernant le grade 1, la tendance est tout à fait inverse : le groupe ESA (n=15) est le plus nombreux, suivi du groupe ESIC (n=11) puis du groupe EE (n=5). Par rapport aux hypothèses formulées au départ de notre travail, nous nous attendions plutôt que les voix des enfants ESA et ESIC soient classées dans les grades 1 (dysphonie légère) et 2 (dysphonie moyenne). Toutefois, le nombre de voix classées dans le grade 1 associé à une « dysphonie légère » n'est pas négligeable (40%). Ce résultat peut être lié à la variabilité observée des jugements perceptifs (cf. Tableau 36). Dans la littérature, plusieurs études ont déjà démontré que le jury d'écoute est davantage variable sur les dysphonies intermédiaires (Kreiman, 1992[174], 1993[Erreur ! Signet non défini.] ; Yu, 2001[162], 2004[173]).

,

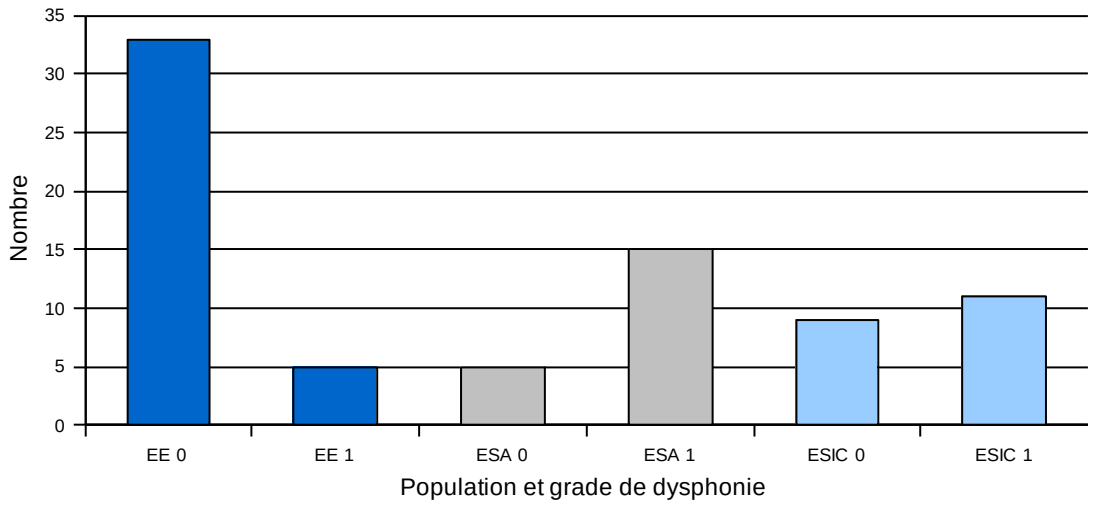


Figure 72 : Répartition des individus selon le grade de dysphonie d'après l'évaluation perceptive et le groupe.

5.3.3. Etude du lien entre les résultats de l'analyse instrumentale et les résultats de l'analyse perceptive

Dans cette dernière partie, nous abordons le lien entre les résultats du classement de l'analyse perceptive et les résultats de l'analyse instrumentale. A cet effet, nous avons réalisé une analyse statistique descriptive et inférentielle des données. L'approche descriptive a consisté à calculer la moyenne et l'écart type des valeurs de tous les paramètres acoustiques (16 paramètres au total) pour l'ensemble des voix en fonction du grade 0 et 1, obtenus dans la section précédente. Le tableau ci-dessous (Tableau 37) présente le résumé des résultats.

Paramètres acoustiques		Grade 0 (n=47)	Grade 1 (n=31)
Voix conv	FoUm (Hz)	236,39 (41,68)	244,88 (49,90)
	Dyn (Hz)	195,03 (83,53)	171,02 (74,22)
Voix Forte	FoUm (Hz)	299,57 (66,22)	312,31 (75,18)
	Dyn (Hz)	156,66 (78,28)	167,94 (124,27)
Voix d'appel	FoUm (Hz)	390,2 (82,10)	372,19 (72,81)
	Dyn (Hz)	185,73 (100,88)	164,03 (78,26)
/a/ tenu	Fo (Hz)	236,35 (45,22)	242,72 (55,65)
	Coef var Fo (%)	1,00 (0,64)	1,26 (0,69)
	Jitter Factor (%)	0,81 (0,73)	1,09 (0,87)
	Intensité (dB)	62,96 (2,99)	60,41 (9,66)
	Coef var Intensité (%)	0,88 (0,47)	1,11 (0,62)
	Shimmer Factor (%)	0,49 (0,21)	0,55 (0,26)
	R H/B (dB)	16,53 (4,34)	14,06 (5,45)
	TMP (sec)	13,89 (4,35)	11,76 (4,53)
Tessiture	Dyn (Hz)	295,35 (152,84)	330,76 (192,41)
Etendue	Dyn (Hz)	433,58 (171,53)	427,15 (179,11)

Tableau 37: Valeurs des moyennes et des écart-types (entre parenthèses) de chacun des paramètres en fonction des grades perceptifs.

Nous observons dans le tableau que tous les paramètres, sauf les paramètres de FoUm de la voix conversationnelle et de la voix forte, la dynamique de la voix forte, la Fo et la tessiture (en gris), varient de façon incohérente par rapport au grade de dysphonie. On s'attend que les valeurs de ces paramètres décroissent inversement au grade car l'aspect fréquentiel du signal vocal est toujours compromis dans la dysphonie. En pratique clinique, on perçoit par principe une diminution de la fréquence fondamentale ainsi qu'une réduction de la dynamique de la parole (courbe mélodique) chez le patient dysphonique. De même, il a été montré une réduction de la dynamique de la hauteur chez l'enfant sourd (Boone, 1966[179] ; Stathopoulos et col., 1986[180]). Néanmoins, d'autres études ont montré que la hauteur de la voix chez l'enfant sourd peut être « monotone », ou bien au contraire, excessivement instable ou diplophonique (Monsen, 1978[181] ; Parkhurst et Levitt, 1978[182] ; Smith, 1975[183] ; Subtelny et col., 1980[184]).

Ainsi, l'incohérence de certains paramètres (cités ci-avant) par rapport au grade de dysphonie réside dans le fait que les enfants de notre étude utilisent une hauteur de parole plutôt excessivement stable. Quant aux autres paramètres, nous constatons une cohérence qui valide les résultats de l'analyse acoustique.

Par ailleurs, les « paramètres incohérents » manifestent de façon systématique un écart-type plus important dans le grade 1 que dans le grade 0. De ces résultats, nous induisons que ces paramètres sont sensibles à l'utilisation de l'échelle à 4 niveaux afin d'effectuer la conversion des valeurs analogiques initialement obtenus. Cette échelle, relativement grossière, n'est pas adaptée pour la discrimination des différences fines dans le degré de dysphonie. De plus, il est probable que certaines voix du grade G1 ont été mal classées par le jury. D'ailleurs, les moyennes des écart-types des jugements perceptifs en témoignent (cf. Tableau 36).

Egalement, nous avons comparé les résultats du classement de l'analyse perceptive et les résultats de l'analyse instrumentale pour les paramètres aérodynamiques (5 paramètres au total) en fonction du grade 0 et 1 pour le groupe ESIC (Tableau 38).

Paramètres aérodynamiques	Grade 0 (n=9)	Grade 1 (n=11)
PSGE (hPa)	9,02 (2,65)	9,24 (2,48)
Débit d'air oral (dm3/s)	0,14 (0,05)	0,12 (0,05)
Intensité (dB)	71,58 (4,90)	73,27 (4,57)
Efficacité glottique (dB/hPa)	8,5 (2,04)	8,38 (1,69)
Efficacité laryngée (dB/hPa*dm3/s)	71,89 (32,04)	83,4 (33,93)

Tableau 38 : Valeurs des moyennes et des écart-types (entre parenthèses) des paramètres aérodynamiques pour le groupe ESIC en fonction des grades perceptifs.

Nous remarquons que les paramètres de débit d'air oral (DAB) et d'efficacité laryngée (en gris) ne manifestent pas une variation cohérente par rapport au grade de dysphonie. D'une part, on attendrait que la valeur de DAB augmente en fonction du grade ; et d'autre part, que la valeur de l'efficacité laryngée diminue (Guerrier et col., 2004[168]). Dans la littérature, le DAB est un paramètre largement utilisé pour décrire la voix pathologique et corrélé à l'intensification de la dysphonie (Hirano, 1988[185] ; Dejonckere, 1990[186] ; Hillenbrand, 1994[187]). Par ailleurs, 55% de la population ESIC a été classée dans le grade 1 (dysphonie légère) par le jury d'écoute. Le nombre non négligeable des voix classées dans un grade intermédiaire nous laisse supposer qu'il y a certaines caractéristiques vocales particulaires aux

enfants implantés cochléaires dans ce cas là qui désorientent le jugement perceptif des membres du jury.

De manière générale, nous constatons qu'un bon nombre de paramètres acoustiques et aérodynamiques révèlent une concordance entre les résultats obtenus de l'évaluation instrumentale et les résultats obtenus de l'évaluation perceptive

La deuxième approche que nous avons abordée pour comparer les résultats de l'analyse instrumentale et les résultats de l'analyse perceptive, était de caractère inférentiel. Nous avons élaboré un modèle statistique basé sur le principe de la *« Régression logistique »*. Notre modèle avait pour but de prédire le grade 1 (ou dysphonie légère) à l'aide des paramètres acoustiques des groupes EE, ESA et ESIC; et des paramètres aérodynamiques du groupe ESIC. Dans ce modèle, la significativité de ces paramètres a alors été testée. Le seuil de rejet de l'hypothèse nulle a été choisi à α=0.05. Nous avons d'abord estimé ce modèle de la probabilité de manifestation d'une dysphonie légère (grade 1) dans l'échantillon total (n=78). Ensuite, nous avons estimé ce modèle dans chaque groupe (EE, ESA, ESIC). Les résultats de l'analyse sont présentés dans le tableau ci-dessous pour les paramètres les plus significatifs (Tableau 39). Les trois dernières colonnes présentent les valeurs de l'estimation; l'erreur standard qui accompagne la valeur de l'estimation et la valeur de la probabilité critique (p.c.). On note en gris, les paramètres sensibles qui mettent en évidence une dysphonie légère chez la population. Le seuil de significativité a été fixé à α=0.05.

	Paramètre	Estimation	SE	p.c.
EE+ESA+ESIC	TMP	-0.100509	0.058658	0.0866
	Jitter	-1.904358	1.056871	0.0716
EE	Dyn Voix conv	-0.03422	0.01737	0.0489
	TMP	-0.6148	0.3101	0.0474
	R H/B	-0.6150	0.3209	0.0553
	Etendue	-0.0111073	0.0049648	0.0253
ESA	Fo /a/	-0.03226	0.01852	0.0815
	Jitter	-1.63188	0.98101	0.0962
ESIC	FoUm Voix forte	0.11664	0.06382	0.0676
	Dyn Voix forte	-0.05912	0.03077	0.0547
	Dyn Voix d'appel	-0.05736	0.03445	0.0959
	Etendue	0.007594	0.003584	0.0341

Tableau 39 : Résultats du modèle de Régression logistique estimé pour l'ensemble de l'échantillon (deux premières lignes) et pour chacun des groupes.

Pour interpréter les paramètres de ces modèles et leur significativité, on se pose la question : dans quelle mesure et de quelle façon le paramètre en question influence le grade de dysphonie ? Si la valeur d'estimation est positive, cela veut dire que lorsque la valeur du paramètre augmente, alors la probabilité ou tendance à la dysphonie légère augmente (Grade 1). A l'inverse, si la valeur d'estimation est négative, quand la valeur du paramètre augmente, alors la probabilité ou tendance à la dysphonie légère diminue. Ainsi, des paramètres ayant un impact sur la manifestation d'une dysphonie légère se manifestent.

Lorsqu'on applique le test à l'ensemble de la population tout groupe confondu, aucun paramètre ne se révèle significatif. Seulement le TMP et le Jitter se montrent à la limite de la valeur de la probabilité critique (2 premières lignes du Tableau 39). En revanche, quand nous effectuons le test pour chacun des groupes, nous constatons que certains paramètres tels que la dynamique de la voix conversationnelle, le TMP et l'étendue vocale se mettent en évidence. L'étendue vocale est le paramètre significatif commun aux groupes EE et ESIC. Néanmoins, la façon dont ce paramètre influence la probabilité de tendance à la dysphonie est opposée dans chaque groupe. Chez le groupe EE, si la valeur du paramètre augmente, la probabilité ou tendance à la dysphonie légère diminue ; ce qui paraît normal. En revanche, chez le groupe ESIC, si la valeur du paramètre augmente, la probabilité ou tendance à la dysphonie légère augmente.

Egalement, nous avons appliqué le test de Régression Logistique pour comparer les résultats de l'analyse objective des paramètres aérodynamiques et les résultats de l'analyse perceptive pour le groupe ESIC (Tableau 40). Le seuil de significativité a été fixé à α=0.05.

Paramètre	Estimation	SE	p.c.
PSGE	3,3979	1,8904	0,0723
Eff glott	5,5857	2,7408	0,0416

Tableau 40 : Résultats du modèle de Régression logistique estimé pour le groupe ESIC à l'aide des paramètres aérodynamiques.
On note en gris,

On observe que la pression sous-glottique se montre à la limite du seuil de significativité et l'efficacité glottique, quant à elle, s'avère significative indiquant une certaine sensibilité à révéler une « dysphonie légère » chez la population ESIC. Pour ce dernier paramètre, la

valeur estimée est positive. Cela signifie que si la valeur du paramètre augmente, la probabilité ou tendance à la dysphonie légère augmente. Néanmoins, cette interprétation n'est pas tout à fait la même en clinique : si la valeur de l'efficacité glottique augmente, la probabilité ou tendance à la dysphonie légère diminue.

5.4. Conclusion

Les résultats de l'analyse perceptive effectuée par un jury d'écoute fiable révèlent que :

- ✓ les voix de la population d'enfants sourds (ESA et ESIC) ne sont pas dysphoniques. En effet, les voix de ces groupes avec les voix d'enfants entendants ont été classées dans les deux premiers niveaux du paramètre "G" (grade global de dysphonie) de l'échelle GRBAS de Hirano qui ne signalent pas une dysphonie à proprement parler,

- ✓ d'une manière globale, l'évaluation instrumentale a été validée par l'analyse perceptive. Les résultats de l'évaluation acoustique et aérodynamique s'avèrent cohérents.
- ✓ un nombre non négligeable de voix classées perceptivement dans le grade intermédiaire (G1) témoignent de la difficulté d'évaluation des caractéristiques vocales chez les enfants sourds.

- ✓ les paramètres acoustiques ayant un impact sur la manifestation d'une dysphonie légère sont : dynamique de la voix conversationnelle, TMP et étendue vocale pour le groupe EE ; et, étendue vocale pour le groupe ESIC.

- ✓ le paramètre aérodynamique ayant un impact sur la manifestation d'une dysphonie légère est l'efficacité glottique pour le groupe ESIC.

- ✓ L'échelle GRBAS n'est pas une échelle pertinente pour évaluer et analyser la voix des enfants déficients auditifs.

- ✓ Il est nécessaire de mettre au point une échelle d'évaluation perceptive de la voix avec des critères bien définis (p.e. nasalité, forçage), tenant compte les particularités de la voix de la population infantile à déficience auditive. Futurs travaux devraient explorer le degré pour lequel les différentes mesures physiologiques influencent les jugements des auditeurs de l'intelligibilité des enfants sourds et des enfants entendant.

- ✓ Ainsi, l'ensemble des résultats suggère qu’effectivement la voix des enfants appareillés et implantés cochléaires de notre étude ne présente pas les caractéristiques traditionnellement retenues pour déterminer la voix dysphonique.

CHAPITRE IV : PROPOSITION DE MODÉLISATION

Nous présentons ici la deuxième partie de la méthodologie expérimentale. Cette partie présente une expérience menée dans le domaine de la modélisation et de la simulation numérique. Elle a consisté à formaliser d'une certaine façon les raisonnements conduits au cours de nos travaux de thèse sur l'action de la boucle audio-phonatoire, objet de notre étude. Etant donné que nous n'avons pas trouvé de modèles acoustiques nous permettant de mieux comprendre la dynamique intriquée de ce mécanisme régulateur de la voix, nous avons envisagé d'en proposer un. De même, nous avons voulu débuter dans l'exercice de la simulation numérique car c'est un outil qui est de plus en plus couramment utilisé dans la pratique scientifique. Dans ce chapitre, le modèle théorique conçu et le simulateur numérique développé. Ce simulateur a été développé dans le cadre d'un « projet industriel » en collaboration avec l'Ecole Centrale de Lyon.

1. Proposition d'un modèle représentant le rôle de la boucle audio-phonatoire chez l'enfant déficient auditif

La compréhension des processus physiologiques impliqués dans la régulation de la parole et de la voix chez l'enfant à déficience auditive et implanté cochléaire représente un énorme défi pour les orthophonistes, phoniatres, phonéticiens et autres spécialistes intéressés au sujet. En effet, divers processus physiologiques intriqués ont lieu afin de contrôler la production de la parole et de la voix via une boucle de rétroaction dénommée « boucle audio-phonatoire » (chapitre I, section 3). Depuis quelques décennies, la recherche sur la production de la parole et de la voix est abordée par une approche analytique de modélisation des phénomènes physiques sous-jacents. Dans ce sens, on vise ainsi à déterminer plus finement le rôle des mécanismes et des structures participant à ce processus. Dans cette perspective, nous avons tenté de développer un modèle explicatif sur la régulation des différents paramètres acoustiques et aérodynamiques de la voix émise par un enfant à déficience auditive. Notre modèle prétend contribuer à la compréhension du processus complexe et dynamique de contrôle audio-phonatoire, tout en tenant compte des avancées de la recherche dans le domaine de la modélisation de la phonation. Notre réflexion s'inscrit dans la lignée de plusieurs années de recherches menées sur le vaste sujet de la production et de la perception de la voix et de la parole.

1.1. Objectifs du modèle

L'objectif principal de cette partie de nos travaux de thèse est d'élaborer un modèle théorique et de le tester ensuite à travers la conception d'un simulateur numérique. La création de cet outil s'inscrit dans une démarche de mise en évidence et de compréhension de l'action de régulation que l'audition exerce sur la parole et la voix, via une boucle de rétroaction dénommée « boucle audio-phonatoire ». En effet, nous souhaitons comprendre avant tout le lien indissociable existant entre la perception et la production orale, tout en ouvrant de nombreuses perspectives de recherche. Ce modèle édité sur un support informatique nous permettra de simuler la régulation de la voix chez l'enfant sourd afin de nous faire une idée de l'incidence du traitement du signal acoustique sur la qualité de la voix émise par l'enfant appareillé et l'enfant implanté cochléaire. Nous tenterons de répondre aux questions formulées au départ de nos travaux de thèse : 1) quel est l'effet du traitement du signal (stratégie de codage) de la prothèse auditive classique et de l'implant cochléaire sur les paramètres acoustiques de la voix de l'ESA et de l'ESIC ?, 2) peut-on établir une relation entre la quantité et qualité d'informations acoustiques entrantes (représentant la perception du son) et sortantes (représentant l'émission vocale) dans un modèle macro-systémique en *« boucle de rétroaction négative »* ?

1.2. Caractéristiques du modèle proposé

Nous ne ferons pas ici une argumentation détaillée et exhaustive des concepts épistémologiques sous-jacents de notre modèle, mais nous exposerons seulement les principes qui le régissent et les caractéristiques qui le définissent, à savoir : a) un modèle descriptif du phénomène étudié, b) un modèle basé sur une approche systémique, c) un modèle systémique de type asservi, d) un modèle représentant les structures et l'organisation du système.

a) Un modèle descriptif du phénomène étudié

D'une part, on sait qu'un modèle est le prolongement de la théorie et que derrière l'idée de la conception d'un modèle, nous avons celles de *reproduction, représentation et explication* (Algoud, 2001)[188]. Notre modèle a été élaboré comme une sorte de « maquette » afin de représenter le phénomène de la perception et de l'émission du signal sonore dans une étroite interaction physique. D'autre part, quand on construit un modèle, c'est généralement pour réaliser des expérimentations soit impossibles à exécuter sur l'objet réel (par exemple, les centrales nucléaires) soit trop coûteuses pour leur reproductibilité (par exemple, dans la

discipline de la biologie génétique). Dans notre cas, la conception d'un modèle théorique était nécessaire car l'objet d'étude portait sur des enfants. Or, les expérimentations chez les enfants ou de manière générale chez l'être humain sont restreintes[*]. Nous étions intéressés à étudier l'action d'un mécanisme physiologique chez l'être humain et à manipuler certains paramètres afin d'observer l'influence sur l'ensemble du mécanisme. Pour ces raisons, nous avons estimé convenable de concevoir un modèle nous permettant de *« reproduire »* autant de fois que nécessaire et de façon satisfaisante le phénomène réel étudié. Une fois la modélisation établie, l'objet et son modèle peuvent être étudiés ou comparés, et ce de façon plus simple qu'en manipulant les données réelles.

Par ailleurs, il existe deux approches généralement opposées lors de l'élaboration d'un modèle. Un modèle peut être de caractère *prédictif* ou de caractère *descriptif* (Ghio, 2007[189]). Le modèle relevant d'une approche prédictive a pour principe d'obtenir un résultat inconnu à partir de données explicatives (par exemple, les modèles prédictifs météorologiques qui prévoient le temps à partir de données atmosphériques). Les données explicatives sont considérées dans ce cas-là comme les paramètres primordiaux d'entrée du modèle. De ce fait, une erreur de prédiction de ce modèle sera estimée comme une faiblesse de la modélisation.

Le modèle relevant d'une approche descriptive a pour principe de représenter, de façon organisée et synthétique, des données réelles. L'objectif est de rendre compte, si possible de manière interprétable, de tout un ensemble d'informations (par exemple, les modèles descriptifs économiques qui décrivent la dynamique de transformation économique observée dans un pays donné). Les données observées ici sont considérées comme des données d'entrée, primordiales et incontestables. Les paramètres du modèle sont alors ajustés pour rendre compte au mieux de ces données. Notre modèle se situe dans cette perspective de représentation.

Généralement, les modèles sont représentés sous la forme de schémas fonctionnels composés de blocs connectés par des lignes d'action. On y retrouve des variables d'entrée qui peuvent être les commandes du modèle, et des variables de résultats en sortie. Les blocs sont le moteur du modèle, intégrant la connaissance ou la description du phénomène modélisé. Ces boîtes

[*] Il existe un code d'étique universel qui veille à ce que les expérimentations sur les humains respectent certaines normes. Dès lors, il relaie nécessairement les valeurs inhérentes à la recherche médicale. Ce code dénommé « code de bioéthique » est formulé par des organisations internationales comme l'Organisation Mondiale de la Santé ou le Comité de Protection de Personnes Sud-Méditerranée II en France, s'incarnant parfois dans des codes déontologiques quasi-juridiques.

peuvent s'avérer transparentes quand les connaissances sont explicitées ou noires, sans capacité d'interprétation des mécanismes abordés quand les connaissances ne le sont pas.

b) Un modèle basé sur une approche systémique

Nous nous sommes portés sur la conception d'un modèle systémique car nous considérions le phénomène étudié, c'est-à-dire le processus physiologique humain de régulation de la voix et de la parole, comme un phénomène s'inscrivant dans la dynamique d'un *« système »*. Par ailleurs, nous avons également estimé le phénomène étudié comme étant un *« système dynamique »*. En effet, il a auparavant été identifié un agencement interne de caractère observable dans le processus de régulation de la voix et de la parole. Il existe une répartition de rôles et une exécution de diverses activités par les mécanismes participant à la régulation de la voix (chapitre I). Le système, objet de notre observation, présente des propriétés qui s'avèrent nécessaires de répertorier et d'analyser. Comme dans tout système, il existe un réseau de communication, véhiculant des informations. Nous tenterons de l'identifier et d'en montrer le rôle dans l'existence du système selon les objectifs de notre étude. Par exemple, nous tenterons d'identifier le flux d'information entrant et sortant des divers blocs composant notre système et nous analyserons en quoi le flux de ces informations caractérisent le comportement des éléments.

c) Un modèle systémique de type « asservi »

Notre modèle s'inspire fondamentalement des principes formulés par Forrester (1958)[190] qui sont évoqués par Algoud (2001)[188] dans son ouvrage *« Systémique »*. Forrester était automaticien. Son œuvre a consisté à transférer en sciences humaines les acquis de l'automatique. Au cœur de cette dernière se situe la théorie des systèmes asservis, ou de *« feedback »* pour les anglo-saxons. Il applique les techniques de simulation à l'étude des phénomènes économiques et donne à son modèle une représentation plus adaptée aux techniques de simulation utilisées en sciences humaines. Nous partons de l'hypothèse que notre objet d'étude, à savoir, la boucle audio-phonatoire, se comporte comme un système asservi, nous avons donc basé la conception de notre modèle sur les principes de la *« théorie de la dynamique des systèmes »*. La dynamique des systèmes étudie les systèmes en mouvement et les forces qui conduisent à ce mouvement. Dans un système de premier ordre, c'est-à-dire le plus simple, les variables de niveau sont représentées par des rectangles, et les flux circulant sont contrôlés par des vannes. Forester représente évidemment, en conformité avec l'étude des systèmes asservis, un retour de la variable de niveau (SR) sur l'entrée où elle est confrontée à la variable-objectif (SO) (Figure 73).

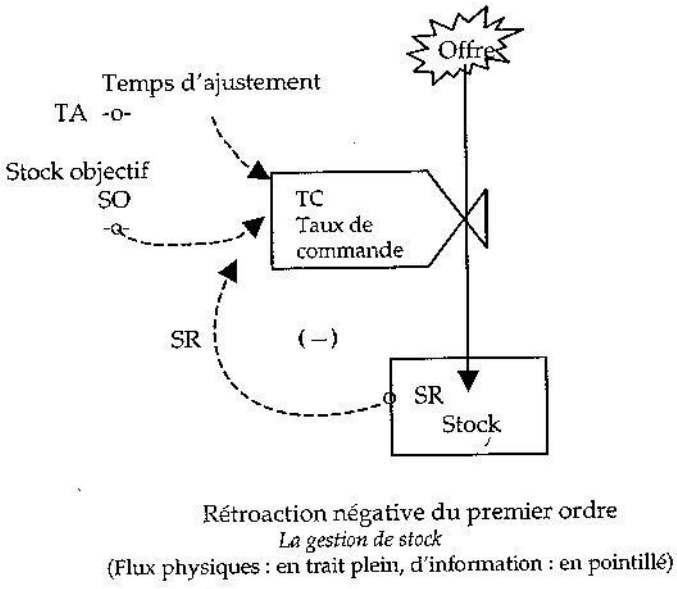


Figure 73 : Schéma d'un exemple de modèle économique de 1er ordre selon Forester (d'après Algoud, 2001).

Afin de mieux comprendre la dynamique d'un système asservi ou de *« rétroaction négative »*, Algoud (2001) donne l'exemple du fonctionnement d'une chasse d'eau (Figure 74).

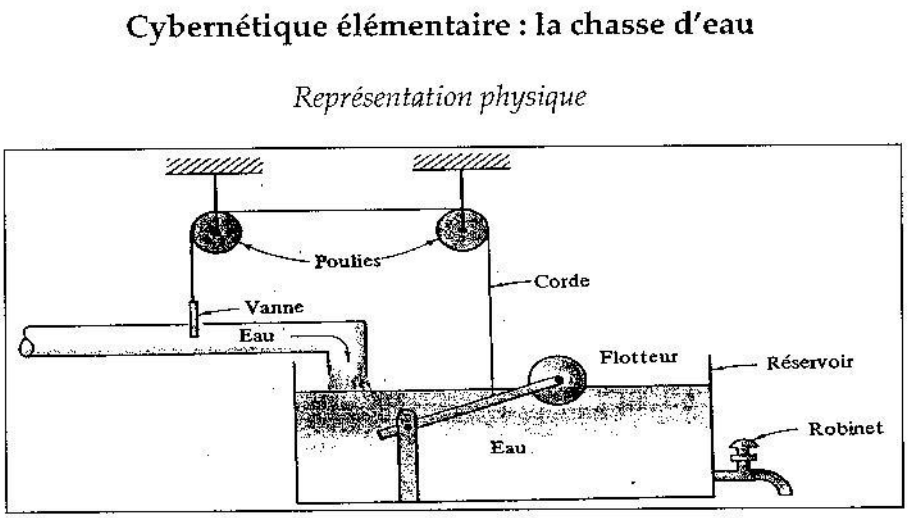


Figure 74 : Schéma d'un système asservi : la chasse d'eau (d'après Algoud, 2001).

Quand la chasse d'eau est tirée, l'eau du réservoir est libérée, ce qui entraîne l'ouverture du robinet d'eau de remplissage. Tout de suite, l'eau en remplissant le réservoir (ou chasse) fait remonter un flotteur, qui va, en remontant, entraîner une tige qui ferme l'arrivée de l'eau. Cette

phase est facilement identifiable au bruit que l'eau fait en remplissant le réservoir. Il diminue peu à peu jusqu'à la fermeture totale de l'arrivée d'eau. Le flotteur, dans sa remontée, réduit l'écart entre le niveau d'eau-objectif, qui correspond à sa position haute, et le niveau à un moment donné qui est celui donné par la position occupée par le flotteur dans sa remontée. C'est la différence entre le niveau objectif et le niveau courant qui va commander l'arrivée d'eau. Plus on va se rapprocher du niveau d'eau final et plus le niveau d'eau va se réduire par l'intermédiaire de la tige. La finalité de ce système est figée, c'est-à-dire, à chaque fois que le niveau du flotteur est en position haute (ou le niveau objectif) le robinet d'eau est coupé. Ce genre de système ne fait que réaliser un programme déterminé. C'est dans cette dynamique que nous concevons le modèle qui représenterait le fonctionnement de la boucle audio-phonatoire. Cette dynamique sera expliquée en détail ultérieurement. Par ailleurs, nous avons intégré la variable « temps » car, d'une part, il est impossible d'atteindre un état final instantanément, et d'autre part, tout système évolue. De ce fait, le système devient dynamique. La poursuite d'un but caractérise les systèmes dynamiques. Sans but, il n'y a pas de dynamique ou de mouvement.

d) Un modèle représentant les structures et l'organisation du système

Algoud (2001[188]) déclare qu'*« un système [élémentaire] est un ensemble cohérent de structures et de lois-institutions organisé suivant une finalité »* (p. 772), (Figure 75).

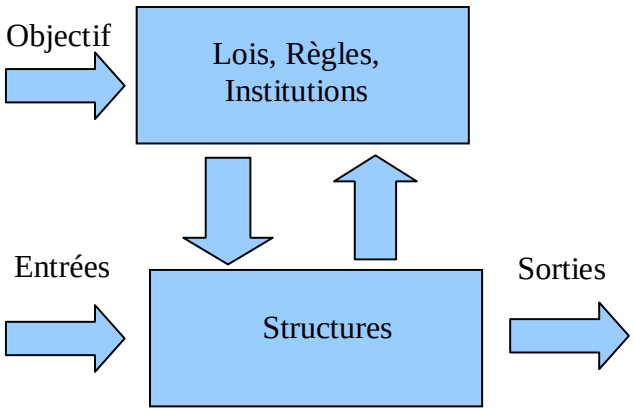


Figure 75 : Schéma représentant les composants d'un système élémentaire ou micro-système. Cette catégorie de système est dotée d'un objectif et d'un niveau d'organisation (d'après Algoud, modifié, 2001).

En effet, les structures sont des relations spécifiquement établies par rapport à un système doté d'une finalité. L'étape finale du cycle d'un système consiste à rapprocher l'état-objectif de cet état final pour en déduire d'éventuelles actions correctrices. Un système élémentaire ou *« micro-système »* peut s'assembler avec d'autres formant ainsi des systèmes beaucoup plus

complexes ou « *macro-systèmes* ». Notre modèle tente de décrire les relations existant entre les micro-systèmes de production et de perception constituant le macro-système de la boucle audio-phonatoire. Dans la section suivante, nous allons présenter le modèle conçu sur la base de ces concepts théoriques de systémique évoqués.

1.3. Description du modèle proposé

Le modèle théorique conçu ici représente les phénomènes physiques et physiologiques ayant lieu lors de l'action de la boucle audio-phonatoire, comme des événements qui s'inscrivent dans un macro-système. En effet, l'action de la boucle audio-phonatoire implique l'intervention d'autres mécanismes tels que celui de la phonation, de l'audition et du cerveau, nous les appellerons « *micro-systèmes* ». Par ailleurs, la voix est sous le contrôle des divers centres auditifs comme le montre la voix des déficients auditifs profonds. Ce contrôle audio-phonatoire semblerait comporter deux composants : un composant automatique involontaire et un composant volontaire (Guerrier et col., 2004[168]). Le composant volontaire est probablement commandé par les voies cortico-bulbaires en réponse aux informations acoustiques provenant de l'oreille interne qui sont acheminées jusqu'au cortex auditif. De même, il existerait un composant involontaire définit par un ensemble de reflexes acoustico-laryngés. Ce contrôle s'exerce en synergie avec le contrôle proprioceptif qui permet l'anticipation pré-phonatoire. C'est ce qui permet aux chanteurs d'émettre d'emblée une note à la hauteur voulue sans attendre le contrôle auditif (Mürbe et cols., 2002[191], 2003[192]). Cet ensemble proprioceptif expliquerait pourquoi la voix des déficients auditifs est relativement peu perturbée de nos jours.

Pour aboutir à cette modélisation, nous avons tout d'abord commencé par identifier les principaux micro-systèmes correspondant aux instances physiologiques impliquées dans le contrôle audio-phonatoire. Ces micro-systèmes ont été représentés en blocs ou modules décrits ci-après, à savoir : i) le bloc vibrateur, ii) le bloc résonateur, iii) le bloc récepteur, iv) le bloc comparateur (Figure 76).

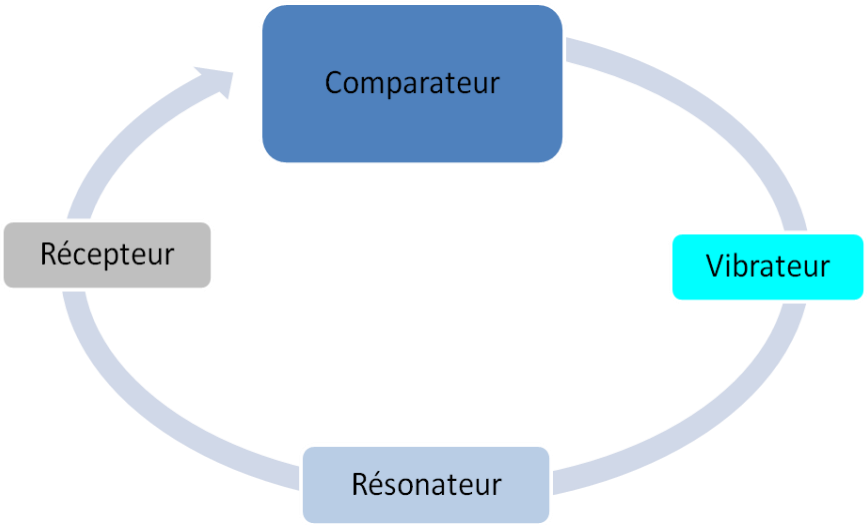


Figure 76 : Schéma simplifié de la disposition des micro-systèmes dans le macro-modèle proposé.

1.3.1. Bloc « Vibrateur »

Le bloc « vibrateur » représente le mécanisme d'émission sonore ou le mécanisme de vibration des cordes vocales à proprement parler. Il a été conçu à partir de l'analyse des critiques formulées et des résultats fournis par plusieurs modèles déjà existant dans la littérature. Nous avons alors eu à faire un choix entre trois grands axes de modélisations existant dans la bibliographie :

1) D'une part, il existe des modèles basés sur les techniques et méthodes de traitement du signal acoustique. Ces modèles caractérisent basiquement un cycle de la dérivée de l'onde glottique à l'aide d'au moins cinq paramètres acoustiques (Sciamarella et D'Alessandro, 2002)[196], comme le modèle de Liljencrants et Fant (1985)[193].

2) D'autre part, il existe une série de modèles purement mécaniques représentant les cordes vocales comme un système discrétisé à *n masses* reliées entre elles et au larynx par des liaisons « ressort-amortisseur ». Dans ces modèles, chaque corde vocale est représentée comme un système d'auto-oscillation mécanique composé d'un couplage de deux masses-ressorts. Les paramètres de ces liaisons (raideur, coefficient d'amortissement) sont ajustés afin de recaler le modèle à la réalité. Seules les caractéristiques géométriques des cordes vocales sont prises en compte pour produire un son. L'auto-entretien de la vibration cordale est attribué à une modification de la géométrie glottique qui crée des distributions de pression glottique différentes pendant les phases d'ouverture et fermeture du cycle d'oscillation des

cordes vocales. La déformation non-uniforme du tissu des cordes vocales est assurée par un modèle mécanique ayant au moins deux degrés de liberté. Cependant, cette modélisation rencontre ses limites car elle ne possède pas assez de degrés de liberté pour modéliser l'ensemble des mouvements des cordes vocales. Elle est assez simple pour fournir une bonne représentation de la réalité. Malgré sa simplicité, cette modélisation se révèle appropriée pour la synthèse de la parole en temps réel utilisant un grand nombre de paramètres. Par exemple, nous citons le premier modèle largement connu, à savoir celui développé par Ishizaka et Flanagan (1972)[194] qui présentent plus de dix-neuf paramètres. Ce modèle a été récemment révisé par Pelorson et col. (1994)[195].

3) Enfin, il existe des modèles mixtes développés sur les principes de la modélisation du traitement du signal acoustique et de la modélisation mécanique et géométrique des cordes vocales. En effet, il apparaît qu'on ne peut pas simuler fidèlement la production vocale en utilisant seulement un modèle mécanique. Les modèles de production sonore peuvent prendre en compte les particularités subtiles qui ne sont pas reproduites par les modèles de traitement du signal acoustique. Ces modèles prétendent dévoiler la relation entre les paramètres acoustiques et physiques pour expliquer le phénomène de génération de la voix. Comprendre cette relation permettrait de comprendre de façon approfondie le fonctionnement du mécanisme responsable des différentes caractéristiques de la voix. Parmi ces modèles, nous avons retenu celui de Sciamarella (2002)[196] que nous présentons ci-dessous.

Modèle retenu pour le bloc de génération du son

Notre travail a donc consisté ici à trouver un modèle qui soit à la fois fidèle à la réalité dans ses résultats, et à la fois convenable par rapport aux objectifs de notre étude. Le modèle de Sciamarella (2002[196], 2003[197], 2004[198]), retenu dans notre étude, permet de coupler les effets dus à la mécanique des cordes vocales et à ceux dus à l'écoulement de l'air. Le modèle de Sciamarella se veut bien plus complet que les modèles mécaniques simples car il prend en compte un autre aspect important qui est l'écoulement de l'air et la façon dont il est perturbé par les cordes vocales. A partir de l'onde de synthèse du flux glottique générée par les différents paramétrages des paramètres physiques du modèle à deux masses, Sciamarella élabore un algorithme pour le calcul des paramètres acoustiques constituant un signal à reproduire. Cet algorithme considère diverses configurations du tract vocal, à savoir : convergente, divergente, collision. Ces configurations seront plus tard expliquées. Sciamarella trouve que la gamme de valeurs des paramètres acoustiques d'un signal sonore obtenues par

le modèle est assez large pour reproduire tout genre de comportement du flux glottique observé.

Ainsi, le modèle de Sciamarella se révèle plus adapté et plus fiable par rapport à la réalité physique du phénomène analysé par notre étude. De plus, ce modèle est un bon compromis entre la complexité de représentation du micro-système (ou bloc) de génération du signal sonore et son implémentation dans notre modèle systémique de régulation audio-phonatoire (macro-système). Certes, il existe d'autres modèles de génération du son plus précis que celui-ci. Cependant, il ne faut pas oublier que les autres blocs (ou micro-systèmes) constituant le modèle systémique ont une précision limitée, et donc il ne sert à rien de prendre un modèle de cordes vocales complexe si la précision supplémentaire qu'il apporte n'est pas utile.

Le modèle symétrique à deux masses avec point de séparation mobile que nous avons retenu est représenté schématiquement ci-dessous (Figure 77).

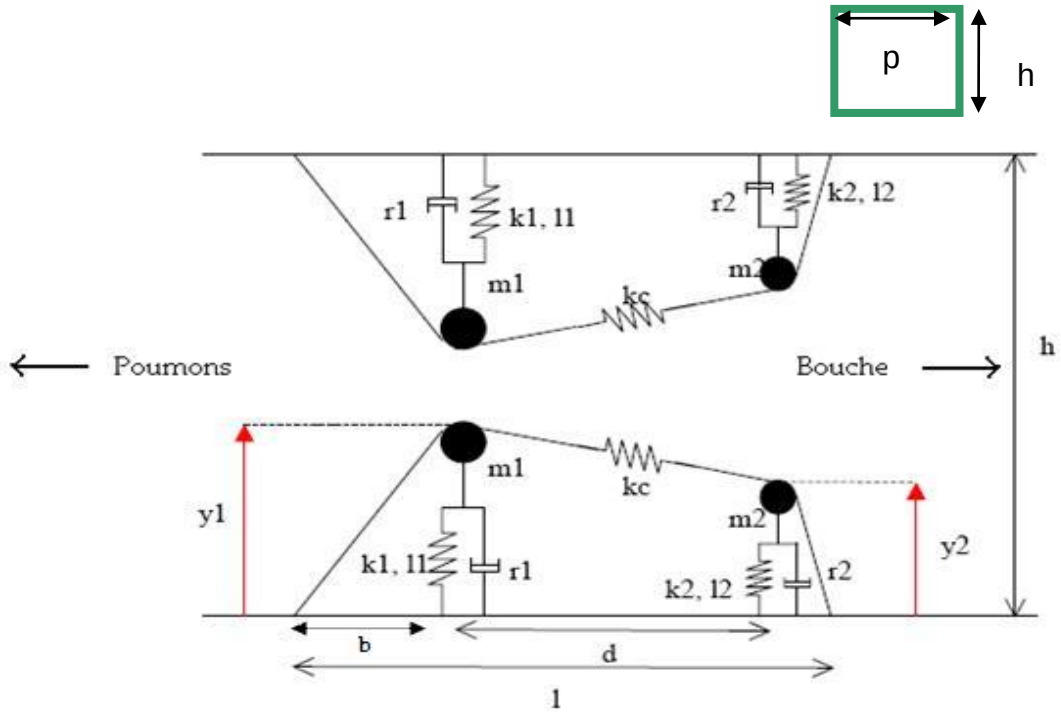


Figure 77 : Schéma représentant les différents paramètres et la structure géométrique du modèle symétrique à deux masses retenu pour notre simulateur

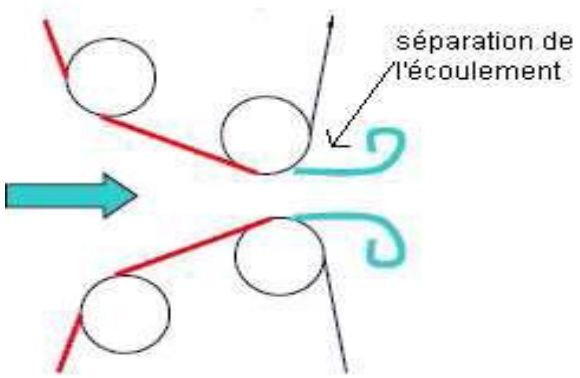
Le modèle peut se présenter en coupe transversale représentant la structure des cordes vocales en deux dimensions disposée horizontalement. Il est à noter que les cordes vocales sont modélisées symétriquement, c'est-à-dire que la corde vocale supérieure aura les mêmes caractéristiques et les mêmes mouvements que la corde vocale inférieure. Cette symétrie est réalisée par rapport à l'axe du canal du fluide aérien. Chaque corde vocale est modélisée par

deux masses *m1* et *m2* qui ne peuvent bouger que dans une seule direction, sens vertical, grâce à un système masse-ressort amorti. Elles sont liées entre elles par une raideur *kc* (contraction des cordes vocales), et liées chacune à la paroi du larynx par un amortisseur de raideur *k1* et *k2* (tension des cordes vocales), de longueur à vide *l1* et *l2* et d'amortissement *r1* et *r2*. Le système masse-ressort amorti est caractéristique de systèmes qui ont un mouvement oscillatoire, et dont l'amplitude des oscillations diminue avec le temps. La hauteur du canal est représentée par *h (x, t)* et est considérée comme un paramètre de fonction linéaire, *l* représente l'épaisseur des cordes vocales et *d* la distance prise entre les deux masses. On a ainsi une dimension supplémentaire *p* du canal (cf. Figure 77), qui est la dimension prise orthogonalement par rapport au schéma du conduit glottique. Ce conduit est alors un tube de section rectangulaire. Le flux d'air s'écoule sur ce schéma de gauche à droite. Comme on l'a dit précédemment, des phénomènes liés à l'écoulement rentrent aussi en jeu. On considère qu'en amont de ces cordes, les poumons fournissent une certaine pression relative *Pe* (surpression par rapport à la pression atmosphérique) et donnent une certaine vitesse *Ve* au flux. Dans le cadre de ce modèle, on se place dans une hypothèse de fluide localement incompressible et quasiment stationnaire existant depuis la trachée jusqu'à un point où le flux aérien sépare les parois de la corde vocale pour former une colonne d'air libérée.

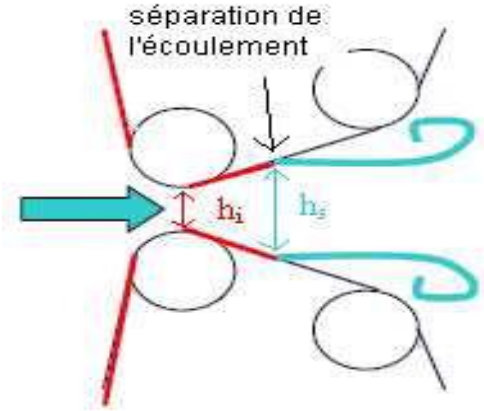
Ce modèle permet de modéliser diverses configurations géométriques des cordes vocales en fonction des paramètres précédemment décrits. Il existe une première configuration, dite « *convergente* » (Figure 78a), l'air pénètre entre les cordes sans perturbation majeure (comme indiqué dans la figure avec des flèches en bleu). L'écoulement d'air ne se sépare de la paroi de la corde vocale qu'à la fin du canal. Avant ce point, on dit que le long des parois (signalées en rouge), il y a formation d'une couche limite où la pression normale reste constante.

La deuxième configuration est dite « *divergente* » (Figure 78b). Ici, il existe un point le long de la corde où l'écoulement se décolle de la paroi. Au-delà de ce point, comme précédemment, l'écoulement est dit turbulent, et on considérera que la pression dans cette zone est négligeable devant celle présente en amont. La position du point de séparation est normalement le fruit d'un calcul de mécanique des fluides, cependant le critère géométrique de Liljencrants (1995)[199] a été retenu. On prendra ce point comme celui où *hs (en bleu)/hi (en rouge)=1,2*. Ceci permet de s'affranchir de longs calculs.

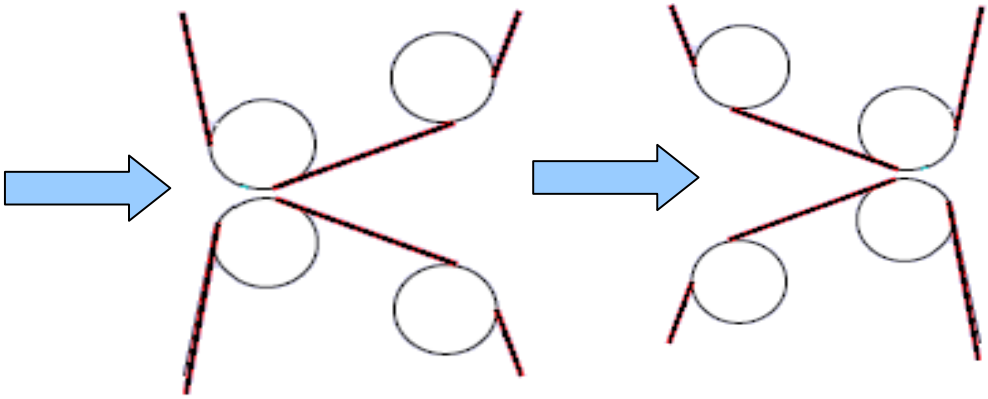
Enfin, une dernière configuration, où deux des masses disposées symétriquement sont suffisamment proche l'une de l'autre *distance=hcollision* pour que l'on ait collision (Figure 78c). Dans ce cas, les raideurs *k1*, *k2*, *kc* sont multipliées par un coefficient *C*, et on considère que l'écoulement est stoppé tant que les masses ne sont pas séparées d'une distance supérieure à $h_{collision}$.



6a) Configuration convergente



6b) Configuration divergente



6c) Configuration de collision

Figure 78 : Schémas des différentes configurations glottiques pouvant être représentées selon le modèle symétrique à deux masses.

En résumé, les poumons vont fournir en amont une certaine pression qui varie au cours du temps. Cette pression va exercer une force sur les parois des cordes vocales. Cette force est à l'origine du mouvement des cordes vocales. On peut alors prévoir le comportement dynamique des cordes vocales en étudiant le système masse-ressort amorti.

L'écriture des équations de ce bloc est présentée dans l'annexe 14 *(« Ecriture d'équations des blocs du modèle »*). La résolution des équations du modèle s'est effectuée en plusieurs étapes (Figure 79). Ainsi à chaque pas de temps *t*, on calcule la pression au niveau des cordes vocales (à l'origine du son) et la position des cordes vocales.

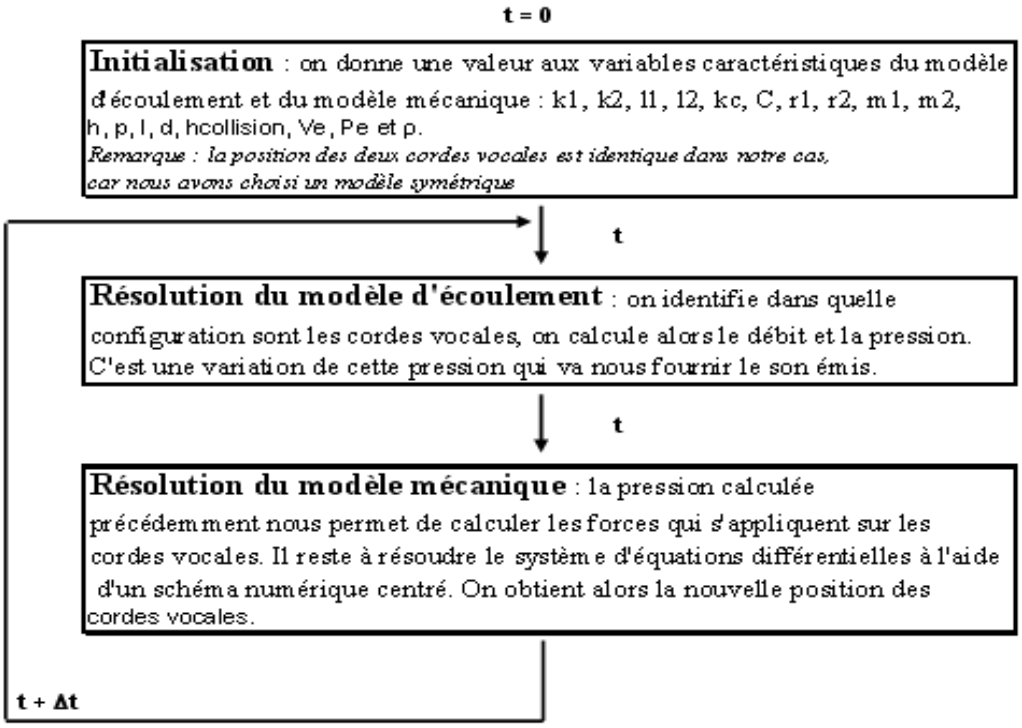


Figure 79 : Architecture de l'algorithme utilisé pour le bloc « vibrateur » ou des cordes vocales.

Dans un premier temps, on a une phase d'initialisation des données et de définition des constantes. Celle-ci permet de rentrer d'une part les différentes caractéristiques du modèle mécanique et de l'écoulement de l'air qui sont les suivantes : *k1, k2, l1, l2, kc, C, r1, r2, m1, m2, h, p, l, d, h collision, Ve, Pe et ρ masse volumique de l'air*. D'autre part, on fixe aussi une géométrie initiale des cordes vocales en entrant une position des masses *y1, y2*[*].

Dans un deuxième temps, grâce aux données de géométrie fournies, l'algorithme de résolution prévoit l'emplacement du point de séparation des cordes vocales selon la configuration glottique sollicitée, à savoir : en fin de conduit s'il y a *convergence*, dans le conduit en cas de

[*] On pose une durée de simulation *T* et un nombre de bouclages *n* pour la simulation numérique. Le nombre de bouclages a une influence directe sur la qualité de la simulation, plus *n* est grand, plus l'incrément de temps est faible, meilleure est la simulation. Ce nombre de bouclages est le nombre de points. Il faut être toutefois attentif au problème de suréchantillonage lorsque le nombre de points est trop important.

divergence, et pas de point ni d'écoulement s'il y a *collision*. A partir des hypothèses faites sur le fluide, le modèle permet aussi de calculer le débit d'air et la pression régnant dans le conduit. Ces deux données sont stockées à chaque incrément, car ce sont elles qui vont fournir le son voisé. Connaissant la pression, il est alors possible de connaître quelles forces *F1* et *F2* vont s'appliquer respectivement sur *m1* et *m2*, en répartissant les forces de pression suivant la Figure 80. Sur m1 vont s'appliquer les forces de pression s'exerçant sur la partie rouge, sur m2 celles s'exerçant sur la partie verte.

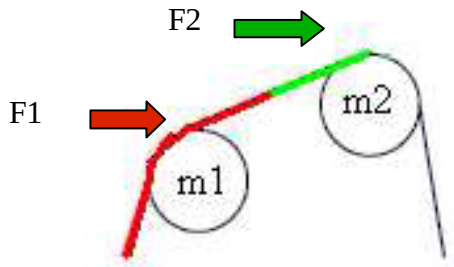


Figure 80 : Répartition des forces de pression sur les masses.

Enfin, dans la dernière étape de résolution, l'algorithme effectue le bilan des forces s'exerçant sur les deux masses *m1* et *m2*, comprenant les forces *F1* et *F2* et celles liées aux différentes raideurs et amortissements. On va ainsi avoir un déplacement des masses pendant l'incrément de temps. La nouvelle position des masses est alors donnée au modèle d'écoulement qui va répéter ses opérations avec cette nouvelle géométrie. L'ensemble de cette boucle s'arrête lorsque la durée de simulation est écoulée, et fournit donc les variations de pression et de débit sur cette durée. Il est important de noter que cet algorithme a été développé pour s'intégrer parfaitement à l'ensemble du modèle du macro-système représentant la boucle audio-phonatoire. A présent, l'algorithme peut être développé pour que s'insèrent les différents blocs ou micro-systèmes « résonateur », « récepteur » et « comparateur » dans le bouclage. Le principe est le suivant : à chaque incrément de temps, le son voisé fourni ici sera transformé par un bloc « résonateur », puis ce son sera écouté par un bloc « récepteur » dont les informations permettront au bloc « comparateur » d'agir sur les paramètres *k1, k2, kc, r1, r2, C, Ve, Pe,* paramètres qui sont pour le moment fixés dans la phase d'initialisation.

1.3.2. Bloc « Résonateur »

Le bloc « résonateur » représente le mécanisme de résonance qui transforme et achemine le son produit par la vibration des cordes vocales jusqu'à la cavité orale pour y rendre effective l'émission vocale. Les cordes vocales génèrent des sons que les cavités résonantes supra-glottiques vont moduler pour en faire un son plus riche. Ces résonateurs modifient le contenu harmonique du signal en modifiant l'amplitude des différentes composantes fréquentielles du signal source. Nous avons tenté de modéliser le bloc résonateur, censé représenter la cavité pharyngale, la cavité buccale et les lèvres[*].

La modélisation du conduit vocal a été notamment développée dans les travaux de recherches sur la production de la parole, champ d'investigation qui a donné lieu à de nombreux travaux (Boë et col., 1994)[200]. La modélisation du conduit vocal se porte d'avantage sur une approche prédictive de type articulatoire la plus fidèle possible de la réalité, ceci dans le but de mieux appréhender les mécanismes de production de la parole avec sa projection acoustique. On trouve dans la littérature plusieurs niveaux de modélisation du conduit vocal (Ghio, 2007)[189] proposant des modèles acoustiques, géométriques, articulatoires, neurophysiologiques et électriques (Stevens et col., 1953 in Ghio, 2007)[201].

Modèle retenu pour le bloc de résonance du signal sonore

Le modèle que nous avons retenu pour représenter la résonance dans le conduit vocal s'inscrit dans le niveau de modélisation géométrique. Le modèle choisi est celui de Fant (1960)[202]. Le principe de cette modélisation postule que les sons de la parole résultent du filtrage réalisé par le conduit vocal, du son glottique et/ou du son provenant d'une source de bruit de friction. Grâce à la connaissance des lois qui régissent la relation entre géométrie et ses conséquences, il est possible de générer un signal acoustique en sortie.

Le modèle de Fant (1960) « à quatre tubes » est un des modèles précurseurs dans le domaine (Figure 81). Le conduit vocal est modélisé en quatre tubes représentant les lèvres, la cavité orale (cavité avant), la constriction de la langue et le pharynx (cavité arrière). Le modèle de Fant est considéré comme étant capable de prédire et de générer un résultat acoustique fidèle à la réalité. Il décrit le conduit vocal comme un conduit composé de quatre tubes de dimensions variables et connectés entre eux. Des travaux plus récents ont proposé

[*] Il est impératif de signaler que nous n'avons pas considéré la modélisation de la cavité nasale dans ce travail, bien qu'elle soit capitale dans la modélisation de production de la voix et de la parole. Ceci car nous nous sommes initialement intéressés à l'émission vocalique. Cependant, nous estimons nécessaire de compléter ultérieurement notre modèle du bloc « résonateur » en tenant compte de la cavité nasale.

d'augmenter le nombre de tubes et de tenir compte des phénomènes aérodynamiques tels que la viscosité de l'air entre autres pour que les modèles soient plus réalistes. Par exemple, Boersma (1993)[203] modélise l'appareil de production de la parole, du diaphragme aux lèvres, comme une série de tubes reliés entre eux. Pour Boersma, chaque tube possède une section à géométrie variable dont la surface est réglée par un ensemble masse-ressort amorti, semblable aux modèles de cordes vocales.

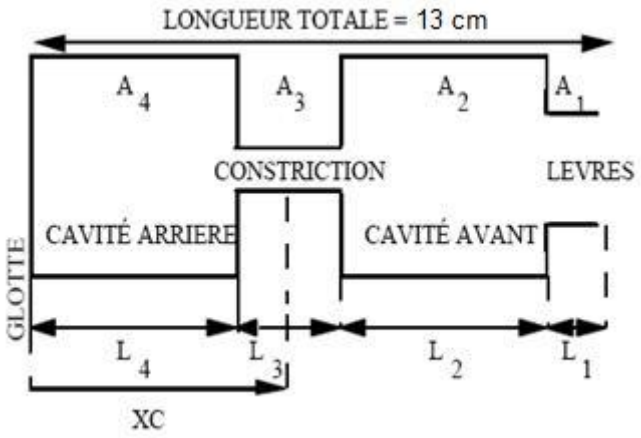


Figure 81 : Modèle de Fant (1960) à quatre tubes simulant la section des lèvres (A1), une cavité frontale (A2), l'articulation linguale (A3) et une cavité arrière (A4) (d'après Fant , 1960, modifié).

Le résultat acoustique est calculé à partir de la dérivée temporelle du flux d'air aux lèvres et aux narines en fonction du mouvement des parois mobiles. Bien que les modèles comme celui de Boersma et de Maeda et col. (1996)[204] semblent très réalistes, nous ne les avons pas retenus car ce sont des modèles nécessitant un temps de calcul important pour la simulation numérique. Comme nous l'avons précédemment signalé pour le bloc « vibrateur », les autres blocs (ou microsystèmes) constituant notre modèle systémique ont une précision limitée, et donc il ne sert à rien de prendre un modèle de résonance dans le conduit vocal complexe si la précision supplémentaire qu'il apporte n'est pas utile.

L'écriture des équations du bloc « Résonateur » est présentée dans l'annexe 14 *(1.3.2. « Ecriture d'équations des blocs du modèle »*). La résolution des équations du modèle nous a permis d'une part, de mettre en relation les caractéristiques géométriques des tubes (surface de la section) en fonction de la distance par rapport à la glotte. D'autre part, elles nous ont permis de définir les propriétés de filtrage du conduit transformant un signal source (ou son glottique) en signal de sortie. Nous avons calculé les fonctions de transfert des tubes

modélisés à partir des données d'entrée de la fonction d'aire. Ensuite, nous avons vérifié la validité des équations du modèle en testant différents paramétrages. On obtient finalement des courbes comme celles ci-dessous (Figure 82 et Figure 83). La taille des tubes va influencer la position centrale de pics des courbes de chaque tube (Figure 83). De même, on observe que les variations des coefficients de réflexion et de transmission retentissent sur la largeur des pics. Enfin, nous avons intégré le calcul du bloc résonateur à l'ensemble du modèle systémique asservi.

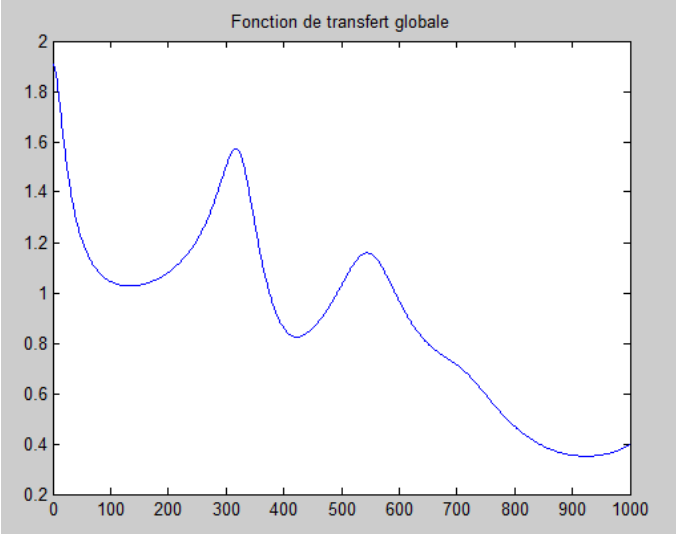


Figure 82 : Courbe représentant la fonction de transfert obtenu pour un résonateur modélisé. En X, les fréquences (Hz); en Y, l'amplitude.

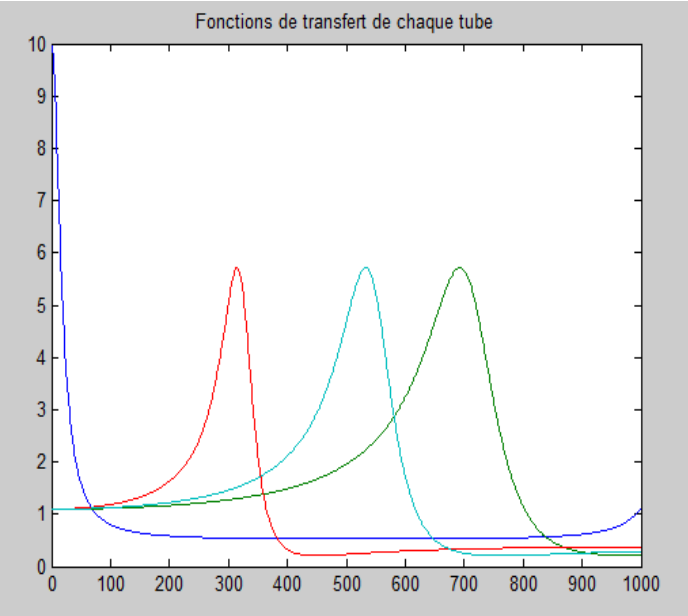


Figure 83 : Courbe représentant la fonction de transfert des différents tubes modélisés.

1.3.3. Bloc « Récepteur »

Le bloc « récepteur » ou « cochlée » a pour but de modéliser la perception du son par l'oreille humaine, c'est-à-dire, modéliser la transformation d'un signal acoustique en un signal électrique. La modélisation de la perception auditive a été depuis longtemps développée. On trouve dans la littérature de nombreux modèles conçus depuis divers points de vue : mécanique ou hydromécanique, électrique ou informatique, ou combinant ces points de vue.

Plusieurs modèles auditifs sont déjà implémentés sur diverses plateformes informatiques. Les plus utilisés sont HUTear[*], DSAM[**] et Auditory Toolbox[***]. HUTear est une boite à outils de Matlab. Elle simule chaque étage de l'audition : l'oreille externe, interne, la cochlée, les cellules ciliées externes, et les cellules du nerf auditif. Pour chaque étage, plusieurs modèles sont disponibles. Pour utiliser HUTear, il faut choisir un modèle par niveau d'analyse, un signal d'entrée, et le programme calcule l'état à chaque étage. Des pré et post-traitements sont également disponibles.

La bibliothèque DSAM est implémentée en langage C. Cette bibliothèque simule de nombreux modèles issus de la recherche sur la physiologie de l'oreille, en particulier le banc de filtre auditif « Gammatone », que nous allons présenter dans la suite de cette section, le modèle de cellules ciliées externes de Meddis (1986[205], 1990[206]), et l'outil d'imagerie auditive Auditory Image Model (AIM) de Patterson (1992[207], 1995[208], 2000[209]).

Enfin, la bibliothèque Auditory Toolbox est une autre boite à outil de Matlab développée par Malcolm Slaney. C'est également une compilation de plusieurs modèles auditifs issus de la recherche, mais il faut noter l'importante part réservée à la représentation. En plus du modèle en banc de filtres Gammatone, et du modèle de cellules cillées externes de Meddis, la bibliothèque contient le modèle auditif de Lyon (1988)[210]. Celui-ci repose sur une approche de la membrane basilaire en ligne de transmission, suivi de plusieurs étages d'adaptation. Egalement, on trouve dans cette bibliothèque le modèle de Seneff (1988)[211] qui décrit la cochlée en combinant un banc de filtre à bandes critiques et des modèles de détection. Tous ces modèles présentent en fait une architecture et une approche très similaires, cependant aucun n'est implémenté sur des plateformes qui fonctionnent en temps réel.

---

[*] http://www.acoustics.hut.fi/software/HUTear/
[**] http://www.pdn.cam.ac.uk/groups/dsam/
[***] http://cobweb.ecn.purdue.edu/~malcolm/interval/1998-010/

Modèle retenu pour le bloc de perception du signal sonore

Pour modéliser la perception auditive, nous avons retenu le modèle du *« banc de filtres Gammatone »* et nous nous sommes essentiellement basés sur les travaux réalisés par Hohmann (2002)[212] et Patterson (1987[213], 1993[214]).

La perception du signal sonore est achevée grâce à la conversion de la pression acoustique en impulsions neuronales et, ceci par l'action des cellules ciliées disposées le long de la membrane basilaire. Chacune des sous-structures spatialement ordonnées au sein de la cochlée (cellules ciliées, membrane tectoriale et membrane basilaire) se comportent comme un filtre auditif, ne laissant passer qu'une gamme limitée de fréquences. En effet, les caractéristiques du système sont très sélectives. L'idée de la modélisation est donc de considérer la membrane basilaire et les cellules ciliées comme un banc de filtres passe-bande, appelés *« filtres auditifs »*. Au fil du temps, les filtres auditifs ont été caractérisés en premier lieu par leur largeur de bande, sous l'hypothèse qu'ils pouvaient être modélisés par des filtres passe-bande idéaux, c'est-à-dire rectangulaires. Etant donné la complexité des phénomènes mis en jeu au niveau de la cochlée, cette hypothèse est évidemment très simplificatrice. En effet, des mesures plus précises des filtres auditifs ont montré leur caractère non linéaire et en particulier leur dépendance vis-à-vis du niveau des stimuli. En réalité, les filtres auditifs ont une bande passante plus étroite pour les niveaux modérés que pour les niveaux forts (Glasberg, 2000)[215].

Dans ce cadre très complexe, une modélisation par des filtres passe-bande idéaux constitue une première approche pratique et permet de quantifier approximativement la résolution fréquentielle du système auditif. Alimentés par un bruit blanc, le filtre auditif réel et le filtre approximé fournissent la même énergie en sortie.

Les filtres auditifs sont définis pour toute fréquence audible et la largeur des filtres augmente avec la fréquence, de telle sorte que cette largeur oscille entre 11% et 17% de la valeur de la fréquence centrale. Ces caractéristiques influent sur la résolution fréquentielle du système auditif. En effet, la discrimination fréquentielle de l'oreille est meilleure pour les basses fréquences que pour les hautes. Les propriétés générales des filtres auditifs peuvent se résumer ainsi :

- les filtres auditifs ont une allure exponentielle,
- le sommet de chaque filtre est un peu arrondi et à peu près symétrique en fréquence,

- les bords sont larges et asymétriques ; en particulier, leur pente du côté des basses fréquences est moins forte,
- la bande passante augmente avec la fréquence : plus le niveau du stimulus est élevé, plus les filtres sont larges.

Les filtres Gammatone sont souvent utilisés pour la modélisation du traitement du son dans la cochlée (Patterson et col., 1987[213]). Le principe des filtres « Gammatone » considère que ce sont des filtres passe-bande linéaires idéaux, dont la largeur évolue avec la fréquence. Il faut savoir que les filtres auditifs réels ne sont pas linéaires. Néanmoins, une modélisation par des filtres passe-bande idéaux constitue une première approche pratique et permet de quantifier approximativement la résolution fréquentielle du système auditif. Nous avons modélisé la cochlée, par un banc de filtres Gammatone *« tout pôle » d'ordre 4* tel que proposé par Pressnitzer et Gnansia (2005)[216]. Le nombre de filtres a été fixé au nombre de 24 afin d'effectuer une modélisation analogue à celle utilisée par l'implant cochléaire. Le principe de ce dispositif est d'exciter directement le nerf auditif à l'aide d'électrodes placées dans la cochlée par une opération chirurgicale. Une partie de l'implant est situé en dehors du conduit auditif (le plus souvent derrière l'oreille), c'est là que se trouve le microphone et le processeur réalisant les calculs nécessaires à la production du signal excitateur envoyé aux électrodes. Les électrodes placées dans la cochlée réalisent une excitation ponctuelle le long de la membrane basilaire, il y a donc un nombre fini de points d'excitation pour un implant donné. Nous avons mentionné précédemment que la cochlée agit comme un banc de filtre passe-bande : chaque électrode excite donc un de ces filtres. Avec les technologies actuelles, les implants cochléaires vont de 4 à 24 bandes ou intervalles de fréquences.

L'écriture et les calculs des équations du bloc « Récepteur » ont été effectués en plusieurs étapes lesquelles sont décrites comme suit.

1) D'abord, nous avons calculé la transformé en *Z* pour les filtres Gammatone. La transformée en *Z*[*] est un outil mathématique de traitement du signal qui est l'équivalent discret de la transformée de *Laplace*. On considère un signal temporel $x(t)$, à partir duquel on peut définir le signal échantillonné $x.(t)$ en ne prenant du signal $x(t)$ que des valeurs à des instants définis.

[*] On l'utilise notamment pour la mise en place de filtres discrets. C'est notre cas ici puisque le modèle cherche à incrémenter sur le temps. Sa définition mathématique est la suivante : la transformée en z est une application qui transforme une suite *S* (définie sur les entiers) en une fonction *S* d'une variable complexe nommée *Z*, telle que : $S(z) = Z(s^*(n)) = \sum s(n) z^{-n}$. La variable *n* représente en général le temps échantillonné. La transformée en *Z* présente l'avantage d'avoir certaines propriétés comparables à la transformée de Laplace, laquelle s'utilise dans le cas continu.

2) Dans un deuxième temps, nous avons modélisé le filtre Gammatone par la fonction de transfert.

3) Ensuite, nous avons effectué la synthèse d'un filtre Gammatone à partir d'une fréquence centrale et d'une largeur de bande. Nous avons utilisé la méthode de mesure de la largeur des filtres appelée *largeur de bande rectangulaire équivalente* ou *Equivalent Rectangular Bandwidth (ERB)* proposé par Moore et Glasberg (1990)[217].

4) Puis, nous avons construit le banc de filtres Gammatone complet. Notre modèle a implémenté 24 filtres branchés en parallèle. Pour la réalisation du banc de filtres complet, nous avons donc dû définir les fréquences centrales et les largeurs de bande des filtres Gammatone (ERB) selon Moore et Glasberg (1990)[217]. Cette largeur conditionne le degré de précision de l'analyse fréquentielle du signal : plus le filtre est étroit et plus l'analyse est fine. Nous avons d'abord choisi une fréquence minimale, une fréquence maximale et un *« pas de largeur de filtres »* : f_{min} = 50 Hz, f_{max} = 5000 Hz et un *pas* de 1,5. Le *pas* ERB_{scale} a été choisi de façon à répartir au mieux les *N* filtres dans l'intervalle de fréquences sélectionnée : différents tests ont conduit à prendre ERB_{scale} = 1.5. *N*, le nombre de filtres à été fixé à 24. Enfin, la fréquence d'échantillonnage a été fixée à f_s = 30 000 Hz pour que le théorème de Shannon soit respecté et pour éviter des temps de calculs trop importants. On rappelle que, d'après le théorème de Shannon, pour pouvoir restaurer un signal continu dont est issu un signal échantillonné, la fréquence d'échantillonnage doit être au moins deux fois supérieure à la fréquence maximale contenue dans le signal échantillonné. Nous avons alors pu caractériser chaque filtre. Etant donné que nous avons choisi d'effectuer l'implémentation numérique de notre modèle en langage C++, nous avons traité directement des signaux temporels.

5) Enfin, nous avons effectué la reconstitution et retard du signal. Une fois le banc de filtres créé, le signal temporel a été recomposé après passage dans le bloc « Récepteur ». En reconstruisant le signal temporel, on peut obtenir alors le son tel qu'il est perçu par « l'oreille ». Puisque les calculs sont faits en temps réel, on peut utiliser le banc de filtres dans une boucle temporelle et le faire fonctionner dans un simulateur comme nous le souhaitions. Nous décrirons ultérieurement le simulateur numérique *ARTIVox* qui a été ainsi conçu sur la base de notre modèle macro-systémique de la boucle audio-phonatoire.

Pondération

Nous avons effectué trois types de pondération : 1) une « pondération avant », 2) une « pondération après », et 3) une « pondération extérieure ». La « pondération avant » permettait d'affecter un gain précédant pour sommer les participations de chaque filtre et ainsi réduire au maximum la distorsion du signal. Initialement, les maximas des signaux sont alignés. La « pondération après » permettait de compenser les défauts dus aux interférences constructives afin d'obtenir une réponse fréquentielle la plus plate possible en sortie. Les coefficients de pondération ont tous été donnés égaux à 100. Enfin, la « pondération extérieure » facilitait le paramétrage et la modification du bloc « récepteur » de façon à adapter la modélisation aux défauts de perception désirés. Les coefficients de ce vecteur sont des valeurs comprises entre *0* et *100*, *0* si le filtre correspond à un intervalle de fréquences non perçues, *100* si au contraire il s'agit d'un des intervalles de fréquences les mieux perçues. Ces valeurs peuvent être alors saisies directement.

Nous avons effectivement implémenté le banc de filtres complet. La structure du modèle pour le bloc « Récepteur » est représentée dans la figure ci-dessous (Figure 84) :

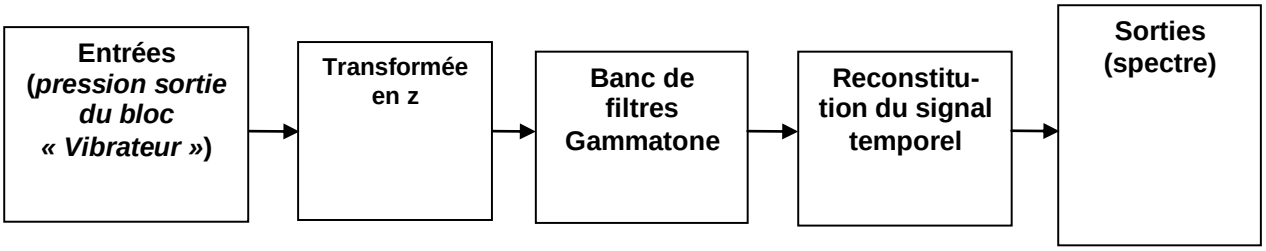


Figure 84 : Schéma de la structure du modèle du bloc « Récepteur » développé dans nos travaux.

1.3.4. Bloc « Comparateur »

Le bloc « Comparateur » ou « cerveau »* est destiné à recevoir en entrée un signal sonore de *x* secondes issu du bloc « Récepteur ». Le bloc « Comparateur » doit alors comparer ce signal sonore à un signal « Consigne », auquel le signal, perçu par le bloc « Récepteur » et reproduit par les blocs « Vibrateur » et « Résonateur », doit idéalement se rapprocher. En effet, on cherche à déterminer si le macrosystème (ou boucle audiophonatoire) fait converger ou non le signal produit par la boucle elle-même vers le signal « Consigne ». Pour cela, nous nous

* Dénommé ainsi dans les trois versions précédentes de notre modèle.

baserons sur un algorithme *de comparaison de transformées de Fourier (cf. Annexe 10, section 1.3.3)*. L'avantage de cet algorithme est qu'il peut être implémenté en langage C++ et est plus rapide que l'algorithme de calcul basé sur l'algorithme de comparaison dynamique couramment utilisé en reconnaissance vocale que nous avions essayé d'implémenter. Malgré sa très grande efficacité, nous avons observé que cet algorithme présente un coût élevé en termes de temps de calcul. Il possède en effet une complexité très grande qui ralentit son exécution de calcul et risque de saturer la mémoire vive de l'ordinateur qui effectue le calcul. Ceci à cause de la taille de construction des matrices.

L'algorithme *de comparaison de transformées de Fourier* propose d'établir une différence des transformées de Fourier du signal émis et du signal consigne perçu[*]. Cet algorithme permet de calculer un coefficient scalaire représentant l'écart existant entre les aires des courbes des spectres fréquentiels, si on les superpose et de ce fait, une erreur entre les deux signaux. Plus l'aire des spectres comparés est grande, plus l'erreur d'émission vocale par rapport au signal perçu est importante. Inversement, afin de rapprocher les deux signaux, il suffit de manipuler les différentes données des divers blocs pour obtenir des valeurs proches de zéro, et au mieux égales à zéro. Ce coefficient représenterait en fait les modifications à effectuer sur le signal émis issu des blocs « Vibrateur » et « Résonateur » pour le rapprocher à un signal sonore Consigne perçu. Si le coefficient est nul, il y a une correspondance parfaite entre les deux signaux (émis et perçu) ; si le coefficient est élevé, il y a une discordance entre les deux signaux. Nous appellerons ce coefficient *« CCTF »*. Nous avons implémenté graphiquement l'affichage d'une courbe représentant la différence des deux signaux (cf. Figure 87), ce qui permet notamment de suivre son évolution avec les modifications de paramétrage pour les différents blocs constituant le modèle macro-systémique. L'affichage de résultats sera expliqué ultérieurement.

1.3.5. Banc d'essai dans son ensemble dit : « Modèle macro-systémique de la boucle audio-phonatoire »

A partir des réflexions de Lombard, Levelt, Lackner et Tuller présentés dans la section consacrée à la problématique de notre étude (chapitre II, section 2), nous avons élaboré un modèle dénommé « Modèle macro-systémique de la boucle audio-phonatoire » qui a été

[*] Rappelons ici que la cochlée est un organe qui permet l'analyse et la décomposition fréquentielle d'un signal sonore en série de Fourier. Par exemple, à l'aide la méthode de décomposition de Fourier, on sait que la voyelle /i/ est la somme d'une trentaine d'ondes sinusoïdales de fréquences comprises entre 125 et 3500 Hz. Etudier les sons vocaliques par exemple revient donc simplement à étudier une onde sinusoïdale.

utilisé pour construire un banc d'essai de simulation numérique. Les évidences fournies dans des travaux précédents montrent la solidarité des mécanismes impliqués dans la production et la perception vocale. Notre modèle met en évidence un fonctionnement physique très subtil et coordonné des micro-systèmes qui sous-tendent des performances très dépendantes entre elles et permettant la mise en relation permanente entre ce qui est produit et ce qui est perçu.

Après avoir présenté les différents blocs ou microsystèmes correspondant aux instances physiologiques impliquées dans le contrôle audio-phonatoire, nous exposerons maintenant la façon dont ces éléments s'articulent et interagissent dans un macro-système dynamique en forme de boucle. La première étape de la conception de notre modèle macro-systémique a consisté à modéliser la structure interne du modèle. Pour y parvenir, nous avons estimé pertinent de modéliser les quatre principaux mécanismes physiologiques impliqués directement dans la régulation de la voix, à savoir : un mécanisme émetteur d'une source vocale représenté par le bloc « Vibrateur » ; un mécanisme modificateur de la source vocale représenté par le bloc « Résonateur »[*] ; un mécanisme récepteur-contrôle de la source vocale représenté par le bloc « Récepteur » et, un mécanisme de comparaison des signaux émis et perçu par rapport au signal consigne représenté par le bloc « Comparateur ». La deuxième étape de notre démarche dans la conception de notre modèle a consisté à établir des relations spécifiques entre les différents micro-systèmes modélisés par rapport à la finalité du macro-système. Cette dernière est de rapprocher l'état-objectif (signal sonore consigne) de cet état final (signal perçu et émis) pour en déduire d'éventuelles actions correctrices.

Pour construire l'architecture du modèle théorique et l'algorithme de programmation, nous avons établi quelques hypothèses de départ :

i) notre modèle considère que le processus de régulation de la voix et de la parole se fait par le mécanisme auditif via une boucle de rétroaction,
ii) nous ne tenons pas compte des caractéristiques et du rôle du cerveau et de l'environnement sonore extérieur dans notre modèle,
iii) nous considérons l'existence d'un déficit de perception et de perte de l'information acoustique à partir du moment où le signal sonore entre dans la boucle d'asservissement.

[*] Ces deux blocs représentent physiquement le rôle de certaines structures anatomiques pour accomplir le processus physiologique de la phonation. Nous prenons en compte dans notre modélisation l'action de la soufflerie pulmonaire, le vibrateur laryngé et les résonateurs. En revanche, nous n'avons pas modélisé l'action du mécanisme articulatoire. Ceci car nous nous sommes concentrés sur l'analyse de la régulation de la voix comme support physique de la parole ainsi que sur les paramètres acoustiques et aérodynamiques la constituant.

Sur la base de ces hypothèses, nous avons conçu un modèle de la boucle audio-phonatoire comme un système d'asservissement dynamique. Par l'implémentation d'un coefficient comparateur, nous avons pu quantifier le flux d'information entrant et sortant dans cette boucle. Le banc d'essai est représenté sur le schéma bloc ci-dessous (Figure 85).

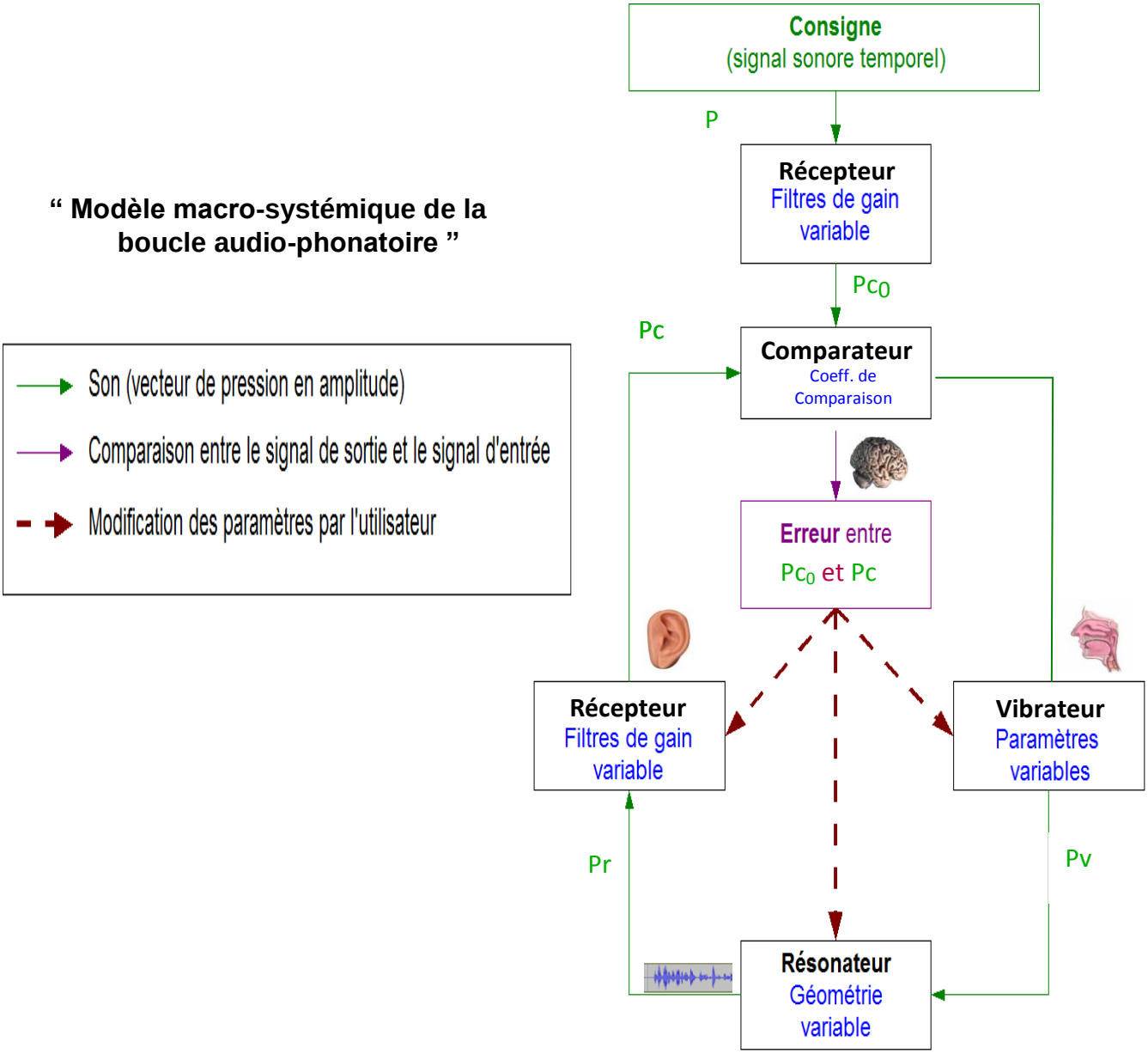


Figure 85 : Schéma détaillé d'architecture du modèle systémique asservi représentant l'action de la boucle audio-phonatoire lors de l'émission d'un signal vocalique (voyelle /a/).

En entrée de ce banc d'essai, on devra saisir un certain nombre de données physiologiques caractérisant un sujet. Pour plus de détails concernant le module de calcul du banc d'essai se reporter à l'annexe 14 (section 1.3.4.).

Tout d'abord, il s'agit de paramétrer le *bloc Cochlée*. Il faut notamment choisir quels sont les coefficients de pondération des filtres (24 au total) modélisant le degré de perception d'un signal acoustique (section *1.3.3*). Chaque filtre modélisé correspond à une certaine gamme de fréquences perçue. Nous avons établi le coefficient de pondération d'un filtre à *100* pour représenter une perception auditive totale et *0* pour une perception auditive nulle. Ici, on tient compte du « degré de perception » d'un signal sonore et on représente par là les divers cas de perception auditive existant chez un sujet. De cette façon, nous avons quantitativement représenté la qualité de perception du *bloc Récepteur*.

Ensuite, il s'agit de saisir les données concernant les blocs Vibrateur et Résonateur, à savoir : les paramètres du modèle mécanique de vibration (masse, coefficients de raideur des ressorts et coefficients d'amortissement) ; les caractéristiques géométriques des cordes vocales et des tubes de résonance (longueur et profondeurs des cavités) ; les paramètres aérodynamiques (densité de l'air, pression sous-glottique et vitesse d'écoulement) ; les caractéristiques des cavités de résonance (longueur et profondeurs des cavités), ainsi que la configuration initiale des cordes vocales pour l'émission d'un son.

Une fois que toutes les données physiologiques ont été saisies dans le modèle du banc d'essai, la simulation peut être lancée. Un *« signal sonore Consigne P »* provenant d'une source sonore extérieure est traité par les filtres gammatone préalablement paramétrés (selon les coefficients de pondération saisis) du bloc Récepteur qui fournit un signal *« Pc_0 »* en sortie. Ce dernier est reconnu par le bloc Comparateur comme étant un signal sonore Consigne.

Le bloc Comparateur transmet le signal *« Pc_0 »* au bloc Vibrateur afin de le reproduire à partir de deux critères déterminants du signal perçu, à savoir : la fréquence fondamentale et l'amplitude. Ensuite, le signal *« Pv »* généré par le bloc Vibrateur est modifié selon les caractéristiques de la forme des tubes de résonance du bloc Résonateur. Le signal acoustique modifié sortant du bloc Résonateur devient *« Pr »*. Ce dernier est traité à nouveau par le bloc Récepteur qui transmet le signal *« Pr »* au bloc Comparateur comme signal *« Pc »* afin de comparer le signal « *émis-perçu Pc* » au signal *« consigne perçu Pc_0 »*. A ce moment-là, le

bloc Comparateur sera en mesure d'effectuer une comparaison des signaux « Pc_0 » par l'algorithme de comparaison des transformées de Fourier que nous avons implémenté.

Il est possible de lancer le banc d'essai ou de recommencer le bouclage autant de fois que nécessaire, à partir des nouvelles informations saisies en modifiant les paramètres d'initialisation de chaque bloc de façon à obtenir une erreur proche de zéro ou une diminution de l'erreur.

En résumé :

- Le modèle « *macro-systémique de la boucle audio-phonatoire* » est de type descriptif et manifeste une dynamique en boucle de « *rétroaction négative* ».
- Ce modèle est constitué de quatre microsystèmes ou blocs : Vibrateur, Résonateur, Récepteur et Comparateur.
- Le banc d'essai que nous avons conçu permet d'intégrer les blocs dans une macrostructure afin de reproduire « vocalement » un signal sonore « perçu ».

2. Conception d'un simulateur numérique : ARTIVox 4.0

Le logiciel dénommé « ARTIVox 4.0 »[*] est le simulateur numérique de la fonction « boucle audio-phonatoire » que nous avons conçu en collaboration avec l'Ecole Centrale de Lyon et qui implémente notre modèle macro-systémique décrit ci-avant.

Notre objectif était ici de concevoir puis de tester un logiciel de simulation de l'action de la boucle audio-phonatoire à partir de notre modèle « macro-systémique » basé sur les quatre blocs moteurs fondamentaux décrits ci-avant.

Dans cette démarche, par la conception d'un simulateur numérique, nous avons tenté de relier les entrées et les sorties d'informations de notre modèle par des équations plus ou moins nombreuses et complexes décrivant les comportements du phénomène réel étudié, à savoir : l'action de la boucle audio-phonatoire.

[*] Le nom de code « *ARTIVox* » de notre simulateur est construit sur l'adjectif « *artificielle* » et sur le nom latin « *Vox* » signifiant « *voix* ».

2.1. Matériel et méthode

Ce travail a été développé en lien avec quatre *« Projets Industriels »* de l'Ecole Centrale de Lyon (ECL). Le projet industriel est un dispositif pédagogique qui permet de mettre les élèves-ingénieurs dans une situation opérationnelle de résolution d'un problème en réponse à la demande de prestation d'un *« commanditaire »* qui provient de l'extérieur, et qui propose un sujet industriel. L'équipe participant à notre projet était donc constitué de :

- un « Commanditaire industriel », représenté par *moi-même*, définissant les objectifs, les livrables du projet, et garantissant la mise en œuvre des moyens et le suivi technique du travail prévus pour le bon déroulement du projet (Guerrero et col., 2009)[218],
- une Equipe-projet constituée de 5 ou 6 élèves-ingénieurs responsables de la programmation informatique et de l'exécution du cahier des charges définis par le commanditaire du projet,
- deux Suiveurs ECL (professeurs, ingénieurs, experts) qui assuraient la mise en œuvre des moyens prévus pour le développement du projet par l'Ecole Centrale de Lyon et le suivi technique du travail réalisé.

Les livrables de ces projets industriels - dont l'objectif était la conception du logiciel ARTIVox de simulation de la boucle audio-phonatoire - étaient : un CD-Rom regroupant l'ensemble des fichiers informatiques développés (scripts de programmation, application ARTIVox) et un manuel d'utilisation.

2.2. Historique des versions du logiciel ARTIVox

En raison de la durée courte accordée pour l'exécution d'un projet industriel sur une année (5 mois), nous avons reconduit le projet pendant 4 ans. De ce fait, quatre versions du logiciel ARTIVox ont été produites (Tableau 41).

Tout d'abord, nous avons construit les différents blocs constituant notre modèle. Par la suite, nous avons défini l'assemblage. Dans un premier temps, la programmation a été réalisée sous Matlab[*]. Dans un second temps, nous avons exécuté une réécriture du logiciel ARTIVox en langage C++. En effet, le temps de calcul d'un bouclage et d'une exécution du programme

[*] Langage de programmation utilisé dans la recherche, l'éducation et l'industrie pour le calcul numérique mais aussi dans les phases de développement de projets.

ARTIVox v3.0 (implémentée sous Matlab) était extrêmement long (10 minutes environ), un tel délai empêchant une quelconque rétroaction.

Date (Janvier – Mai)	Version	Description
2006	v1.0	Bloc Vibrateur (Matlab)
2007	v2.0	Amélioration Bloc Vibrateur (Matlab)
2008	v3.0	Bloc Récepteur + Résonateur + Comparateur + Banc d'essai « Bouclage » (Matlab)
2009	v4.0	Amélioration Banc d'essai « Bouclage » (C++) : IHM, son, graphes, coefficient

Tableau 41 : Tableau de l'évolution des versions produites du logiciel ARTIVox.

D'autre part, Matlab est un langage de programmation qui demande une licence d'exploitation payante. Ceci implique alors un investissement financier important pour chaque ordinateur équipé. Puisque C++ est un langage de programmation répandu dans le monde qu'il peut être exécuté sous différents systèmes d'exploitation tel que Microsoft Windows, qu'il permet la réduction du temps de calcul du bouclage et qu'il ne demande pas le paiement de droits d'utilisation, nous avons opté pour ce dernier.

2.3. Interface ARTIVox v4.0

Le logiciel ARTIVox v4.0 est la dernière version de notre simulateur et est compatible avec tout ordinateur équipé de Windows98 ou plus, moyennant l'installation de petits outils qui seront ensuite détaillés. Ceci en fait donc un logiciel autonome.

Ce logiciel a été conçu comme un outil d'assistance à l'étude de la régulation de la voix et de la parole par le mécanisme auditif. Nous l'avons voulu basique et facile à utiliser. Les paramètres modifiables pour chaque bloc et leur signification par rapport au modèle théorique sont détaillés par le bouton Aide. Les blocs à paramétrer sont identifiés chacun avec une couleur différente dans l'interface (Figure 86).

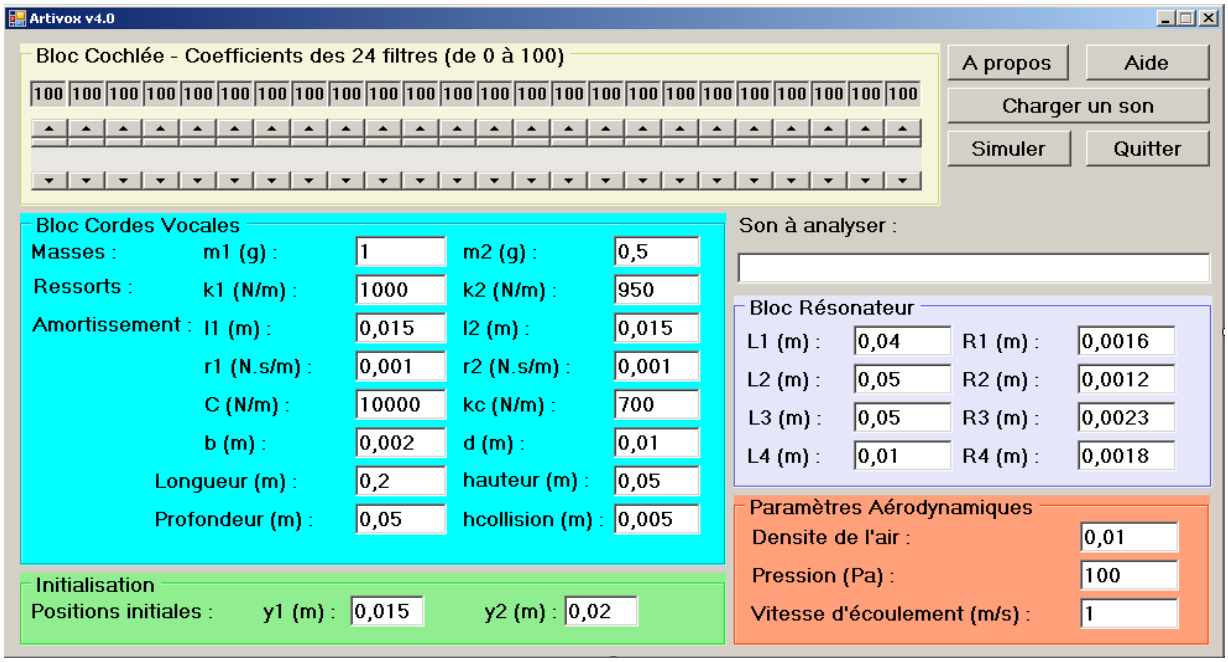


Figure 86 : Interface du logiciel de simulation de la boucle audio-phonatoire ARTIVox v4.0

En effet, ARTIVox v4.0 propose une partie didactique destinée aux professionnels de la rééducation de la voix et de la parole en clinique ayant un niveau de compréhension physique et physiologique élémentaire du système modélisé. Le simulateur propose donc d'effectuer des expérimentations virtuelles grâce à la possibilité de saisir les différentes valeurs facilement modulables pour chaque bloc du système. On peut complexifier et par la suite prendre en compte plus de phénomènes caractéristiques du bouclage désiré. Ainsi, on peut constituer toute une banque de données selon les différents paramétrages grâce aux résultats qui peuvent être sauvegardés.

De plus, on peut corréler les données expérimentales cliniques avec les données théoriques lors des différentes simulations. Notre postulat était que le logiciel ARTIVox v4.0 puisse mettre en évidence que la qualité et la quantité de la production vocale sont soumises au contrôle du mécanisme auditif, ce contrôle s'inscrivant dans la dynamique d'un bouclage.

2.3.1. Interface Homme-Machine ARTIVox v4.0

Lorsqu'on lance le programme ARTIVox v4.0, une fenêtre d'interface s'ouvre conformément à la Figure 86. Pour exécuter le programme, on charge d'abord un *« son consigne »* en cliquant sur le bouton **Charger un son**. Une fenêtre apparaît pour sélectionner dans un répertoire le son souhaité comme consigne. Le son Consigne doit être un fichier son au format **.wav* contenant un échantillon de l'émission d'une voyelle ayant une durée supérieure à cinq secondes. Les détails sur l'import et l'export de sons sont présentés ultérieurement.

On paramètre ensuite les différents blocs. Pour que le programme soit cohérent avec le modèle théorique sous-jacent, nous l'avons aussi conçu en bloc de différentes couleurs : un premier bloc correspond au bloc « Cochlée » ou « Récepteur » (en gris), un deuxième au bloc « Cordes Vocales » ou « Vibrateur » (en cyan), et un troisième bloc correspondant au bloc « Résonateur » (en bleu). Par rapport au schéma-bloc de la régulation ou bouclage, nous avons rajouté un bloc dénommé « Initialisation » (en vert) et un bloc « Paramètres aérodynamiques » (en rouge). Il est possible de modifier tous les paramètres de ces blocs, de façon tout à fait indépendante les uns des autres. On trouvera dans la rubrique **Aide** l'explication succincte de chacun de ces paramètres qui ont déjà été présentés dans la *section 1.3*.

Pour démarrer le calcul, il suffit de cliquer sur **Simuler**. Le calcul d'une simulation est de l'ordre de la dizaine de secondes. Lorsque le calcul de simulation est terminé, une fenêtre de résultats s'ouvre.

2.3.2. Affichage de résultats de la simulation numérique ARTIVox v4.0

L'affichage de résultats présente simultanément cinq graphiques des transformées de Fourier du signal sonore analysé (Figure 87). En l'axe d'abscisses, les fréquences et en l'axe de coordonnées, l'amplitude. Le premier représente l'enveloppe du spectre fréquentiel du « son émis » à la sortie du bloc Résonateur (vecteur Pr,Figure 85) ; le deuxième représente le « son consigne » (vecteur P) ; le troisième représente le « son émis-perçu » (vecteur Pc) ; le quatrième visualise le « son consigne perçu » (vecteur Pc_0) ; enfin, le cinquième graphique représente les écarts fréquentiels entre le « son consigne perçu » (vecteur Pc_0) et le « son émis-perçu » (vecteur Pc). De plus, l'affichage des résultats présente :

- le coefficient CCTF (Coefficient de Comparaison des Transformées de Fourier) pour les courbes des graphiques 3 et 4 (Figure 87).

- un bouton **Jouer le son émis** (Figure 87, en bas à gauche), sur lequel on peut cliquer et déclencher l'écoute du son émis d'après le paramétrage de la simulation de la boucle.

Figure 87 : Fenêtre d'affichage des résultats de la simulation par le logiciel ARTIVox 4.0.

Chacun de ces graphiques présente un pic maximal d'amplitude *« 1 »* quelle que soit la simulation. En effet, la représentation graphique est en pourcentage par rapport à l'amplitude du pic maximal : cette présentation règle ainsi les problèmes liés aux variations d'échelle.

2.3.3. Export et import des sons

D'une part, nous avions besoin de comparer un son *« émis »* et un son *« consigne »* et, d'autre part, nous étions intéressés à apprécier le résultat de comparaison, ou le *« son émis »*. Pour ce faire, nous avons réalisé une procédure de chargement d'un son et une procédure d'émission d'un son au format **.wav* (format standard), (Annexe 14). Cette procédure a été automatisée pour n'importe quel fichier **.wav* d'entrée.

3. Simulation numérique : ARTIVox 4.0

Dans cette section, nous présentons les simulations effectuées à l'aide du simulateur ARTIVox v4.0. D'abord, nous fournissons les données d'entrée que nous avons utilisées pour les nombreuses simulations. Ensuite, nous procédons à la simulation de l'action de la boucle audio-phonatoire pour le cas d'un enfant normo-entendant puis d'un enfant déficient auditif.

3.1. Données d'entrée pour la simulation numérique

L'interface de l'application ARTIVox v4.0 permet à l'utilisateur de saisir différents paramètres de simulation. Les données d'entrée pour chaque module de l'interface sont présentées comme suit.

a) Bloc « cochlée » ou « récepteur »

On sélectionne d'abord les coefficients de pondération des filtres gammatone du bloc « Récepteur » ou « Cochlée » dans les 16 premières cases prévues à cet effet (Figure 88). L'impact des filtres est ajustable par des curseurs qui glissent des valeurs comprises entre 0 et 100, ce qui permet de saisir différents coefficients et de simuler divers « degrés de perception » d'un signal acoustique pour chaque filtre.

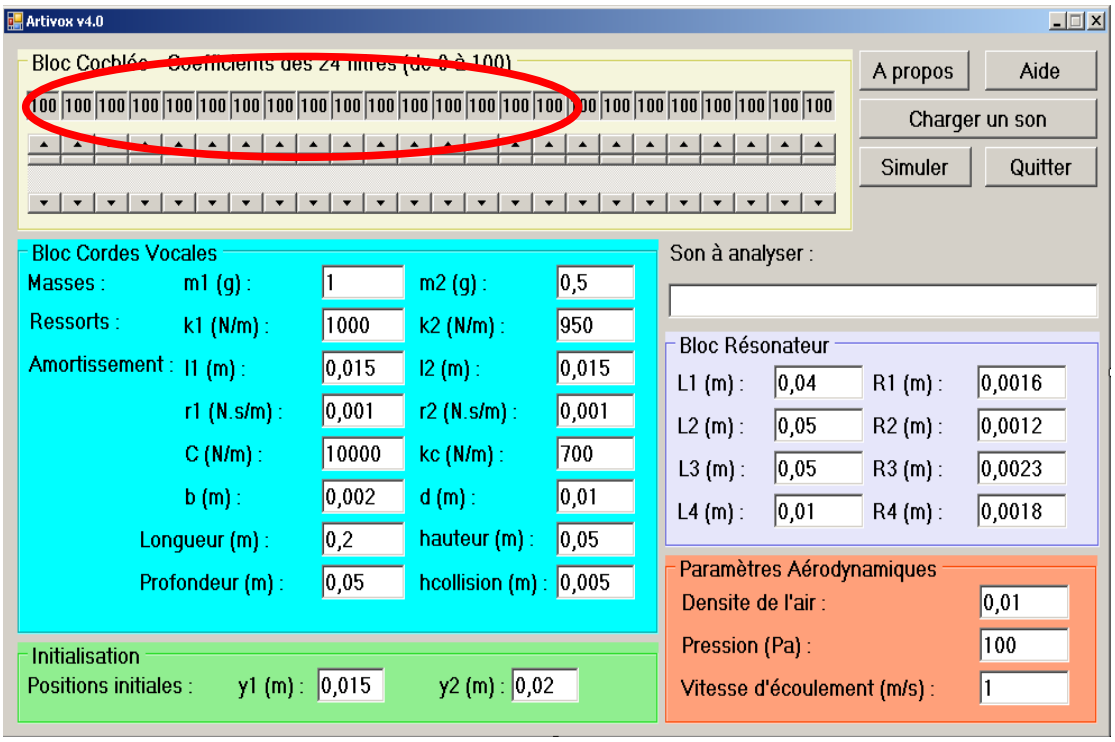


Figure 88 : Interface du logiciel ARTIVox v4.0. : paramétrage du bloc Cochlée.

Actuellement, la version ARTIVox v4.0 ne permet pas le traitement des fréquences de 5 000 jusqu'à 20 000 Hz pour des raisons de programmation (cf. Annexe 14, *1.3.3*). Par conséquent, les filtres 17 à 24 ne sont pas actifs. Le tableau suivant représente les fréquences centrales de chacun des filtres que nous avons modélisés (Tableau 42).

Numéro du filtre	Fréquence centrale (Hz)	Numéro du filtre	Fréquence centrale (Hz)
1	298	**13**	3239
2	388	**14**	3825
3	494	**15**	4510
4	617	**16**	5310
5	761	**17**	6246
6	930	**18**	7340
7	1127	**19**	8618
8	1357	**20**	10111
9	1626	**21**	11856
10	1940	**22**	13896
11	2307	**23**	16280
12	2737	**24**	19065

Tableau 42 : Tableau contenant les fréquences centrales (Hz) de chaque filtre modélisé dans le bloc « cochlée ». Les filtres 16 à 24 ne sont pas encore actifs dans la version ARTIVox v4.0.

b) Bloc « Cordes Vocales » ou « Vibrateur »

Puisque nous n'avions pas à disposition des données expérimentales physiologiques des cordes vocales, les valeurs des paramètres pour ce bloc ont été inspirées des travaux présentées par Pelorson et col. (1994)[195], Garrel, Nicollas et col. (2006)[219], Horáček et Švec (2002[222], 2005[220]). Ces données concernent les aspects dimensionnels des cordes vocales relatifs aux caractéristiques physiques du système masse-ressort. Nous voulions effectuer des simulations numériques en utilisant des valeurs dimensionnelles et des propriétés semblables à celles des cordes vocales humaines. Les données que nous avons utilisées sont résumées dans le

Tableau *43* (les valeurs figurant sur l'interface de cette figure ne correspondent pas aux valeurs que nous avons utilisées pour les simulations). Pour mémoire, dans ce bloc, nous avons simulé la vibration des cordes vocales à partir d'un modèle symétrique à deux masses : chaque corde vocale était modélisée par deux masses et un système de ressorts-amortisseurs (Figure 89).

Paramètre	Mesure
Longueur de la corde vocale	0,014 < *profondeur*[*] < 0,022 m
Longueur vibrante	0,003 < *d* < 0,006 m
Raideur (k1 ou k2)	5 < *k* < 100 N/m
Amortissement (r1 ou r2)	0,010 < *r* < 0,1 N.s/m
Masse	0,01 < *m* < 0,4 g

Tableau 43 : Données physiologiques utilisées d'autres auteurs

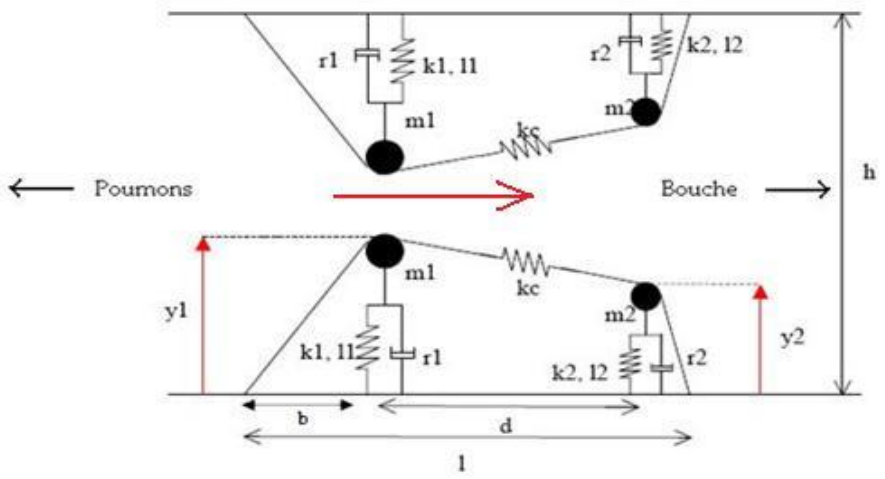


Figure 89 : Modèle à deux masses-ressort symétrique du bloc « cordes vocales ».

Avec :

- **m1** et **m2** : masses de la corde vocale.

- **k1** et **k2** : raideurs des ressorts
- **r1** et **r2** : amortisseurs

} Tension de la corde vocale.

- **l1** et **l2** : distance de l'élongation des ressorts.

[*] Dans notre travail, nous avons interprété la mesure de longueur de la corde comme une mesure de profondeur depuis une coupe transversale où les cordes vocales sont disposées horizontalement selon les modèles mécaniques présentés par les ingénieurs.

- **kc** : raideur du ressort entre les masses **m1** et **m2,** ou contraction du muscle lors du changement de configuration glottique (convergente, divergente, collision).

- **d** : distance entre les deux ressorts ou partie vibrante.

- **b** : distance entre le point d'attache et le ressort **r1.**

- **l** : épaisseur de la corde vocale.

- **h** : hauteur (ou diamètre) du conduit vocal*, désignée par « hauteur » dans l'interface.

- **Profondeur** : longueur de corde vocale appréciée depuis une vue supérieure.

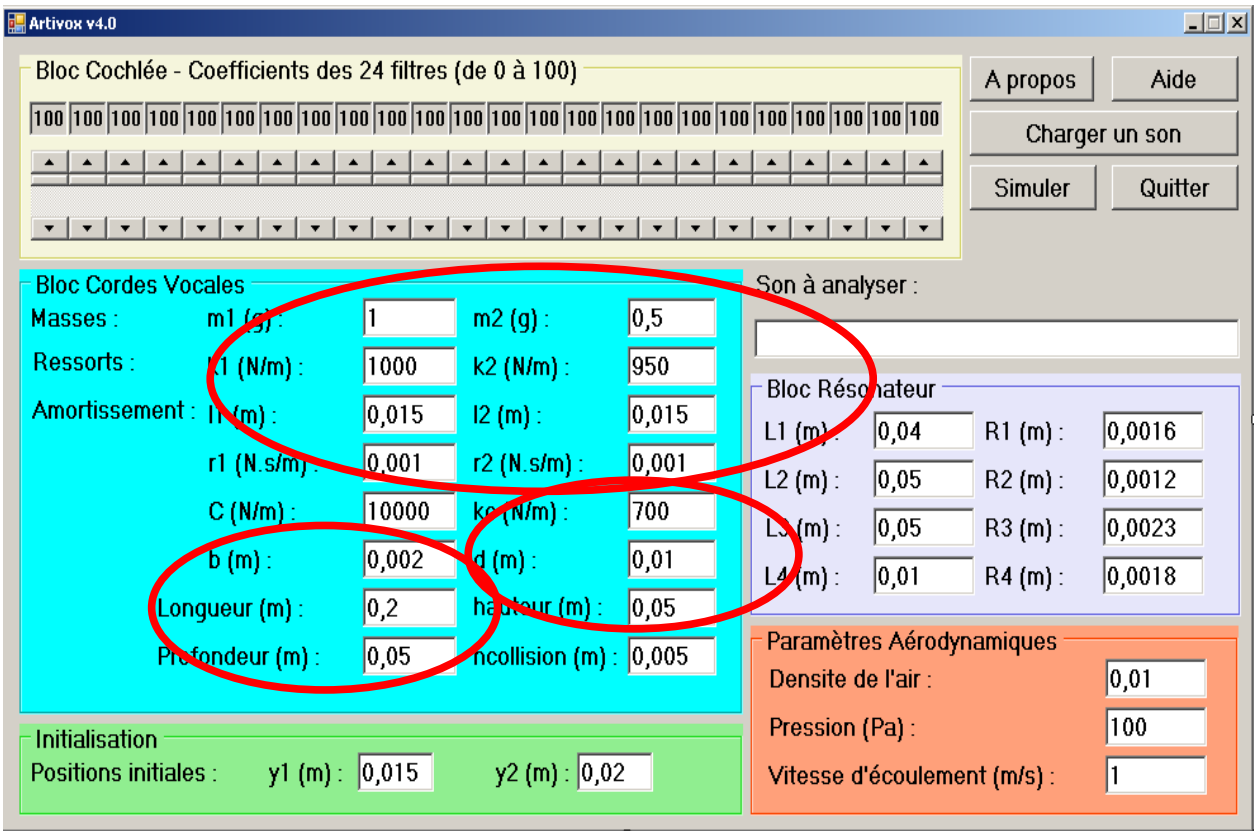


Figure 90 : Interface du logiciel ARTIVox v4.0 : paramétrage du bloc Cordes vocales.

c) Bloc « Résonateur »

Dans ce bloc, nous pouvons saisir les paramètres dimensionnels *L* (longueur) et *R* (rayon) des tubes représentant le conduit vocal (Figure 91) d'après le modèle à tubes de Fant (Figure 92). Les valeurs figurant sur l'interface de cette figure ne correspondent pas aux valeurs que nous avons utilisées pour les simulations. Nous avons recalculé les valeurs présentées par Fant pour

* PELORSON X. et col., 1994.

le conduit vocal d'un adulte soit entre 15 et 18 cm environ. La mesure de la longueur totale du conduit vocal que nous avons prise comme référence est celle proposée par Heim et col. (2002)[221] pour le conduit vocal d'un enfant, soit 13 cm environ.

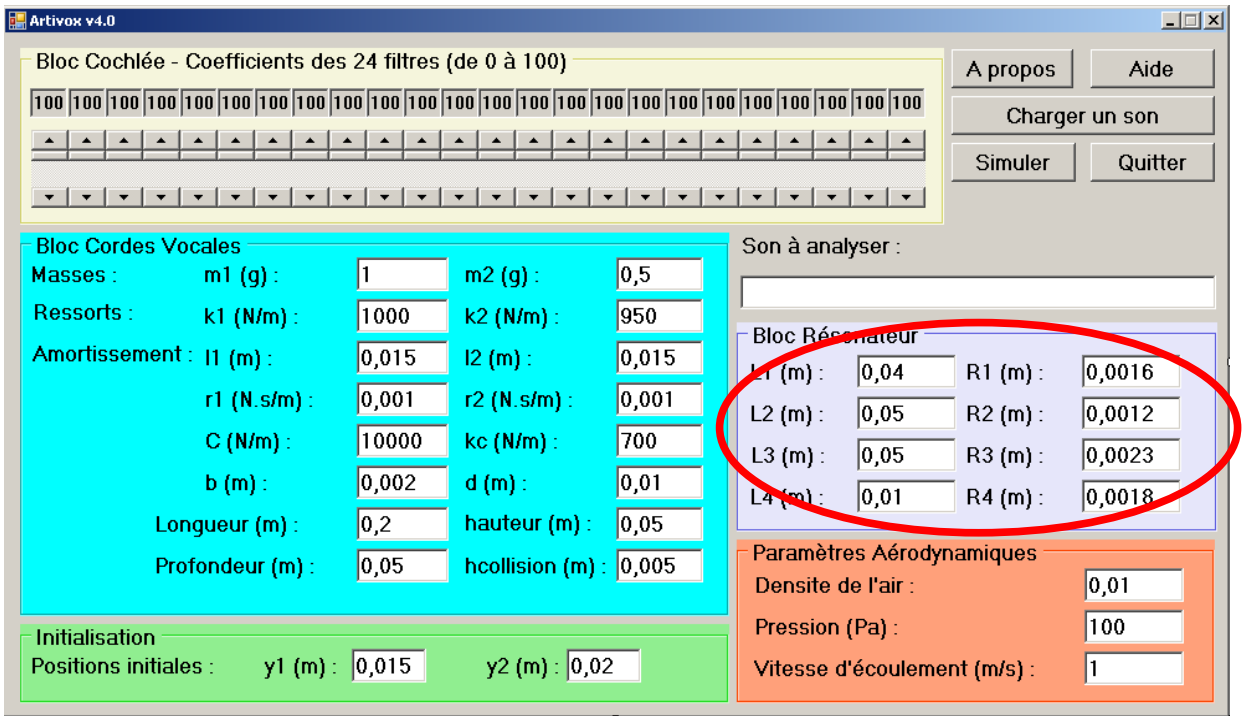


Figure 91 : Interface du logiciel ARTIVox v4.0 : paramétrage du bloc Résonateur.

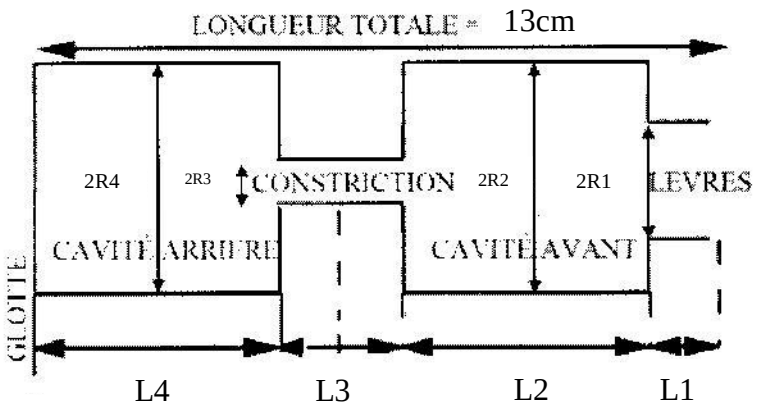


Figure 92 : Modèle du conduit vocal de Fant représentant le bloc « Résonateur ».

d) Paramètres d'initialisation de la simulation

Dans ce bloc, on trouve les paramètres pour modifier la position initiale des cordes vocales et obtenir ainsi trois des configurations glottiques les plus courantes dans la modélisation : convergente, divergente et de collision (Figure 93, les valeurs figurant sur l'interface de cette figure ne correspondent pas aux valeurs que nous avons utilisées pour les simulations).

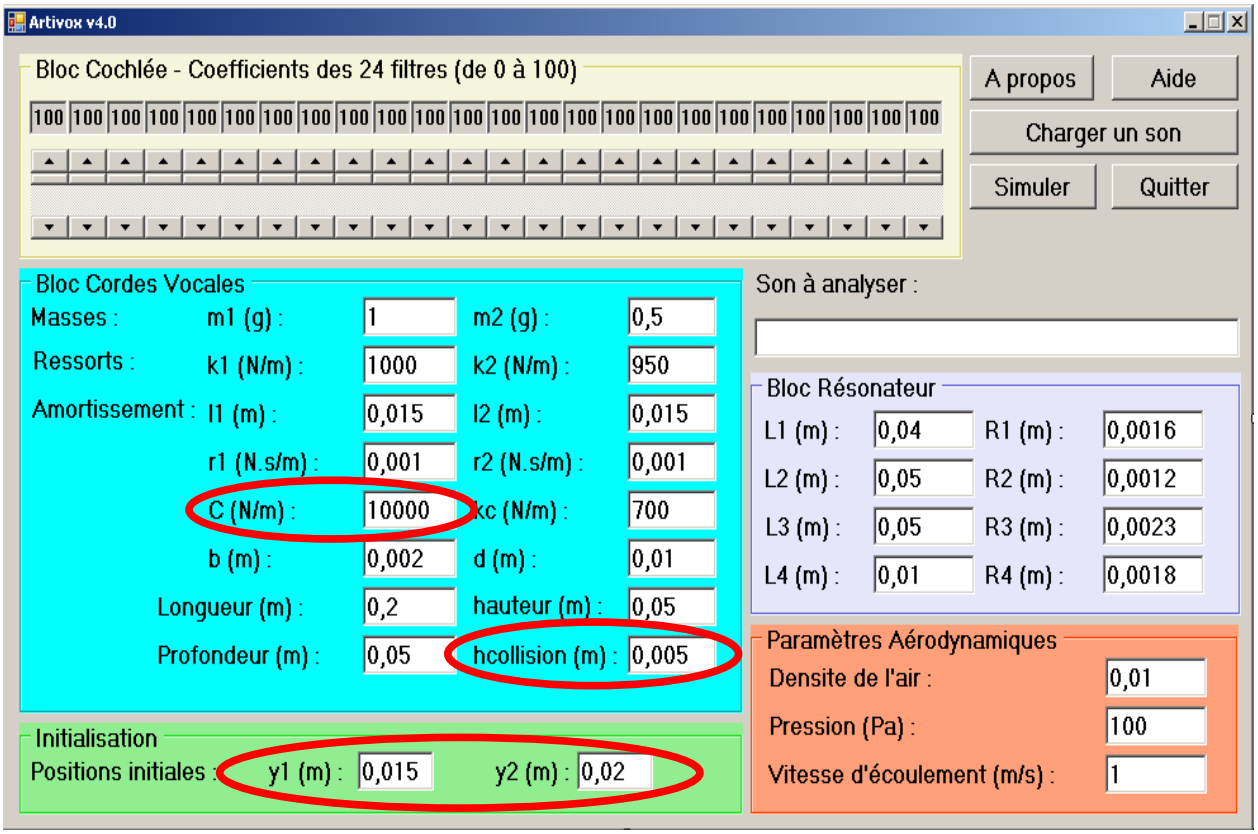


Figure 93 : Interface du logiciel ARTIVox v4.0 : paramétrage du bloc Initialisation.

Avec :

- **C** : coefficient de collision.

- **hcollision** : espace glottique existant pendant la configuration de collision.

- **y1** et **y2** : positions initiales des masses des cordes vocales (Figure 93).

Pour les différentes configurations glottiques, il est rappelé que :

- si $y1 < y2$: la **configuration** est **convergente**

- si y1 > y2 : la **configuration** est **divergente**

- si *(h-2)*y1 < hcollision* ou *(h-2)*y2 < hcollision*; alors *k1'= k1*C*, *k2'=k2*C* et *kc'= kc*C* : la **configuration** est en **collision**.

e) Paramètres aérodynamiques

Les données d'entrée de ce bloc sont relatives à l'air expiré par l'individu. Elles concernent la pression, la densité de l'air et la vitesse d'écoulement de l'air (Figure 94, les valeurs figurant sur l'interface de cette figure ne correspondent pas aux valeurs que nous avons utilisées pour les simulations). On peut modifier *la pression de l'air* $\mathcal{P}$ sortant des poumons. Elle doit être supérieure à la pression atmosphérique. Nous avons choisi la valeur de référence de cette « surpression » comme étant $\mathcal{P}_{poumons} - \mathcal{P}_{atm} = 100\ Pa$. Cette valeur correspond à la valeur moyenne arrondie de la PSGE que nous avons mesurée chez la population d'enfants implantés cochléaires de notre étude (cf. chapitre III, *section 5.3*). On peut modifier également la *densité de l'air* ρ. Elle dépend de la température et de l'altitude. Nous avons pris comme valeur de référence ρ=1,2 kg/m^3 correspondant à une température de 20°C[222]. Enfin, on peut modifier la *vitesse de sortie de l'air U*. Nous avons opté pour une vitesse de U=20 m/s, ce qui correspond à un débit moyen, mais elle peut aller jusqu'à 40 m/s dans nos simulations.

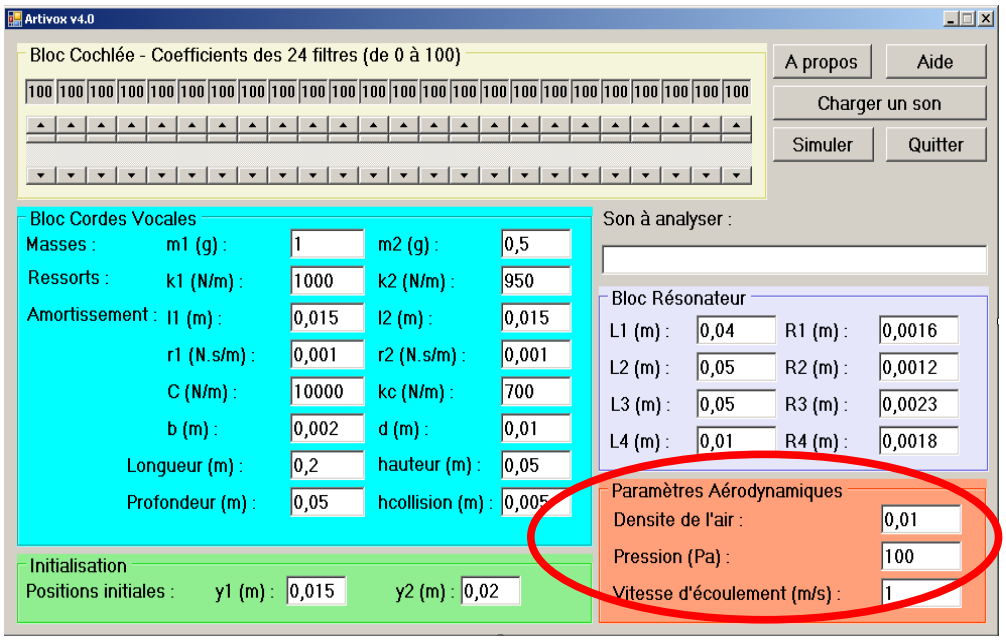


Figure 94 : Interface du logiciel ARTIVox v4.0 : paramétrage du Paramètres aérodynamiques.

3.2. Résultats de la simulation numérique

Etant donné que l'objectif principal de la simulation numérique était d'observer la dynamique de l'ensemble du système de bouclage pour la perception et l'émission d'un signal sonore, nous avons seulement réalisé des simulations dans une configuration glottique : *convergente*. Ceci car nous étions intéressés par l'observation du comportement du mécanisme glottique juste au moment du rapprochement des cordes vocales permettant de produire un son, et par l'analyse ainsi de son rôle dans l'ajustement mécanique pour atteindre le signal sonore consigne perçu.

Puisque l'application ARTIVox v4.0 nécessite un signal sonore consigne correspondant au son à analyser pour démarrer la simulation, nous avons utilisé quelques extraits d'un /a/ tenu de la voix des enfants de notre étude. Nous avons utilisé également des tons purs équivalents aux fréquences fondamentales des extraits de voix sélectionnés. Les fréquences analysées étaient égales à : 160 Hz, 206 Hz et 295 Hz. Ces fréquences ont été sélectionnées car elles sont représentatives de la hauteur des émissions orales recueillies sur le groupe contrôle de notre étude.

Le paramétrage des divers blocs de l'application ARTIVox v4.0 pour les bouclages a été réalisé en fonction des résultats obtenus, notamment du coefficient scalaire de comparaison CCTF* qui nous apportait des informations à propos du rapprochement entre le son consigne perçu** Pc_0 et le son émis perçu*** Pc (Figure 95) : plus le coefficient CCTF diminuait, plus l'écart entre les transformées de Fourier des signaux se réduisait.

Pour ce faire, nous avons commencé la tâche de paramétrage à partir de quelques valeurs présentées dans la littérature. Ensuite, nous avons analysé le comportement des variables de chaque bloc du simulateur et leur dynamique dans le système de bouclage asservi. A partir des résultats obtenus, notamment du coefficient CCTF, nous avons déterminé les valeurs d'entrée pour la simulation bloc par bloc.

a) Nous avons commencé par le bloc *« Cochlée »*. Nous avons supposé que la perception auditive d'un enfant était d'efficacité maximale, c'est-à-dire que son profil auditif

---

* Coefficient de Comparaison des Transformées de Fourier. Quand CCTF=0, alors les deux Transformées de Fourier comparées sont identiques.
** Un signal quelconque que l'enfant perçoit et souhaiterait reproduire.
*** Le signal reproduit par l'enfant

correspondait à celui d'un enfant ayant une audition normale. Pour reproduire cet effet, nous avons alors activé les 16 premiers filtres du « bloc cochlée » à 100% de perception.

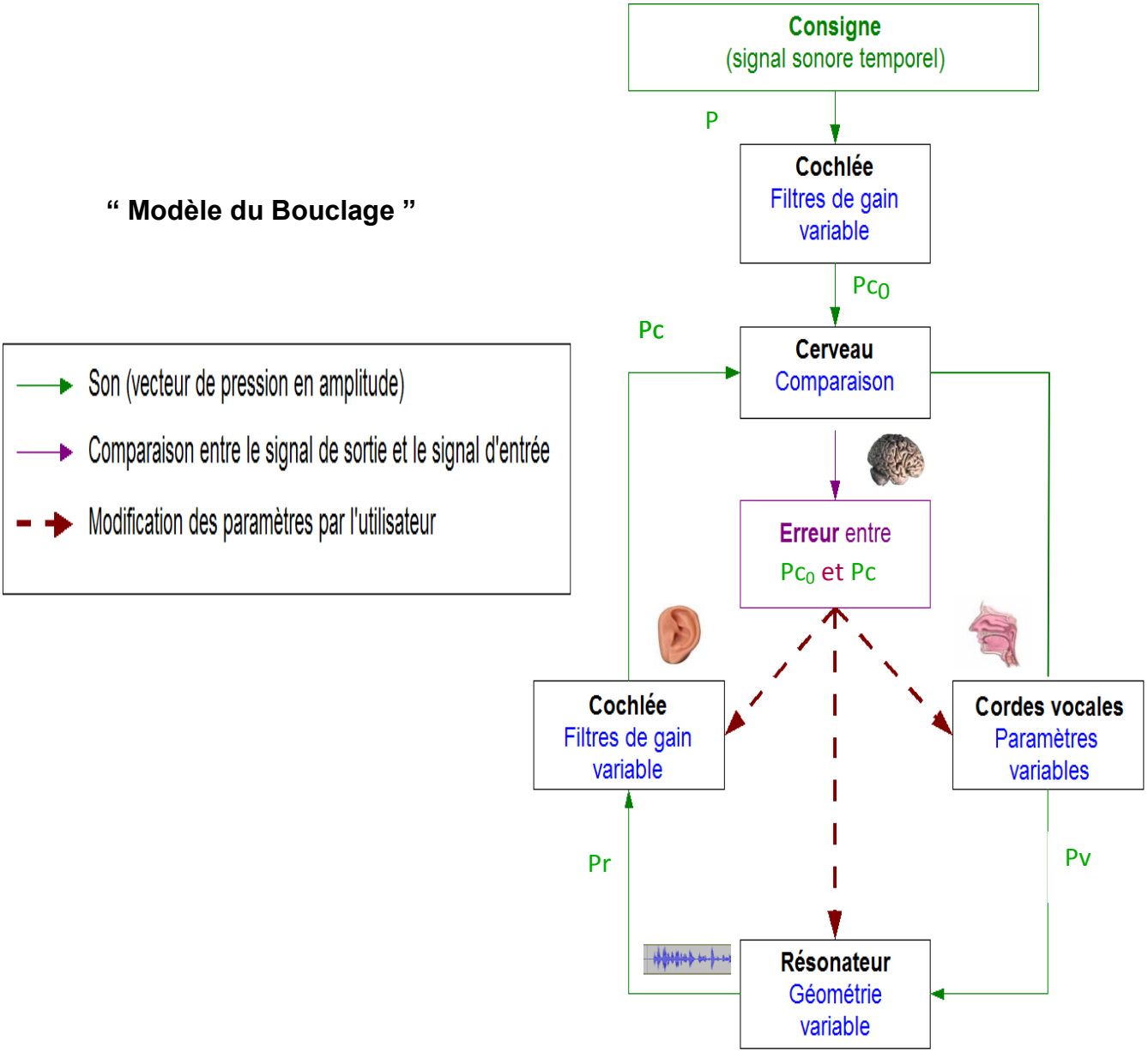


Figure 95 : Schéma de l'architecture du modèle systémique asservi ARTIVox v4.0 représentant l'action de la boucle audio-phonatoire lors de l'émission d'un signal vocalique (voyelle /a/).

b) Concernant le paramétrage du bloc *« Cordes Vocales »*, nous avons observé que :

- **m1** et **m2** : Quand la valeur de *m1* augmentait par rapport à la valeur de *m2* (Figure 96), la fréquence du signal émis *Pv* (Figure 95) diminuait. Quand la valeur de *m2* diminuait par rapport à la valeur de *m1* (Figure 96), la fréquence émise *Pv* (Figure 95) augmentait. Nous avons constaté une corrélation négative entre le paramètre de masse *m* et de fréquence *Fo :* plus la masse augmente et plus la Fo diminue (Garrel, Nicollas et col., 2006[223]). Nous avons considéré le paramètre *m* comme une variable indépendante. Les valeurs de *m* ont été fixées à m1= 0,03 et m2=0,01 pour les simulations car ce sont avec ces valeurs que nous avons obtenu les valeurs minimales du « *coefficient de comparaison des Transformées de Fourier CCTF*» indiquant le rapprochement des signaux.

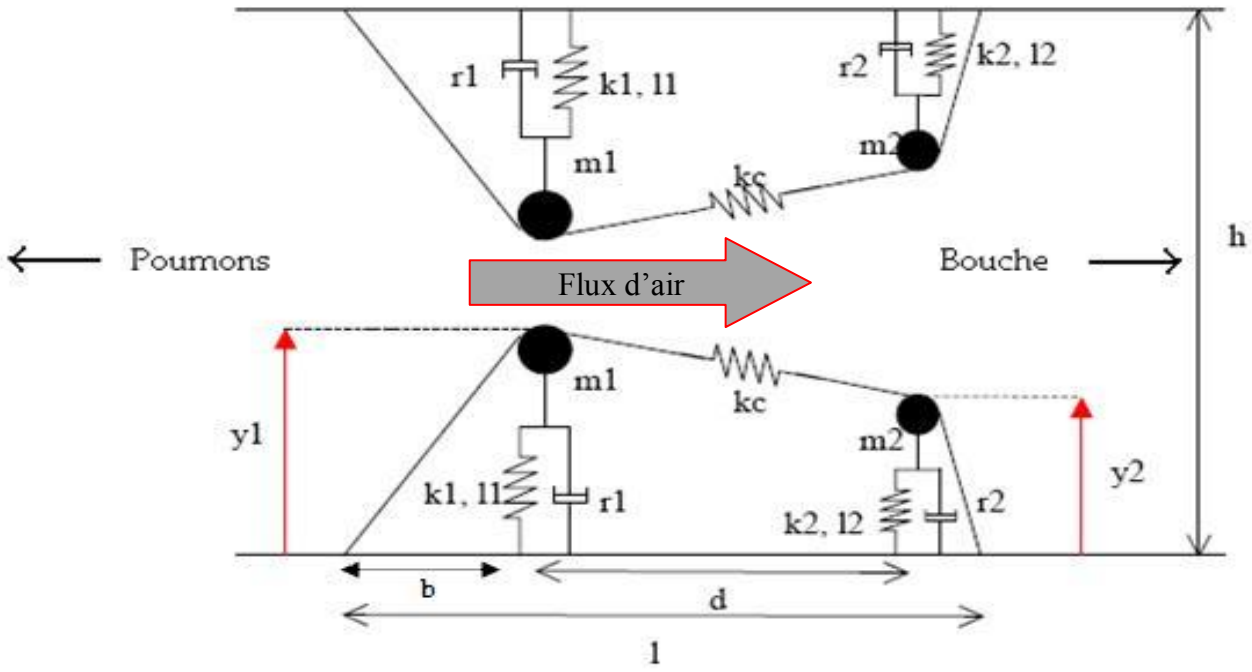


Figure 96 : Modèle à deux masses-ressort symétrique du bloc « cordes vocales »

- **k1** et **k2** : Ces paramètres représentent la tension des cordes vocales. D'une part, nous avons noté que lorsque la valeur de k1 oscille entre 2 N/m et 8 N/m, le coefficient CCTF reste à peu près stable entre les valeurs de 9 et 11 (Figure 97, à gauche). Quand la valeur de k2 oscille entre 4 N/m et 10 N/m, le coefficient CCTF reste stable dans le même rang de valeurs (Figure 97, à droite). En revanche, nous avons observé des coefficients CCTF très élevés quand les valeurs de k1 étaient au-dessus de 8 N/m et quand les valeurs de k2 étaient en-dessous de 4 N/m. D'autre part, nous avons remarqué que la valeur de kc ou raideur du ressort entre les

masses m1 et m2, n'avait pas d'incidence sur le coefficient dans une configuration glottique convergente. C'est pourquoi, nous avons fixé sa valeur à 5 et supposé que le coefficient de raideur serait équivalent à celui de k2. Enfin, nous avons constaté qu'une diminution de la raideur de k1 par rapport à k2 et de k2 par rapport à k1 entrainait une augmentation légère de la fréquence. Ainsi, nous avons fixé les valeurs de raideur pour ces ressorts à : k1=5,5 N/m ; k2=5 N/m ; kc=5 N/m.

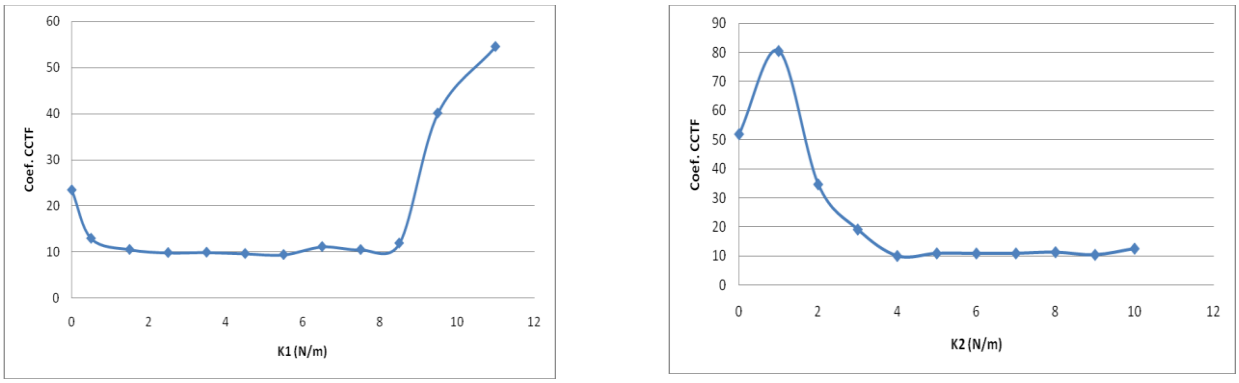


Figure 97 : Courbes de la liaison de k1 (à gauche) et k2 (à droite) avec le coefficient de comparaison CCTF.

- **r1** et **r2** : Les valeurs de ces paramètres représentent les coefficients d'amortissement, c'est-à-dire, la force proportionnelle à la vitesse de sollicitation de l'amortisseur de la corde vocale. Nous avons observé que le coefficient CCTF reste à peu près stable entre les valeurs de 9 et 11 quand 0,002 N.s/m < r1< 0,025 N.s/m (Figure 98, à gauche).

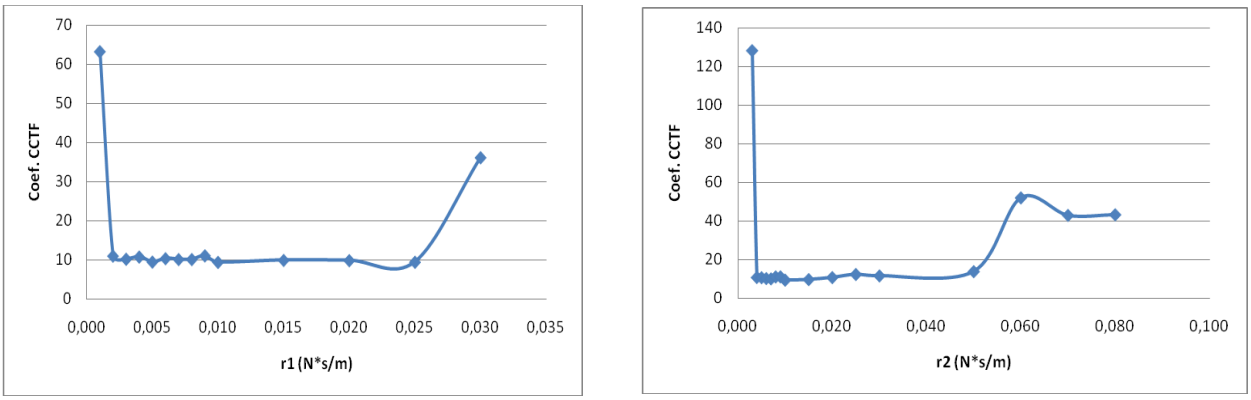
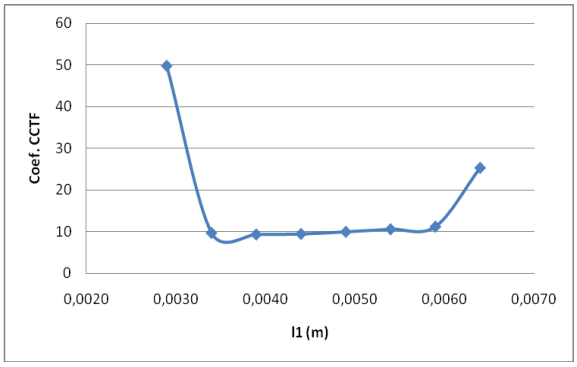


Figure 98 : Courbes de la liaison de r1 (à gauche) et r2 (à droite) avec le coefficient de comparaison CCTF.

Par ailleurs, le coefficient CCFT reste stable entre les valeurs de 9 et 13 quand 0,004 N.s/m < *r2< 0,05* N.s/m (Figure 98, à droite). En revanche, nous avons observé des coefficients CCTF très élevés quand *r1 < 0,002 N.s/m* et *r1 > 0,025 N.s/m* ; et, *r2 < 0,004 N.s/m* et *r2 > 0,050 N.s/m*. Nous avons constaté qu'une diminution de l'amortissement entrainait une diminution légère de la fréquence. Nous avons fixé les valeurs des amortisseurs : *r1=0,01 N.s/m ; r2=0,01 N.s/m*.

- **l1** et **l2** : Le coefficient CCTF se stabilise entre les valeurs de 9 et 12 quand 0,0034 m < *l1< 0,0059* m (Figure 99, à gauche). En revanche, le coefficient CCFT n'atteint pas tout à fait une stabilisation pour *l2*. Les valeurs varient entre 9 et 12 quand 0,0079 m < *l2< 0,0098* m (Figure 99, à droite). Nous avons constaté qu'une diminution de la valeur d'élongation du ressort *l1* par rapport à *l2* entrainait une légère diminution de la fréquence et qu'une augmentation de l'élongation de *l1* produisait l'effet contraire. Pour *l2*, nous avons constaté que la diminution ou l'augmentation de *l1* entrainait un effet inverse sur la hauteur du son. Les valeurs de ces paramètres ont été fixées à : *l1=0,0039 m ; l2=0,0088 m*.



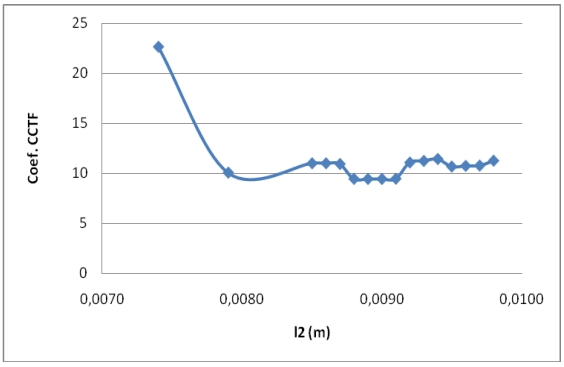


Figure 99 : Courbes de la liaison de l1 (à gauche) et l2 (à droite) avec le coefficient de comparaison CCTF.

- **b, d, et l** : Nous avons supposé dans nos simulations que *b=d/2* et *l=b+d*. Quand 0,00075 m < *l < 0,006* m, le coefficient CCTF se stabilise légèrement entre 9 et 11 (Figure 100). Nous avons noté qu'une diminution de la distance entre les masses *m1 et m2 (ou épaisseur de la partie vibrante de la corde)* entrainait une légère augmentation de la fréquence. Les valeurs de ces paramètres ont donc été fixées à : *b=0,0015 m ; d=0,0030 m ; l=0,0045 m*.

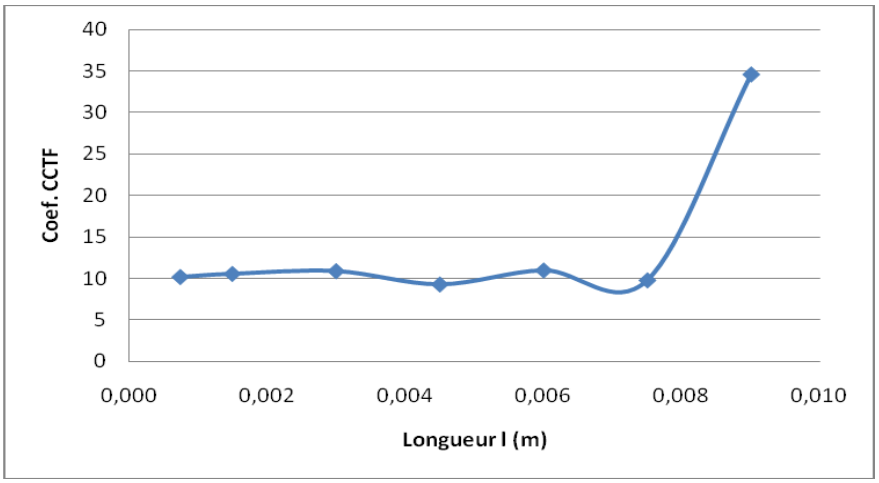
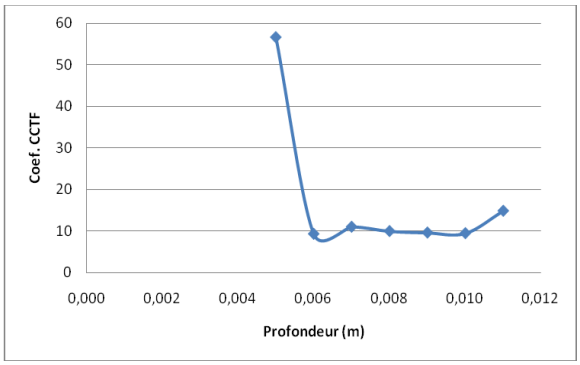


Figure 100 : Courbe présentant la liaison de l avec le coefficient de comparaison CCTF.

- **h et Profondeur** : Quand *0,006 m < Profondeur < 0,010 m* et *h=0,018 m*, le coefficient CCTF reste stable entre 9 et 11 (Figure 101, à gauche). Quand *0,017 m < h < 0,024 m* et *Profondeur=0,006 m*, le coefficient CCTF reste stable entre 9 et 13 (Figure 101, à droite). Nous avons retenu la valeur de *h* tel que proposé par Horacek et Sveck (2002)[222]. Les valeurs de ces paramètres ont donc été fixées à : *h=0,018 m* ; *Profondeur=0,006 m*.



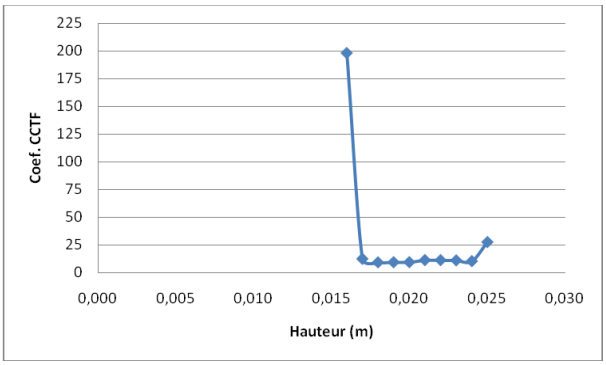


Figure 101 : Courbes de la liaison de la profondeur (à gauche) et de la hauteur (à droite) avec les valeurs des coefficients de comparaison CCTF obtenues par la simulation numérique.

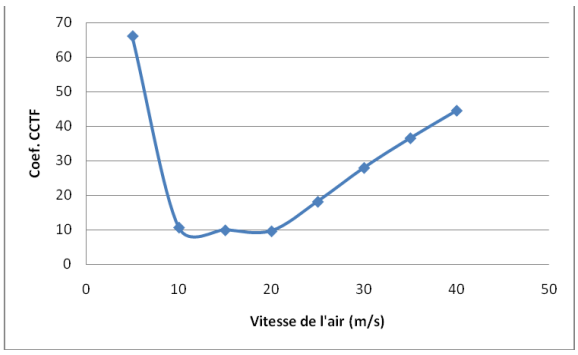
c) Concernant le paramétrage du bloc « *Résonateur* », nous nous sommes basés sur la valeur du Coefficient CCTF et le signal sonore émis en sortie par le simulateur pour trouver des valeurs cohérentes. De même, nous avons pris en compte les valeurs présentées par Guerti (2005)[224] pour les longueurs L1 et L3. Il a été assez laborieux de fixer les valeurs pour les 8 paramètres du bloc résonateur tout en surveillant la cohérence entre la qualité du son émis par le simulateur et la valeur minimale du coefficient CCTF. Comme nous l'avons dit

précédemment, nous avons utilisé une longueur totale de conduit vocal infantile égale à 13 cm. Après maintes simulations, nous avons fixé les valeurs pour les longueurs (L) et les rayons (R) des différents tubes comme suit :

- L1=0,010 m ; R1=0,005 m
- L2=0,035 m ; R2=0,018 m
- L3=0,050 m ; R3=0,005 m
- L4=0,035 m ; R4=0,018 m

Nous avons noté que le modèle à quatre tubes de Fant permet de modéliser d'une façon simple le comportement acoustique du conduit vocal. Même si la forme du conduit vocal est assez complexe pour être représenté schématiquement par des formes de tubes, le modèle de Fant permet de comprendre la façon dont les variations de la géométrie du conduit vocal retentissent sur les sons de la parole. En effet, les caractéristiques acoustiques des signaux sonores que nous avons tenté de reproduire sont liées d'une part à la source, mais surtout à la configuration géométrique du conduit vocal. Celle-ci est très dynamique car la moindre modification géométrique entraine une variation acoustique précise et particulière.

d) Concernant le paramétrage du bloc *« Paramètres Aérodynamiques »*, la valeur de la densité de l'air a été fixée à 1,2 kg/m^3. Nous avons observé le comportement de la pression et de la vitesse de l'air en fonction du coefficient CCTF. Quand *650 Pa < Pression < 1300 Pa* et *vitesse de l'air=20 m/s*, le coefficient reste à peu près stable entre 9 et 12 (Figure 102, à gauche). Quand *10 m/s < Vitesse de l'air < 20 m/s et Pression=700 Pa*, le coefficient reste à peu près stable entre 9 et 11 (Figure 102, à droite).



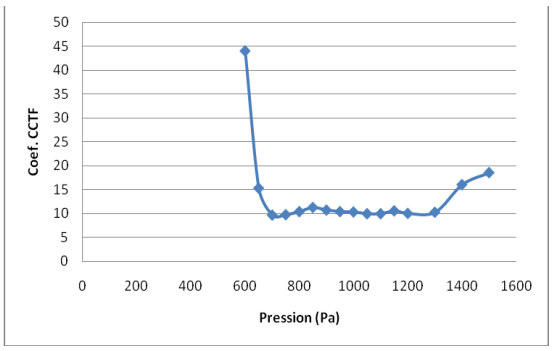
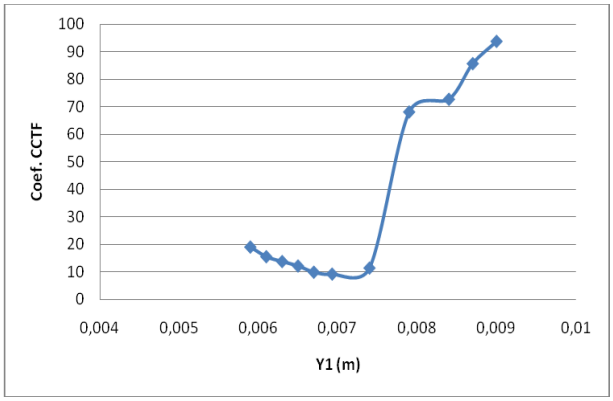


Figure 102 : Courbes de la liaison de la pression (à gauche) et de la vitesse de l'air (à droite) avec le coefficient de comparaison CCTF.

Par ailleurs, nous avons noté qu'une augmentation de la pression provoquait une augmentation de la fréquence. En revanche, une diminution de la vitesse de l'air entrainait une légère augmentation de la fréquence. Les valeurs de ces paramètres ont été fixées à : *pression=700 Pa*, *vitesse=20 m/s*.

e) Concernant le paramétrage du bloc *« Paramètres d'initialisation »*, nous avons effectué un paramétrage pour une configuration glottique convergente : *y1 < y2*. Quand *0,0059 m < y1 < 0,00693 m* et *y2=0,0087 m*, le comportement entre *y1* et le coefficient CCTF est linéaire décroissant : plus la valeur de *y1* augmente, plus le coefficient diminue (Figure 103), à gauche). Quand *0,007 m < y2 < 0,0088 m et y1=0,00693 m*, le coefficient CCTF reste stable entre les valeurs 9 et 11 (Figure 103, à droite). Les valeurs de ces paramètres ont été fixées à : *y1=0,00693 m* ; *y2=0,0087 m* ; *C=1N/m* ; *hcollision=0,000198 m*.



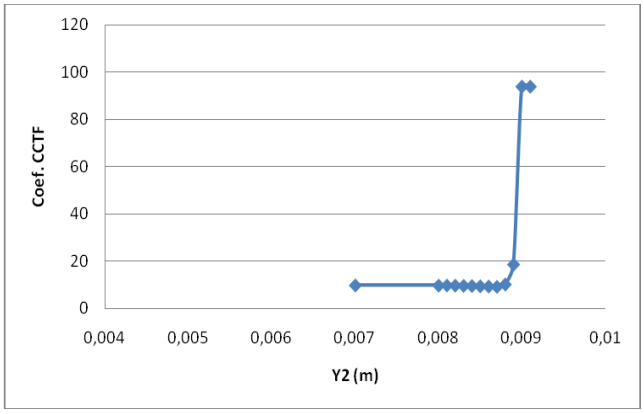


Figure 103 : Courbes de la liaison de y1 (à gauche) et de y2 (à droite) avec les valeurs des coefficients de comparaison CCTF obtenues par simulation numérique.

Une fois défini le paramétrage de chaque bloc, nous avons effectué les simulations numériques ARTIVox pour deux cas : enfant entendant et enfant déficient auditif.

3.2.1 Cas de l'enfant entendant

On suppose ici que la perception auditive de l'enfant est d'efficacité maximale correspondant à une audition normale. Pour reproduire cet effet, nous avons activé les 16 filtres du bloc « Cochlée » à 100% de perception. Nous avons paramétré les autres blocs de l'interface ARTIVox à partir des données présentées dans le Tableau 44. Après avoir exécuté la

simulation, les résultats étaient affichés dans une fenêtre qui présentait les graphiques et le coefficient CCTF calculé (Figure 104). Dans la figure, on observe un exemple de simulation dans lequel les courbes du « signal sonore Consigne » et le « signal sonore Consigne perçu » sont semblables car les filtres auditifs sont à 100% fonctionnels. Par ailleurs, les courbes du son « émis » (son émis avant le réglage dans la boucle de rétroaction) et du son « émis perçu » (son émis après le réglage des différents paramètres dans la boucle de rétroaction) sont légèrement différentes.

Bloc cordes vocales	Bloc résonateur	Paramètres aérodynamiques	Initialisation
m1=0,03 g ; m2=0,01 g	L1=0,010 m ; R1=0,005 m	Densité de l'air=1,2 kg/m^3	y1=0,00693 m
k1=5,5 N/m ; k2=5 N/m	L2=0,035 m ; R2=0,018 m	Pression=700 Pa	y2=0,0087 m
C=0 N/m ; kc=5 N/m	L3=0,050 m ; R3=0,005 m	Vitesse d'écoulement=20 m/s	
r1=0,01 N.s/m ; r2=0,01 N.s/m	L4=0,035 m ; R4=0,018 m		
l1=0,0039 m ; l2=0,0088 m			
b=0,0015 m ; d=0,0030 m			
L=0,0045 m ; h=0,018 m			
Profondeur=0,006 m ;			
hcollision=0, 000198 m			

Tableau 44 : Valeurs d'entrée pour les différents blocs du modèle de simulation ARTIVox correspondant au cas de l'enfant entendant.

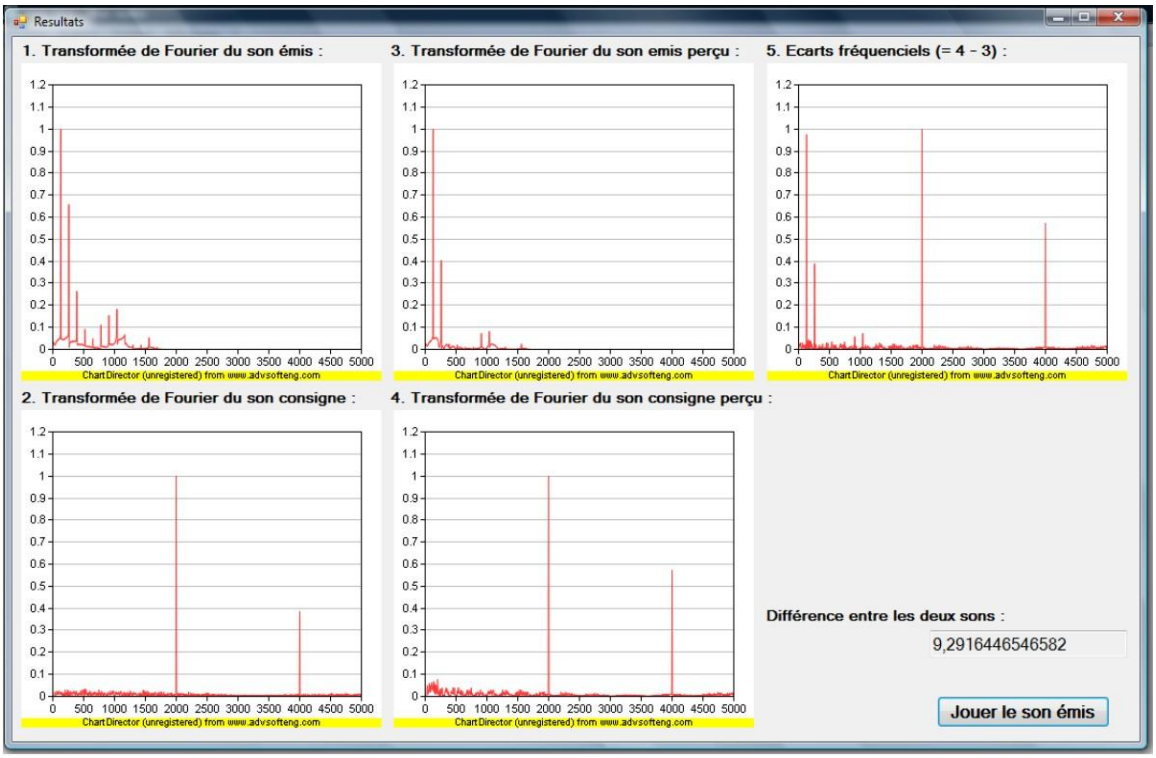


Figure 104 : Exemple d'un bouclage : filtres du bloc Cochlée à 100%.

Nous avons effectué une simulation pour trois fréquences fondamentales les plus représentatives de la hauteur des émissions vocales de notre population contrôle pour la tenue vocalique /a/ : $Fo_1 \approx 160$ *Hz* ; $Fo_2 \approx 206$ *Hz* ; $Fo_3 \approx 295$ *Hz*. Egalement, nous avons effectué des simulations à partir d'extraits de tons purs pour les mêmes fréquences. Nous avons observé une différence de résultats concernant le coefficient CCTF. Ce coefficient est plus bas lorsqu'on effectue les simulations à partir des signaux acoustiques de tons purs plutôt qu'avec des signaux acoustiques de l'émission d'un /a/ tenu enregistré (Tableau 45). Nous en déduisons que les caractéristiques de conception du logiciel ARTIVox sont plus performantes pour la simulation de signaux acoustiques simples que pour la simulation de signaux acoustiques complexes tels que ceux extraits de la voix enregistrée de notre population contrôle.

Fréquences	CCTF pour un /a/ tenu	CCTF pour un ton pur
$Fo_1 \approx 160$ *Hz*	9,42	7,23
$Fo_2 \approx 206$ *Hz*	9,25	7,10
$Fo_3 \approx 295$ *Hz*	9,29	5,90

Tableau 45 : Valeurs du coefficient CCTF pour une tenue vocalique et un ton pur.

Etant donné que nous n'avons pas développé l'automatisation de la rétroaction[*], c'est-à-dire la correction automatique des différents paramètres pour que le graphique de comparaison entre le signal « émis-perçu » et le signal « consigne » se ressemblent et que le coefficient CCTF converge le plus possible vers 0, nous avons réalisé cette manipulation manuellement. En effet, nous avons tenté de rapprocher le plus possible le signal « émis-perçu » au signal « consigne » en saisissant des valeurs possibles pour les divers paramètres de chaque bloc précédemment présenté.

3.2.2 Cas de l'enfant déficient auditif

On suppose ici que la perception du signal sonore à produire est défaillante, par conséquent la qualité et la quantité de filtrage perceptif ne sont pas à 100% optimales. Pour reproduire cet effet, nous avons activé partiellement certains filtres du bloc « cochlée ». Nous rappelons que

[*] Ceci mériterait le développement d'une version v5.0 du simulateur ARTIVox.

seulement les filtres 1 à 16 ont été utilisés car les filtres 17 à 24 ne sont pas encore actifs dans la version 4.0 d'ARTIVox. Concernant les blocs « Cordes Vocales », « Résonateur », « Paramètres Aérodynamiques » et « Initialisation », nous avons conservé le même paramétrage que pour le cas d'un enfant entendant (Tableau 44).

Afin de mieux comprendre d'une part, le rôle des filtres auditifs du bloc de perception et d'autre part, l'effet de la quantité d'information acoustique entrant dans la boucle de rétroaction de notre modèle, nous avons observé les valeurs du coefficient CCTF résultant de la modification du paramétrage pour le bloc « Cochlée ».

Premièrement, nous avons observé les valeurs du coefficient CCTF lorsque l'on diminuait le taux de perception pour chacun des filtres. Pour ce faire, on diminuait le taux de perception d'un seul filtre à 50% puis à 0% à l'aide du curseur de l'interface ARTIVox tandis que les autres filtres étaient fixés à 100%. La simulation a été effectuée à partir du signal sonore d'un ton pur (son simple) *F≈295 Hz*. Un tableau de résultats est présenté ci-dessous (Tableau 46).

Filtre	100%	50%	0%
1	5,90	6,63	7,70
2	5,90	6,25	7,05
3	5,90	6,16	6,66
4	5,90	6,05	6,40
5	5,90	6,01	6,25
6	5,90	6,12	6,50
7	5,90	6,43	7,24
8	5,90	6,20	7,06
9	5,90	5,91	6,09
10	5,90	5,88	6,00
11	5,90	5,87	5,89
12	5,90	5,90	5,94
13	5,90	5,91	5,97
14	5,90	5,92	6,00
15	5,90	5,93	6,04
16	5,90	5,90	6,01

Tableau 46 : Valeurs du coefficient CCTF en fonction de la pondération des filtres du bloc Cochlée pour un ton pur perçu.

De manière générale, nous remarquons que les valeurs du coefficient CCTF augmentent au fur et à mesure que nous diminuons le taux d'entrée du signal sonore « consigne » de 50% à 0% pour le filtrage perceptif de certaines fréquences. C'est-à-dire, plus la quantité d'information acoustique entrant par un filtre dans la boucle de rétroaction diminue, plus

l'écart entre le signal sonore « consigne » et le signal « émis-perçu » augmente de manière globale. Le taux d'information acoustique entrant par un seul filtre peut altérer globalement le taux d'information acoustique sortant de la boucle de rétroaction.

Par ailleurs, lorsqu'il s'agit du signal sonore d'un /a/ tenu *Fo≈295 Hz* (son complexe), nous remarquons une amélioration du coefficient pour les filtres 10, 11 et 12 quand le taux de perception est réduit à 50% et à 0% (Tableau 47). Les valeurs en rouge signalent une diminution du coefficient pour les filtres 10, 11 et 12 quand on diminue à 0 le taux de perception (perception nulle). Les valeurs en rouge signalent une diminution du coefficient pour les filtres 10, 11 et 12 quand on diminue le taux de perception à 0 (perception nulle).

Filtre	100%	50%	0%
1	9,29	9,66	10,31
2	9,29	9,38	9,81
3	9,29	9,38	9,58
4	9,29	9,44	9,73
5	9,29	9,46	9,83
6	9,29	9,45	9,84
7	9,29	9,4	9,78
8	9,29	9,25	9,59
9	9,29	8,93	9,03
10	9,29	7,99	7,78
11	9,29	9,69	7,26
12	9,29	9,87	9,37
13	9,29	9,18	9,43
14	9,29	9,24	9,78
15	9,29	8,55	9,8
16	9,29	8,67	9,64

Tableau 47 : Valeurs du coefficient CCTF en fonction de la pondération des filtres du bloc Cochlée pour un /a/ tenu perçu.

Quand on diminue ou que l'on supprime l'action des filtres traitant des fréquences aiguës, le signal sonore d'entrée ayant pour fréquence fondamentale *Fo≈295 Hz* est mieux reproduit, comparé et ajusté au signal « consigne ». Dans ce cas là, nous remarquons que la diminution de la fonctionnalité des filtres perceptifs ne traitant pas directement la fréquence fondamentale du signal sonore consigne complexe n'interfère ni dans le traitement perceptif par le bloc « Cochlée » ni dans la comparaison du signal par le bloc « Comparateur ».

De plus, la différence de résultats obtenus concernant les valeurs du coefficient CCTF pour les filtres 10, 11 et 12 (Tableau 46 et Tableau 47) avec un son pur et avec un son complexe

suggère que la quantité et la qualité d'information acoustique sortant de la boucle de rétroaction sont conditionnées par la nature même du signal sonore analysé. Le logiciel ARTIVox v4.0 est donc sensible à ces deux types de signaux pour effectuer les simulations.

Deuxièmement, nous avons regroupé les filtres 1 à 16 selon leurs fréquences centrales (Tableau 47) en trois groupes. Les filtres 1 à 3 ont été regroupés comme des filtres traitant les fréquences graves, les filtres 4 à 6 comme des filtres traitant les fréquences moyennes et les filtres 7 à 16 comme des filtres traitant les fréquences aiguës. Ensuite, nous avons effectué diverses simulations en modifiant le taux de perception du signal consigne pour chaque groupe de filtres.

De manière générale, lorsqu'on diminue simultanément le taux de perception à 50% puis à 0% des filtres d'un même groupe tandis que les autres filtres restent fixés à 100%, le coefficient CCTF augmente. D'autant plus lorsque le taux de perception pour l'ensemble des filtres percevant les fréquences aiguës est fixé à 0%, le coefficient augmente de manière considérable. Encore une fois, quand la quantité d'information acoustique d'entrée est partielle ou nulle, l'écart entre les signaux « consigne perçu » et « émis perçu » augmente. L'allure des courbes de la Transformée de Fourier du signal « consigne » et du signal « consigne perçu » entrant dans la boucle est tout à fait différente (Figure 105). On observe que les courbes du son « consigne » et son « consigne perçu » sont différentes quand le taux de perception est fixé à 0% pour les filtres auditifs percevant les fréquences aiguës (7 à 16).

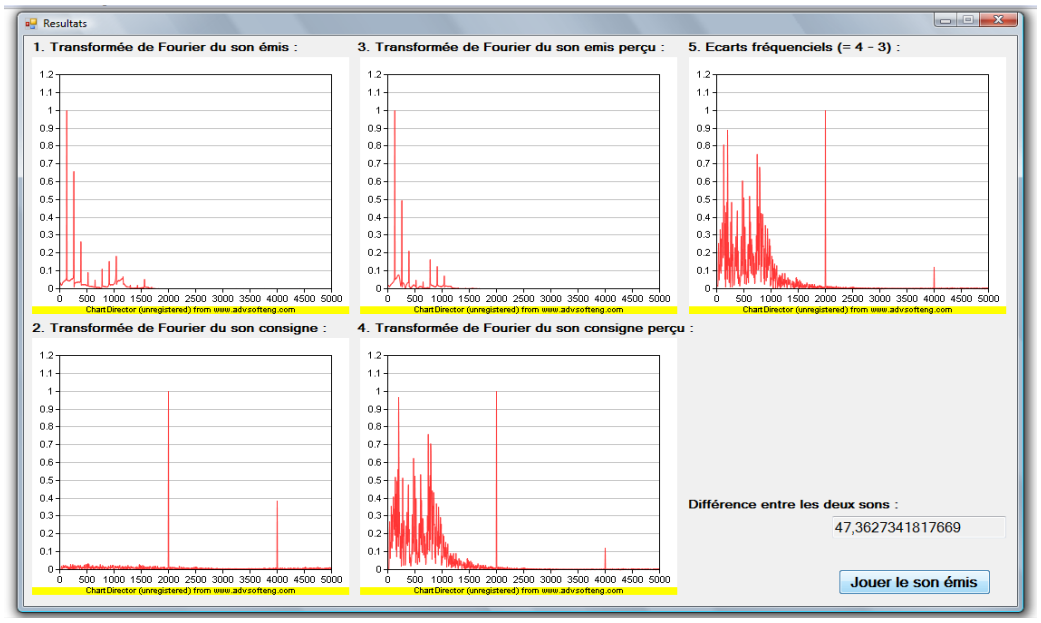


Figure 105 : Exemple d'un bouclage : filtres du bloc Cochlée à 100% (f≈295 Hz, /a/ tenu).

C'est là où l'ajustement des divers paramètres des autres blocs s'impose par le biais du mécanisme de la boucle de rétroaction. Un nombre défini de bouclages devra avoir lieu afin d'atteindre le signal consigne.

Enfin, nous avons calculé les valeurs du coefficient CCTF pour l'ensemble des filtres percevant les fréquences aiguës (filtres 7 à 16) lorsque le taux de perception était total (100%), partiel (50%) ou nul (0%). Ceci car nous nous intéressions aux différents paramétrages des filtres associés à une mauvaise perception dans les fréquences aiguës. Dans la surdité de perception profonde, les fréquences aiguës sont perdues. C'est pourquoi, nous avons simulé la perception des fréquences aiguës quand les filtres des fréquences graves et moyennes étaient paramétrés à 100%, 50% et 0%. Nous présentons le tableau de résultats des simulations ci-dessous (Tableau 48).

	Fréquences graves (Filtres 1 à 3)			Fréquences moyennes (Filtres 3 à 6)			Fréquences aiguës (Filtres 7 à 16)		
Simulation	100%	50%	0%	100%	50%	0%	100%	50%	0%
1							9,29	11,61	47,36
2							9,75	11,18	48,31
3							11,34	12,35	34,24
4							9,58	10,27	32,68
5							10,3	11,43	26,93
6							9,86	9,29	47,36
7							11,81	11,4	34,24
8							10,47	10,3	26,93
9							12,85	12,85	****

Tableau 48 : Valeurs du coefficient CCTF pour les filtres 7 à 16 résultant du paramétrage des filtres pour les fréquences graves et moyennes.

Dans les six premières colonnes, on trouve les combinaisons de paramétrage des taux de perception (100%, 50% et 0%) pour les groupes de filtres 1 à 6. Dans les trois dernières colonnes, on trouve les valeurs du coefficient CCTF calculés pour les diverses combinaisons de paramétrage en modifiant le taux de perception du groupe de filtres 7 à 16. Par exemple, dans la première simulation, nous avons obtenu un coefficient CCTF de 9,29 lorsque le taux de perception de l'ensemble des filtres percevant les fréquences aiguës, moyennes et graves était à 100% (simulation 1, Tableau 48). Le coefficient CCTF vaut 11,61, lorsque le taux de perception de l'ensemble de filtres percevant les fréquences aiguës est à 50% et le taux de perception pour les fréquences moyennes et graves était toujours à 100%. Le coefficient CCTF vaut 47,36, lorsque le taux de perception de l'ensemble des filtres percevant les fréquences aiguës est à 0% et le taux de perception pour les fréquences moyennes et graves est toujours à 100%, et de même pour les autres simulations. Dans le tableau, les cases en bleu

signalent le taux de perception fixé pour les filtres 1 à 6 dans chaque simulation. Les coefficients en vert indiquent les meilleurs valeurs obtenues et en orange les valeurs les plus mauvaises. La valeur du coefficient CCTF est très élevée signalant une « erreur » quand tous les filtres sont à 0%. La simulation était réalisée à partir d'un /a/ tenu Fo≈295 Hz.

A travers les simulations précédemment effectuées, nous avons reproduit la fonctionnalité partielle ou totale des filtres du bloc « Cochlée » pour la perception des fréquences aiguës et calculé les valeurs du coefficient CCTF. De ces simulations à partir d'un /a/ tenu Fo≈295 Hz, il ressort que :

- Quand le taux de perception des filtres pour les fréquences aiguës (filtres 7 à 16) est à 100%, à 50% et à 0%, le coefficient CCTF le plus mauvais résulte d'un paramétrage dans lequel les fréquences graves et les fréquences moyennes ne sont pas du tout perçues (0%). Ces résultats déterminent quels sont les paramétrages des filtres associés à une mauvaise perception des fréquences aiguës.

- De manière générale, les coefficients CCTF augmentent quand le taux de perception du groupe de filtres pour les fréquences aiguës diminue partiellement (50%) ou totalement (0%). L'augmentation du coefficient est beaucoup plus importante quand le taux de perception est nul. Nous constatons une liaison inverse entre les valeurs du coefficient CCTF et le taux de perception d'entrée dans la boucle de rétroaction. Les valeurs du coefficient sont plutôt cohérentes par rapport aux paramétrages des filtres.

4. Conclusions

Ce chapitre présente la réflexion théorique menée à partir des résultats obtenus dans la partie expérimentale. Nous cherchions à mieux comprendre les différences acoustiques observées entre la voix des enfants ESA et ESIC et la voix des enfants EE (expérience 1). En effet, il était important pour nous d'apprécier les résultats acoustiques obtenus sur un signal simulé en fonction de la variation de certains paramètres. Le dispositif que nous présentons ici permet de contrôler un grand nombre de paramètres indépendamment les uns des autres. Dans un premier temps, nous avons élaboré un modèle « macro-systémique » représentant, à notre sens, l'action de la boucle audio-phonatoire comme un système asservi de régulation de la voix et contrôlé par un micro-système sensoriel : le mécanisme auditif. Le modèle permet de calculer et interpréter les actions exécutées par différents micro-systèmes ou blocs que le

constituent. Nous avons tenté de représenter l'action de la boucle audio-phonatoire résultant de la mise en jeu d'un processus physiologiques conscient et volontaire. Nous représentons dans le modèle l'action du mécanisme vocal, auditif et du cerveau. Dans un deuxième temps et avec l'aide de l'Ecole Centrale de Lyon, nous avons conçu le logiciel ARTIVox permettant la simulation numérique de l'action de la boucle audio-phonatoire à partir de notre modèle systémique de rétro-action. Ce simulateur s'est voulu suffisamment paramétrable et relativement proche de la réalité pour rendre compte de la diversité des phénomènes existant chez un enfant entendant et un enfant à déficience auditive. Le logiciel ARTIVox v4.0 fournit un coefficient représentant l'écart existant entre un « son consigne » et un « son émis » ainsi que des graphiques des transformées de Fourier représentant un signal sonore à différentes instances : signal consigne, signal consigne-perçu, signal émis, signal émis-perçu et signal comparé. Ce modèle implémenté dans un simulateur numérique nous a permis de répondre à certaines questions et hypothèses formulées au départ de notre travail, notamment concernant l'incidence d'une défaillance de la boucle audio-phonatoire.

Nous avons pu observer le rôle des filtres du bloc « Cochlée » dans notre modèle de la boucle audio-phonatoire simulée comme une boucle asservie. Ces filtres agissent de manière sélective sur la perception des signaux acoustiques entrant dans la boucle. Quand il s'agit d'un signal sonore simple, la fréquence du signal est perçue par un groupe de filtres désigné pour le traitement. En revanche, quand il s'agit d'un signal sonore complexe, plusieurs groupes de filtres traitent les fréquences (fondamentale et harmoniques) constituant le signal. Ainsi, la nature du signal acoustique d'entrée joue un rôle important pour déterminer la quantité et la qualité de l'information acoustique sortant de la boucle.

Egalement, nous avons apprécié l'effet de la quantité d'information acoustique d'entrée dans la boucle sur la quantité d'information acoustique de sortie. Par le biais d'un coefficient scalaire de comparaison CCTF calculé par notre modèle lors d'un bouclage, nous avons constaté que la quantité d'information acoustique entrant était inversement proportionnelle à la valeur du coefficient CCTF. Les valeurs de ce coefficient semblent cohérentes avec les simulations effectuées. Ce coefficient peut s'avérer utile pour estimer par exemple la perte d'information acoustique traitée dans une boucle défaillante et ainsi analyser le cas d'un enfant malentendant.

Nous n'avons pas observé ici le lien entre le degré de perte auditive (dB SPL) et la valeur du coefficient CCTF car nous n'avons pas exploité le paramètre de l'intensité. Pour l'instant, le

modèle ARTIVox v4.0 développé tente de reproduire la régulation de l'émission vocale par le mécanisme auditif à partir du paramètre de la fréquence d'un signal acoustique. Nos futurs travaux tenteront d'analyser d'autres paramètres portant une richesse d'information acoustique comme l'est l'intensité.

Enfin, bien que présentant de nombreuses limites dues aux hypothèses de modélisation, le simulateur numérique ARTIVox v4.0 permet d'approcher la compréhension de la dynamique de régulation, en boucle fermée, de la voix humaine. Ce phénomène de régulation, assez contre-intuitif, peut ainsi être expérimenté à volonté avec notre simulateur dont les caractéristiques des sorties produites sont assez fidèles avec les mesures expérimentales que nous avons pu réaliser dans le cadre du chapitre III.

La difficulté majeure ici est de définir les relations existant entre les blocs et l'action du bloc Comparateur. Quel paramètre modifier et dans quelle mesure ? Quel jeu de paramètres est en cause ? Quel bloc du modèle est le plus compromis pour atténuer ou augmenter l'écart entre les deux signaux comparés? Pour répondre à ces questions, il nous a fallu avoir à disposition toute une banque de données des caractéristiques physiologiques des blocs modélisés et des échantillon sonore consigne. Les données ont été saisies dans notre modèle macro-systémique du bouclage pour lancer une simulation numérique et vérifier ainsi la cohérence de notre modèle par rapport à la réalité observée. Les valeurs obtenus des différentes simulations ont été présentées ici (section 3.2 de ce chapitre).

CHAPITRE V : CONCLUSION GENERALE ET PERSPECTIVES

L'étude que nous avons effectuée s'inscrit dans la continuité des travaux de recherche axés sur l'exploration de la voix des enfants sourds et l'apport des appareils de correction auditive sur la qualité de la voix. Dans cette étude, nous nous sommes attachés à évaluer et à caractériser la voix des enfants sourds appareillés et implantés cochléaires sur le plan acoustique, aérodynamique et perceptif. Le but général de ce travail était d'identifier et de caractériser des troubles vocaux persistant chez l'enfant sourd appareillé et implanté cochléaire.

L'appréciation auditive de la voix « particulière » des enfants sourds appareillés et implantés cochléaires nous a permis d'orienter la réflexion de notre démarche scientifique vers ce sujet d'étude. Nous nous sommes posés la question si les caractéristiques vocales de ces enfants seraient associées à des troubles vocaux. Or, la littérature sur la caractérisation physique de la *voix* de l'enfant sourd appareillé et implanté cochléaire n'est pas abondante. Les résultats présentés dans les quelques études existantes sont parfois contradictoires.

C'est pourquoi, nous avons effectué une étude transversale comparative de la voix de cette population infantile, à l'aide d'un protocole d'évaluation multiparamétrique d'analyse instrumentale et perceptive adapté. Nous avons mené cette étude en tenant compte de l'équipement et du matériel mis à disposition de l'orthophoniste dans le cadre de sa pratique clinique. Nous avons utilisé des procédés cliniques employés à l'heure actuelle en Orthophonie et Phoniatrie ainsi qu'en Phonétique expérimentale. Ces procédés sont présentés ici comme étant deux approches complémentaires : une approche acoustique complétée par une analyse aérodynamique (expérience 1 et 2, chapitre III) et une approche perceptive (expérience 3, chapitre III). Les deux approches nous ont permis de recueillir une série de données qui ont été exploitées afin de concevoir un modèle théorique expliquant le rôle que la boucle audio-phonatoire accomplit dans la production d'un signal sonore chez l'enfant sourd avec un canal auditif rétablit. Ainsi, nous pouvons expliquer les caractéristiques vocales « particulières » de sa voix.

Dans **l'approche expérimentale acoustique** (expérience 1), nous avions pour objectif de constituer une base de données sur les paramètres acoustiques pour trois groupes de population appariés en âge : enfants entendant, enfants appareillés et enfants implantés cochléaires. L'évaluation multiparamétrique de la voix a inclus essentiellement des mesures acoustiques afin d'analyser quantitativement la qualité vocale et de déterminer l'éventuelle manifestation d'une dysphonie. Nous avons établi d'une part une comparaison entre la

population contrôle et les groupes d'enfants sourds ; et d'autre part, une comparaison entre les groupes d'enfants appareillés et implantés présentant des caractéristiques de surdité et de port de l'aide auditive plutôt homogènes, à savoir : des enfants manifestant une surdité sévère à profonde, ayant été appareillés/implantés avant l'âge de trois ans et portant l'appareillage ou l'implant depuis trois ans ou plus.

De cette expérience, il ressort que *i) trois paramètres acoustiques sont discriminants entre le groupe d'enfants sourds et le groupe d'enfants entendants*, à savoir : coefficient de variation de F0, Jitter et Coefficient de variation d'intensité. Ces paramètres témoignent d'une instabilité de la fréquence du signal à court et moyen terme mais également d'une instabilité de l'intensité à moyen terme. Nous avons trouvé que les valeurs du coefficient de variation de la Fo et du Jitter sont à la frontière de la normalité pour le groupe d'enfants sourds (ESA et ESIC). Ces valeurs sont plus élevées chez le groupe ESA que chez le groupe ESIC. L'augmentation de ces paramètres chez le groupe ESA peut expliquer la présence d'une légère raucité correspondant bien à la perception de dysphonie. Campisi et col. (2005)[93] ont observée une augmentation de ces paramètres dans les profils vocaux des enfants sourds avant l'implantation. Campisi souligne que le coefficient de variation de la Fo peut rester perturbé jusqu'à six mois après l'activation de l'implant cochléaire. Nos résultats suggèrent également un effet inter-groupe sur la qualité de la voix qui réside dans la différence de prothèse auditive portée par chaque groupe. Les voix du groupe qui porte un implant cochléaire sont sensiblement comparables à celles du groupe contrôle.

ii) En outre, *trois paramètres acoustiques sont discriminants entre le groupe d'enfants appareillés et le groupe d'enfants implantés cochléaires*, à savoir : Jitter, Shimmer et Rapport H/B. Ces paramètres signalent une légère instabilité de la fréquence et de l'intensité à court terme ainsi qu'une présence de bruit superposé au signal vocal. Ils sont plus accentuées chez le groupe d'appareillés que chez le groupe d'implantés. Ces trois paramètres ensemble ont été corrélés à la perception de raucité vocale mais également à la perception d'une voix soufflée (Dekrom, 1994[137] ; Eskenazi et col., 1990[225] ; Yumoto et col., 1984[144] ; Martin et col., 1995[226]).

iii) Enfin, l'ensemble de résultats de l'analyse acoustique suggère que *la voix des enfants sourds de notre étude ne présente pas les caractéristiques traditionnellement retenues pour déterminer la voix dysphonique*. En effet, nous n'avons pas identifié de paramètres acoustiques altérés parmi ceux qui ont été mesurés (Fo, Jitter, Coefficient de variation F0,

Intensité, Shimmer, Coefficient de variation d'intensité, Rapport H/B, Etendue vocale, Tessiture, TMP). De plus, la littérature signale divers problèmes de fiabilité qui remettent en question la mesure de ces paramètres acoustiques. Par exemple, la mesure du Jitter pose de nombreux problèmes. Il en existe plusieurs définitions mathématiques et sa valeur dépend dans des proportions qui peuvent être très importantes de la technique de mesure de la Fo. Il en va de même pour la mesure du Shimmer (Guerrier et col., 2004[168]). Par ailleurs, le paramètre Rapport H/B basé sur l'énergie du spectre harmonique et celle du spectre de bruit ne donne une véritable estimation du bruit de souffle qu'avec une fréquence de vibration stable. Nous constatons que les paramètres acoustiques classiquement utilisés en clinique et en recherche, et que nous avons choisis dans notre étude pour l'évaluation de la stabilité du vibrateur laryngien, manquent d'efficacité clinique par absence d'une normalisation.

A partir de ce constat et des conclusions de la première expérience, nous avons considéré nécessaire d'effectuer une **analyse des paramètres aérodynamiques** (PSGE, DAB, Intensité, Efficacité glottique et Efficacité laryngée) (expérience 2) comme proposé par la Société Européenne de Laryngologie. L'analyse de ces paramètres apportent un complément d'information concernant la qualité du vibrateur laryngé en tant que transducteur d'énergie aérienne en énergie sonore. Nous nous sommes inspirés des travaux de Higgins et col. (1996[129], 2000[130], 2003[82]) qui ont analysé la voix et la parole des enfants implantés cochléaires à l'aide des mesures acoustiques et aérodynamiques. Les travaux de Higgins et col. avaient révélé la persistance de troubles vocaux après l'implantation cochléaire. Higgins et col. suggèrent que les mesures physiologiques peuvent mettre en évidence des anomalies qui ne seraient pas identifiés par des mesures perceptives ou acoustiques chez les enfants implantés.

Puisque nous n'avons pu évaluer les paramètres aérodynamiques pour les groupes EE et ESA, cette expérience de l'étude expérimentale nous a permis d'accomplir partiellement les objectifs proposés. De l'ensemble des résultats obtenus pour le groupe ESIC, nous constatons que la fonction laryngée chez l'enfant implanté cochléaire s'accomplit manifestement sans comportements pathologiques et évolue au cours du temps. Les valeurs des paramètres aérodynamiques de la PSGE et du DAB s'avèrent cohérentes et normales pour une population ayant été implantée avant l'âge de 3 ans et portant l'implant cochléaire après huit ans ou plus. On sait bien que la PSGE est un autre paramètre d'évaluation des dysphonies qui est en relation directe avec le grade de pathologie, c'est-à-dire, si la valeur du paramètre augmente, elle traduit davantage un grade de pathologie. Par ailleurs, la durée de l'implantation a une

incidence significative sur la PSGE : plus la durée de l'implantation augmente, plus la PSGE diminue. Dans ce sens là, nous constatons que la PSGE diminue et oscille entre 7 et 9 hPa après huit ans ou plus d'implantation. En somme, nous avons pu caractériser la voix de l'enfant implanté cochléaire à partir des mesures physiologiques. En termes de faisabilité et reproductibilité, la technique de recueil de mesures physiologiques de Smitheran & Hixon (1981[169]) initialement proposée aux adultes a pu être adaptée à une population infantile en prenant en compte certaines contraintes d'ordre matériel (taille du masque, position de l'enfant pendant l'enregistrement, manipulation d'une sonde, stress psychologique de l'enfant). Des moyens techniques existants peuvent être adaptés afin d'exploiter ces mesures auprès de cette population.

Les conclusions des deux expériences précédemment menées dans notre approche acoustique nous ont rappelé la nécessité d'effectuer une analyse perceptive de façon standardisée et contrôlée. Dans l'**approche perceptive** (expérience 3), les résultats de l'analyse perceptive effectuée par un jury d'écoute fiable ont révélé que les voix de la population d'enfants sourds (ESA et ESIC) ne sont pas dysphoniques. Ainsi, l'une des hypothèses de notre étude étant que les voix des enfants sourds manifestaient des troubles vocaux n'a pas été confirmée. En effet, les voix de ces groupes avec les voix d'enfants entendants ont été classées dans les deux premiers niveaux (G0=voix normale, G1=voix légèrement dysphonique) du paramètre "G" (grade global de dysphonie) de l'échelle GRBAS de Hirano. 75% des voix de la population ESA contre 55% des voix de la population ESIC ont été perçues comme étant des voix légèrement dysphoniques. Globalement, les résultats de l'analyse acoustique et aérodynamique (expérience 1 et 2) ont été validés par l'analyse perceptive. Les valeurs des paramètres acoustiques marquant des différences inter-groupes (Coefficient de variation Fo, Jitter, Coefficient de variation d'intensité, Shimmers, Rapport H/B), sont cohérentes par rapport au classement du jury d'écoute. Toutefois, un nombre non négligeable de voix classées perceptivement dans le grade intermédiaire (G1) témoigne de la difficulté d'évaluation des caractéristiques vocales d'enfants sourds par les membres du jury d'écoute. Dans la littérature, plusieurs étude ont déjà montré que le jury d'écoute est davantage variable sur l'évaluation des dysphonies intermédiaires (Kreiman, 1992[174], 1993[227] ; Revis, 2001[162], 2004[132]). Bien que les résultats de l'évaluation instrumentale des enfants sourds ne manifestent pas de signes d'anomalie vocale, l'évaluation perceptive met en évidence une variabilité élevée du jury pour évaluer les voix de ces groupes. Ce résultat suggère que la voix des enfants sourds comporte des caractéristiques qui la rendent difficile à saisir perceptivement par rapport aux qualités de la voix à proprement parler. Il est possible que des

caractéristiques relatives à la parole (nasalité, articulation, pauses, rythme,…) soient ancrées dans l'ensemble d'éléments physiques de la voix des enfants sourds. Nous avons constaté que l'échelle GRBAS n'est pas une échelle pertinente pour évaluer et analyser la voix des enfants déficients auditifs. Il est nécessaire de mettre au point une échelle d'évaluation perceptive de la voix avec des critères bien définis (p.e. nasalité, forçage), tenant compte des particularités de la voix de la population infantile sourde.

Un autre travail a été envisagé afin de mieux comprendre les différences acoustiques observées entre la voix des enfants ESA et ESIC et la voix des enfants EE (expérience 1). Nous nous sommes attachés à la conception et au test d'un logiciel de simulation de l'action de la boucle audio-phonatoire afin de reproduire élémentairement la dynamique de régulation à partir de quatre mécanismes physiologiques y participant : auditif, vibrateur, résonateur et comparateur (cerveau). Nous avons tenté d'élaborer un modèle vibro-acoustique afin d'expliquer l'effet de la perception du signal acoustique sur l'émission vocale lorsque la boucle audio-phonatoire est défaillante. En effet, l'approche résulte d'un besoin d'acquérir des compétences dans le domaine de la simulation numérique nous permettant de comprendre la dynamique de régulation de la boucle audio-phonatoire. La régulation de la boucle est possible grâce à la différence quantitative établie par le mécanisme comparateur entre l'information acoustique entrant et l'information acoustique sortant du système. Cette comparaison est basée sur les caractéristiques fréquentielles de l'information acoustique entrant dans le système. Nous avons voulu représenter cette régulation comme étant contrôlée.

Nous avons implémenté le modèle proposé dans le simulateur numérique ARTIVox v4.0. Ceci afin de disposer d'un banc d'essai permettant de manipuler théoriquement différentes données et de les corréler ultérieurement avec les données expérimentales. Certaines données d'entrées ont été prises de notre base de données expérimentales précédemment présentée, à savoir : la pression, les extraits vocaux, la Fo. Egalement, nous avons du consulter d'autres bases de données dans la littérature.

Les valeurs d'initialisation ont été basées sur la simulation analogue de l'action de la boucle audio-phonatoire d'un enfant entendant et sourd bénéficiant d'une prothèse auditive. Nous avons mené de nombreuses simulations en manipulant manuellement les différents paramètres de notre modèle afin d'obtenir le signal le plus proche possible du signal consigne. A l'issue de ces simulations, les résultats obtenus montrent que la nature du signal acoustique d'entrée joue un rôle important pour déterminer la quantité et la qualité de l'information acoustique

sortant de la boucle. Nous avons pu observer le rôle des filtres gammatones du bloc « cochlée/récepteur » de notre modèle. Ces filtres agissent de manière sélective sur la perception des signaux acoustiques entrant dans la boucle. Quand il s'agit d'un signal sonore simple, la fréquence du signal est perçue par *un groupe* de filtres désigné pour le traitement. En revanche, quand il s'agit d'un signal sonore complexe, *plusieurs groupes* de filtres traitent les fréquences (fondamentale et harmoniques) constituant le signal. Si le groupe de filtres désigné est fonctionnel pour traiter la fréquence fondamentale du signal, ce dernier est perçu et comparé. Si la fonctionnalité des filtres traitant les harmoniques du signal sonore est compromise, le signal est pourtant perçu et comparé par le bloc « comparateur ». Les modifications en sortie étaient effectuées manuellement car l'automatisation du bouclage et de correction n'est pas encore implémentée.

Nous avons observé l'effet de la quantité d'information acoustique d'entrée sur la quantité d'information acoustique de sortie dans la boucle. Par le biais d'un coefficient scalaire CCTF (Coefficient de Comparaison des Transformées de Fourier) calculé par notre modèle lors d'un bouclage, nous avons constaté que la quantité d'information acoustique entrant était inversement proportionnelle à la valeur du coefficient CCTF. Les valeurs de ce coefficient semblent cohérentes aux simulations effectuées avec les données de notre étude expérimentale. Ce coefficient s'avère utile pour estimer la quantité d'information acoustique entrant et sortant de la boucle de rétroaction.

Nous n'avons pas corrélé le degré de perte auditive (dB SPL) à la valeur du coefficient CCTF car nous n'avons pas exploité le paramètre de l'intensité. Pour l'instant, notre modèle reproduit la régulation de l'émission vocale par le mécanisme auditif en exploitant le paramètre de hauteur du signal acoustique.

Nous avons observé une corrélation du coefficient CCTF avec le taux de perception du signal d'entrée. Ceci pourrait expliquer qu'avec l'augmentation de la surdité (ou degré de perte auditive) s'observe non seulement une diminution de la production orale (perte en quantité), mais également une perte en qualité avec apparition de distorsions auditives (dues à la perturbation de l'analyse des sons par les cellules ciliées). Cette perte de qualité explique les difficultés qui peuvent se rencontrer dans l'adaptation des aides auditives. Elle pourrait expliquer également l'intolérance paradoxale, aux bruits rencontrés chez des sujets sourds.

La boucle audio-phonatoire ne peut pas s'analyser uniquement du point de vue vibro-acoustique et physiologique. L'influence de l'environnement sonore externe et d'autres

aspects psycho-acoustiques y intervenant nécessitent d'être pris en compte pour modéliser l'action de la boucle audio-phonatoire de manière plus fidèle. Néanmoins, un modèle aussi simple que celui que nous proposons a permis d'effectuer des observations sur l'effet de la qualité et de la quantité d'information acoustique entrant sur la qualité et la quantité d'information acoustique sortant dans un système de bouclage.

Cet exercice de modélisation a abouti à la production d'un logiciel et sa documentation (manuel d'utilisation et support théorique). En outre, cette expérience nous a permis d'interagir avec des professionnels d'autres disciplines nous apportant des expériences dans le domaine de l'acoustique, de la mécanique, de la physique et de l'informatique qui nous ont été indispensables pour comprendre le phénomène de la régulation de la voix par le mécanisme auditif via la boucle audio-phonatoire.

L'originalité de notre thèse réside dans la proposition d'un modèle implémenté dans un outil informatique permettant de simuler l'action de la boucle audio-phonatoire défaillante à partir de nos données expérimentales. Notre étude contribue aux réflexions menées dans le vaste et complexe domaine de recherche de la production et perception de la parole et de la voix. Particulièrement, ce travail contribue à une connaissance de la caractérisation de la voix des enfants malentendants : celle-ci nécessite d'être abordée de divers points de vue. Il s'agit d'une étude transdisciplinaire proposant un éclairage clinique sur la voix des enfants sourds oralisés afin d'améliorer sa prise en charge rééducative. Nous avons décrit la nature acoustique de la voix des enfants sourds à partir d'un nombre d'effectifs statistiquement fiable (expérience 1) ainsi que la fonction laryngée des enfants implantés cochléaires (expérience 2). De plus, nous avons constaté l'utilité et la reproductibilité de la technique de Smitheran et Hixon pour la mesure de la PSGE chez l'enfant en pratique clinique. Nous avons également conduit une étude perceptive (expérience 3) aboutissant au rejet de notre hypothèse que les voix des enfants sourds appareillés et implantés d'un certain profil (en termes de surdité et appareillage) sont pathologiques.

Limites et perspectives

De nombreux points n'ont pu être abordés dans cette étude et mériteraient d'être développés ultérieurement :

- A la suite de l'analyse des paramètres acoustiques des groupes de population (expérience 1), nous avons constaté une incohérence des résultats concernant le paramètre d'intensité pour les différents types de voix (conversationnelle, forte et d'appel). Il est très probable que cette incohérence soit associée à un problème de calibrage du matériel. De ce fait, nous n'avons pu intégrer dans nos analyses ce paramètre important pour comprendre le rôle de l'adaptation de la puissance sonore au type de voix parlée utilisé. Il serait donc souhaitable de mesurer ce paramètre tout en prenant les précautions techniques nécessaires afin de rendre les résultats fiables. Par ailleurs, l'analyse acoustique effectuée sur la parole spontanée n'a pas montré de différences significatives inter-groupe du fondamental usuel (FoUm, dynamique) alors que sur la voyelle tenue, nous avons mis en évidence des différences. Nous en déduisons que : soit l'extrait de parole utilisé n'est pas adéquat pour ce type d'analyse instrumentale, soit le matériel d'analyse employé dans notre étude n'est pas performant avec l'extrait de parole traité. Par conséquent, il serait nécessaire de considérer un autre type de matériau phonétique proposant simultanément les avantages d'une tenue vocalique mais également les avantages d'un extrait de parole afin de mieux analyser instrumentalement la qualité acoustique de la voix au cours de la parole.

- D'autres méthodes d'analyse acoustique comme l'analyse du spectogramme permettant l'observation simultanée de la fréquence, de l'intensité et du temps et le phonétogramme s'avèrent utiles dans ce genre d'étude. Cependant, cette dernière méthode normalisée et proposée par l'UEP* demande une procédure lourde bien qu'automatisée par ordinateur. Or, il s'agit d'une population infantile.

- A la suite de l'analyse des paramètres aérodynamiques de la population ESIC (expérience 2), nous avons constaté la nécessité d'une base de données normative sur ces paramètres pour effectuer une comparaison. Il conviendrait d'effectuer une étude transversale et dans la mesure du possible longitudinale, sur un nombre d'effectifs statistiquement fiable et sur une population infantile bénéficiant d'un implant cochléaire bilatéral.

* Union de Phoniatres Européens

- Il serait raisonnable d'envisager la standardisation d'un protocole d'évaluation multiparamétrique de la voix des enfants sourds afin de faciliter par exemple la comparaison de résultats de suivi d'un enfant. Cela serait intéressant par exemple pour les différentes équipes d'implantation cochléaire qui évaluent régulièrement les effets de l'implantation. De même, une étude des paramètres acoustiques et aérodynamiques serait pertinente.

- A la suite de l'analyse perceptive des groupes de population, nous avons constaté que l'utilisation de l'échelle GRBAS n'était pas adéquate pour évaluer perceptivement la voix des enfants sourds. Nous l'avons inclus dans notre étude car nous étions intéressés à déterminer si les voix étaient ou non dysphoniques. En effet, cette échelle semble insuffisante pour déterminer des troubles vocaux sur une population infantile avec des caractéristiques aussi particulières que celle des enfants sourds. Il serait convenable d'envisager la technique « phonetic labelling » proposée et validée par Revis (2004[132], 2006[228]) consistant à écouter et à caractériser chaque phonème d'une phrase en normal/rauque/soufflé/dévoisé afin d'affiner l'appréciation de la voix tout en tenant compte des occurrences éventuellement pathologiques au cours de la parole et en fonction des contextes phonétiques et linguistiques.

- Nous n'avons pu analyser les résultats obtenus de l'observation clinique du comportement vocal pour des raisons de temps. Il serait intéressant d'effectuer une étude sur l'aspect postural et de sa corrélation aux mesures instrumentales. Afin de tirer profit de cet aspect qui porte sur la façon dont un sujet produit la voix, il est utile d'envisager des méthodes objectives comme l'analyse kinématique qui consiste à déterminer les coordonnées de points cibles d'une image vidéo, rapportées à un référentiel déterminé lors de la procédure de calibration.

- A la suite de la simulation numérique, nous avons constaté que les paramètres proposés dans le modèle sont nombreux et assez interdépendants pour réaliser une rétroaction automatique basée sur quelques paramètres (par exemple raideur, vitesse de l'air). Si nous formulons comme hypothèse que *« le cerveau régule la voix grâce à des centres de stockage d'information sur les modifications à effectuer pour obtenir certains sons »*, l'automatisation de la rétroaction pourrait avoir lieu en insérant des blocs de mémoire dans le modèle pour reconnaitre les sons « consignes ». Une autre solution envisageable est d'implémenter un algorithme de comparaison dynamique utilisé en reconnaissance vocale afin de modifier en temps réel les paramètres, et d'observer les variations d'un signal sur les graphiques. Cette solution nécessiterait un temps de calcul plus court envisageable par la programmation d'une

transformée de Fourier rapide.

BIBLIOGRAPHIE

BIBLIOGRAPHIE

1. SHANNON, C.E., « A Mathematical Theory of Communications », *Bell System Technical Journal, July, 1948.*

2. MEILLET, Antoine, *Linguistique historique et linguistique générale*, Vol. I, Paris, Champion, 1921.

3. DE SAUSSURE, Ferdinand, *Cours de linguistique générale*, 1ère éd. 1916, éd. Charler Bally et Albert Sechehaye, trad. Wade Baskin, New York, McGraw-Hill Book Company, 1966.

4. HUMBOLDT, W. Von, *Sur le caractère national des langues*, Ière éd. 1822, Paris, Seuil Essais, 2000.

5. KONOPCZYNSKI, Gabrielle, « La voix : mono-support ou multi support ? » in *A la recherche de voix du dialogisme*, Cahiers de praxématique, n° 49, Montpellier, Presses Universitaires de la Méditerranée, 2007, p. 33-56.

6. HIRST, Daniel et DI CRISTO, Albert, *Intonation systems: a survey of twenty languages*, Cambridge, Cambridge University Press, 1998.

7. YOYOTTE J., « La réligion égyptienne » in *Histoire des religions I. Encyclopédie de La Pléiade*, Paris, Gallimard, 1956.

8. DUROZOI, Gérard et ROUSSEL, André, *Dictionnaire de philosophie*, Paris, Nathan, 1997.

9. ARISTOTE, Livre II, Chapitre VIII, *De l'âme*, trad. et notes J. Barthélemy-Saint-Hilaire, Paris, Libraire Philosophique de Ladrange, 1846, [En ligne : http://remacle.org/bloodwolf/philosophes/Aristote/tableame.htm]. Consulté le 26 juin 2009.

10. ARISTOTE, Organon I, Catégories II, *De l'interprétation*, trad. et notes J. Tricot, Paris, Vrin, Bibliothèque de textes philosophiques, 1984.

11. ARISTOTE, Livre I, *Politique*, trad. et notes J. Barthélemy-Saint-Hilaire, Paris, Libraire Philosophique de Ladrange, 1824, [En ligne : http://remacle.org/bloodwolf/philosophes/Aristote/tablepolitique.htm]. Consulté le 26 juin 2009.

12. PARRET, H., *La voix et son temps*, De Boeck Université, Bruxelles, 2002. ISBN 2-8041-3963-8.

13. LAFON, Jean-Claude, « Pathologie vocale, surdité et éducation de l'enfant. De l'audition de l'enfant », *Bulletin d'Audio-phonologie : Annales scientifiques de l'Université de Franche-Comté*, Vol. 12, n° 5-6, 1996, p. 529-543.

14. « Approches historique et épistémologiques des recherches sur la parole et la voix », *in La voix et ses troubles chez les enseignants*, Centre d'expertise collective de l'INSERM, Les éditions INSERM, 2006, p. 15-90.

15. TESTON, Bernard, « L'œuvre d'Etienne-Jules Marey et sa contribution à l'émergence de la phonétique dans les sciences du langage », *Travaux interdisciplinaires du Laboratoire Parole et Langage*, Vol. 23, p. 237-266.

16. PERELLO, J., *Trastornos del habla*, Barcelona, Masson, 5a edición, 1990.

17. MYSAK, Edward D., *Patología de los mecanismos del habla*, Barcelona, Salvat Editores S.A., 1980.

18. WOISARD, V., PERCODANI, J., SERRANO, E., PESSAY, JJ., « L'évolution de l'appareil phonatoire et la voix de l'enfant », *Données Actuelles*, Vol. 194, 1998, p. 39-44.

19. TITZE IR., *Principles of voice production*, Englewood Cliffs, NJ, Prentice Hall, 1994.

20. HACKI, T., et HEITMÜLLER, S., « Development of child's voice : premutation, mutation », *Pediatr Otorhinolaryngol*, Vol. 49, Suppl. 1, p. 141-144.

21. NICOLLAS, R., « Normal voice in children between 6 to 12 year of age : database and nonlinear analysis », *Journal of Voice*, Vol. 22, n° 6, p. 671-675.

22. PAPEZ, J.W., « A proposed mechanism of emotion », *Arch Neurol Psychiatr*, Vol. 38, 1937, p. 725-743.

23. MAC KEAN, P.D., « Psychosomatic disease and the 'viseral brain' recent developments bearing on the Papez theory of emotion », In *Basic Reading in Neuropsychologics*, R.L. Isaacson ed., New York, Harper and Row, 1964, p. 181-211.

24. MESULAM, M.M., « Patterns in behavioral neuroanatomy », In *Principles of Behavioral*, F.A. Davies, Philadelphia, ed. Neurology, 1985, p. 1-70.

25. DAMASIO, A., « The Feeling of What Happens : Body and Emotion in the Making of Consciousness », New York, Harcourt Brace, 1999, in *Anatomy of emotions*, F. Staub, L. Brugglmann, P. Maglstrettl, J. Bogousslavsky, *Schweiz Arch Neurol Psychiatr*, Vol. 153, 2002, p. 344-353.

26. SONNINEN, A., « Phoniatric viewpoints on hoarseness », *Acta Otolaryngol*, Vol. 263, 1970, p. 68-81.

27. IWARSSON, J. et SUNDBERG, J., « Effects of lung volume on vertical larynx position during phonation », *J Voice*, Vol. 12, 1998, p. 159-165.

28. MC KINNEY, J., *Diagnostic and correction of vocal faults*, Nashville, Broadman Press, 1982.
»
29. FAURE, M.A., « Dynamique respiratoire et qualités acoustiques de la voix », *Rev Laryngol Otol Rhinol (Bord)*, Vol. 108, n° Spec, 1987, p. 369-372.

30. BOUISSET, S. et MALTON, B., *Muscle, posture et mouvement : bases et applications de la méthode électromyographique*, Paris, Hermann, 1995.

31. DORLAND, W.A., *The American Illustrated Medical Dictionnary*, Philadelphia, Saunders, 1951, p.1222.

32. STARK, J. *Bel Canto*, Toronto, University of Toronto Press, 1999, p.104.

33. GOULD, W.J., « Voice Production and Postural Mechanisms », Actes du15e Congrès Phoniatrique, Buenos Aires, Août, 1971.

34. LE HUCHE, François et ALLALI, André, *La voix. 3 tomes*, Paris, Masson Ed., 1990, (rééd. Paris, Masson Ed., 2001).

35. THIEBAULD, C.M. et SPRUMONT, S., *L'enfant et le sport : Introduction à un traité de médecine du sport chez l'enfant*, Bruxelles, De Boeck Université, 1997, p. 117. ISBN 2804-126374

36. AMY DE LA BRETQUE, Benoit, *Le chant : contraintes et liberté*, Paris, Ed. Fuzeau Courlay, 1991.

37. FRIED, M.P., *The Larynx, a multidisciplinary approach*, St Louis, Mo, Mosby Ed. (2nd edition), 1996.

38. BURT, A.M., *A Textbook of neuroanatomy*, Philadelphia, WB Saunders, Vol.1, 1993.

39. LECHEVALIER, B. et IGLESIAS, S., « Structures anatomiques de la perception auditive : bruits, langage, musique », p. 73-83, in *Perception auditive et compréhension du langage : état initial, état stable et pathologie*, J. Lambert, J.L. Nespoulous, Marseille, Ed. Solal, 1997.

40. CHEVALLIER, Jean-Marc et BONFILS, Pierre, *Anatomie, Tome 3 : ORL*, Paris, Flammarion Médecine, 1998.

41. MATRAS, Jean-Jacques, *Le son*, Paris, Ed. Que sais-je?, 1982, p. 34.

42. HELMHOLTZ, Hermann Von, « Théorie physiologique de la musique fondée sur l'étude des sensations auditives », Paris, Masson, 1868 in *Le son*, Jean-Jacques Matras, Ed. Que sais-je?, Paris, 1982, p. 34.

43. MAVLOV, L., « Amusia due to rythm agnosia in a musician with left hemisphere damage : a non auditory supramodal defect », *Cortex*, Vol. 16, 1980, p. 331-338.

44. LECHEVALIER, B., PLATEL, H. et EUSTACHE, F., « Neurophysiologie de l'identification musicale », *Revue Neurologique, vol. 151*, 1995, p. 505-510.

45. PERETZ, I., « Processing of local and global musical information by unilateral braindamaged patients », *Brain*, Vol. 113, 1990, p. 1185-1205.

46. WISE, R., CHOLLET, F., HADAR, U., FRISTON, K., HOFFNER, E. et FRACKOWIAK, RSJ., « Distribution of cortical neural networks involved in word comprehension and word retrieval », *Brain*, Vol. 114, 1991, p. 1803-1817.

47. ZATORRE, R.J., EVANS A.C., METTER, E. et GJEDDE, A., « Lateralisation of phonetic and pitch discrimination in pitch processing », *Science*, Vol. 256, 1992, p. 846-848.

48. CAMBIER, J., ELGHOZI, D., SIGNORET, J.L. et HENIN, D., « Contribution de l'hémisphère droit au langage des aphasiques, Disparition de ce langage après lésion droite », *Revue Neurologique (Paris), Vol. 139,* 1983, p. 55-63.

49. BELIN, Pascal, « Hearing voices : the cognitive neuroscience of human voice perception », Tours, Symposium IRIA, 2009.

50. HANNEQUIN, D., GOULET, P. et JOANETTE, Y., *Hémisphère droit et langage*, Paris, Masson, 1987.

51. LIBERMAN, A.M. et MATTINGLY, I.G., « The Motor Theory of Speech Perception Revised », *Cognition*, 21, 1985, p. 1-36.

52. OHALA, J.J., « Against the direct realist view of speech perception », *Journal of Phonetics*, Vol. 14, 1986, p. 75-82.

53. SEGUI, J., « La perception du langage parlé : données et théories », in *Neuropsychologie. Perception auditive et compréhension du langage : état initial, état stable et pathologie*, J. Lambert, J.L. Nespoulous, Marseille, Ed. Solal, 1997.

54. REVIS, J., *L'analyse perceptive des dysphonies : approche phonétique de l'évaluation vocale*, Thèse de doctorat, Université de la Méditerranée, Marseille, 2004.

55. LOMBARD, Etienne, « Le signe de l'élévation de la voix », *Annales des maladies de l'oreille et du larynx*, Vol. 37, n° 2, 1911, p. 101-119.

56. ALLEN, Jont B, « Harvey Fletcher's role in the creation of communication acoustics », *J. Acoust. Soc. Am.*, Vol. 99, n° 4, 1996, p. 1826-1839.

57. LANE, H., et TRANEL B, « The Lombard sign and the role of hearing in speech », *Speech and Hearing*, Vol. 14, 1971, p. 677-709.

58. GARDE E.J., « Le réflexe cochléo-phonatoire : état actuel de la question après quelques faits expérimentaux et pathologiques », *Revue Lar. Otol. Rhonol.*, Vol. 86, 1965, p. 631-638.

59. LANE H, et SISSON C, « Regulation of voice communication by sensory dynamics », *J. Acoust. Soc. Am.*, Vol. 47, n° 2, 1970, p. 618-624.

60. HUSSON, R., *Le chant*, Paris, Presses Universitaires de France, 1962.

61. VINTER, Shirley, « L'émergence du langage de l'enfant déficient auditif : Des premiers sons aux premiers mots », Collection d'Orthophonie, Ed. Masson, Paris, 1994.

62. LOCKE J.L. et PEARSON D.M., « Vocal learning and the emergence of phonological capacity. A neurobiological approach » in *Phonological development : models, research and implications*, Ferguson, C., Menn, L., Stoel-Gammon, C. eds., Maryland, Timomium, 1992, p. 91-129.

63. DEGGOUJ, Naima et DEVRESSE, Pierre-Yves, « Comprendre les troubles de l'audition », *La Revue de la Médecine Générale*, n° 231, Mars, 2006, p. 112.

64. LINA-GRENADE, G., TRUY E., POROT, M., COLLET, L. et DISANT, F., « Surdité de l'enfant : un diagnostic précoce est impératif », *Pédiatrie au quotidien*, Vol. 7, 2000, p. 991.

65. FRANÇOIS, M., « Le dépistage des surdités de perception chez l'enfant », in *Les surdités de perception*, Collection ORL, A. ROBIER, Paris, Ed. Masson, 2001, p. 23-32.

66. DUMONT A, *L'orthophoniste et l'enfant sourd*, Paris, Ed. Masson, 1988, p. 163.

67. VAN LEM, N., et PEAULT, D, *La surdité chez l'enfant,* Paris, Coll. Médecine actuelle Bibliogr. Index Le Sycomore, 1984, p. 200, ISBN 5818687.

68. DOWNS, M. et YOSHINAGA-ITANO, C., « The efficacy of early identification for children with hearing impairment », *Pediatric Clinics of North America*, Vol. 46, 1999, p. 79-87.

69. YOSHINAGA-ITANO, C., SEDEY, A.L., COULTER, D.K., MEHL, A.L., « Language of early – and later – identified children with hearing loss », *Pediatrics*, Vol. 102, 1998, p. 1161-1171.

70. DAUMAN, R., ROUSSEY, M. et GARABEDIAN, N., « La surdité permanente de l'enfant relève-t-elle du dépistage néonatal ? », *Pratiques et Organisation des soins*, Vol. 40, 2009, n° 3, p. 207-212.

71. PARVING, Agnete, « Hearing screening – aspects of epidemiology and identification of hearing impaired children », *Int J Pediatr Otorhinolaryngol*, Vol. 49, Suppl. 1, 1999, p. 287-292.

72. FRANÇOIS, M., « Dépistage des surdités néonatales », *Archives de pédiatrie*, Vol. 10, Suppl. 1, 2003, p. 152-154.

73. DENOYELLE, F, « Surdités de Perception d'origine génétique », in *Les surdités de perception*, Collection ORL, A. ROBIER, Paris, Ed. Masson, 2001, p. 33-40.

74. MORZARIA, S., WESTERBERG, B. et KOZAK, F., « Systematic review of the etiology of bilateral sensorineural hearing loss in children », *Int J Pediatr Otorhinolaryngol*, Vol. 68, 2004, p. 1193-1198.

75. OUELLET, C., LE NORMAND, M.T. et Cohen, H., « Language evolution in children with
cochlear implants », *Brain and Cognition*, Vol. 46, n° 1 et 2, 2001, p. 231-235.

76. LITOVSKY, R.Y. et col., « Bilateral cochlear implants in adults and children », *Arch Otolaryngol Head Neck Surg*, 2004, Vol. 130, p. 648-655.

77. ITO, K., MOMOSE T., OKU, S., ISHIMOTO, S., YAMASOBA, T., SUGASAWA, M. et KAGA K., « Cortical activation shortly after cochlear implantation », *Audiol Neurootol*, Vol. 9, n° 5, 2004, p. 282-293.

78. GARABEDIAN, E.N., LOUNDON, N., MONDAIN, M., PIRON, J.P., ROMAN, S. et TRIGLIA, J.M., « Les implants cochléaires pédiatriques », *Ann Otolaryngol Chir Cervicofac*, Vol. 120, n° 3, 2003, p. 139-151.

79. COCHARD, N. et col., « Développement des différents systèmes linguistiques chez les enfants porteurs d'un implant cochléaire », *Rev Laryngol Otol Rhinol*, Vol. 119, n° 4, 1998, p. 277-279.

80. GILBERT, H.R. et CAMPBELL, M.L., « Speaking fundamental frequency in three groups of hearing-impaired individuals », *Journal of Communication Disorders*, Vol. 13, 1980, p. 195-205.

81. HOCEVAR-BOLTEZAR, Irena et col., « Change of phonation control after cochlear implantation », *Otology and Neurotology*, Vol. 27, n° 4, 2006, p. 499-503.

82. HIGGINS, M.B., MCCLEARY, E.A., CARNEY, A.E., SCHULTE, L., « Longitudinal changes in children's speech and voice physiology after cochlear implantation », *Ear-and-hearing*, Vol. 24, n° 1, 2003, p.48-70. ISSN: 0196-0202.

83. LEPOT-FROMENT, C., et CLEREBAUT, N., *L'enfant sourd : communication et langage*, De Boeck Université, 1996.

84. STOEL-GAMMONT, C., "The acquisition of segmental phonology by normal and hearing-impaired children", 1983, In I. Hochberg, H. Levitt, M.J. Osberger eds., *Speech of the hearing impaired*, Baltimore, University Park Press, 1988.

85. STOEL-GAMMONT, C., "Prelinguistic vocalizations of hearing-impaired and normally hearing subjects : A comparison of consonantal inventories", *Journal of Speech and Hearing Disorders*, Vol. 53, 1988, p. 302-315.

86. STOEL-GAMMONT, C., *Premeaningful vocalizations of hearing-impaired and normally-hearing subjects*, Actes du XIIème Congrès International des Sciences Phonétiques, Aix-en-Provence, 1991, p.154 – 157.

87. STOEL-GAMMONT, C. et OTOMO, K., "Babbling development of hearing-impaired and normally hearing subjects", *Journal of Speech and Hearing Disorders*, Vol. 51, 1986, p. 33–41.

88. HIGGINS, M.B., MCCLEARY, E.A. et SCHULTE, L., "Articulatory changes with short-term deactivation of the cochlear implants of two prelingually deafened children", *Ear-and-hearing*, Vol. 22, n° 1, 2001, p. 29-45.

89. SEIFERT, E., OSWALD, M., BRUNS, U., VISCHER, M., KOMPIS, M. et HAEUSLER, R., "Changes of voice and articulation in children with cochlear implants", *Int J Pediatr Otorhinolaryngol*, Vol. 66, n° 2, 2002, p. 115 -123. IS-ISSN: 0165-5876.

90. MOUNIER-KUHN, P., LAFON, J.C., ROSSI, C., « La voix du jeune enfant sourd », *Journal Français d'Oto-Rhino-Laryngologie*, Vol. 16, n° 2, 1967, p. 121–123.

91. MANOLSON, A., *Comparative study of intonation patterns in normal hearing and hearing impaired infants,* Actes du 7ème Congrès International des Sciences Phonétiques, La Haye, Mouton, 1972.

92. VINTER, Shirley, « Appareillage et vocalisations », *Rééducation Orthophonique*, Vol. 27, n° 157, 1989, p.17-25.

93. CAMPISI, P, et col., "Acoustic analysis of the voice in pediatric cochlear implant recipients : A longitudinal study", *The Laryngoscope*, Vol. 115, 2005, p. 1046-1050.

94. PERIER, O., « L'enfant à audition déficiente. Aspects médicaux, éducatifs, sociologiques et psychologiques », *Acta Oto-Rhino-Laryngologica Belgica*, Vol. 41, 1987, p. 125-420.

95. VINTER, Sylvie, « Contrôle des premières productions vocales du bébé sourd », *Bulletin d'Audiophonologie. Annales Scientifiques de l'Université de Franche-Comté*, Vol. 3NS, n° 6, 1987, p. 659–670.

96. BIZAGUET, E., *La perspective de l'audioprothèse,* Colloque International d'Audiophonologie « Clinique et Recherche en Audiophonologie », Besançon, 1992.

97. KENT, R., OBERGER, M.J., NETSELL, R. et HUSTEDDE, C.G., "Phonetic development in identical twins dg in audition function", *Journal of Speech and Hearing Disorders*, Vol. 52, 1987, p. 64–75.

98. PORTMANN, M., « Les prothèses acoustiques implantées : état actuel et perspectives d'avenir », *Bulletin d'Audiophonologie. Annales Scientifiques de l'Université de Franche-Comté*, Vol. 9, 1992, p. 371–382.

99. DEJEAN, Y., UZIEL, A., REUILLARD, F., « L'implantation cochléaire chez le jeune enfant un an après », *Bulletin d'Audiophonologie. Annales Scientifiques de l'Université de Franche-Comté*, Vol. 9, 1992, p. 401 – 412.

100. HEATH A., *Cochlear implants, A collection of experiences of users of all ages,* High Wycombe (GB), National Cochlear Implant Users Association (NCIUA), , 2001, IB ISBN: 0954125703.

101. JAMES, A.L., et PAPSIN, B.C., **"**Cochlear implant surgery at 12 months of age or younger", *Laryngoscope*, Vol. 114, n° 12, 2004, p. 2191-2195.

102. SPENCER, L.J., GANTZ, B.J. et KNUTSON, J.F., "Outcomes and achievement of students who grew up with access to cochlear implants", *Laryngoscope,* Vol. 114, n° 9, 2004, p.1576-1581.

103. MANRIQUE M., CERVERA-PAZ, F.J., HUARTE, A. et MOLINA, M., Advantages of cochlear implantation in prelingual deaf children before 2 years of age when compared with later implantation, *Laryngoscope,* 2004 Vol. 114, n° 8, p. 1462-1469.

104. LEE, H.J., KANG, F., OH, S.H., KANG, H., LEE, D.S., LEE, M.C. et KIM, C.S., "Preoperative differences of cerebral metabolism relate to the outcome of cochlear implants in congenitally deaf children", *Hear Res*, Vol. 203, 2005, n° 1-2, p. 2-9.

105. LESINSKI-SCHIEDAT, A., ILLG, A., HEERMANN, R., BERTRAM B. et LENARZ T., “Paediatric cochlear implantation in the first and the second year of life : A comparative study”, *Cochlear Implants International*, Vol. 5, n° 4, 2004, p. 146–159.

106. LESINSKI-SCHIEDAT, A., ILLG, A., HEERMANN, R., BERTRAM, B. et LENARZ, T., “Paediatric cochlear implantation in the first year of life : Preliminary results”, *HNO*, Vol. 54, n° 7, 2006, p. 565-572.

107. HAENSEL, J., ENGELKE, J.C., OTTENJANN, W. et WESTHOFEN, M., **“**Long-term results of cochlear implantation in children”, *Otolaryngol Head Neck Surg*, Vol. 132, n° 3, 2005, p. 456-458.

108. COLLETTI, V., CARNER, M., MIORELLI, V., GUIDA, M., COLLETTI, L. et FIORINO, F.G., “Cochlear implantation at under 12 months : report on 10 patients”, *Laryngoscope*, Vol. 115, n° 3, 2005, p. 445-449.

109. HARRISON, R.V., GORDON, K.A. et MOUNT, R.J., “Is There a Critical Period for Cochlear Implantation in Congenitally Deaf Children? Analyses of Hearing and Speech Perception Performance after Implantation”, *Developmental-Psychobiology*, Vol. 46, n° 3, 2005, p. 252-261.

110. FLIPSEN, Peter Jr. et COLVARD, Lana G., “ Intelligibility of conversational speech produced by children with cochlear implants ”, *Journal of Communication Disorders*, Vol. 39, 2006, p. 93-108.

111. HEID, S., HARTMANN, R. et KLINKE, R., “A model for prelingual deafness, the congenitally deaf white cat-population statistics and degenerative changes”, *Hear Res*, Vol. 115, n° 1-2, 1998, p. 101-112.

112. SUMMERFIELD, Q., et TOMLINSON, J., *The use of cochlear implantation, Trent Inst. for Health Services Research*, Sheffield (GB), 1996. IB ISBN: 190073304.8

113. LANE, H. et GRODIN, M., “Ethical issues in cochlear implant surgery : an exploration into disease, disability, and the best interests of the child”, *Kennedy-Institute-of-Ethics-Journal*, Vol. 7, 1997, p. 231-251.

114. BARTON, G.R., BLOOR, K.E., MARSHALL, D.H. et SUMMERFIELD, A.Q., “Health service costs of paediatric cochlear implantation: influence of the scale and scope of activity”, *Int J Audiol*, Vol. 43, n° 7, 2004, p. 369-376.

115. SACH, T.H., WHYNES, D.K., ARCHBOLD, S.M. et O’DONOGHUE, G.M., “Estimating time and out-of-pocket costs incurred by families attending a pediatric cochlear implant programme”, *Int. J. Pediatr. Otorhinolaryngol*., Vol. 69, 2005, p. 929-936.

116. DUMONT, A., « Attention et mémorisation chez les sujets sourds implantes cochléaires », *GLOSSA*, Vol. 39, 1994, p. 26-28. ISSN: 0298-6477.

117. YANG, H.M., LIN, C.Y., CHEN, Y.J. et WU, J.L., « The auditory performance in children using cochlear implants : effects of mental function », *Int J Otorhinolaryngol*., Vol. 68, n° 9, 2004, p.1185-1188.

118. HANLEY, T.D. et STEER, M.D, « Effect of distracting noise upon speaking rate, duration, and intensity », *J Speech Hear Disord*, Vol. 14, 1949, p. 363-368.

119. DRAEGERT, G.L, « Relationships between voice variables and speech intelligibility in high level noise », *Speech Monograph*, Vol. 18, 1951, p. 272–278.

120. DREHER, J. et O'NEILL, J, "Effects of ambient noise on speaker intelligibility for words and phrases", *J. Acoust. Soc. Am.*, Vol. 29, 1957, p. 1320-1323.

121. LADEFOGED, P, *Three areas of experimental phonetics*, London, Oxford University Press, 1967.

122. SUMMERS, W.V., PISONI, D.B., BERNACKI, R.H., PEDLOW, R.I. et STOKES, M.A., "Effects of noise on speech production : Acoustic and perceptual analyses", *J Acoust Soc Am*, Vol. 84, n° 3, 1988.

123. HAGE, S.R., JÜRGENS, U. et EHRET, G., « Audio-vocal interaction in the pontine brainstem during self-initiated vocalization in the squirrel monkey », *Eur J Neurosci*, Vol. 23, n° 12, 2006, p. 3297-308.

124. LANE, H. et TRANEL, B, « The Lombard sign and the role of hearing in speech », *J Speech Hear Res*, Vol. 14, 1971, p. 677-709.

125. JÜRGENS, U., « The neural control of vocalization in mammals : a review », *J Voice*, Vol. 23, n° 1, 2009, p. 1-10.

126. LEVELT, W.J.M., Speaking : From intention to articulation, Cambridge, MA, MIT Press, 1989.

127. LACKNER, J. R., et TULLER, B. H., "The role of efference monitoring in the detection of self-produced speech errors", in *Sentence processing: Psycholinguistic studies presented*, W. E. Cooper et E. C. T. Walker (Eds.), *Merrill Garrell*, Hillsdale, NJ, Erlbaum, 1979, p. 281-294.

128. HOCEVAR-BOLTEZAR, I., VATOVEC, J., GROS, A. et ZARGI, M., "The influence of cochlear implantation on some voice parameters", *Int J Pediatr Otorhinolaryngol*, Vol. 69, 2005, p. 1635-1640.

129. HIGGINS, M.B., CARNEY, A.E., MCCLEARY, E.A., ROGERS, S., "Negative intraoral air pressures of deaf children with cochlear implants : Physiology, phonology and treatment", *Journal of Speech and Hearing Research*, Vol. 39, 1996, p. 957-967.

130. HIGGINS, M.B., MCCLEARY E.A. et SHULTE, L., "Use of visual feedback to treat negative intraoral air pressures of preschoolers with cochlear implants", *American Journal of Speech Language Pathology*, Vol. 9, 2000, p. 21-35.

131. SVIRSKY, M.A., JONES, D., OSBERGER, M.J. et MIYAMOTO, R.T., "The effect of auditory feedback on the control of oral-nasal balance by pediatric cochlear implant users", *Ear-and-hearing*, Vol. 19, 1998, p. 385-393.

132. REVIS, Joana, *L'analyse perceptive des dysphonies: approche phonétique de l'évaluation vocale*, Thèse de doctorat, Université de la Méditerranée, Marseille, 2004.

133. PERRIN, C., BERGER-VACHON, A., TOPOUZKHANIAN, A., TRUY, E. et MORGON, A., « Evaluation of cochlear implanted children's voices », *International Journal of Pediatric Otorhinolaryngology*, n° 47, 1999, p. 181-186.

134. LE NORMAND, Marie-Thérèse, « Evaluation de ma production spontanée du langage oral et de l'activité sémantique du récit chez l'enfant d'âge préscolaire », *Donnes Actuelles, Rééducation Orthophonique*, n° 231, 2007, p. 53-71.

135. GIOVANNI, Antoine, *Analyse objective des dysphonies avec l'appareillage EVA : études préliminaires à la définition d'un protocole normalisé*, DEA Langage et Parole, Université de Provence, 1995.

136. GIOVANNI, A., ROBERT, D., ESTUBLIER, N., TESTON, B., ZANARET, M. et CANNONI, M., « Objective evaluation of dysphonia : preliminary results of a device allowing simultaneous acoustic and aerodynamic measurements », *Folia Phoniatr Logop*, n° 48, 1996, p. 175-185.

137. DE KROM, G., « Consistency and reliability of voice quality ratings for different types of speech fragments », *J Speech Hear Res*, n° 37, 1994, p. 985-1000.

138. HAMMARBERG, B, et cols., "Perceptual and acoustic correlates of abnormal voice qualities", *Acta Otolaryngol*, n° 90, 1980, p. 441-451.

139. WOLFE, V., FITCH, J., CORNELL, R., « Acoustic prediction of severity in commonly occurring voice problems », *J Speech Hear Res*, n° 38, 1995, p. 273-279.

140. HARTL, D.M., HANS, S., CREVIER-BUCHMAN, L., LACCOURREYE, O., VAISSIERE, J. et BRASNU, D., « Méthodes actuelles d'évaluation des dysphonies », *Ann Otolaryngol Chir Cervicofac*, Vol. 122, n° 4, 2005, p. 163-172.

141. YU, Ping, *Méthodes instrumentales d'analyse de la dysphonie : Corrélation avec l'analyse perceptive*, Thèse de doctorat, Université de la Méditerranée, 2001.

142. GHIO, Alain, « Evaluation acoustique », in *Les dysarthries.* P. Auzou, V. Rolland, S. Pinto, C. Ozsancak éds., Marseille, Solal, 2007, p. 236-247.

143. YAN, Yuling, AHMAD, Kartini, KUNDUK, Melda et BLESS, Diane, "Analysis of Vocal-Fold vibrations from High-Speed Laryngeal Images using a Hilbert Transform-Based Methodology", *Journal of Voice*, Vol. 19, n° 2, 2005, p. 161-175.

144. YUMOTO, E. et GOULD, W, « Harmonics-to-noise ratio as an index of the degree of hoarseness », *J Acoust Soc Am*, n° 71, 1982, p. 1544-1550.

145. PICCIRILLO, J., COLIN, P., DENNIS, F. et FREDERICKSON, J., « Multivariate analysis of objective vocal function », *Ann Otol Rhinol Laryngol*, n° 107, 1998, p. 107-112.

146. GIOVANNI, Antoine, *Contribution à l'élaboration d'un protocole multiparamétrique d'évaluation objective des dysphonies*, Thèse en Science, Université de la Méditerranée, 1998.

147. GORDON, M., MORTON, F., SIMPSON, I., “Air flow measurements in diagnosis, assesment and treatment of mechanical dysphonia”, *Folia Phoniatr*, n° 30, 1978, p. 161-174.

148. LENDEN, J.M. et FLIPSEN, P. Jr., “Prosody and voice characteristics of children with cochlear implants”, *Journal of Communication Disorders*, Vol. 40, 2007, p. 66-81.

149. PETERSON, G.E., et BARNEY, H.L., “Control methods used in a study of the vowels”, *J Acoust Soc Am*, Vol. 24, 1952, p. 175-184.

150. SCHAUWERS, K., GILLIS, S., DAEMERS, K., DE BEUKELAER, C., et GOVAERTS P.J., “Cochlear implantation between 5 and 20 months of age : the onset of babbling and the audiologic outcome”, *Otol Neurotol*, Vol. 25, n° 3, 2004, p. 263-270.

151. HAMZAVI, J., DEUTSCH, W., BAUMGARTNER, W.D., DENK, D.M., ADUNKA ,O. et GSTOETTNER, W., “Cochlear implantation und Auditive Rückkopplung”, *Wien. Klin. Wochenschr*, Vol. 112, 2000, p. 515-518.

152. SZYFTER, A., PRUSZEWICZ, A., WOZNICA, B., SWIDZINSKI, E., SZYMIEC, E., KARLIK, M., “The acoustic analysis of voice in patients with multi-channel cochlear implant”, *Rev. Laryngol. Otol. Rhinol.*, Vol. 117, 1996, p. 225-227.

153. LEDER, S.B., SPITZER, J.B. et KIRCHNER, J.C., “Immediate effects of cochlear implantation on voice quality”, *Arch. Otorhinolaryngol.*, Vol. 244, 1987, p. 93-95.

154. KIRK, I.K. et EDGERTON, B.J., “The effects of cochlear implant use on voice parameters”, *Otolaryngol., Clin. Noth. Am.*, Vol. 16, 1983, p. 281-292.

155. PERRIN, M.E., BERGER-VACHON, C., LEDISSEZ, C., KAUFFMANN, I. et MORGON, A., “The normality of the voice of cochlear implant children”, *Adv. Otorhinolaryngol.*, Vol. 50, 1995, p. 167-173.

156. MONINI, S., BANCI G., BARBARA, M., ARGIRO, M.T. et FILIPO, R., “Clarion cochlear implant : short term effects on voice parameters”, *Am. J. Otol.*, Vol. 18, 1997, p. 719-725.

157. TITZE, I., “Toward standards in acoustic analysis of voice”, *J Voice*, Vol. 8, 1994, p. 1-7.

158. VAN LIERDE, Kristiane M., BART, M. Vinck, BAUDONCK, Nele, DE VEL Eddy et DHOOGE, Ingeborg, “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*, Vol. 44, n° 8, 2005, p. 452-465.

159. DALLEAS, B., et DALLEAS, A., « Qualités acoustiques de la voix », *Revue de Laryngologie*, Suppl., 1987, p. 373-377.

160. FINNEGAN, Denis E., “Maximum Phonation Time for children with Normal Voices”, *Folia Phoniat.*, Vol. 37, 1985, p. 209-215.

161. SMITHERAN, J.R., et HIXON, T.J., “A clinical method for estimating laryngeal airway resistance during vowel production”, *J Speech Hear Dis*, n° 46, 1981, p. 138-146.

162. YU, Ping, OUAKNINE, Maurice, REVIS, Joana et GIOVANNI, Antoine, « Objective Voice Analysis for Dysphonic Patients: A Multiparametric Protocol Including Acoustic and Aerodynamic Measurements », *Journal of Voice*, n° 4, vol. 15, 2001, p.529-542.

163. ARENDS, N., POVEL, D., VAN OS, E. et SPETH ,L., “Predicting voice quality of deaf speakers on the basis of glottal characteristics”, *Journal of Speech and Hearing Research*, Vol. 33, 1990, p. 116-122.

164. DAGENAIS, P.A. et CRITZ-COSBY, P., “Consonant lingual-palatal contacts produced by normal-hearing and hearing-impaired children”, *Journal of Speech and Hearing Research*, Vol. 34, 1991, p. 1423-1435.

165. DAGENAIS, P.A. et CRITZ-COSBY, P., “Comparing tongue positioning by normal-hearing and hearing-impaired children during vowel production”, *Journal of Speech and Hearing Research*, Vol. 35, 1992, p. 35-44.

166. LEEPER, H.A., PEREZ, D.M. et MENCKE E.O., “The influence of utterance length upon bilabial closure duration of selected deaf children”, *Journal of Communication Disorders*, Vol. 13, 1980, p. 373-383.

167. SVIRSKY, M.A., et CHIN, S.B., “Speech production”, In *Cochlear implants*, S.B. Waltzman et N.L. Cohen eds., New York, Thieme, 2000, p. 293-309.

168. GUERRIER, Bernard, GIOVANNI, Antoine et REMACLES, Marc, *Pathologie de la corde vocale de l’adulte*, Société Française d’Oto-laryngologie et de Chirurgie de la Face et du Cou, 2004, p. 97.

169. SMITHERAN, J.R., et HIXON, T.J., “A clinical method for estimating laryngeal airway resistance during vowel production”, *J Speech Hear Dis*, Vol. 46, 1981, p. 138-146.

170. KREIMAN, J. et GERATT, B., « The perceptual structure of pathologic voice quality », *J Acous Soc Am*, n° 100, 1996, 1787-1795.

171. DE BODT, M., WUYTS, F., VAN DE HEYNING, P. et CROUX, C., « Test-retest study of the GRBAS scale : influence of experience and professional background on perceptual ratings of voice quality », *J Voice*, n° 1, 1997, p. 74-80.

172. GIOVANNI, A., REVIS, J., TRIGLIA, J.M., « Objective aerodynamic and acoustic measurement of voice improvement after phonosurgery », *Laryngoscope*, n° 4, Vol. 109, 1999, p. 656-60.

173. YU, P., REVIS, J., WUYTS, F., ZANARET, M. et GIOVANNI, A., « Correlation of instrumental voice evaluation with perceptual voice analysis using a modified visual analog scale », *Folia Phoniatr Logop*, n° 6, Vol. 54, 2002, p. 274-281.

174. KREIMAN, J., GERRATT, B.R., PRECODA, K. et BERKE, G.S., « Individual differences in voice quality perception », *J Speech Hear Res*, n° 35, 1992, p. 512-520

175. GOLDSTONE, R.L., « Perceptual learning », *Ann Rev Psychol*, n° 49, 1998, p. 585-612.

176. ANDERS, L., HOLLIEN, H., HURME, P., SONNINNEN, A. et WENDLER, J., « Perceptual evaluation of hoarseness by several classes of listeners », *Folia Phoniatr*, n° 40, 1988, p. 91-100.

177. DEJONCKERE, P., OBBENS, C., DE MOOR, G.M. et WIENEKE, G.H., « Perceptual evaluation of dysphonia : reliability and relevance », *Folia Phoniatr*, n° 45, 1993, p. 76-83.

178. DROESBEKE, J.J., LEJEUNE, M. et SAPORTA, G., *Modèles statistiques pour données qualitatives*, Paris, Ed. Technip, 2005.

179. BOONE, D.R., « Modification of the voices of deaf children », *The Volta Review*, Vol. 68, 1966, p. 686-694.

180. STATHOPOULOS, E., DUCHAN, J., SONNENMEIER, R., et BRUCE, N., « Intonation and pausing in deaf speech », *Folia Phoniatrica*, Vol. 38, 1986, p. 1-12.

181. MONSEN, R., « Toward measuring how well hearing-impaired children speak », *Journal of Speech and Hearing Research*, Vol. 21, 1978, p. 197-219.

182. PARKHURST, B, et LEVITT, H, "The effect of selected prosodic errors on the intelligibility of deaf speech", *Journal of Communication Disorders*, Vol. 11, 1978, p. 249-256.

183. SMITH, C., « Residual hearing and speech produced in deaf children », *Journal of Speech and Hearing Research*, Vol. 18, 1975, p. 795-811.

184. SUBTELNY, J., WHITEHEAD, R. et ORLANDO, N., « Description and evaluation of an instructional program to improve speech and voice diagnosis of the hearing impaired », *The Volta Review*, Vol. 82, 1980, p. 85-95.

185. HIRANO, M., HIBI, S., YOSHIDA, T., HIRADE, Y., KASUYA, H. et KIKUCHI, Y., "Acoustic analysis of pathological voice. Some results of clinical application", *Acta Otolaryngol*, Vol. 105, 1988, p. 432-438.

186. DEJONCKERE, P., « Bruit de turbulence et apériodicité dans la voix pathologique. Une approche
Multifactorielle », *Rev Laryngol*, Vol. 111, 1990, p. 353-357.

187. HILLENBRAND, J., CLEVELAND, R.A. et ERICKSON, R.L., "Acoustic correlates of breathy vocal quality", *J Speech Hear Res*, Vol. 37, 1994, p. 769-778.

188. ALGOUD, J.P., *Systémique : vie et mort de la civilisation occidentale, Vol. 1 « La Désintégration », Vol. 2 « L'Extinction »*, Lyon, L'interdisciplinaire, 2001.

189. GHIO, Alain, « Modélisation du conduit vocal », in *Les dysarthries.* P. Auzou, V. Rolland, S. Pinto, C. Ozsancak éds., Marseille, Solal, 2007, p. 236-247.

190. FORRESTER, J., « Industrial dynamics : a major breakthrough for decision makers », *Harvard Business Review*, Vol. 36, n° 4, 1958, p. 37-66.

191. MÜRBE, D., PABST, F., HOFMANN, G. et SUNDBERG, J., « Significance of auditory and kinesthetic feedback to singers' pitch control », *J Voice,* Vol. 16, 2002, p. 44–51.

192. MÜRBE, D., PABST, F., HOFMANN, G. et SUNDBERG, J., « Effects of a professional solo singer education on auditory and kinesthetic feedback : a longitudinal study of singers' pitch control », *J Voice,* Vol. 18, n° 2, 2003, p. 236–241.

193. FANT, G., LILJENCRANTS J. et LIN, Q., « A four parameter model of the glottal flow », *STL-QPSR*, Vol. 85, n° 2, 1985, p. 1-13.

194. ISHIZAKA, K. et FLANAGAN, J. L., « Synthesis of Voiced Sounds form a two-mass model of the vocal cords », *Bell. Syst. Tech. J.*, Vol. 51, 1972, p. 1233-1268.

195. PELORSON, X, HIRSCHBERG, A, VAN HASSEL, R.R. et WIJNANDS, A.P.J., « Theoretical and experimental study of quasisteady-flow separation within the glottis during phonation. Application to a modified two-mass model », *J. Acoust. Soc. Am.*, Vol. 96, n° 6, 1994, p. 3416-3431.

196. SCIAMARELLA, D. et D'ALESSANDRO, C., *A study of the Two-Mass Model in terms of Acoustic parameters*, International Conference on Spoken Language Processing ICSLP, 2002, p. 2313-2316.

197. SCIAMARELLA, D. et D'ALESSANDRO, C., *Reproducing laryngeal mechanisms with a two-mass model,* Proc. of the 8th European Conference on Speech Communication and Technology, Eurospeech 2003, 2003, p. 2385-2388.

198. SCIAMARELLA, D. et D'ALESSANDRO, C., "On the acoustic sensitivity of a symmetrical two-mass model of the vocal folds to the variation of control parameters", *Acta Acustica,* Vol. 90, 2004, p. 746-761.

199. LILJENCRANTS, J., « Airflow visualizations in a model of the glottal passage », in *From speech signal to acoustic source,* Periodic Progress Report 2, European ESPRIT/BR 6975 : Speech Maps project, C. Shadle editor, 1995.

200. BOË, L.J., MAEDA, S. et PERRIER, P., *La modélisation articulatoire : un demi-siècle d'évolution entre fonctionnel, physique et biomécanique,* Actes des XXèmes Journées d'Etudes sur la Parole, Trégastel, France, 1994, p. 41-54.

201. STEVENS, K.N., KASOWSKI, S. et FANT, C.G.M., « An electrical analog of the vocal tract », 1953, in J. Flanagan et L. Rabiner (eds), Speech Synthesis, Dowden, Hutchinson, Ross, Inc. Stroudsburg, USA, 1973, p. 108-116.

202. FANT, G., *Acoustic theory of the speech production*, The Hague, Netherlands, Mouton, 2nd edition, 1960.

203. BOERSMA, P., « An articulatory synthesizer for the simulation of consonants », *Proceedings of Eurospeech*, Berlin, Germany, 1993, p. 1907-1910.

204. MAEDA, S., CARRÉ, R., « Modèles de production », in *Fondements et Perspectives en traitement automatique de la parole*, H. Meloni ed., Paris, Coll. Universités Francophones, AUPELF-UREF, 1996, p. 31-53.

205. MEDDIS, R, « Simulation of mechanical to neural transduction in the auditory recepter », *Journal of the Acoustical Society of America,* Vol.79, n° 3, 1986, p. 702-711.

206. HEWITT, M.J. et MEDDIS, R., « Implementation details of a computation model of the inner hair-cell/auditory-nerve synapse », *Journal of the Acoustical Society of America,* Vol.87, n° 4, 1990, p. 1813-1816.

207. PATTERSON, R.D., ROBINSON, K., HOLDSWORTH, J., McKEOWN, D., ZHANG, C., et ALLERHAND M., « Complex sounds and auditory images, Auditory physiology and perception », *Proceedings of the 9h International Symposium on Hearing,* Y. Cazals, L. Demany, K. Horner (eds), Pergamon, Oxford, 1992, p. 429-446.

208. PATTERSON, R.D., ALLERHAND, M., et GIGUERE C., « Time-domain modeling of peripheral auditory processing : A modular architecture and a software platform », *J. Acoust. Soc. Am.,* Vol. 98, 1995, p. 1890-1894.

209. PATTERSON, R.D., « Auditory images: How complex sounds are represented in the auditory system », *J Acoust Soc Japan (E), Vol.* 21, n° 4, 2000, p. 183-190.

210. SLANEY, M., « Lyon's Cochlear Model », *Apple Computer Technical Report,* n° 13, 1988.

211. SENEFF, S., « A joint synchrony/mean-rate model of auditory speech processing », *Journal of Phonetics,* Vol. 16, 1988, p. 55-76.

212. HOHMANN, V., « Frequency analysis and synthesis using a Gammatone filterbank », *Acta Acustica,* Vol. 88, 2002, p. 433 – 442.

213. PATTERSON, R.D, NIMMO-SMITH, J., HOLDSWORTH, J. Et RICE, P., *An efficient auditory filterbank based on the Gammatone function, IOC Speech Group on Auditory Modeling, RSRE,* December, 1987.

214. SLANEY, M., « An Efficient Implementation of the Patterson-Holdsworth Auditory Filter Bank », *Apple Computer Technical Report,* Perception Group, Advanced Technology Group, n° 35, 1993.

215. GLASBERG, B.R. et MOORE B.C., J., « Frequency selectivity as a function of level and frequency measured with uniformly exciting notched noise », *J. Acoust. Soc. Am.,* Vol. 108, 2000, p. 2318–2328.

216 PRESSNITZER, D. et GNANSIA, D., *Real time auditory model*, ICMC Conference, 1134, Barcelona, September, 2005.

217. GLASBERG, B.R. et MOORE, B.C.J., « Derivation of auditory filter shapes from notched noise data », *Hearing Research,* Vol. 47, 1990, p. 103–138.

218. GUERRERO, Harold, SERRAFERO, Patrick et BARKAT-DEFRADAS, Melissa, *Pour l'insertion professionnel des BAC+8 en entreprise : exemple d'une expérience de gestion des projets industriels au cours d'une thèse*, Ecole d'été, L'université dans la société, Actes du 8e Colloque Européen UNiSO, Amiens, 2009.

219. GARREL, R., SCHERER. R., NICOLLAS, R., GIOVANNI, A. et OUAKNINE, M., « Using the relaxation oscillations principle for simple phonation modeling », *Journal of Voice*, Vol. 22, n° 4, 2006, p. 385-398.

220. HORACEK, J., SIDLOF, P., SVEC, J.G., "Numerical simulations of self-oscillations of human vocal folds with Hertz model of impact forces", *Journal of Fluids and Structures*, Vol. 20, 2005, p. 853-869.

221. HEIM, Jean-Louis, BOE, Louis-Jean et ABRY, Christian, « La parole à la portée du conduit vocal de l'Homme de Neandertal : Nouvelles recherches, nouvelles perspectives », *C.R. Palevol*, Vol. 1, 2002, p. 129-134.

222. HORACEK, J. et SVEC, J.G., "Aeroelastic model of vocal-fold-shaped vibrating element for studying the phonation threshold", *Journal of Fluids and Structures*, Vol. 16, n° 7, 2002, p. 931-955.

223. GARREL, R., SCHERER, R., NICOLLAS, R., GIOVANNI, A. et OUAKNINE, M., « Using the relaxation oscillations principle for simple phonation modeling », *Journal of Voice*, Vol. 22, n° 4, 2006, p. 385-398.

224. GUERTI, Mhania, *Cibles résonantiques des voyelles orales*, 3nd International Conference : Sciences of Electronic, SETIT 2005, Technologies of Information and Telecommunications, Tunisia, 2005.

225. ESKENAZI, L., CHILDERS, DG. et HICKS DM., "Acoustic correlates of vocal quality", *J Speech Hear Res*, Vol. 33, 1990, p. 298-306.

226. MARTIN, D., FITCH, J. et WOLFE, V., "Pathological voice type and the acoustic prediction of severity", *J Speech Hear Res*, Vol. 38, 1995, p. 765-771.

227. KREIMAN, J., GERRATT, B.R., KEMPSTER, G.B., ERMAN, A. et BERKE, G.S., "Perceptual evaluation of voice quality: review, tutorial and a framework for future research", *J Speech Hear Res*, Vol. 36, 1993, p. 21-40.

228. REVIS, Joana, GHIO, Alain et GIOVANNI, Antoine, *Phonetic labeling of dysphonia : a new perspective in perceptual voice analysis*, Proceedings of the 7th International Conference Advances in Quantitative Laryngology, Voice and Speech Research, October, 2006.

TABLE DES FIGURES

Figure 1 : Place de la voix dans la communication orale humaine. 22
Figure 2 : Le larynx et les plis vocaux dessinés par Léonard De Vinci. 27
Figure 3 : Vue de profil de la posture debout dite « commode ». 32
Figure 4 : Schéma des étages de l'appareil phonatoire. 34
Figure 5 : Mouvement d'inspiration. 35
Figure 6 : Muscles expiratoires accessoires (d'après Le Huche, 2001, p. 38, 47). 36
Figure 7 : Vue supérieure antéro-postérieure des cartilages du larynx. 38
Figure 8 : Vue de profil des muscles extrinsèques responsables du maintien de la position du larynx dans le cou (d'après Bonfils et Chevallier, 1998[40]). 39
Figure 9 : Schéma de la structure « feuilletée » de la corde vocale proposée par Hirano (Coupe transversale) 41
Figure 10 : Configuration glottique d'un cycle vibratoire des cordes vocales. 42
Figure 11 : Etages constituant le pharynx et position de la cavité pharyngée dans le cou (d'après Le Huche, 2001, p.17[34]) 44
Figure 12 : Différence morphologique du larynx du nouveau-né et de l'adulte. 48
Figure 13 : Variation de position du plan glottique avec l'âge (d'après Fried, 1996[37]). 49
Figure 14 : Fo moyenne de la voix parlée en fonction de la longueur de la portion membraneuse des cordes vocales (d'après Titze, 1994, p.180). 50
Figure 15 : Schéma des trois compartiments qui constituent l'appareil auditif 53
Figure 16 : L'oreille moyenne et l'oreille interne, cochlée dépliée. 54
Figure 17 : Coupe d'un tour de la cochlée. 56
Figure 18 : Vue détaillée de l'organe de Corti. 56
Figure 19 : Schéma de l'organisation tonotopique de la membrane basilaire. 58
Figure 20 : Schéma simplifié des voies auditives centrales à quatre niveaux. 60
Figure 21 : Vue latérale du cerveau : aires ou zones corticales primaires et secondaires. 61
Figure 22 : Exemple d'une onde sonore simple. 64
Figure 23 : Exemple d'une onde sonore complexe et des ondes simples la constituant 64
Figure 24 : Prothèses auditives classiques. 78
Figure 25 : Prothèse auditive classique intra-conduit. 78
Figure 26 : Composants de l'implant cochléaire 80
Figure 27 : Le porte-électrodes de l'implant dans la cochlée. 81
Figure 28 : Répartition du groupe EE en fonction du sexe. 107
Figure 29 : Répartition de l'âge du groupe EE en fonction du sexe. 108
Figure 30 : Répartition du groupe ESA en fonction du sexe. 108
Figure 31 : Répartition de l'âge du groupe ESA en fonction du sexe 109
Figure 32 : Répartition du groupe ESIC en fonction du sexe. 109
Figure 33 : Répartition de l'âge du groupe ESIC en fonction du sexe. 110
Figure 34 : Répartition de la population selon l'âge d'appareillage. 112
Figure 35 : Répartition de la population selon la durée du port de l'appareillage. 113
Figure 36 : Répartition du groupe selon le modèle de l'implant cochléaire utilisé. 114
Figure 37 : Distribution du groupe selon la stratégie de codage de l'implant cochléaire. 114
Figure 38 : Répartition du groupe en fonction du côté implanté. 115
Figure 39 : Répartition du groupe en fonction du nombre d'électrodes actives de l'implant cochléaire. 115
Figure 40 : Répartition du groupe selon l'âge à l'implantation. 116
Figure 41 : Répartition du groupe selon la durée d'implantation. 116
Figure 42 : Distribution des groupes ESA et ESIC en fonction du degré de surdité. 117
Figure 43 : Distribution des groupes ESA et ESIC en fonction de la période de début de la surdité. 118

Figure 44 : Distribution des groupes ESA et ESIC en fonction de l'âge de diagnostic de la surdité. 118
Figure 45 : Distribution des groupes ESA et ESIC en fonction de l'étiologie de la surdité. ..119
Figure 46 : Contexte d'enregistrement de la voix d'un enfant.....120
Figure 47 : Fenêtre d'affichage de l'application « PROSODIE » du dispositif EVA™2 126
Figure 48 : Fenêtre d'affichage de l'application « PROFIL VOCAL » du dispositif EVA™2. 127
Figure 49 : Fenêtre d'affichage de l'application « PROSODIE » du dispositif EVA™2 pour l'analyse de la tessiture. 129
Figure 50 : Fenêtre d'affichage de l'application « PROSODIE » du dispositif EVA™2 pour l'analyse de l'étendue vocale. 130
Figure 51 : Diagramme des moyennes et graphe en boites des médianes de FoUm (Hz) de la voix conversationnelle pour les groupes EE, ESA et ESIC.....132
Figure 52 : Diagramme des moyennes et graphe en boites des médianes de la dynamique (Hz) de la voix conversationnelle pour les populations EE, ESA et ESIC.....133
Figure 53 : Diagramme des moyennes et graphe en boites des médianes de FoUm (Hz) de la voix forte pour les groupes EE, ESA et ESIC.....134
Figure 54 : Diagramme des moyennes et graphe en boites des médianes de la dynamique (Hz) de la voix forte pour les groupes EE, ESA et ESIC. 135
Figure 55 : Diagramme des moyennes et graphe en boites des médianes de FoUm (Hz) de la voix d'appel pour les groupes EE, ESA et ESIC. 136
Figure 56 : Diagramme des moyennes et graphe en boites des médianes de la dynamique (Hz) de la voix d'appel pour les groupes EE, ESA et ESIC.....136
Figure 57 : Diagramme des moyennes et graphe en boites des médianes de Fo (Hz) d'un /a/ tenu pour les groupes EE, ESA et ESIC. 139
Figure 58 : Diagramme des moyennes et graphe en boites des médianes du Coefficient de variation de Fo (%) d'un /a/ tenu pour les groupes EE, ESA et ESIC. 140
Figure 59 : Diagramme des moyennes et graphe en boites des médianes de Jitter factor (%) d'un /a/ tenu pour les populations EE, ESA et ESIC. 141
Figure 60 : Diagramme des moyennes et graphe en boites des médianes de l'intensité moyenne (dB) d'un /a/ tenu pour les populations EE, ESA et ESIC. 142
Figure 61 : Diagramme des moyennes et graphe en boites des médianes de Coefficient de variation de l'intensité moyenne (%) d'un /a/ tenu pour les groupes EE, ESA et ESIC. 142
Figure 62 : Diagramme des moyennes et graphe en boites des médianes de Shimmer factor (%) d'un /a/ tenu pour les groupes EE, ESA et ESIC. 143
Figure 63 : Diagramme des moyennes et graphe en boites des médianes de Rapport Harmoniques/Bruit (dB) d'un /a/ tenu pour les groupes EE, ESA et ESIC. 144
Figure 64 : Graphe en boite des médianes et graphe en ligne de l'étendue vocale (Hz) pour les groupes EE, ESA et ESIC.....145
Figure 65 : Graphes en boites des médianes et graphe en ligne de la tessiture vocale (Hz) pour les groupes EE, ESA et ESIC.....146
Figure 66 : Diagramme des moyennes et graphe en boites des médianes de Temps Maximal de Phonation (TMP en sec) pour les groupes EE, ESA et ESIC.....147
Figure 67 : Fenêtre d'affichage de l'application « EFFICACITE GLOTTIQUE » du dispositif EVA™2 pour l'analyse des paramètres aérodynamiques. 156
Figure 68 : L'enfant pendant l'enregistrement.....157
Figure 69 : Histogrammes de la distribution d'effectifs du groupe ESIC pour les différents paramètres aérodynamiques. 160
Figure 70 : Nuages de points et droite de régression dans les cas où le coefficient de corrélation linéaire est significatif dans le groupe ESIC. 165
Figure 71 : Échelle visuelle analogique. 169

Figure 72 : Répartition des individus selon le grade de dysphonie d'après l'évaluation perceptive et le groupe. 176
Figure 73 : Schéma d'un exemple de modèle économique de 1er ordre selon Forester (d'après Algoud, 2001). 189
Figure 74 : Schéma d'un système asservi : la chasse d'eau (d'après Algoud, 2001). 189
Figure 75 : Schéma représentant les composants d'un système élémentaire ou micro-système. 190
Figure 76 : Schéma simplifié de la disposition des micro-systèmes dans le macro-modèle proposé. 192
Figure 77 : Schéma représentant les différents paramètres et la structure géométrique du modèle symétrique à deux masses retenu pour notre simulateur 194
Figure 78 : Schémas des différentes configurations glottiques pouvant être représentées selon le modèle symétrique à deux masses. 196
Figure 79 : Architecture de l'algorithme utilisé pour le bloc « vibrateur » ou des cordes vocales. 197
Figure 80 : Répartition des forces de pression sur les masses. 198
Figure 81 : Modèle de Fant (1960) à quatre tubes simulant la section des lèvres (A1), une cavité frontale (A2), l'articulation linguale (A3) et une cavité arrière (A4) (d'après Fant , 1960, modifié). 200
Figure 82 : Courbe représentant la fonction de transfert obtenu pour un résonateur modélisé. En X, les fréquences (Hz); en Y, l'amplitude. 201
Figure 83 : Courbe représentant la fonction de transfert des différents tubes modélisés. 201
Figure 84 : Schéma de la structure du modèle du bloc « Récepteur » développé dans nos travaux. 206
Figure 85 : Schéma détaillé d'architecture du modèle systémique asservi représentant l'action de la boucle audio-phonatoire lors de l'émission d'un signal vocalique (voyelle /a/). 209
Figure 86 : Interface du logiciel de simulation de la boucle audio-phonatoire ARTIVox v4.0 214
Figure 87 : Fenêtre d'affichage des résultats de la simulation par le logiciel ARTIVox 4.0. ..216
Figure 88 : Interface du logiciel ARTIVox v4.0. : paramétrage du bloc Cochlée. 217
Figure 89 : Modèle à deux masses-ressort symétrique du bloc « cordes vocales ». 219
Figure 90 : Interface du logiciel ARTIVox v4.0 : paramétrage du bloc Cordes vocales. 220
Figure 91 : Interface du logiciel ARTIVox v4.0 : paramétrage du bloc Résonateur. 221
Figure 92 : Modèle du conduit vocal de Fant représentant le bloc « Résonateur ». 221
Figure 93 : Interface du logiciel ARTIVox v4.0 : paramétrage du bloc Initialisation. 222
Figure 94 : Interface du logiciel ARTIVox v4.0 : paramétrage du Paramètres aérodynamiques. 223
Figure 95 : Schéma de l'architecture du modèle systémique asservi ARTIVox v4.0 représentant l'action de la boucle audio-phonatoire lors de l'émission d'un signal vocalique (voyelle /a/). 225
Figure 96 : Modèle à deux masses-ressort symétrique du bloc « cordes vocales » 226
Figure 97 : Courbes de la liaison de k1 (à gauche) et k2 (à droite) avec le coefficient de comparaison CCTF. 227
Figure 98 : Courbes de la liaison de r1 (à gauche) et r2 (à droite) avec le coefficient de comparaison CCTF. 227
Figure 99 : Courbes de la liaison de l1 (à gauche) et l2 (à droite) avec le coefficient de comparaison CCTF. 228
Figure 100 : Courbe présentant la liaison de l avec le coefficient de comparaison CCTF. 229
Figure 101 : Courbes de la liaison de la profondeur (à gauche) et de la hauteur (à droite) avec les valeurs des coefficients de comparaison CCTF obtenues par la simulation numérique. ..229
Figure 102 : Courbes de la liaison de la pression (à gauche) et de la vitesse de l'air (à droite) avec le coefficient de comparaison CCTF. 230

Figure 103 : Courbes de la liaison de y1 (à gauche) et de y2 (à droite) avec les valeurs des coefficients de comparaison CCTF obtenues par simulation numérique. 231
Figure 104 : Exemple d'un bouclage : filtres du bloc Cochlée à 100%. 232
Figure 105 : Exemple d'un bouclage : filtres du bloc Cochlée à 100% (f≈295 Hz, /a/ tenu).236

LISTE DES TABLEAUX

Tableau 1 : Synopsis de muscles du larynx, leur fonction, innervation et points anatomiques d'origine et d'insertion. 39
Tableau 2 : Synopsis de muscles sous-hyoïdiens, leur fonction, innervation et points anatomiques d'origine et d'insertion. 40
Tableau 3 : Synopsis de muscles sus-hyoïdiens, leur fonction, innervation et points anatomiques d'origine et d'insertion. 40
Tableau 4 : Classification de la surdité selon le degré (d'après le BIAP). 73
Tableau 5 : Synopsis de quelques études réalisées sur la voix des enfants implantés cochléaires.....96
Tableau 6 : Modèles des prothèses auditives classiques portées par les enfants de l'étude. ..112
Tableau 7 : Processeurs vocaux des implants cochléaires portés par les enfants. 113
Tableau 8 : Fondamental Usuel Moyen (FoUm en Hz) de la voix conversationnelle pour les différents groupes. 132
Tableau 9 : Dynamique (Hz) de la voix conversationnelle pour les différents groupes. 133
Tableau 10 : Fondamental Usuel Moyen (FoUm en Hz) de la voix forte pour les différents groupes.....134
Tableau 11 : Dynamique (Hz) de la voix forte pour les différents groupes. 134
Tableau 12 : Fondamental Usuel Moyen (FoUm en Hz) de la voix d'appel pour les différents groupes.....135
Tableau 13 : Dynamique (Hz) de la voix d'appel pour les différents groupes.....136
Tableau 14 : Fréquence fondamentale ou Fo (Hz) d'un /a/ tenu pour les groupes.....139
Tableau 15 : Coefficient de variation de Fo (%) d'un /a/ tenu pour les trois populations. 140
Tableau 16 : Jitter factor (%) d'un /a/ tenu pour les trois groupes de population. 140
Tableau 17 : Intensité moyenne (dB) d'un /a/ tenu pour les trois populations.....141
Tableau 18 : Coefficient de variation de l'intensité moyenne (%) d'un /a/ tenu pour les trois populations. 142
Tableau 19 : Shimmer factor (%) d'un /a/ tenu pour les trois populations. 143
Tableau 20 : Rapport Harmoniques/Bruit (dB) d'un /a/ tenu pour les trois populations. 143
Tableau 21 : Dynamique (Hz) de l'étendue vocale pour les trois populations.....145
Tableau 22 : Dynamique (Hz) de la tessiture pour les trois populations.....146
Tableau 23 : Temps Maximal de Phonation (TMP en sec) pour les trois groupes de population. 146
Tableau 24 : Classement des paramètres acoustiques avant l'application des tests ANOVA. 150
Tableau 25 : Résultats du test ANOVA 1.....151
Tableau 26 : Résultats du test ANOVA 2.....152
Tableau 27 : ANOVA 1 et 2. 152
Tableau 28 : Classement des paramètres acoustiques semblables ou différents après l'application des tests ANOVA 1 et 2. 153
Tableau 29 : Valeur des principaux indices statistiques descriptifs pour les paramètres aérodynamiques du groupe ESIC (n=20). 159
Tableau 30 : Résumé des résultats du test de significativité du Coefficient de corrélation. ..164
Tableau 31 : Tableau contenant la définition des paramètres de l'échelle GRBAS. 170
Tableau 32 : Conversion des valeurs continues d'une échelle visuelle analogique en valeurs ordinales d'une échelle mixte à 4 niveaux. 171
Tableau 33 : Interprétation des valeurs du coefficient Kappa (d'après De Bodt, 1997). 172
Tableau 34 : Concordance intra-auditeur pour chacun des paramètres de l'échelle INGRBAS. 173
Tableau 35 : Concordance inter-auditeur pour chacun des paramètres de l'échelle 174

Tableau 36 : Valeurs moyennes des écart-types en fonction du grade global de la dysphonie. 175
Tableau 37: Valeurs des moyennes et des écart-types (entre parenthèses) de chacun des paramètres en fonction des grades perceptifs. 177
Tableau 38 : Valeurs des moyennes et des écart-types (entre parenthèses) des paramètres aérodynamiques pour le groupe ESIC en fonction des grades perceptifs. 178
Tableau 39 : Résultats du modèle de Régression logistique estimé pour l'ensemble de l'échantillon (deux premières lignes) et pour chacun des groupes. 179
Tableau 40 : Résultats du modèle de Régression logistique estimé pour le groupe ESIC à l'aide des paramètres aérodynamiques. 180
Tableau 41 : Tableau de l'évolution des versions produites du logiciel ARTIVox. 213
Tableau 42 : Tableau contenant les fréquences centrales (Hz) de chaque filtre modélisé dans le bloc « cochlée ». Les filtres 16 à 24 ne sont pas encore actifs dans la version ARTIVox v4.0. 218
Tableau 43 : Données physiologiques utilisées d'autres auteurs.....219
Tableau 44 : Valeurs d'entrée pour les différents blocs du modèle de simulation ARTIVox correspondant au cas de l'enfant entendant. 232
Tableau 45 : Valeurs du coefficient CCTF pour une tenue vocalique et un ton pur.....233
Tableau 46 : Valeurs du coefficient CCTF en fonction de la pondération des filtres du bloc Cochlée pour un ton pur perçu. 234
Tableau 47 : Valeurs du coefficient CCTF en fonction de la pondération des filtres du bloc Cochlée pour un /a/ tenu perçu.....235
Tableau 48 : Valeurs du coefficient CCTF pour les filtres 7 à 16 résultant du paramétrage des filtres pour les fréquences graves et moyennes. 237

ANNEXES

ANNEXES

Annexe 1 : Notice d'information pour les parents – Groupe EE **p. 279**

Annexe 2 : Autorisation des parents pour l'enregistrement vocal **p. 280**

Annexe 3 : Questionnaire de dépistage pour les enfants EE **p. 281**

Annexe 4 : Questionnaire QPEMVox pour les enfants EE **p. 283**

Annexe 5 : Questionnaire QPEMVox pour les enfants ESA/ESIC **p. 285**

Annexe 6 : Fiche d'informations sur la prothèse/surdité – ESA/ESIC **p. 289**

Annexe 7 : Consignes données aux enfants pour l'enregistrement vocal **p. 290**

Annexe 8 : Protocole d'évaluation PEMVox – EE/ESA **p. 292**

Annexe 9 : Protocole d'évaluation PEMVox – ESIC **p. 295**

Annexe 10 : Résultats de l'analyse du questionnaire QPEMVox **p. 298**

Annexe 11 : Base de données sur les paramètres aérodynamiques pour le groupe d'enfants implantés cochléaires (ESIC) **p. 315**

Annexe 12 : Notice d'information pour le jury sur la passation de l'évaluation perceptive **p. 316**

Annexe 13 : Exemple de fiche de cotation d'une séance d'écoute **p. 320**

Annexe 14 : Ecriture d'équations des blocs du modèle – Chapitre IV **p. 321**

Annexe 15 : Résultats de l'observation clinique du comportement vocal et de la qualité de l'émission sonore d'un /a/ tenu **p. 339**

Annexe 1 : Notice d'information pour les parents - Groupe EE

Notice d'information et de consentement pour l'Evaluation de la Voix d'un enfant

Laboratoire PRAXILING UMR 5267 CNRS - Université Montpellier III
17 rue abbé de l'épée, 34000 - Montpellier

Madame, Monsieur,

Les enfants avec problèmes de voix sont nombreux en consultation et en rééducation orthophonique. Les motifs de consultation sont fréquemment associés à des épisodes de « perte de la voix » dus à un comportement vocal abusif (ex : pendant les récréations, les activités de loisirs, les colères, …). L'enfant utilise sa voix avec un effort accompagné d'une contraction musculaire de tout l'appareil vocal. La voix devient alors rauque, faible et grave. Le problème vocal de l'enfant peut s'installer dans le temps, passer inaperçu et s'aggraver par la suite car l'enfant essaie de compenser les mauvaises performances de sa voix. Grâce à la rééducation orthophonique, l'enfant arrive à maîtriser sa voix et à s'en servir correctement par l'apprentissage de bonnes pratiques vocales.

Dans le cadre de mes travaux de recherche sur **les problèmes de la voix de l'enfant de 5 à 13 ans**, je souhaiterais réaliser une **évaluation de la voix** des enfants de *l'école Jean Sibelius* de Montpellier. Cette évaluation permettra de déterminer l'existence éventuelle d'un problème vocal chez certains enfants, d'en analyser les caractéristiques et d'en améliorer la prise en charge thérapeutique.

L'évaluation consiste à enregistrer - à l'aide d'un magnétophone pendant **une séance de 15 minutes environ** - la voix de l'enfant prononçant quelques phrases standardisées. Nous vous proposons donc de participer à cette étude et de bénéficier **gratuitement** de cette évaluation. **Ce bilan aura lieu dans les locaux de l'école de votre enfant**, la date d'évaluation étant communiquée ultérieurement.

Les données recueillies dans le cadre de cette étude resteront strictement anonymes et confidentielles. Nous n'autoriserons leur consultation que par des personnes collaborant à la recherche, désignées par l'organisateur de cette étude et un représentant des Autorités de Santé. Les conclusions de cette étude seront publiées de manière anonyme dans des revues et manifestations scientifiques.

Enfin, le Directeur de l'Ecole vous remettra un compte-rendu de l'évaluation de la voix de votre enfant.

La présente notice comprend donc les précisions nécessaires pour que vous puissiez donner, si vous le souhaitez, votre accord afin de participer - en toute connaissance de cause et sans obligation de votre part - à cette étude. Si vous acceptez que votre enfant soit évalué, je vous remercie de bien vouloir remplir le **questionnaire** ci-joint. Merci également de remplir et signer le **formulaire parental de consentement et d'autorisation**, l'ensemble de ces 2 documents (le questionnaire et le formulaire) étant à rendre à l'enseignant de votre enfant avant le 30 novembre 2006.

Je reste à votre disposition pour toute information complémentaire et vous remercie sincèrement de votre collaboration.

Harold A. GUERRERO-LOPEZ
Orthophoniste

Annexe 2 : Autorisation des parents pour l'enregistrement vocal

Formulaire parental de consentement et d'autorisation de participation à une étude sur les problèmes de la voix de l'enfant de 5 à 13 ans

Laboratoire PRAXILING UMR 5267 CNRS - Université Montpellier III
17 rue abbé de l'épée, 34000 - Montpellier

(fait en 2 exemplaires : un exemplaire remis aux parents,
un exemplaire conservé par l'organisateur de la recherche)

Je soussigné, Mr./Mme ……...
parent de ……..,[a] élève de la classe ……..……..
de l'école élémentaire Jean Sibélius de Montpellier, autorise que mon enfant participe à l'étude sur les problèmes de la voix de l'enfant de 5 à 13 ans réalisée par Harold A. GUERRERO-LOPEZ, orthophoniste doctorant du *Laboratoire PRAXILING UMR 5267 du CNRS et de l'Université Montpellier III* et autorise :

> ***« l'exploitation scientifique anonyme des données recueillies lors des enregistrements et la diffusion éventuelle des enregistrements audio dans le cadre de conférences exclusivement scientifiques ».***

Par ailleurs, j'ai bien noté que :
- *je suis libre d'accepter ou de refuser de participer à cette recherche,*
- *j'ai la possibilité de poser toutes les questions qui me paraissent utiles pour la bonne compréhension de la notice d'information et de recevoir des réponses claires et précises,*
- *je dispose d'un délai de réflexion suffisant avant de prendre ma décision,*
- *j'accepte librement et volontairement de faire participer mon enfant à cette recherche décrite dans la notice d'information,*
- *je peux faire arrêter à tout moment la participation de mon enfant à cette recherche sans supporter aucune responsabilité. J'en informerai alors directement M. Harold A. Guerrero-Lopez,*
- *toutes les données et informations qui concernent mon enfant resteront STRICTEMENT CONFIDEN-TIELLES. Je n'autorise leur consultation que par des personnes compétentes désignées par M. Harold A. Guerrero-Lopez,*
- *j'ai pris connaissance de mon droit d'accès et de rectification des informations nominatives concernant mon enfant et qui sont traitées de manière automatisée,*
- *je pourrai à tout moment demander toute information complémentaire à M. Harold A. Guerrero-Lopez.*

Fait à ……..……., le ……

Signature de l'organisateur de la recherche :
Harold A. GUERRERO-LOPEZ , orthophoniste

Signature de l'un des parents de l'enfant,
précédée de la mention "Lu et Approuvé"

[a] Inscrire le nom et le prénom de l'enfant.

Annexe 3 : Questionnaire de dépistage pour les enfants EE

Questionnaire à l'attention des parents d'enfants de l'Ecole Elémentaire *Jean Sibélius*

Date : _____ Classe : _____
Nom de l'enfant : _____ Age : _____
Sexe : _____
Nom des parents : _____Tél : _____

Donnez-vous votre accord pour que votre enfant participe à une étude sur la voix des enfants de 5 à 13 ans de la ville de Montpellier ? OUI____ NON____ Si OUI, veuillez répondre aux questions suivantes.

1) Avez-vous déjà consulté un médecin O.R.L., neurologue, psychologue, orthophoniste et/ou autre spécialiste ces 12 deniers mois pour votre enfant ?
OUI____ NON____ *Si OUI, quel spécialiste ?*
Pour quel(s) motif(s) ?

2) Votre enfant a-t-il eu une intervention chirurgicale au cours de sa vie ?
OUI____ NON____ *Si OUI, pour quel(s) motif(s) ?*

3) Votre enfant suit-il actuellement une rééducation avec un orthophoniste ?
OUI____ NON____ *Si OUI, pour quel(s) motif(s) ?*

4)Votre enfant suit-il actuellement une rééducation avec un spécialiste autre qu'orthophoniste (ex : kinésithérapeute, psychologue,…) ?
OUI____ NON____ *Si OUI, pour quel(s) motif(s) ?*

5) Considérez-vous que votre enfant a des difficultés pour :
- comprendre ce qu'on lui dit ? OUI____ NON____ *Si OUI, précisez :*

- s'exprimer devant les autres ? OUI____ NON____ *Si OUI, précisez :*

6) Pensez-vous que l'expression orale de votre enfant est en décalage avec celle des autres enfants ? OUI____ NON____ *Si OUI, précisez :*

7) Pensez-vous que la compréhension orale de votre enfant est en décalage avec celle des autres enfants ? OUI____ NON____ *Si OUI, précisez :*

8) Votre enfant a-t-il souvent des difficultés pour prononcer un son, une syllabe, un mot ou une phrase ? OUI____ NON____ *Si OUI, précisez :*

Tournez la page SVP pour la suite

9) Votre enfant modifie, ajoute ou enlève-t-il des sons dans les mots ?
OUI____ NON____ *Si OUI, précisez :*

10) Votre enfant a-t-il des difficultés pour lire (ex : lecture lente, déchiffrée, pleine d'ajouts, d'omissions, d'inversions, de modification de sons/lettres) ?
OUI____ NON____ *Si OUI, précisez :*

11) Votre enfant a-t-il des difficultés pour écrire (ex : écriture lente, déchiffrée, pleine d'ajouts, d'omissions, d'inversions, de modification de sons/lettres) ?
OUI____ NON____ *Si OUI, précisez :*

12) Votre enfant se plaint-il souvent de maux de gorge (ex : sinusites, angines, rhino-pharyngites, rhumes ou allergies respiratoires,…) ?
OUI____ NON____ *Si OUI, combien de fois par an ?*_____

13) La voix de votre enfant est-elle souvent rauque, accompagnée d'un « souffle » ou très faible ? OUI____ NON____ *Si OUI, précisez :*

14) Votre enfant a-t-il souvent des problèmes auditifs (ex : douleurs à l'oreille, mauvaise audition, demande de répétition de ce qu'on lui dit) ? OUI____ NON____ *Si OUI, précisez :*

15) Votre enfant a-t-il des difficultés pour bouger les lèvres, la mâchoire, la langue ?
OUI____ NON____ *Si OUI, précisez :*

16) Votre enfant manifeste-t-il des difficultés pour manger (ex : mâcher des aliments solides ou semi-solides ; boire des liquides) ? OUI____ NON____ *Si OUI, précisez*

17) Votre enfant a-t-il des difficultés :
- motrices ? OUI____ NON____ *Si OUI, précisez :*
- mentales ? OUI____ NON____ *Si OUI, précisez :*
- comportementales ? OUI____ NON____ *Si OUI, précisez :*
- sensorielles (vue, audition, autre) ? OUI____ NON____ *Si OUI, précisez :*

18) Votre enfant a-t-il des difficultés scolaires ? OUI____ NON____ *Si OUI, précisez :*

Merci de remettre ce questionnaire dûment rempli au maître d'école avant le 30 novembre 2006

Annexe 4 : Questionnaire QPEMVox pour les enfants EE

Laboratoire PRAXILING UMR 5267 CNRS
Université Montpelier III
Ecole Jean Sibelius

Madame, Monsieur,

Suite à l'analyse du questionnaire précédent que vous avez rempli, nous vous informons que votre enfant a été sélectionné pour participer à l'évaluation vocale qui aura lieu dans les locaux de l'école Jean Sibelius. Le bilan vocal sera réalisé le 26 mars et le 23 avril. Pour cela, nous vous remercions de bien vouloir remplir un deuxième **questionnaire (« PEMVox »)** ci-joint et de le remettre à M. GUERRERO (Orthophoniste évaluateur) le 26 mars lors de la séance d'évaluation. Votre présence n'est pas nécessaire à cette séance. L'évaluation de la voix de votre enfant s'effectuera pendant les heures de cours. Je reste à votre disposition pour toute information complémentaire et vous remercie sincèrement de votre collaboration.

Harold A. GUERRERO-LOPEZ,
Orthophoniste

QUESTIONNAIRE PEMVox (EE)

Date : _____
Nom de l'enfant : _____
Date et lieu de naissance : ____/____/____ _____
Sexe : M_____ ***F***_____ ***Age :*** _____ ***Code postal domicile :*** _____
Profession de la mère/père : _____/_____

Le questionnaire PEMVox recueille un nombre de renseignements sur les habitudes vocales et les facteurs socio-environnementaux qui peuvent être corrélés à la qualité et à la fonction vocale de l'enfant de manière directe ou indirecte. Toute l'information recueillie restera anonyme et sera traitée de manière confidentielle. Lisez attentivement et cochez ou écrivez la réponse qui correspond au cas de votre enfant.

1. Environnement familial

1.1. Y a-t-il des fumeurs dans la famille ? Oui ____ Non ____

1.2. Y a-t-il des personnes atteintes de difficultés vocales dans la famille ? Oui ___ Non___

1.3. Y a-t-il d'autres personnes atteintes de surdité dans la famille ? Oui ____ Non ____

1.4. Dans votre foyer, parlez-vous dans une autre langue ? Oui ____ Non ____
*Si Oui, laquelle ?*_____

2. Niveau sonore vocal utilisé

2.1. Votre enfant parle le plus souvent :
plutôt faiblement ____ normalement ____ plutôt fortement ____

2.2. Votre enfant arrive à se faire entendre dans un groupe et/ou dans un milieu bruyant :
très difficilement ____ normalement ____ très facilement ____

2.3. Dans des situations d'urgence (ex : appel au secours, …), votre enfant arrive à se faire entendre :
très difficilement ____ normalement ____ très facilement ____

3. Activités de loisir

3.1. Votre enfant écoute-t-il de la musique ou regarde-t-il la télévision ? : Oui ____ Non ____
Si Oui, le niveau sonore est : plutôt faible ____ normal ____ plutôt fort ____

3.2. Quelle est la régularité de ces activités ? : plutôt rare ____ normale ____ plutôt fréquente ____

3.3. Quelles sont les activités de loisirs de votre enfant ? :

Activités extrascolaires	**Jamais**	**Occasionnellement** (au moins 1 fois/mois)	**Souvent** (au moins 1 fois/semaine)	**Type d'activité**
Musicales				
Sportives				
Artistiques				
Autre (scouts, etc)				

3.4. Quel temps votre enfant dépense-t-il pour ses activités de loisir ?
peu de temps ____ le temps nécessaire ____ beaucoup de temps ____

3.5. Pendant les activités de jeux, votre enfant parle-t-il :
plutôt faiblement ____ normalement ____ plutôt fortement ____

4. Manifestation des signes d'anomalie vocale

4.1. Considérez-vous que la voix de votre enfant manifeste des difficultés ? Oui ____ Non ____
Si *NON*, passez à la question No. 6 de cet item.

4.2. Considérez-vous que les difficultés vocales de votre enfant sont associées au(x) situation(s) de :
Changements de température ____ Stress (personnalité de l'enfant) ____
Abus de la voix ____ Manque de maîtrise de la voix ____
Défaillance auditive_____ Maux de gorge (rhumes, infections, allergies) ____
Autres, lesquels ? _____

4.3. A quel moment de la journée, la voix de votre enfant semble être la plus altérée ?
le matin ____ l'après-midi ____ le soir ____

4.4. Les difficultés vocales de votre enfant pour parler et/ou chanter apparaissent de façon :
spontanée ____ progressive ____

4.5. Si des difficultés vocales se manifestent, apparaissent-elles au cours de la journée de façon :
intermittente ____ constante ____

4.6. Votre enfant se plaint-il de maux de gorge (ex : rhumes, infections ou allergies respiratoires) :
jamais ____ 1 à 3 fois par an ____ plus de 3 fois par an ____

4.7. La voix de votre enfant devient-elle rauque (ou accompagnée d'un souffle) lorsqu'il se plaint de maux de gorge ?
jamais ____ occasionnellement ____ toujours ____

4.8. Votre enfant a t-il eu une intervention chirurgicale au cours de sa vie ? Oui ____ Non ____
Si Oui, pour quel(s) motif(s) ?

4.9. Votre enfant se plaint-il de douleurs d'oreille :
jamais ____ 1 à 3 fois par an _____ plus de 3 fois par an ____

Commentaires

Merci de prendre le temps nécessaire pour compléter ce questionnaire avec des commentaires que vous considérez pertinents à formuler et qui n'ont pas été traités dans les questions précédentes :

Merci pour votre collaboration précieuse.

Annexe 5 : Questionnaire QPEMVox pour les enfants ESA/ESIC

Laboratoire PRAXILING – UMR 5267 CNRS – **Université Montpellier III**
C.H.U. Gui de Chauliac – Service O.R.L. et Chirurgie Cervico-faciale – Unité de Phoniatrie
2007 - 2008

QUESTIONNAIRE PEMVox (ESA/ESIC)

Date : _____
Nom de l'enfant : _____
Date et lieu de naissance : ____/____/____ _____
Sexe : M_____ ***F***_____ ***Age :*** _____ ***Code postal domicile :*** _____
Profession de la mère/père : _____/_____

Le questionnaire PEMVox recueille un nombre de renseignements sur les habitudes vocales et les facteurs socio-environnementaux qui peuvent être corrélés à la qualité et à la fonction vocale de l'enfant de manière directe ou indirecte. Toute l'information recueillie restera anonyme et sera traitée de manière confidentielle. Lisez attentivement et cochez ou écrivez la réponse qui correspond au cas de votre enfant.

1. Environnement familial

1.1. Y a-t-il des fumeurs dans la famille ? Oui ____ Non ____

1.2. Y a-t-il des personnes atteintes de difficultés vocales dans la famille ? Oui ___ Non___

1.3. Y a-t-il d'autres personnes atteintes de surdité dans la famille ? Oui ____ Non ____

1.4. Dans votre foyer, parlez-vous dans une autre langue ? Oui ____ Non ____
*Si Oui, laquelle ?*_____

2. Niveau sonore vocal utilisé

2.1. Votre enfant parle le plus souvent :
plutôt faiblement ____ normalement ____ plutôt fortement ____

2.2. Votre enfant arrive à se faire entendre dans un groupe et/ou dans un milieu bruyant :
très difficilement ____ normalement ____ très facilement ____

2.3. Dans des situations d'urgence (ex : appel au secours, …), votre enfant arrive à se faire entendre :
très difficilement ____ normalement ____ très facilement ____

3. Activités de loisir

3.1. Votre enfant écoute-t-il de la musique ou regarde-t-il la télévision ? : Oui ____ Non ____
Si Oui, le niveau sonore est : plutôt faible ____ normal ____ plutôt fort ____

3.2. Quelle est la régularité de ces activités ? : plutôt rare ____ normale ____ plutôt fréquente ____
3.3. Quelles sont les activités de loisirs de votre enfant ? :

Activités extrascolaires	**Jamais**	**Occasionnellement** (au moins 1 fois/mois)	**Souvent** (au moins 1 fois/semaine)	**Type d'activité**
Musicales				
Sportives				
Artistiques				
Autre (scouts, etc)				

3.4. Quel temps votre enfant dépense-t-il pour ses activités de loisir ?
peu de temps ____ le temps nécessaire ____ beaucoup de temps ____

3.5. Pendant les activités de jeux, votre enfant parle-t-il :
plutôt faiblement ____ normalement ____ plutôt fortement ____

4. Manifestation des signes d'anomalie vocale

4.1. Considérez-vous que la voix de votre enfant manifeste des difficultés ? Oui ____ Non ____
Si *NON*, passez à la question No. 6 de cet item.

4.2. Considérez-vous que les difficultés vocales de votre enfant sont associées au(x) situation(s) de :
Changements de température ____ Stress (personnalité de l'enfant) ____
Abus de la voix ____ Manque de maîtrise de la voix ____
Défaillance auditive_____ Maux de gorge (rhumes, infections, allergies) ____
Autres, lesquels ? _____

4.3. A quel moment de la journée, la voix de votre enfant semble être la plus altérée ?
le matin ____ l'après-midi ____ le soir ____

4.4. Les difficultés vocales de votre enfant pour parler et/ou chanter apparaissent de façon :
spontanée ____ progressive ____

4.5. Si des difficultés vocales se manifestent, apparaissent-elles au cours de la journée de façon :
intermittente ____ constante ____

4.6. Votre enfant se plaint-il de maux de gorge (ex : rhumes, infections ou allergies respiratoires) :
jamais ____ 1 à 3 fois par an ____ plus de 3 fois par an ____

4.7. La voix de votre enfant devient-elle rauque (ou accompagnée d'un souffle) lorsqu'il se plaint de maux de gorge ?
jamais ____ occasionnellement ____ toujours ____

4.8. Votre enfant a t-il eu une intervention chirurgicale au cours de sa vie ? Oui ____ Non ____
Si Oui, pour quel(s) motif(s) ?

4.9. Votre enfant se plaint-il de douleurs d'oreille :
jamais ____ 1 à 3 fois par an _____ plus de 3 fois par an ____

5. Education et intégration scolaire

5.1. Classe : ____ Etablissement : _____

5.2. Type d'établissement scolaire :
ordinaire (école, collège) ____ centre spécialisé ____ autre _____

5.3. Type d'intégration scolaire : normale ____ adaptée (enseignants, locaux, matériel) ____
Si adapté, précisez le type d'adaptation dont votre enfant bénéficie dans l'établissement scolaire :

Si adaptée, considérez-vous que l'établissement scolaire est :
mal adapté ____ moyennement adapté ____ bien adapté ____

5.4. Type de scolarisation :
conventionnelle (plein temps) ____ partielle (mi-temps) ____

5.5. Les résultats de la scolarité de votre enfant sont actuellement :
insuffisants ____ sans remarques spécifiques ____ satisfaisants ____

5.6. Votre enfant est-il en relation régulière avec des enfants :
moins âgés que lui ____ de son âge ____ plus âgés que lui ____
entendants ____ sourds ____ handicapés (autre que audition) ____

5.7. Votre enfant porte-t-il l'appareillage ou l'implant (en position ON) dans l'école :
quelques heures (1-4h) ____ demi-journée (4-7h) ____ toute la journée (7h ou plus) ___

5.8. Considérez-vous que l'équipe éducative est soucieuse du bon fonctionnement de la prothèse de votre enfant durant les cours ? Oui ____ Non ____ *Si Non, pourquoi ?*

5.9. Votre enfant est-il à l'aise dans son école ? Oui ___ Non ____ *Si Non, pourquoi ?*

5.10. Etes-vous satisfait(e) de l'accueil donné par l'école à votre enfant ?
Oui ____ Non ____ *Si Non, pourquoi ?*

5.11. Comment votre enfant communique-t-il dans l'école ?

- Langage oral uniquement ____
- Langage oral + L.P.C.* ____
- Langage oral + signes ____
- L.S.F.** uniquement ____
- Autre moyen ____, lequel ? _____

6. Mode de communication de l'enfant

6.1. Comment vous adressez-vous à votre enfant ? :

- Langage oral uniquement ____
- Langage oral + L.P.C. ____
- Langage oral + signes ____
- L.S.F. uniquement ____
- Autre moyen ____, lequel ? _____

6.2. Comment votre enfant s'adresse-t-il aux membres de la famille ? :

- Langage oral uniquement ____
- Langage oral + L.P.C. ____
- Langage oral + signes ____
- L.S.F. uniquement ____
- Autre moyen ____, lequel ? _____

6.3. Comment votre enfant s'adresse-t-il aux personnes extérieures de la famille ? :

- Langage oral uniquement ____
- Langage oral + L.P.C. ____
- Langage oral + signes ____
- L.S.F. uniquement ____
- Autre moyen ____, lequel ? _____

6.4. Votre enfant s'adresse-t-il aux autres de manière :
spontanée ____ stimulée (par vous ou quelqu'un d'autre) ____

6.5. Les gens semblent-ils être gênés avec la voix de votre enfant ? Oui ____ Non ____
Si oui, pourquoi ?

7. Port de la prothèse auditive

7.1. Votre enfant porte-t-il l'appareillage ou l'implant (en position ON) à la maison :
quelques heures (1-4h) ____ demi-journée (4-7h) ____toute la journée (7h ou plus) ___

7.2. La communication de votre enfant est-elle facile et efficace avec la prothèse ?
Oui ____ Non ____ *Si Non, pourquoi ?*

---

* L.P.C. : Langage Parlé Complété
** LSF : Langue de Signes Française

7.3. Considérez-vous que la qualité de la voix de votre enfant a été améliorée :
avant 3 ans d'appareillage ou d'implantation ___ après 3 ans d'appareillage ou d'implantation ___

7.4. Pensez-vous que la qualité de la voix de votre enfant change après chaque séance de réglage de l'appareillage ou de l'implant ?
Oui ____ Non ____ *Si Oui, précisez ?*

7.5. Pensez-vous que le comportement de votre enfant ait changé depuis l'appareillage ou l'implan-tation ?
Oui ____ Non ____ *Si Oui, précisez ?*

7.6. La qualité de la voix de votre enfant vous inquiète toujours ? Oui ____ Non ____
Si Oui, pourquoi ?

7.7. Votre enfant manifeste-t-il actuellement des sensations d'inconfort auditif avec l'appareillage ou l'implant :
jamais ____ occasionnellement ____ très régulièrement ____

7.8. Avez-vous toujours de contraintes concernant l'audition de votre enfant ? Oui ____ Non ____
Si Oui, lesquelles ?

7.9. Dans la globalité, êtes-vous satisfait(e) des résultats obtenus par l'appareillage ou l'implantation cochléaire de votre enfant sur sa voix ? Oui ____ Non ____ *Si Non, pourquoi ?*

Commentaires
Merci de prendre le temps nécessaire pour compléter ce questionnaire avec des commentaires que vous considérez pertinents à formuler et qui n'ont pas été traités dans les questions précédentes :

Merci pour votre collaboration précieuse.

Annexe 6 : Fiche d'informations sur la prothèse/surdité – ESA/ESIC

Laboratoire PRAXILING – UMR 5267 CNRS – **Université Montpellier III**
C.H.U. Gui de Chauliac – Service O.R.L. et Chirurgie Cervico-faciale – Unité de Phoniatrie
2007 - 2008

Informations sur la prothèse/surdité ESA/ESIC

Nom de l'enfant : _____
Date de chirurgie (pour l'I.C.) :_____
Date d'activation (pour l'I.C.) :_____

A. Prothèse auditive

1. L'enfant porte-t-il une prothèse auditive :
   - Conventionnelle (voir question No. 2) _____ Unilatérale ____ Bilatérale ____
   - Implant cochléaire (voir question No. 3) _____ Unilatérale ____ Bilatérale ____

2. Type de prothèse auditive conventionnelle :
   - Marque/fabricant : _____
   - Modèle : _____
   - Forme : Contour d'oreille _____
     Boîtier _____
     Intra-auriculaire (intra-coque, intra-conduit) _____
     Autre _____
   - Type de conduction : Aérienne _____ Osseuse _____ Mixte _____

3. Type d'implant cochléaire :
   - Marque/fabricant : Nucleus _____ Clarion _____ Autre _____
   - Modèle : _____
   - Nombre d'électrodes : _____
   - Processeur : _____
   - Stratégie de codage : _____

4. Durée d'appareillage/implantation (an; mois) : _____

5. Age d'appareillage/implantation (an; mois) : _____
Lieu : _____

B. Antécédents auditifs :

1. Période de début de la surdité : Congénitale _____ Prélinguale _____ Périlinguale _____

2. Age de diagnostique de la surdité (an; mois) : _____
   Lieu : _____

3. Cause(s) de la surdité : _____

4. Degré de surdité :
   *Sévère T I (perte auditive de 71 à 80 dB) _____
   *Sévère T II (perte auditive de 81 à 90 dB) _____
   *Profonde T I (perte auditive de 91 à 100 dB) _____
   *Profonde T II (perte auditive de 101 à 110 dB) _____
   *Profonde T III (perte auditive de 111 à 119 dB) _____

Commentaires :

Annexe 7 : Consignes données aux enfants pour l'enregistrement

Laboratoire PRAXILING – UMR 5267 CNRS – **Université Montpellier III**
C.H.U. Gui de Chauliac – Service O.R.L. et Chirurgie Cervico-faciale – Unité de Phoniatrie
2007 - 2008

CONSIGNES DONNEES A L'ENFANT PENDANT L'ENREGISTREMENT PROTOCOLE PEMVox

I. EVALUATION ET ENREGISTREMENT DES PARAMETRES ACOUSTIQUES DE LA VOIX

1. Types de voix

Voix conversationnelle (parole spontanée)

1) Demander à l'enfant *: « Raconte-moi ce que tu as fait le week-end dernier » ou « Qu'est-ce que tu as fait le week-end dernier ? »*
2) Enregistrer sur DAT
3) Demander à l'enfant *: « Dis-moi ce que tu aimes faire avec tes copains/copines ? » ou « Qu'est-ce que tu aimes faire avec tes copains/copines ?*

4) Noter *la position et la posture* observées

Voix forte (parole guidée : comptage projeté, Le Huche)

1) Demander à l'enfant : « *Compte à haute voix jusqu'à 20 comme si tu voulais que ta maman, qui se trouve au fond du couloir, t'entende* »
2) Donner l'exemple
3) Enregistrer sur DAT

Voix d'appel (parole guidée)

1) Demander à l'enfant : *« Appelle ta maman comme si elle était loin (à l'entrée de la salle d'attente) »*
2) Donner l'exemple
3) Enregistrer sur DAT

2. Tenues vocaliques (/a/ tenu)

1) Demander à l'enfant : *« Dis la voyelle /a/ le plus longtemps possible sans faire effort. Pour cela, tu vas respirer profondément avant de la dire. On fera ceci 3 fois »*
2) Donner l'exemple
3) Enregistrer sur DAT
4) Répéter l'exercice 2 fois de plus.

3. Tessiture (vocalise de la sirène /a/) et Etendue vocale (vocalise de la fusée /i/)

1) Présenter le déssin et demander à l'enfant : *« Tu vas imiter la sirène d'une ambulance sans forcer avec la voyelle /a/ »*
2) Donner l'exemple
3) Enregistrer sur DAT
4) Demander à l'enfant : *« Maintenant, tu vas imiter le son d'une fusée avec la voyelle /i/ »*
5) Donner l'exemple

6) Enregistrer sur DAT

II. EVALUATION DU COMPORTEMENT RESPIRATOIRE

1. Comportement respiratoire

1) Demander à l'enfant : *« Respire normalement »*
2) Compter le nombre de cycles avec un chronomètre
3) Demander à l'enfant : *« Dis un /a/ tenu et je vais juste poser ce miroir entre ton nez et ta lèvre pendant que tu prononces le /a/ »*
4) Noter ce qui a été observé sur la tonicité musculaire et les réponses aux questions

III. EVALUATION ET ENREGISTREMENT DES PARAMETRES AERODYNAMQIUES DE LA VOIX (Logiciel SESANE : Efficacité glottique)

PSGE

1) Demander à l'enfant : *« Tu vas répéter la syllabe /pa/ 10 fois quand je te donne le signal »*
2) Donner l'exemple
3) Enregistrer sur EVA

Annexe 8 : Protocole d'évaluation PEMVox – EE/ESA

Laboratoire PRAXILING UMR 5267 CNRS - Université Montpellier III
Ecole Jean Sibelius/Collège Jeu de Mail
2006 - 2007

Protocole PEMVox EE/ESA

Date d'évaluation: **Identification :**
Sexe : **Age :**
Enregistrement D.A.T. : **Grade G (échelle GRBAS) : 0 1 2 3**

1. ENREGISTREMENT DE LA VOIX – PARAMETRES ACOUSTIQUES

1.1. Types de voix

1.1.1. Voix conversationnelle (parole spontanée)

- ***Fondamental usuel moyen (FoUm)* :** ____ Hz, ____ (selon la notation musicale)
- ***Fréq. min. (Hz) :*** ____, ____
- ***Fréq. max. (Hz) :*** ____, ____
- ***Dynamique (Hz)*** *:* ____, ____ *(½ ton)*
- ***E-type (Hz) :*** ____, ____ *(½ ton)*
- ***Intensité (dB)*:** ____

- ***Axe de posture (visage, région thoracique, rachis cervico-thoracique)*** :

Vertical ____ vers l'avant ____ vers l'arrière ____ instable ____

- ***Position :***

Tête : relâchée ____ normal ___ étirée ____
Epaules : élevées uni/bi ____ normales ____ étirées AV/AR ____
Dos : vertical/normal ____ étiré (flexion en avant)/malformation ____
Bras : croisés à l'arrière (AR)/l'avant (AV) ____ long du corps ____ instables ____
Jambes (stabilité) : très serrées (TS)/très écartées (TE) ____ normales ____ instables ____
Talons (contact avec le sol) : latéral (L)/pointes (P)/talons (T) ____ normal ____ instables ____
Pieds : très serrés/très écartés ____ normal (tournés vers l'extérieur, axe 30°) ____ Instables ____

1.1.2. Voix forte (parole guidée)

- ***Fondamental usuel moyen (FoUm)* :** ____ Hz, ____ (selon la notation musicale)
- ***Fréq. min. (Hz) :*** ____, ____
- ***Fréq. max. (Hz) :*** ____, ____
- ***Dynamique (Hz) :*** ____, ____ *(½ ton)*
- ***E-type (Hz) :*** ____, ____ *(½ ton)*
- ***Intensité (dB)*:** ____

- ***Axe de posture (visage, région thoracique, rachis cervico-thoracique)*** :

Vertical ____ vers l'avant ____ vers l'arrière ____ instable ____

- ***Position :***

Tête : relâchée ____ normal ___ étirée ____
Epaules : élevées uni/bi ____ normales ____ étirées AV/AR ____
Dos : vertical/normal ____ étiré (flexion en avant)/malformation ____
Bras : croisés à l'arrière /l'avant ____ long du corps ____ instables ____
Jambes (stabilité) : très serrées/très écartées ____ normales ____ instables ____
Talons (contact avec le sol) : latéral/pointes/talons ____ normal ____ instables ____
Pieds : très serrés/très écartés ____ normal (tournés vers l'extérieur, axe 30°) ____ Instables _____

1.1.3. Voix d'appel (parole guidée)

- ***Fondamental usuel moyen (FoUm) :*** ____ Hz, ____ (selon la notation musicale)
- ***Fréq. min. (Hz) :*** ____, ____
- ***Fréq. max. (Hz) :*** ____, ____
- ***Dynamique (Hz) :*** ____, ____ *(½ ton)*
- ***E-type (Hz) :*** ____, ____ *(½ ton)*
- ***Intensité (dB):*** ____
- ***Mécanisme vocal utilisé :*** 1 (voix de poitrine) ___ 2 (voix de tête) ___ Impression de changement __

- ***Axe de posture (visage, région thoracique, rachis cervico-thoracique)*** :
  Vertical ____ vers l'avant ____ vers l'arrière ____ instable ____

1.2. Tenues vocaliques (/a/ tenu)

1.2.1. Paramètres acoustiques : A partir du signal du T.M.P. le plus long (selon les 3 essais précédents), relever les valeurs suivants sur la partie la plus stable du signal

Paramètres	Valeurs mesurées	
Fo moyenne	Hz	(note)
E-Type Fo		Hz
Coeff. de variation Fo		%
Jitter factor		%
Int. Moyenne		dB
E-Type Int.		dB
Coeff. de variation Fo		%
Shimmer factor		%
Rapport Harmoniques/Bruit		dB

1.2.2. Qualité de la tenue vocalique : A partir du profil vocal (selon les 3 essais précédents)

- ***Attaque (début) :*** normal ____ brusque ____ faible ____
- ***Partie stable :*** normal ____ instable ____ interrompu ____
- ***Sortie (fin) :*** normal ____ brusque ____ Faible ____

1.2.3. Tessiture (vocalise de la sirène avec la voyelle /a/)

- ***Fréq. min. (Hz) :*** ____, ____
- ***Fréq. max. (Hz) :*** ____, ____
- ***Dynamique (Hz) :*** ____, ____ *(½ ton)*
- ***E-type (Hz) :*** ____, ____ *(½ ton)*

1.2.4. Etendue vocale (vocalise de la fusée avec la voyelle /i/)

- ***Fréq. min. (Hz) :*** ____, ____
- ***Fréq. max. (Hz) :*** ____, ____
- ***Dynamique (Hz) :*** ____, ____ *(½ ton)*
- ***E-type (Hz) :*** ____, ____ *(½ ton)*

1.2.5. T.M.P.

Essais T.M.P.	Essai 1		Essai 2		Essai 3	
***T.M.P.**		sec		sec		sec
Fo moyenne	Hz	(note)	Hz	(note)	Hz	(note)
Intensité moyenne		dB		dB		dB

2. OBSERVATION DU COMPORTEMENT RESPIRATOIRE

2.2. Tonicité musculaire au repos

Zone corporelle	hyper	normal	hypo
Visage			
Cou			

- ***Type respiratoire*** : scapulaire ____ costal sup. ____ costo-abdom. (mixte) ____ Abdom. ____
- ***Mode respiratoire*** : prédominance nasale ____ prédominance orale ____

- ***Fréquence respiratoire*** : ____ cycles/min
- ***Perméabilité nasale (épreuve au miroir : présence de buée) :*** oui ____ non ____

Commentaires :

Annexe 9 : Protocole d'évaluation PEMVox - ESIC

Laboratoire PRAXILING UMR 5267 CNRS - Université Montpellier III
C.H.U. Gui de Chauliac – Service O.R.L. et Chirurgie Cervico-faciale – Unité de Phoniatrie
2007 - 2008

Protocole PEMVox EE/ESA

Date d'évaluation: **Identification :**
Sexe : **Age :**
Enregistrement D.A.T. : **Grade G (échelle GRBAS) : 0 1 2 3**

1. ENREGISTREMENT DE LA VOIX – PARAMETRES ACOUSTIQUES

1.1. Types de voix

*1.1.1. Voix conversationnelle (parole spontanée)**1.1.1. Voix conversationnelle (parole spontanée)*

- ***Fondamental usuel moyen (FoUm)* :** ____ Hz, ____ (selon la notation musicale)
- ***Fréq. min. (Hz) :*** ____, ____
- ***Fréq. max. (Hz) :*** ____, ____
- ***Dynamique (Hz)*** *:* ____, ____ *(½ ton)*
- ***E-type (Hz) :*** ____, ____ *(½ ton)*
- ***Intensité (dB)*:** ____

- ***Axe de posture (visage, région thoracique, rachis cervico-thoracique)*** :

Vertical ____ vers l'avant ____ vers l'arrière ____ instable ____

- ***Position :***

Tête : relâchée ____ normal ___ étirée ____
Epaules : élevées uni/bi ____ normales ____ étirées AV/AR ____
Dos : vertical/normal ____ étiré (flexion en avant)/malformation ____
Bras : croisés à l'arrière (AR)/l'avant (AV) ____ long du corps ____ instables ____
Jambes (stabilité) : très serrées (TS)/très écartées (TE) ____ normales ____ instables ____
Talons (contact avec le sol) : latéral (L)/pointes (P)/talons (T) ____ normal ____ instables ____
Pieds : très serrés/très écartés ____ normal (tournés vers l'extérieur, axe 30°) ____ Instables ____

1.1.2. Voix forte (parole guidée)

- ***Fondamental usuel moyen (FoUm)* :** ____ Hz, ____ (selon la notation musicale)
- ***Fréq. min. (Hz) :*** ____, ____
- ***Fréq. max. (Hz) :*** ____, ____
- ***Dynamique (Hz) :*** ____, ____ *(½ ton)*
- ***E-type (Hz) :*** ____, ____ *(½ ton)*
- ***Intensité (dB)*:** ____

- ***Axe de posture (visage, région thoracique, rachis cervico-thoracique)*** :

Vertical ____ vers l'avant ____ vers l'arrière ____ instable ____

- ***Position :***

Tête : relâchée ____ normal ___ étirée ____
Epaules : élevées uni/bi ____ normales ____ étirées AV/AR ____
Dos : vertical/normal ____ étiré (flexion en avant)/malformation ____
Bras : croisés à l'arrière /l'avant ____ long du corps ____ instables ____
Jambes (stabilité) : très serrées/très écartées ____ normales ____ instables ____
Talons (contact avec le sol) : latéral/pointes/talons ____ normal ____ instables ____
Pieds : très serrés/très écartés ____ normal (tournés vers l'extérieur, axe 30°) ____ Instables _____

1.1.3. Voix d'appel (parole guidée)

- ***Fondamental usuel moyen (FoUm) :*** ____ Hz, ____ (selon la notation musicale)
- ***Fréq. min. (Hz) :*** ____, ____
- ***Fréq. max. (Hz) :*** ____, ____
- ***Dynamique (Hz) :*** ____, ____ *(½ ton)*
- ***E-type (Hz) :*** ____, ____ *(½ ton)*
- ***Intensité (dB):*** ____
- ***Mécanisme vocal utilisé :*** 1 (voix de poitrine) ___ 2 (voix de tête) ___ Impression de changement __

- ***Axe de posture (visage, région thoracique, rachis cervico-thoracique)*** :
  Vertical ____ vers l'avant ____ vers l'arrière ____ instable ____

1.2. Tenues vocaliques (/a/ tenu)

1.2.1. Paramètres acoustiques : A partir du signal du T.M.P. le plus long (selon les 3 essais précédents), relever les valeurs suivants sur la partie la plus stable du signal

Paramètres	Valeurs mesurées	
Fo moyenne	Hz	(note)
E-Type Fo		Hz
Coeff. de variation Fo		%
Jitter factor		%
Int. Moyenne		dB
E-Type Int.		dB
Coeff. de variation Fo		%
Shimmer factor		%
Rapport Harmoniques/Bruit		dB

1.2.2. Qualité de la tenue vocalique : A partir du profil vocal (selon les 3 essais précédents)

- ***Attaque (début) :*** normal ____ brusque ____ faible ____
- ***Partie stable :*** normal ____ instable ____ interrompu ____
- ***Sortie (fin) :*** normal ____ brusque ____ Faible ____

1.2.3. Tessiture (vocalise de la sirène avec la voyelle /a/)

- ***Fréq. min. (Hz) :*** ____, ____
- ***Fréq. max. (Hz) :*** ____, ____
- ***Dynamique (Hz) :*** ____, ____ *(½ ton)*
- ***E-type (Hz) :*** ____, ____ *(½ ton)*

1.2.4. Etendue vocale (vocalise de la fusée avec la voyelle /i/)

- ***Fréq. min. (Hz) :*** ____, ____
- ***Fréq. max. (Hz) :*** ____, ____
- ***Dynamique (Hz) :*** ____, ____ *(½ ton)*
- ***E-type (Hz) :*** ____, ____ *(½ ton)*

1.2.5. T.M.P.

Essais T.M.P.	Essai 1		Essai 2		Essai 3	
***T.M.P.**		sec		sec		sec
Fo moyenne	Hz	(note)	Hz	(note)	Hz	(note)
Intensité moyenne		dB		dB		dB

2. OBSERVATION DU COMPORTEMENT RESPIRATOIRE

2.2. Tonicité musculaire au repos

Zone corporelle	hyper	normal	hypo
Visage			
Cou			

- ***Type respiratoire*** : scapulaire ____ costal sup. ____ costo-abdom. (mixte) ____ Abdom. ____
- ***Mode respiratoire*** : prédominance nasale ____ prédominance orale ____

- ***Fréquence respiratoire*** : ____ cycles/min
- ***Perméabilité nasale (épreuve au miroir : présence de buée) :*** oui ____ non ____

3. ENREGISTREMENT DE LA SERIE /pa/ - PARAMETRES AERODYNAMQIUES : P.S.G.E.

# cycle	D.A.O. (dm^3/s)	P.I.O. (hPa)	Intensité (dB)	Eff. Glottique (dB/hPa)	Eff. Laryngée ($dB/hPa*dm^3/s$)
2e cycle					
3e cycle					
4e cycle					
5e cycle					
6e cycle					
7e cycle					
8e cycle					
9e cycle					
Moyenne					

Commentaires :

Annexe 10 : Résultats de l'analyse du questionnaire QPEMVox

Afin d'appréhender les facteurs environnementaux, socio-familiaux et les habitudes vocales de la population de l'étude, nous avons recueilli des informations auprès des parents des enfants à l'aide d'un questionnaire lors de l'enregistrement vocal. En clinique, cette information permet d'expliquer depuis une autre perspective, autre que l'organique, l'origine de certains troubles vocaux identifiés chez un patient[b]. L'information recueillie à l'aide de ce type de questionnaire permet également l'orientation de la prise en charge des troubles vocaux du patient. Nous avons inclut cette information dans notre travail à titre informatif et descriptif.

Les résultats de l'analyse du questionnaire intitulé QPEMVox sont présentés en trois parties. La première partie traite des aspects de l'environnement familial, du niveau sonore vocal utilisé dans les différents types de voix, des activités de loisirs et de la manifestation des signes d'anomalie vocale perçus par les parents. Cette enquête a été effectuée auprès des trois groupes de populations : EE, ESA et ESIC. La deuxième partie présente des aspects concernant l'intégration scolaire, le mode de communication et le port de la prothèse par les enfants ESA et ESIC. La troisième partie présente les caractéristiques de l'appareillage et de l'implantation ainsi que de la surdité des enfants.

Première partie

1. Environnement familial. Les réponses des parents au questionnaire sur les caractéristiques de l'environnement familial sont résumées comme suit :

Dans le groupe EE :

- 45% des enfants vivent avec des membres de la famille qui sont fumeurs.
- Aucun membre de la famille de l'enfant ne manifeste de troubles vocaux.
- 11% des enfants proviennent de foyers dont l'un des membres de la famille est atteint de surdité. Par ailleurs, 2% des parents des enfants n'ont pas fourni de réponse.
- 18% des enfants proviennent de foyers parlant une autre langue que le français (Anglais, Allemand, Polonais, Grec, Arabe, Arabe syrien, Berbère algérien, Langues africaines telles que le Fufuldé).

[b] Par example, la dysphonie dysfonctionnelle simple (selon les termes de Le Huche, 2001).

Dans le groupe ESA :

- 45% des enfants vivent avec des membres de la famille qui sont fumeurs.
- Seulement 5% des enfants proviennent de foyers dont l'un des membres de la famille manifeste des troubles vocaux.
- 30% des enfants proviennent de foyers dont l'un des membres de la famille est atteint de surdité.
- 40% des enfants proviennent de foyers parlant une autre langue que le français oral (Anglais, Russe, Arabe, Espagnol, Roumain, dialectes régionaux tels que l'Occitan et le Catalan).

Dans le groupe ESIC :

- 40% des enfants vivent avec des membres de la famille qui sont fumeurs.
- Aucun membre de la famille de l'enfant ne manifeste de troubles vocaux.
- Seulement 5% des enfants proviennent de foyers dont l'un des membres de la famille est également atteint de surdité. Par ailleurs, 15% des parents des enfants n'ont pas fourni de réponse.
- Les parents ne signalent pas l'usage d'une autre langue que le français oral.

Les résultats révèlent l'existence de deux facteurs environnementaux communs aux trois groupes de population : les habitudes tabagiques parentales et la cohabitation avec un déficient auditif dans le foyer. Dans la littérature, la corrélation entre les facteurs environnementaux, socio-familiaux, éducatifs et, la qualité de la voix infantile n'est pas bien cernée (Mac Arthur et Healey, 1995)[c]. D'après Dejonckere (1998)[d], il n'y aurait pas de corrélation significative entre les facteurs tels que les habitudes tabagiques parentales et, la qualité de la voix de l'enfant. Par ailleurs, nous avons noté l'existence accentuée d'un environnement familial bilingue chez le groupe ESA. Bien que le bilinguisme a été signalé comme un facteur potentiel favorisant la raucité de la voix, aucune significativité statistique n'a été reportée (Senderholm, 1995)[e].

---

c MAC ARTHUR, C.J., et HEALY, G.B., « Acquired voice disorders in the pediatric population » in *Diagnosis and treatment of voice disorders*, J.S. Rubin et al. eds., New York/Tokio, Igaku Shoin, 1995, p. 189-202.

d DEJONCKERE, P.H., « Qualité de voix chez l'enfant et facteurs sociaux / environnementaux » in *Les pathologies vocales chez l'enfant*, Paris, Rééducation orthophonique, n° 194, 1998, p. 39-44.

e SENDERHOLM, Elisabeth, MCALLISTER, Anita, DALKVIST, Jan, SUNDBERG, Johan, "Aetiologic factors associated with hoarseness in ten-years-old children", *Folia Phoniatr. Logop.*, Vol. 47, 1995, p. 262-278.

2. Niveau sonore vocal. Concernant l'utilisation de la voix de l'enfant dans les différentes situations communicatives de la vie courante :

Dans le groupe EE :

- 18% des enfants utilisent une voix conversationnelle plutôt forte.
- 8% des enfants ont du mal à se faire entendre dans un groupe et/ou dans un milieu bruyant, 10% des enfants arrivent à communiquer de façon très aisée. Par ailleurs, 3% des parents n'ont pas exprimé leur avis.
- 8% des enfants ont du mal à se faire entendre dans des situations de détresse, 13% des enfants arrivent à communiquer de façon très aisée. Par ailleurs, 5% des parents n'ont pas donné leur opinion.

Dans le groupe ESA :

- 30% des enfants utilisent une voix conversationnelle plutôt forte; 10% des enfants utilisent une voix plutôt faible.
- 25% des enfants ont du mal à se faire entendre dans un groupe et/ou dans un milieu bruyant, 15% des enfants arrivent à communiquer de façon très aisée.
- 15% des enfants ont de mal à se faire entendre dans des situations de détresse, 30% des enfants arrivent à communiquer de façon très aisée. Par ailleurs, 5% des parents n'ont pas donné leur opinion.

Dans le groupe ESIC :

- 20% des enfants utilisent une voix conversationnelle plutôt forte; 5% des enfants utilisent une voix plutôt faible.
- 20% des enfants ont du mal à se faire entendre dans un groupe et/ou dans un milieu bruyant, 10% des enfants communiquent plutôt facilement.
- 5% des enfants ont du mal à se faire entendre dans des situations de détresse, 20% des enfants arrivent à communiquer de façon très aisée.

Il est intéressant d’observer que la perception des parents sur la performance vocale de leurs enfants pour les différents types de voix correspond de manière globale aux résultats obtenus par l’évaluation acoustique instrumentale. Cette correspondance est relative au rapprochement de la performance vocale des enfants entendant (EE) et des enfants implantés (ESIC).

3. Activités de loisir. Concernant les activités de détente de l'enfant, nous observons que :

Dans le groupe EE :

- 42% des enfants participent à des activités sportives, 26% des enfants pratiquent trois activités de loisir différentes (musicales, sportives et artistiques), 13% des enfants pratiquent deux activités différentes (musicales et/ou sportives et/ou artistiques), 11% des enfants pratiquent une activité musicale ou artistique, 5% des enfants ne pratiquent aucune activité de loisir. Par ailleurs, 3% des parents n'ont pas fourni de réponse.
- Tous les enfants écoutent de la musique et regardent la télévision comme une activité principale dans leur vie quotidienne. 10% de ces enfants écoutent de la musique et regardent la télévision à un niveau sonore élevé.
- 71% des parents estiment que leurs enfants passent le temps nécessaire dans les activités de loisir; 19% des parents considèrent que leurs enfants passent beaucoup de temps dans les activités de loisir; 5% des parents estiment que leurs enfants passent peu de temps dans les activités de détente; 5% des parents ne donnent aucune réponse.
- 63% des parents considèrent que leurs enfants utilisent une intensité de voix conversationnelle plutôt normale lors des activités de détente. En revanche, 32% des parents pensent que l'intensité de la voix conversationnelle de leurs enfants utilisée pendant les activités de loisir est trop forte.

Dans le groupe ESA :

- 40% des enfants participent à des activités sportives, 20% des enfants pratiquent trois activités de loisir différentes (musicales, sportives et artistiques), 20% des enfants pratiquent deux activités différentes (musicales et/ou sportives et/ou artistiques), 15% des enfants pratiquent une activité musicale, ou artistique, ou autre (activités associatives, religieuses, socioculturelles); 5% des enfants ne pratiquent aucune activité de loisir. Par ailleurs, 3% des parents n'ont pas fourni de réponse.
- Tous les enfants écoutent de la musique et regardent la télévision comme une activité principale dans leur vie quotidienne. 70% de ces enfants écoutent de la musique et regardent la télévision à un niveau sonore élevé.
- 60% des parents estiment que leurs enfants passent le temps nécessaire dans les activités de loisir; 20% des parents considèrent que leurs enfants passent beaucoup de temps; 20% des parents estiment que leurs enfants passent peu de temps dans les activités de détente.
- 50% des parents considèrent que leurs enfants utilisent une intensité de voix conversationnelle plutôt normale lors des activités de détente. En revanche, 50% des

parents pensent que l'intensité de la voix conversationnelle de leurs enfants utilisée pendant les activités de loisir est trop forte ou trop faible.

Dans le groupe ESIC :

- 40% des enfants participent à des activités sportives, 10% des enfants pratiquent trois activités de loisir différentes (musicales, sportives et artistiques), 30% des enfants pratiquent deux activités différentes (musicales et/ou sportives et/ou artistiques); 5% des enfants ne pratiquent aucune activité de loisir.
- Tous les enfants écoutent de la musique et regardent la télévision comme une activité principale dans leur vie quotidienne. 30% de ces enfants écoutent de la musique et regardent la télévision à un niveau sonore élevé et 5% des enfants à faible intensité.
- 60% des parents estiment que leurs enfants passent le temps nécessaire dans les activités de loisir; 25% des parents considèrent que leurs enfants y passent beaucoup de temps; 15% des parents estiment que leurs enfants passent peu de temps dans les activités de détente.
- 90% des parents considèrent que leurs enfants utilisent une intensité de voix conversationnelle plutôt normale lors des activités de détente. En revanche, 10% des parents pensent que l'intensité de la voix conversationnelle de leurs enfants utilisée pendant les activités de loisir est trop forte.

D'une part, nous remarquons que les enfants portant une aide auditive pratiquent majoritairement des activités de loisir collectives. Le port d'une aide auditive leur permet de participer à diverses activités avec leurs pairs y compris celles demandant des performances auditives et vocales. Les activités de loisir en groupe favorisent la communication orale et le processus de socialisation essentiel dans la construction du langage de l'enfant. Le comportement des enfants appareillés et implantés est plus ouvert à l'échange avec le monde sonore, de ce fait les parents observent un changement positif de leur personnalité (cf. *7. Port de la prothèse auditive*). D'autre part, les enfants appareillés semblent avoir l'habitude de regarder la télévision et d'écouter de la musique à une forte intensité. Sur cette activité quotidienne dans la vie de l'enfant, nous nous interrogeons si la prothèse auditive portée par ces enfants est bien adaptée, ou si c'est plutôt que l'enfant ne porte pas sa prothèse pendant cette activité de loisir. Les réponses concernant le port de l'appareillage de ce questionnaire (cf. *7. Port de la prothèse auditive*) mettent en évidence qu'il y a effectivement des enfants notamment appareillés qui n'utilisent pas leurs prothèses à la maison. Par conséquent,

certaines activités comme l'écoute de la musique ou de la télévision se font dans des conditions bruyantes.

4. Manifestation des signes d'anomalie vocale. Quant à la manifestation des signes de troubles vocaux, nous constatons que :

Dans le groupe EE :

- 5% des parents considèrent que la qualité de la voix de leurs enfants est perturbée au moment de l'étude. Par ailleurs, 3% des parents ne donnent pas de réponse.
- 50% des parents ne se sentent pas concernés par la question sur les situations qui pourraient être associées à la manifestation éventuelle des difficultés vocales chez leurs enfants. Concernant 50% des parents qui ont répondu à la question, ils estiment que les difficultés vocales survenant éventuellement chez leurs enfants sont liées à :

  * une situation de stress (13%) ;
  * des maux de gorge ou des périodes de « fragilité laryngée » (13%) ;
  * une situation de malmenage vocal ou un manque de maîtrise de la voix (8%) ;
  * une situation de stress psychologique ayant lieu pendant une période de fragilité laryngée (i.e. maux de gorge) (8%) ;
  * une situation d'abus vocal ayant lieu pendant une période de fragilité laryngée (5%);
  * une situation d'abus vocal (3%) ;

- Les parents qui ont répondu à la question précédente, estiment que le moment de la journée où la voix de l'enfant semble être la plus abîmée est :

  * le soir (33%) ;
  * le matin (11%) ;
  * l'après-midi (3%) ;
  * le matin et le soir (3%) ;

- Ces difficultés vocales se manifestent soit de façon spontanée (24%) soit de façon progressive (13%). Par ailleurs, 13% des parents n'ont pas répondu.
- 61% des enfants manifestent des maux de gorge, rhumes, infections ou allergies respiratoires moins de 3 fois par an; 26% des enfants manifestent ces signes très rarement; 10% des enfants manifestent ces symptômes plus de 3 fois par an. Par ailleurs, 3% des

parents n'ont pas fourni de réponse.

- 45% des parents signalent que la voix de leur enfant ne devient pas « rauque » ou « soufflée » lors de la manifestation des maux de gorge; 41% des parents indiquent que la voix devient occasionnellement rauque et soufflée; 3% des parents signalent la manifestation inévitable de la raucité ou l'accompagnement d'un souffle de la voix; 11% des parents ne donnent aucune réponse.
- 21% des enfants ont subi une intervention chirurgicale nécessitant une intubation; 3% des parents n'ont pas fourni de réponse.
- 82% des enfants n'ont jamais eu de problèmes auditifs importants; 15% des enfants se plaignent de douleurs d'oreille moins de 3 fois par an; 3% des enfants manifestent des douleurs d'oreille plus de trois fois par an.

Dans le groupe ESA :

- 30% des parents considèrent que la qualité de la voix de leurs enfants est perturbée au moment de l'étude.
- 30% des parents ne se sentent pas concernés par la question sur les situations qui pourraient être associées à la manifestation éventuelle des difficultés vocales chez leurs enfants. Concernant 70% des parents qui ont répondu à la question, ils estiment que les difficultés vocales survenant éventuellement chez leurs enfants sont liées à :

  * un déficit auditif (25%) ;
  * des maux de gorge ou des périodes de « fragilité laryngée » (15%) ;
  * une situation de malmenage vocal ou manque de maîtrise de la voix (15%) ;
  * une situation de stress (10%) ;
  * une situation d'abus vocal (5%) ;

- Les parents qui ont répondu à la question précédente, estiment que le moment de la journée où la voix de l'enfant semble être la plus abîmée est :

  * le soir (21%) ;
  * le matin (16%) ;
  * toute la journée (6%) ;
  * sans réponse (27%) ;

- Ces difficultés vocales se manifestent soit de façon spontanée (30%) soit de façon progressive (10%) ; 20% des parents ne savent pas. Par ailleurs, 10% des parents n'ont pas répondu.
- 50% des enfants manifestent des maux de gorge, rhumes, infections ou allergies respiratoires moins de 3 fois par an; 35% des enfants manifestent ces signes très rarement; 15% des enfants manifestent ces symptômes plus de 3 fois par an.
- 45% des parents signalent que la voix de leur enfant ne devient pas « rauque » ou « soufflée » lors de la manifestation des maux de gorge; 40% des parents indiquent que la voix devient occasionnellement rauque et soufflée; 5% des parents signalent la manifestation inévitable de la raucité ou l'accompagnement d'un souffle de la voix; 10% des parents ne donnent aucune réponse.
- 60% des enfants ont subi une intervention chirurgicale nécessitant une intubation; 5% des parents n'ont pas fourni de réponse.
- 40% des enfants n'ont jamais eu de douleurs d'oreille importantes; 30% des enfants se plaignent de douleurs d'oreille moins de 3 fois par an; 25% des enfants manifestent des douleurs d'oreille plus de trois fois par an; 5% des parents n'ont pas fourni de réponse.

Dans le groupe ESIC :

- 25% des parents considèrent que la qualité de la voix de leurs enfants est perturbée au moment de l'étude.
- 30% des parents ne se sentent pas concernés par la question sur les situations qui pourraient être associées à la manifestation éventuelle des difficultés vocales chez leurs enfants. Concernant 70% des parents qui ont répondu à la question, ils estiment que les difficultés vocales survenant éventuellement chez leurs enfants sont liées à :

  * un déficit auditif (15%);
  * des maux de gorge ou des périodes de « fragilité laryngée » (15%)
  * une situation de stress en présence d'un déficit auditif et/ou d'une période de « fragilité laryngée » (15%)
  * une situation de stress (10%)
  * une situation d'abus vocal en présence d'un déficit auditif et/ou d'une période de « fragilité laryngée » (10%)
  * un manque de maîtrise de la voix en présence d'un déficit auditif et/ou d'une période de « fragilité laryngée » (5%)

- Les parents qui ont répondu à la question précédente, ils estiment que le moment de la journée où la voix de l'enfant semble être la plus abîmée est :

  * le soir (35%)
  * le matin (15%)
  * l'après-midi (10%)
  * sans réponse (10%)

- Ces difficultés vocales se manifestent soit de façon spontanée (35%) soit de façon progressive (30%). Par ailleurs, 5% des parents n'ont pas répondu.
- 50% des enfants manifestent des maux de gorge, rhumes, infections ou allergies respiratoires moins de 3 fois par an; 30% des enfants manifestent ces signes très rarement; 20% des enfants manifestent ces symptômes plus de 3 fois par an.
- 45% des parents signalent que la voix de leur enfant ne devient pas « rauque » ou « soufflée » lors de la manifestation des maux de gorge; 40% des parents indiquent que la voix devient occasionnellement rauque et soufflée; 15% des parents signalent la manifestation inévitable de la raucité ou l'accompagnement d'un souffle de la voix; 11% des parents ne donnent aucune réponse.
- 45% des enfants ont subi avant l'implantation cochléaire une intervention chirurgicale nécessitant une intubation.
- 45% des enfants n'ont jamais eu de douleurs d'oreille importantes; 50% des enfants se plaignent de douleurs d'oreille moins de 3 fois par an; aucun des enfants ne manifeste de douleurs d'oreille plus de trois fois par an; 5% des parents n'ont pas fourni de réponse.

L'ensemble de réponses montre qu'il y a un pourcentage de parents d'enfants appareillés et implantés (25% environ) qui perçoit la voix de leurs enfants comme étant altérée. Nous induisons que cette appréciation d'anomalie vocale est associée sans doute aux caractéristiques particulières de la parole de l'enfant sourd. Les parents des enfants ne sont pas des experts mais « des auditeurs candides ».
Revis (2004)[f] démontre l'effet de la sensibilité des auditeurs dans le jugement de la voix. Il se peut que les parents aient jugé plutôt sur les aspects relatifs à la parole qu'aux caractéristiques de la voix proprement dites. Nous avons inclus dans le questionnaire cet item afin d'avoir un aperçu sur le jugement perceptif que les parents pouvaient donner à la voix de leurs enfants.

---

f REVIS Joana, *L'analyse perceptive des dysphonies : approche phonétique de l'évaluation vocale*, Thèse de doctorat, Université de la Méditerranée, 2004.

Si l'on compare globalement les résultats de l'évaluation perceptive effectuée par notre jury d'écoute (cf. section *1.4. Evaluation perceptive*) et les réponses données par les parents ici, nous pourront remarquer qu'il y a une relative tendance des jugements. Par ailleurs, les parents estiment que la voix altérée de leurs enfants est due essentiellement au déficit auditif. Les signes d'anomalie vocale surviennent de façon spontanée le soir et sont notamment perçus comme une raucité et/ou un accompagnement de souffle.

Résumé des informations obtenues pour les items 1, 2, 3 et 4 :

- Il existe deux facteurs environnementaux communs aux trois groupes de population : les habitudes tabagiques parentales et la cohabitation avec un déficient auditif dans le foyer.
- L'appréciation des parents sur les performances vocales de leurs enfants pour communiquer quotidiennement rejoint la tendance que nous avons trouvée dans l'évaluation acoustique instrumentale : la performance vocale de l'ESIC est plutôt semblable à celle de l'EE.
- Les enfants ESA et ESIC pratiquent des activités de loisir majoritairement collectives.
- Un certain nombre de parents d'enfants ESA et ESIC estiment que la voix de leurs enfants est altérée. Des signes d'anomalie vocale sont référés. Ces signes surviennent de façon spontanée le soir et sont notamment perçus comme une raucité et/ou un accompagnement de souffle de la voix.

Deuxième partie

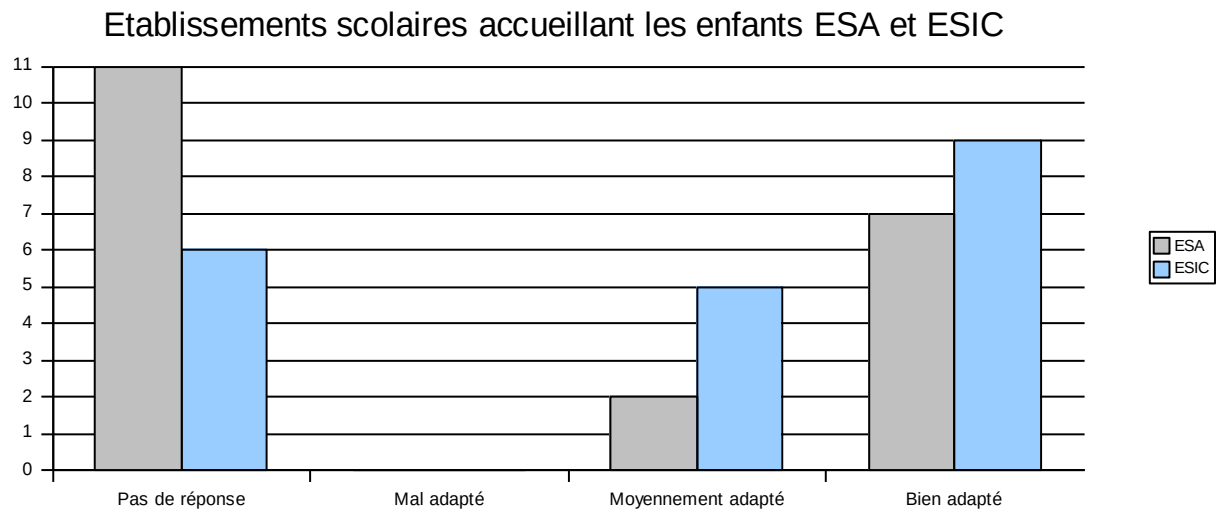
5) Éducation et intégration scolaire. Concernant cet aspect, nous avons seulement enquêté sur les groupes ESA et ESIC.

- Tous les enfants ESA et ESIC sont inscrits dans des établissements scolaires ordinaires.
- 65% des enfants du groupe ESA ainsi que du groupe ESIC suivent un cursus académique classique; 35% des enfants ESA et ESIC reçoivent une éducation partiellement classique avec un aménagement de leur emploi du temps leur permettant de bénéficier de quelques cours adaptés. Les établissements scolaires accueillant les enfants de notre étude mettent à disposition un personnel éducatif spécialisé (éducateurs spécialisés, éducateurs en LPC et en LSF[g], assistants de vie scolaire), un matériel (microphone HF pour les enseignants) et

g LPC = Langage Parlé Complété, LSF = Langue des Signes Française.

des locaux adaptés.

- Un pourcentage non négligeable de parents d'enfants ESA (55%) et ESIC (30%) ne souhaitent pas donner leur avis sur la qualité des moyens mis à disposition par les établissements scolaires pour l'intégration scolaire de leurs enfants (Graphique 7). Les parents des enfants ESIC sont plus nombreux à juger de manière positive les moyens proposés par les établissements scolaires.

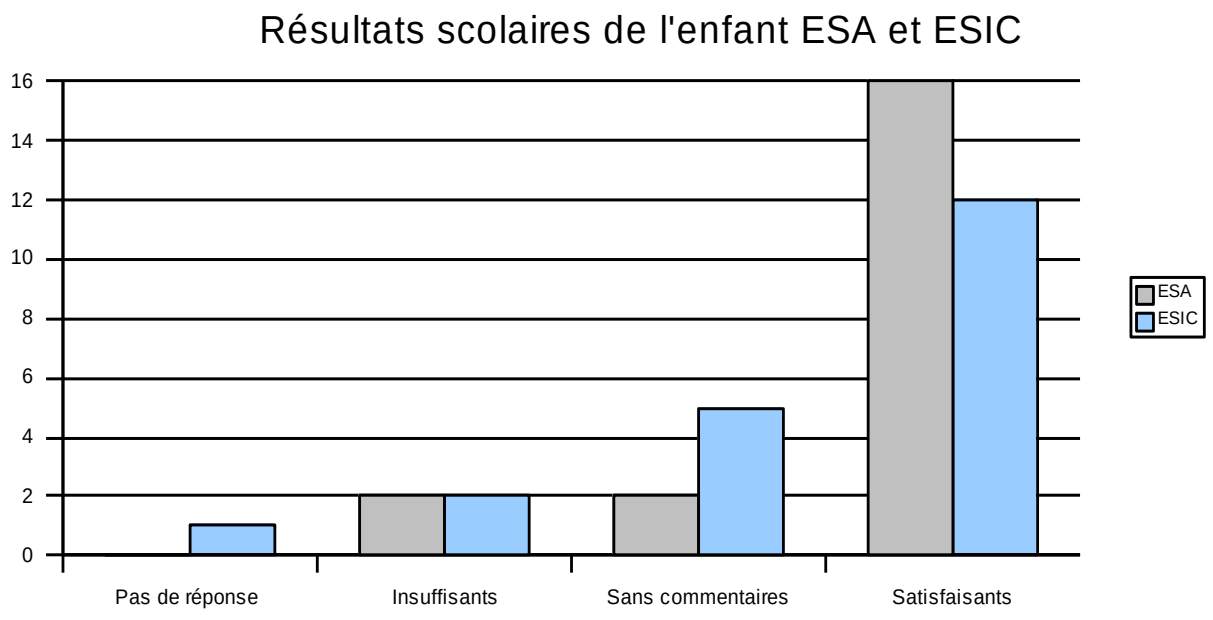


Graphique 7 : Perception des parents des enfants ESA et ESIC sur les établissements scolaires accueillant les enfants.

- D'ailleurs, les parents des enfants ESA (75%) et des enfants ESIC (80%) considèrent que l'équipe éducative est soucieuse en l'occurrence du bon fonctionnement de la prothèse auditive portée par l'enfant durant les cours. Les autres parents ESA (25%) et ESIC (20%) expriment un mécontentement du soin porté par le personnel enseignant à la surveillance du fonctionnement de la prothèse et à l'intégration de l'enfant dans l'établissement scolaire.
- Les parents des enfants ESA (80%) signalent une réussite scolaire de leurs enfants plus élevée que celle signalée par les parents des enfants ESIC (60%), (Graphique 8).
- Tous les enfants et leurs parents semblent être satisfaits de l'accueil et de leur intégration dans les établissements scolaires.

Les résultats montrent que tous les enfants ont été intégrés dans un milieu scolaire entendant. Seulement un tiers de la population bénéficie d'une intégration adaptée. Les parents ne souhaitent pas prononcer de jugement sur la qualité des moyens d'adaptation proposés par les établissements scolaires. Nous nous posons la question sur les motifs de l'absence de réponse à cette question. Il est clair que c'est toujours l'élément éthique et de respect de la vie privée

qui a présidé à la collecte de nos données et qui nous a empêché d'obtenir à tout prix des réponses à notre questionnaire.



Graphique 8 : Perception des parents des enfants ESA et ESIC sur les résultats du travail scolaire et l'intégration éducative de l'enfant dans l'établissement.

Seuls les renseignements livrés en toute simplicité et spontanéité ont été recueillis. Quant à la réussite scolaire, elle est manifestement plus avantageuse chez l'enfant appareillé que chez l'enfant implanté. En quoi cette différence se justifierait-elle ? Une des réponses possibles à cette question réside dans les attentes des parents des enfants bénéficiant de l'implant cochléaire. Les parents oublient parfois la nature et les limites de l'implant. Par conséquent, les parents éprouvent un écart entre leurs attentes et les résultats de leurs enfants en fournissant des jugements sévères.

6) Mode de communication. Comme pour l'aspect de l'éducation et de l'intégration scolaire, nous avons voulu seulement questionner les groupes ESA et ESIC. Puisqu'un des critères de sélection de la population de notre étude était l'utilisation de la langue française orale comme moyen privilégié de communication, tous les enfants ayant participé à l'étude (ESA et ESIC) sont « oralisés ». Nous avons constaté l'emploi d'autres moyens de communication accompagnant l'expression orale :

- Dans l'établissement scolaire, seulement trois enfants ESA (n=20) oralisent à l'aide d'un autre système de communication : un enfant à l'aide du LPC et deux enfants à l'aide de la LSF. Par ailleurs, deux enfants ESIC (n=20) oralisent à l'aide d'un autre système de communication : un enfant à l'aide du LPC et un enfant à l'aide de la LSF.
- Au foyer, la communication est manifestement orale entre les parents et les enfants (80% chez l'ESA et 75% chez l'ESIC). Certains enfants et parents préfèrent communiquer à l'aide d'un autre système de communication : LPC ou LSF (20% chez l'ESA et 25% chez l'ESIC).
- Les enfants ESA et ESIC communiquent avec d'autres membres de la famille et avec d'autres personnes essentiellement par le français oral.
- L'enfant ESA et ESIC s'adresse de manière générale à son interlocuteur dans une conversation de manière spontanée (80% et 85% respectivement).
- 15% des parents des enfants ESA et ESIC déclarent avoir l'impression que les gens sont gênés avec la voix de leurs enfants.

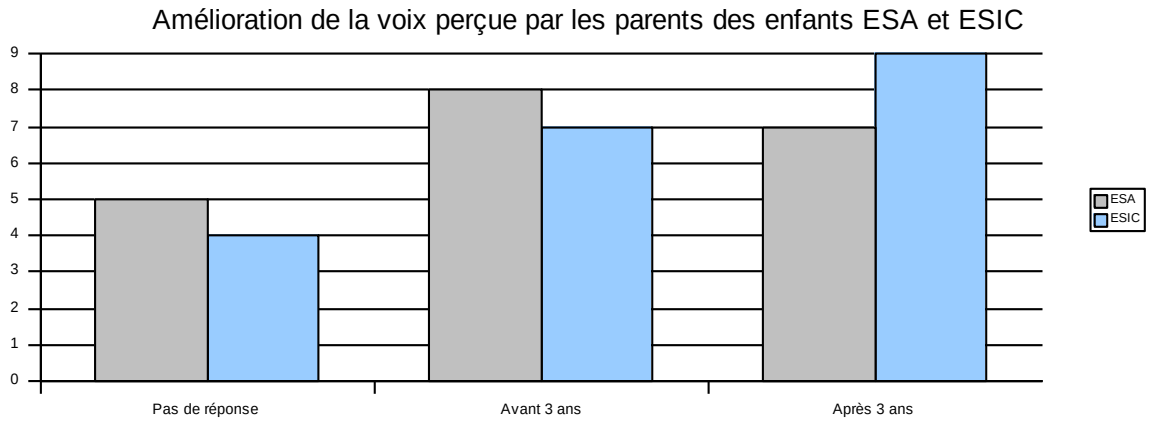
Nous remarquons que l'intégration scolaire des ESA et des ESIC dans le milieu entendant (cf. *5. Education et intégration scolaire*) favorise l'utilisation d'une communication orale dans la vie quotidienne des enfants. Le mode de communication est un facteur qui joue un rôle capital dans le développement des stratégies de communication orale chez l'enfant implanté cochléaire (Tobey et col., 2003***). Certains auteurs affirment que les enfants implantés qui grandissent dans un environnement exclusivement oral développent plus rapidement la parole (Gaul Bouchard, Le Normand et Cohen, 2007[*] ; Moog et Geers, 2003[**] ; Tober et col., 2003[***]). D'autres considèrent que l'usage des « codes manuels » (par exemple, LPC et Langue de Signes) accompagnant la communication ne semble pas avoir un effet préjudiciable sur l'acquisition de la parole (Yoshinaga-Itano, 2000[****]). Uchanski et Geers, (2003)[229] et Spencer (2004)[*****] ont démontré que les programmes d'éducation orale ont un effet bénéfique considérable sur les compétences de production orale. En effet, un mode de

[*] GAUL BOUCHARD, M.E., LE NORMAND, MT., COHEN, H., "Production of consonants by prelingually deaf children with cochlear implants", *Clinical Linguistics and Phonectics*, Vol. 21, 2007, p. 875-884.
[**] MOOG, JS, et GEERS, AE., "Epilogue : Major findings and conclusions and implications for deaf education", *Ear and Hearing*, Vol. 24 Suppl., 2003, p. 121-125.
[***] TOBEY, EA., GEERS, AE, BRENNER, C., ALTUNA, D., et GABBERT, G., "Factors associated with development of speech production skills in children implanted by age five", *Ear and Hearing*, Suppl. 24, 2003, 36-45.
[****] YOSHINAGA-ITANO, C., "Development of audition and speech: Implication for early intervention with infants who are deaf or hard of hearing", *The Volta Review*, Vol. 100, 2000, p. 213-34.
[*****] SPENCER, P.E., "Individual differences in language performance after cochlear implantation at one to three years of age: Child, family, and linguistic factors", *Journal of DeafStudies and DeafEducation*, Vol. 9, 2004, p. 395-412.

communication total utilisé dans la vie quotidienne de l'enfant peut avoir une incidence aussi bien sur la quantité que la qualité de la parole de l'enfant implanté.

7) Port de la prothèse auditive. Quant au port de l'aide auditive dans la vie quotidienne de l'enfant, nous avons observé que :

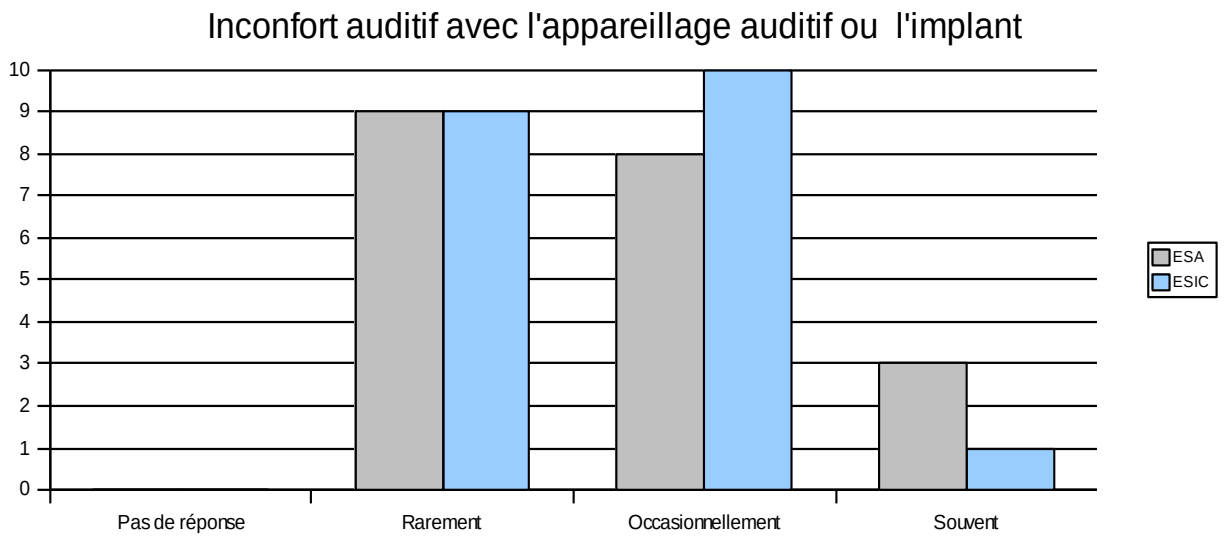
- Tous les enfants ESA et ESIC portent l'appareillage auditif en position ON toute la journée (au moins 7 heures) dans l'établissement scolaire. En revanche, 10% des enfants ESA enlèvent ou débranchent l'appareil auditif pendant une demi-journée (environ entre 4 et 7 heures) à la maison.
- 5% des parents des enfants ESA considèrent que la prothèse auditive classique n'est pas efficace et ne facilite pas la communication orale de l'enfant.
- Les parents des enfants ESA réfèrent que la voix de leurs enfants a été améliorée avant 3 ans d'appareillage alors que les parents des enfants ESIC signalent une amélioration après 3 ans d'implantation (Graphique 9).



Graphique 9 : Perception des parents des enfants ESA et ESIC sur le moment où l'amélioration de la voix de leurs enfants s'apprécie davantage.

- 80% des parents des enfants ESA et ESIC estiment qu'aucun changement ne survient sur la voix de leur enfant quelques temps après les séances de réglage.
- Les parents des enfants ESA (90%) et ESIC (95%) semblent être satisfaits, au moment de l'étude, de l'apport de l'appareillage et de l'implant sur la voix de leur enfant. Les autres parents expriment être toujours inquiets à cause de la qualité légèrement perturbée de la voix de leur enfant. Par ailleurs, 75% des parents ESA et 60% des parents ESIC témoignent que le comportement social et communicatif de leur enfant a changé positivement depuis l'appareillage ou l'implantation.

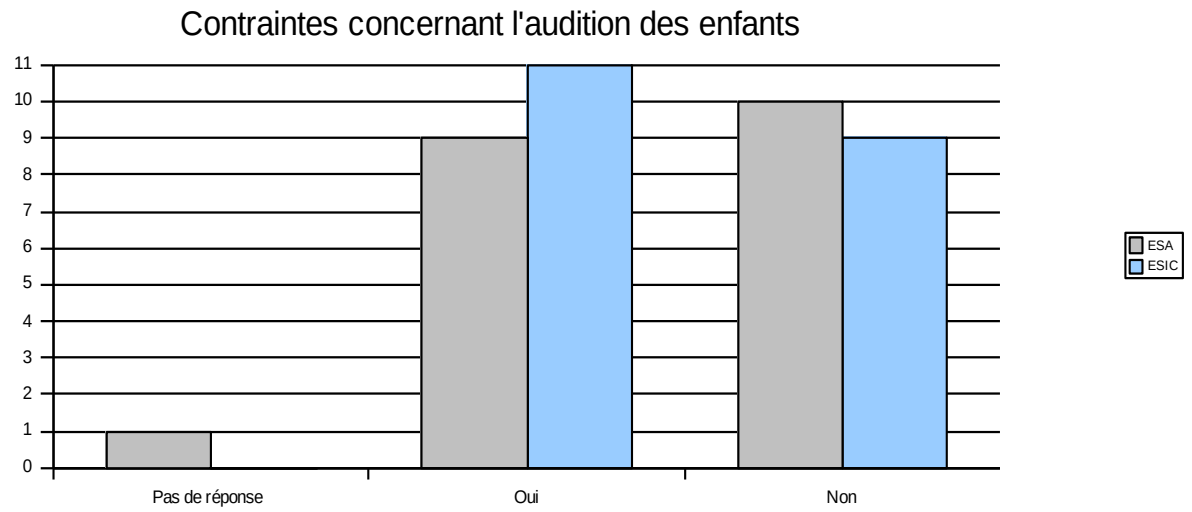
- 45% des enfants ESA et ESIC ne reportent pas de gêne auditive avec l'appareillage ou l'implant (par exemple, l'effet Larsen, le port de la prothèse pendant les activités sportives, entre autres) dans la vie quotidienne (Graphique 10). En revanche, des sensations d'inconfort auditif physique se manifestent occasionnellement chez l'ESA (40%) et chez l'ESIC (50%); l'inconfort auditif se manifeste plus souvent chez l'ESA (15%) que chez l'ESIC (5%).



Graphique 10 : Inconfort auditif occasionné par le port de l'appareillage ou l'implant.

- 45% des parents ESA et 55% des parents ESIC disent avoir toujours des soucis concernant l'audition de leurs enfants. Ces soucis sont liés à la manipulation de l'appareil auditif par l'enfant, les problèmes rencontrés d'ordre technique ainsi que les frais engagés lors des déplacements pour les séances de réglage et suivi orthophonique notamment pour les enfants ESIC (Graphique 11).
- Dans la globalité, les parents des enfants ESIC sont véritablement satisfaits des résultats obtenus avec l'implant cochléaire en ce qui concerne l'amélioration de l'intelligibilité, la compréhension de la parole, ainsi que sur l'aspect esthétique de la voix de leur enfant. Seulement, 15% des parents des enfants ESA expriment un mécontentement de l'apport de la prothèse auditive classique.

Quant au port quotidien de l'aide auditive, il est curieux d'observer que certains enfants ESA se privent volontairement de l'information auditive pendant quelques heures dans la journée. C'est un choix qui est apparemment sous le consentement des parents.



Graphique 11 : Les parents des enfants ESA et ESIC déclarent être confrontés toujours à des difficultés concernant l'audition de leurs enfants.

Lorsque l'enfant ne porte ni active son aide auditive pendant longtemps ou de manière régulière, il perd des informations auditives importantes pour la production vocale. La littérature signale que l'absence d'un retour auditif même après avoir développé un langage et une parole normaux contribue à une détérioration de la production de la parole et de la voix. Cette détérioration concerne la fréquence fondamentale, l'intensité ainsi que la longueur des syllabes et des pauses (Maasen, 1985 ; Lane, 1991 ; Pruszewicz, 1993)[h,i,j]. Higgins et col. (2001) ont observé que le débranchement de l'implant cochléaire pendant 1 heure entrainait une dégradation de la précision articulatoire et du timing de production chez l'enfant sourd prélingual implanté. Il semblerait que certains enfants ayant des compétences excellentes de production et de perception de la parole et de la voix aient une conscience que leur contrôle articulatoire est compromis quand le retour auditif est inaccessible. Ils adapteraient à ce moment là leur comportement articulatoire. Les questions que nous nous posons ici est : pourquoi ces enfants auraient le désir de couper leur lien avec le monde sonore lorsqu'ils se trouvent au foyer ? Est-ce que l'enfant ne prend pas du plaisir à entendre toujours ? Est-ce que cette situation est associée à un problème d'adaptation prothétique ? Nous n'avons pas obtenu d'éléments pour répondre à ces questions.

---

h MAASEN, B., POVEL, D.J., "The effect of segmental and suprasegmental corrections on the intelligibility of deaf speech", *J Acoust. Soc. Am*., Vol. 78, 1985, p. 877-886.
i LANE, H., WEBSTER, J.W., "Speech deterioration in post-lingually deafened adults", *J Acoust. Soc. Am*., Vol. 89, 1991, p. 859-866.
j PRUSZEWICZ, A., DEMENKO, G., WIKA, T., "Variability analysis of Fo parameter in the voice of individuals with hearing disturbances", *Acta Otolaryngol. (Stockh)*, Vol. 113, 1993, 450-454.

Par ailleurs, nous avons observé que les parents d'ESA ont perçu une amélioration de la voix de leurs enfants avant 3 ans d'appareillage alors que les parents d'ESIC ont perçu l'amélioration plus tard. Cet item de notre interrogatoire devrait être posé de manière plus fine afin de déterminer pourquoi cette divergence de perception.

Quant au confort auditif avec la prothèse, ce sont plutôt les parents et non les enfants qui s'expriment sur cet aspect. Un peu plus de la moitié des parents d'ESA et d'ESIC sont soucieux à cause des contraintes techniques et des frais financiers que les aides auditives entrainent notamment l'implant cochléaire. On sait bien que l'implant cochléaire est une prothèse auditive électrophysiologique sophistiquée. Les familles des enfants qui en bénéficient encourent trois types de coûts : financiers médicaux, financiers non médicaux et de temps (Sach et col., 2005)[k]. Ces coûts assez élevés, assumés par les familles, pourraient compromettre éventuellement le suivi post-implant des enfants. Des études sur cet aspect économique de l'implant cochléaire ont été effectuées (Sach, 2002, 2003 ; Cheng, 2000)[l,m,n]. Bien que l'investissement financier pour l'acquisition de ces types de prothèses soit important, les parents des enfants appareillés et implantés de notre étude sont satisfaits des résultats obtenus avec les aides auditives notamment avec l'implant cochléaire.

Résumé des informations obtenues pour les items 5, 6 et 7 :

- La réussite scolaire est plus avantageuse chez l'ESA que chez l'ESIC.
- Certains enfants appareillés et implantés utilisent d'autres modes de communication que l'oral : LPC et LSF.
- Les parents d'ESA ont perçu une amélioration de la voix de leurs enfants avant 3 ans d'appareillage alors que les parents d'ESIC ont perçu une amélioration après 3 ans.
- Les parents des enfants ESA et ESIC sont satisfaits des résultats obtenus avec les aides auditives notamment avec l'implant cochléaire.

---

k SACH, T.H., WHYNES, D.K., ARCHBOLD, S.M., O'DONOGHUE ,G.M., "Estimating time and out-of-pocket costs incurred by families attending a pediatric cochlear implant programme", *Int. J. Pediatr. Otorhinolaryngol.*, Vol. 69, 2005, p. 929-936.

l SACH, T., "Current knowledge and future directions: the economics of cochlear implantation" in *Cochlear Implantation : Cost creating or Cost Saving?*, C. O'Neill ed., Oxford, Proceedings of the Health Technology Assessment Conference, Hughes Associates, 2002, p. 35-70.

m SACH, T.H., WHYNES, D.K., "Measuring indirect costs: is there a problem?", *Appl. Health Econ. Health Policy*, Vol. 2, n° 3, 2003, p. 135-139.

n CHENG, A.K., RUBIN, H.R., POWE, N.R., MELLON, N.K., FRANCIS, H.W., NIPARKO, J.K., "Cost-utiltiy analysis of the cochlear implant in children", *J. Am. Med. Assoc.*, Vol. 284, 2000, p. 850-856.

Annexe 11 : Base de données sur les paramètres aérodynamiques pour le groupe d'enfants implantés cochléaires (ESIC)

Filles					
Groupe d'âge	PSGE (hPa)	DAB (dm³/s)	Intensité (dB)	Eff. glottique (dB/hPa)	Eff. Laryngée (dB/hPa. dm³/s)
5;0 à 6;5 (n=0)	*	*	*	*	*
6;6 à 9,5 (n=2)	10,64 (1,66)	0,145 (0,029)	75 (6,66)	7,07 (0,47)	50,69 (13,73)
9;6 à 12;5 (n=1)	7,48	0,111	70	9,34	84,92
12;6 à 13;5 (n=3)	8,30 (0,75)	0,074 (0,017)	72 (3,74)	8,81 (0,66)	130,77 (25,79)

Tableau 1 : Moyennes et écart-types (entre parenthèses) des paramètres aérodynamiques pour les filles de la population d'enfants implantés cochléaires (ESIC)

Garçons					
Groupe d'âge	PSGE (hPa)	DAB (dm³/s)	Intensité (dB)	Eff. glottique (dB/hPa)	Eff. Laryngée (dB/hPa. dm³/s)
5;0 à 6;5 (n=1)	15,23	0,064	78	5,29	84,14
6;6 à 9,5 (n=7)	9,48 (2,33)	0,132 (0,034)	71 (5,34)	7,88 (1,53)	62,82 (14,19)
9;6 à 12;5 (n=2)	9,91 (3,32)	0,166 (0,081)	77 (6,61)	8,05 (2,25)	62,91 (40,91)
12;6 à 13;5 (n=4)	6,94 (0,96)	0,152 (0,055)	71 (2,67)	10,56 (1,45)	84,04 (35,75)

Tableau 2 : Moyennes et écart-types (entre parenthèses) des paramètres aérodynamiques pour les garçons de la population d'enfants implantés cochléaires (ESIC)

Annexe 12 : Notice d'information pour le jury sur la passation de l'évaluation perceptive

Laboratoire PRAXILING UMR 5267 CNRS - Université Montpellier III
C.H.U. Gui de Chauliac – Service O.R.L. et Chirurgie Cervico-faciale – Unité de Phoniatrie
2007 - 2008

Notice d'information pour le jury sur la passation de l'évaluation perceptive de la voix d'enfants

Cher membre du jury d'écoute,

Vous avez accepté d'intégrer un jury d'écoute en tant qu'auditeur. Ceci ayant pour but d'analyser perceptivement la voix de 78 enfants âgés de 5 à 13 ans. Afin de faciliter le bon déroulement de la séance d'évaluation, je vous demande de lire attentivement la présente notice et de suivre soigneusement les consignes données ci-dessous. Merci par avance pour votre précieuse collaboration et le temps consacré à ces séances d'évaluation.

Informations générales sur la session d'écoute et d'analyse de la qualité de la voix

L'évaluation perceptive est la méthode d'évaluation de la qualité de la voix la plus ancienne. Son principe est d'évaluer à l'oreille et de prononcer des jugements sur la qualité de la voix d'un patient à l'aide d'une échelle de mesure. L'auditeur exprime ses impressions auditives de la voix écoutée en donnant une cotation selon différents paramètres sélectionnés en accord avec les objectifs de l'étude. Le protocole d'évaluation perceptive proposé dans notre étude comprend :

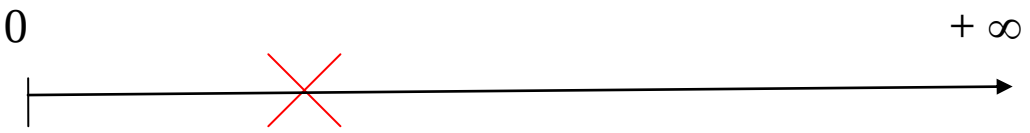
- 3 passations ou sessions d'écoute à effectuer chacune à un intervalle d'une semaine.
- Le mode de passation est individuel. Chaque auditeur évalue séparément les échantillons.
- Le déroulement de chaque séance doit être strictement identique.
- La cotation des voix est effectuée selon l'échelle GRBAS Hirano, l'échelle d'évaluation globale d'intelligibilité de la parole de Svirsky et l'échelle perceptive de résonance nasale de Van Lierde. Les paramètres à évaluer et la technique de cotation sont développés ci-dessous.

Cotation des paramètres évalués

Après l'écoute d'un échantillon de voix, vous devrez noter la qualité de cette voix à l'aide d'une échelle de mesure visuelle analogique selon plusieurs paramètres. Vous exprimerez

votre appréciation par une marque ***« X »*** sur un axe pour chaque paramètre. Les valeurs proches du ***« 0 »*** sont considérées comme des valeurs exprimant la tendance vers la ***normalité*** et les valeurs vers ***« + ∞ »*** exprimant la tendance vers la ***dysfonction, le trouble ou la pathologie*** (cf. ci-dessous).

Voix normale



Vous trouverez un axe pour chaque paramètre. Les paramètres à évaluer sont les suivants :

a. *Concernant l'évaluation globale d'intelligibilité de la parole*, le paramètre **« I »** mesure le grade global d'intelligibilité de production de la parole.

b. *Concernant l'évaluation perceptive de la résonance nasale de la parole*, le paramètre **« N »** mesure le degré de hypernasalité perçue (excès de résonance de voyelles et de consonnes voisées dans les cavités nasales).

c. *Concernant l'échelle GRBAS* :

- **G** : Grade global de dysphonie. Impression globale du degré d'anormalité de la voix.

- **R** : Raucité. Impression d'irrégularité des vibrations de la corde vocale qui correspond aux fluctuations irrégulières de la Fo et/ou à l'amplitude du son glottique.

- **B** : Caractère soufflé. Impression d'une fuite d'air assez importante à travers les cordes vocales, relative à des turbulences.

- **A** : Asthénie. Manque de puissance de la voix relatif à une intensité faible du son et/ou manque d'harmoniques élevés.

- **S** : Forçage. Impression d'un état hyperfonctionnel de phonation relatif à une Fo anormalement haute (bruit dans les hautes fréquences et/ou richesse en harmoniques dans les hautes fréquences).

Les échantillons de voix sont extraits de la parole spontanée émise par les enfants. L'examinateur a demandé à l'enfant de raconter ce que l'enfant a fait pendant le week-end dernier ou les vacances, ainsi que raconter ce qu'il aime faire comme activité de loisir.

Passation d'écoute et d'évaluation des voix

1. Avant de démarrer la session, assurez-vous d'avoir sous la main la fiche de cotation correspondante à la séance d'évaluation à entamer. Ex : Fiche de cotation – Séance No. 1 à remplir que pour la première séance.
2. Ouvrez le fichier *« Matériau_ phonétique_jury d'écoute »*.
3. Ensuite, ouvrez un sous-dossier nommé « 1ère séance_analyse_perceptive ».
4. Vous trouverez tous les *fichiers son* de la séance (78 fichiers= 78 extraits vocaux). Ne renommez ni changez l'ordre d'affichage des extraits car c'est avec ces noms et dans cet ordre qui vous sont présentés les fichiers sur la fiche de cotation pour les évaluer.
5. Cliquez 2 fois sur le fichier et patientez le lancement de votre logiciel pour la lecture du fichier son. Il est possible d'écouter les échantillons via : Windows Media Player, RealPlayer, Quick Time Player ou d'autres logiciels permettant la lecture de fichiers son sous format ***.WAV.** Il est conseillé d'utiliser un logiciel sans une interface de visualisation afin de faciliter votre concentration auditive. Ecoutez l'extrait vocal autant de fois nécessaire avant de noter votre appréciation sur la fiche de cotation.
6. Vérifiez la correspondance du nom du fichier son et du nom de la voix à évaluer. Marquez votre réponse sur l'échelle tel qu'on vous l'a précédemment indiquée.
7. Répétez les manipulations 5-6 pour les fichiers suivants.

L'examinateur conviendra mutuellement avec vous les dates afin de récupérer les fiches de cotation remplies.

<u>Quelques conseils</u> :

- Avant de commencer la session, vérifiez l'état opérationnel de votre matériel informatique (ex : charge électrique de votre batterie si vous travaillez sur un ordinateur portable, le volume de vos enceintes,…).

- Effectuez les séances d'écoute dans un endroit calme et évitez les interruptions pendant la séance afin de garder la concentration.

- La session d'écoute doit impérativement s'effectuer le même jour.

- Respectez l'intervalle d'écoute demandé entre chaque séance d'évaluation (1 semaine). Pour cela, il serait convenable de définir de dates au préalable.

- Si votre séance d'écoute prend plus de 45 minutes, faites une pause de 15 minutes environ avant de reprendre votre session. Autrement dit, faites une pause de 15 minutes après une série d'écoute de 45 minutes.

- Si vous avez besoin éventuellement d'autres explications ou informations complémentaires avant le début de la session, n'hésitez pas à contacter l'examinateur (téléphone portable : 0663965491).

Merci pour votre précieuse collaboration

Harold A. GUERRERO, Orthophoniste
Doctorant en Sciences du Langage, Université Montpellier III

Annexe 13 : Exemple de fiche de cotation d'une séance d'écoute

Laboratoire PRAXILING UMR 5267 CNRS - Université Montpellier III
C.H.U. Gui de Chauliac – Service O.R.L. et Chirurgie Cervico-faciale – Unité de Phoniatrie
2007 - 2008

Séance d'ecoute et d'analyse de la voix

Séance No. 1 **Date de la séance : _____ / _____ / _____**
Nom de l'auditeur : _____

Voix ADJF8b

0 + ∞
I ---------->
0 + ∞
N ---------->
0 + ∞
G ---------->
0 + ∞
R ---------->
0 + ∞
B ---------->
0 + ∞
A ---------->
0 + ∞
S ---------->

Voix ADMF6

0 + ∞
I ---------->
0 + ∞
N ---------->
0 + ∞
G ---------->
0 + ∞
R ---------->
0 + ∞
B ---------->
0 + ∞
A ---------->
0 + ∞
S ---------->

Voix AEBM9

0 + ∞
I ---------->
0 + ∞
N ---------->
0 + ∞
G ---------->
0 + ∞
R ---------->
0 + ∞
B ---------->
0 + ∞
A ---------->
0 + ∞
S ---------->

Annexe 14 : Ecriture d'équations des blocs du modèle - Chapitre IV

1.3.1. Bloc « Vibrateur »

Écriture des équations

Le modèle précédemment décrit va être ici traduit en équations.

a. Modèle d'écoulement de l'air

Pour établir des relations entre la pression venant des poumons, la position des cordes vocales et la pression en sortie des cordes vocales, nous avons fait quelques hypothèses sur l'écoulement de l'air :

- On suppose que le fluide est incompressible et newtonien.
- On suppose qu'on travaille en dehors des couches limites ou des zones de turbulence. Le décollement ou non de la couche limite est prise en compte grâce au paramètre *b* (Figure 1).

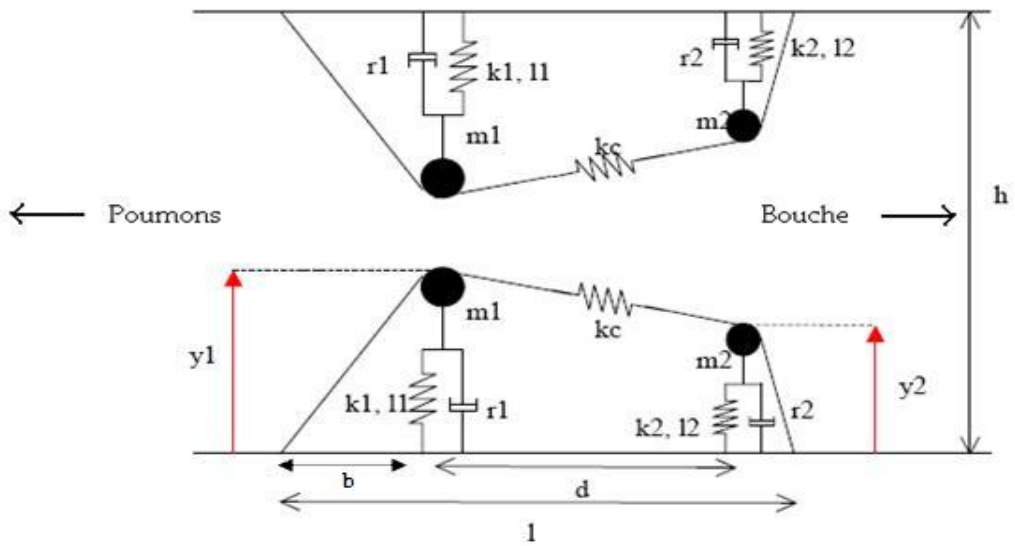


Figure 1 : Schéma représentant les différents paramètres et la structure géométrique du modèle symétrique à deux masses retenu pour notre simulateur.

On utilise deux équations, celle de la conservation du débit, et celle de Bernoulli, applicables d'après les hypothèses faites sur le modèle. Dans ces zones de turbulences on prendra une pression relative nulle car négligeable, et les couches limites n'induisent pas de variations de pression perpendiculairement à la paroi ; la pression sur la paroi sera donc celle du fluide à la même abscisse *x* (Figure 2).

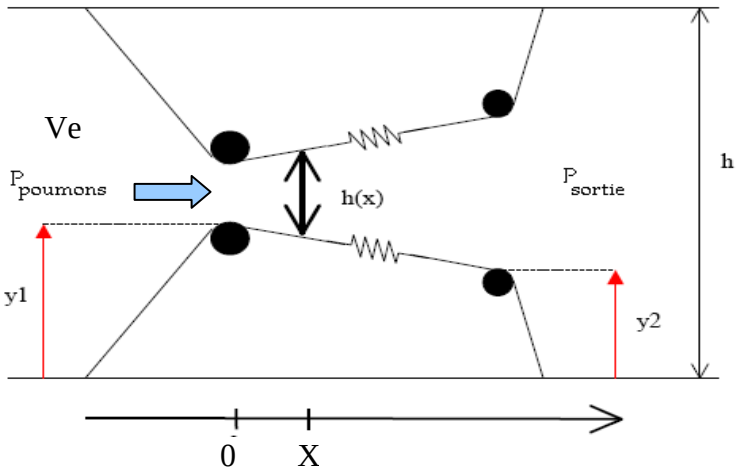


Figure 2 : Schéma du modèle à deux masses (M2M) représentant le sens de l'écoulement de l'air et h(x) pour la mise en équation

La conservation du débit nous permet d'écrire :

$$Ve.h = Ve(x) \ .h(x)$$

et,

l'équation de Bernoulli nous donne :

$$\frac{Ve^2}{2} + \frac{Pe}{\rho} = \frac{V(x)^2}{2} + \frac{P(x)}{\rho}$$

On aboutit ainsi aux équations donnant le débit limité par le conduit, et la pression en toute abscisse *x*. En combinant ces 2 relations, on obtient comme variation de pression :

$$P(x) = \frac{\rho}{2} Ve^2 \left(1 - \frac{h^2}{\left(h - 2y(x)\right)^2}\right) + Pe$$

avec :

$$y(x) = \begin{cases} y_2 + \left(1 - \frac{x}{d}\right)(y_1 - y_2) & si\ y_2 < y_1 \\ y_1 + x\frac{y_2 - y_1}{d} & si\ y_2 > y_1 \end{cases}$$

Ayant l'expression de la pression en tout point de la surface, nous pouvons calculer les forces *F1* et *F2*. La pression va exercer des forces *F1* et *F2* qui vont s'appliquer sur les parois des cordes vocales. On peut calculer ces forces en utilisant la relation :

$$F = p \ . \int_{paroi} P(x)dx$$

Pour rendre possible ce calcul, nous avons discrétisé les parois en *m* éléments de longueur *x*. Sur ces éléments, la pression est considérée comme constante, et la force résulte ainsi de la sommation de l'ensemble des petites forces qui s'exercent sur chacun des éléments.

$$F = \sum_{\text{éléments}} dF = \sum_{i=1}^{m} P(x_i).\Delta x$$

b. Modèle mécanique

Le bilan des forces sur chaque masse nous donne deux équations couplées, à savoir :

$$m1 * \ddot{y}1 + r1 * \dot{y}1 + k1 * (y1 - l1) + kc * \frac{(y1 - y2) * (\sqrt{(y1 - y2)^2 + d^2} - d)^2}{\sqrt{(y1 - y2)^2 + d^2}} = F1$$

$$m2 * \ddot{y}2 + r2 * \dot{y}2 + k2 * (y2 - l2) + kc * \frac{(y2 - y1) * (\sqrt{(y1 - y2)^2 + d^2} - d)^2}{\sqrt{(y1 - y2)^2 + d^2}} = F2$$

Dans la partie à gauche de l'égalité, on reconnaît bien la structure d'une équation différentielle pour un système masse-ressort amorti.

c. Résolution des équations du modèle

La résolution des équations du modèle s'effectue en plusieurs étapes (Figure 3). Ainsi à chaque pas de temps *t*, on calcule la pression au niveau des cordes vocales (à l'origine du son) et la position des cordes vocales.

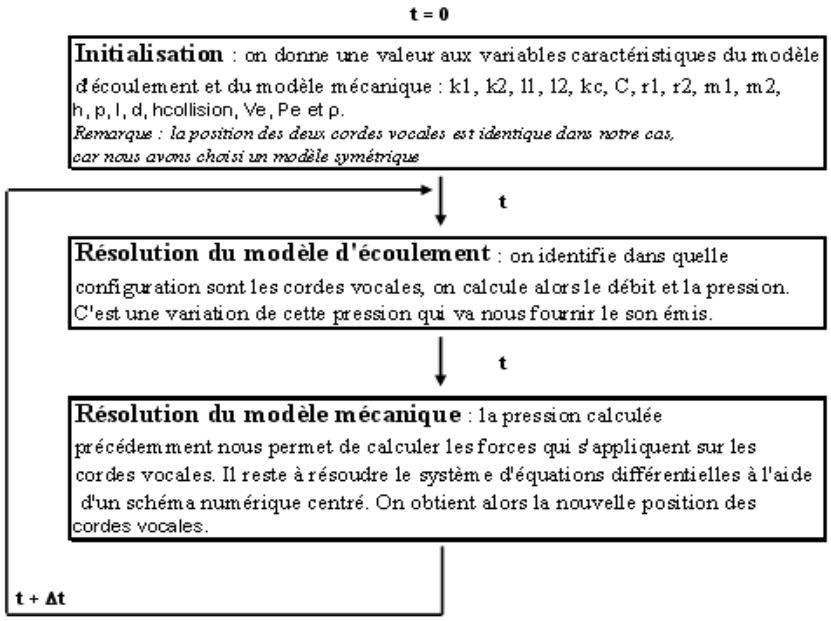


Figure 3 : Architecture de l'algorithme utilisé pour le bloc « vibrateur » ou des cordes vocales.

Dans un premier temps, on a une phase d'initialisation des données et de définition des constantes. Celle-ci permet de rentrer d'une part les différentes caractéristiques du modèle mécanique et de l'écoulement de l'air qui sont les suivantes : *k1, k2, l1, l2, kc, C, r1, r2, m1, m2, h, p, l, d, h collision, Ve, Pe et ρ masse volumique de l'air*. D'autre part, on fixe aussi une géométrie initiale des cordes vocales en entrant une position des masses *y1, y2*[*].

Dans un deuxième temps, grâce aux données de géométrie fournies, l'algorithme de résolution prévoit l'emplacement du point de séparation des cordes vocales selon la configuration glottique sollicitée, à savoir : en fin de conduit s'il y a *convergence*, dans le conduit en cas de *divergence*, et pas de point ni d'écoulement s'il y a *collision*. A partir des hypothèses faites sur le fluide, le modèle permet aussi de calculer le débit d'air et la pression régnant dans le conduit. Ces deux données sont stockées à chaque incrément, car ce sont elles qui vont fournir le son voisé. Connaissant la pression, il est alors possible de connaître quelles forces *F1* et *F2* vont s'appliquer respectivement sur *m1* et *m2*, en répartissant les forces de pression comme suit selon la figure 4.

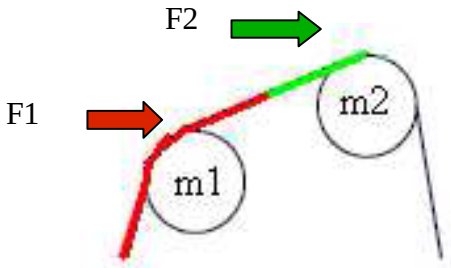


Figure 4 : Répartition des forces de pression sur les masses. Sur m1 vont s'appliquer les forces de pression s'exerçant sur la partie rouge, sur m2 celles s'exerçant sur la partie verte.

Enfin, dans la dernière étape de résolution, l'algorithme effectue le bilan des forces s'exerçant sur les deux masses *m1* et *m2*, comprenant les forces *F1* et *F2* et celles liées aux différentes raideurs et amortissements. On va ainsi avoir un déplacement des masses pendant l'incrément de temps. La nouvelle position des masses est alors donnée au modèle d'écoulement qui va répéter ses opérations avec cette nouvelle géométrie. L'ensemble de cette boucle s'arrête lorsque la durée de simulation est écoulée, et fournit donc les variations de pression et de débit sur cette durée. Il est important de noter que cet algorithme a été développé pour s'intégrer parfaitement à l'ensemble du modèle du macro-système représentant la boucle audio-phonatoire. A présent, l'algorithme peut être développé pour que s'insèrent les

[*] On pose une durée de simulation *T* et un nombre de bouclages *n* pour la simulation numérique. Le nombre de bouclages a une influence directe sur la qualité de la simulation, plus *n* est grand, plus l'incrément de temps est faible, meilleure est la simulation. Ce nombre de bouclages est le nombre de points. Il faut être toutefois attentif au problème de suréchantillonage lorsque le nombre de points est trop important.

différents blocs ou micro-systèmes « résonateur », « récepteur » et « comparateur » dans le bouclage. Le principe est le suivant : à chaque incrément de temps, le son voisé fourni ici sera transformé par un bloc « résonateur », puis ce son sera écouté par un bloc « récepteur » dont les informations permettront au bloc « comparateur » d'agir sur les paramètres *k1, k2, kc, r1, r2, C, Ve, Pe,* paramètres qui sont pour le moment fixés dans la phase d'initialisation.

1.3.2. Bloc « Résonateur »

Écriture des équations

Le conduit vocal modélisé en quatre tubes prend en considération les lèvres, la cavité orale (cavité avant), la constriction de la langue et le pharynx (cavité arrière). Pour le calcul, nous avons utilisé deux types d'équations : fonction d'aire et fonction de transfert. La « fonction d'aire » nous a permis de mettre en relation les caractéristiques géométriques des tubes (surface de la section) en fonction de la distance par rapport à la glotte. La « fonction de transfert » nous a permis de définir les propriétés de filtrage du conduit transformant un signal source (ou son glottique) en signal de sortie. A partir des données d'entrée de la fonction d'aire, il a été possible, par le biais d'équations mathématiques relatives à la propagation des sons et en opérant des simplifications, d'obtenir la fonction de transfert du tube.

On sait que l'équation issue du passage d'une onde acoustique à travers un tube est de la forme:

$$H(w) = \frac{A}{1 + a_1.i.w + a_2.w^2} \quad on\ se\ limitera\ à\ l'ordre\ 2$$

Mais pour savoir à quoi correspondait chaque terme, nous sommes passés par un modèle de schéma-blocs (Figure 5).

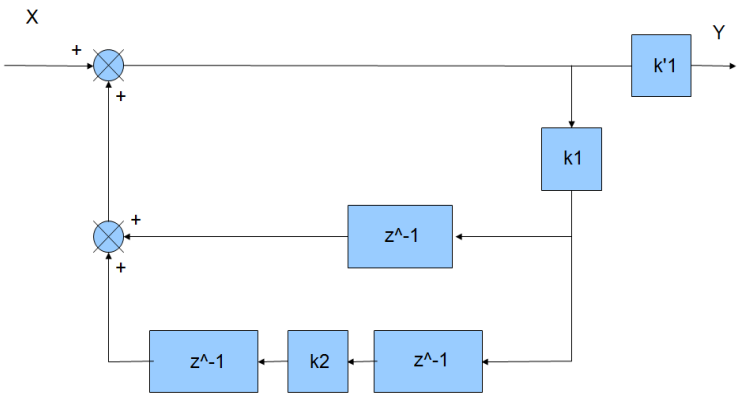


Figure 5 : Schéma-bloc d'un tube modélisé.

On observe sur ce schéma, que l'onde transmise Y est la somme de :

- l'onde incidente X
- l'onde réfléchie sur la paroi 1 avec un coefficient de réflexion $k1$ et subissant un délai $Z\text{-}1$
- l'onde deux fois réfléchie sur les parois 1 et 2 et subissant un délai $Z\text{-}2$

Enfin, Y est transmise au tube suivant avec un coefficient de transmission $k'1=1\text{-}k1$. On obtient ainsi l'équation en *Z* de la fonction de transfert d'un tube. Pour obtenir l'équation de l'ensemble du modèle, c'est-à-dire pour juxtaposer les tubes, il suffit de multiplier les fonctions de transfert. Enfin, les coefficients k et k', correspondent au coefficient de réflexion et de transmission d'un tube à un autre. On les calcule en faisant les rapports des surfaces des tubes et, pour obtenir un modèle de tube continu (Figure 6), on les multiplie par un terme dépendant de $f : (exp^{-cste.f})$[**]. Mais comme nous le verrons plus tard, ces coefficients n'ont pas un impact significatif sur la forme de la solution et rallonge le temps de calcul. Ces coefficients k et k' peuvent prendre des valeurs triviales (0 ou 1) pour le premier et le dernier (ou quatrième) tube. Ainsi, on obtient théoriquement le modèle.

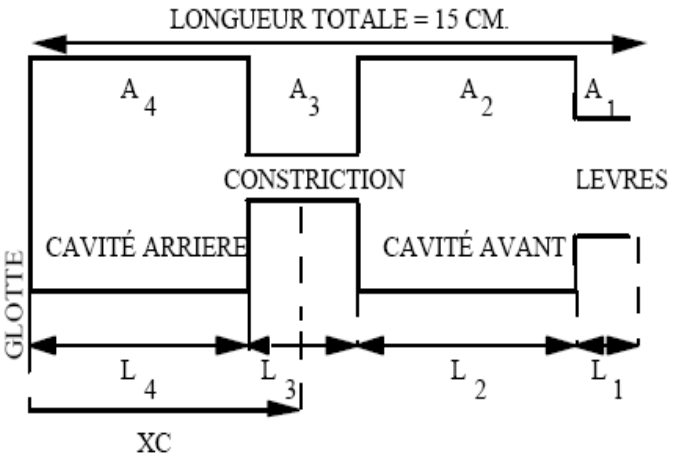


Figure 6 : Modèle de Fant à quatre tubes simulant la section des lèvres (A1), une cavité frontale (A2), l'articulation linguale (A3) et une cavité arrière (A4) (d'après Fant , 1960[*] modifié).

[**] TEFFAFI H, GUERIN B, DJERADI A, « Une nouvelle méthode de mesure de la fonction d'aire du conduit vocal : cas des voyelles du français », *Revue électronique internationale de sciences du langage Sudlangue*, n°4, p. 139-160.

[*] FANT G., *Acoustic theory of the speech production*, The Hague, Netherlands, Mouton, 2nd edition, 1960.

Nous avons cherché à peaufiner et à vérifier la validité des équations du modèle reproduit. Pour cela, nous avons créé une procédure Matlab traçant la fonction de transfert du bloc « résonateur » (Figure 7).

```
k1=0.1; k2=0.1; k3=0.1; k4=0.1; L1=0.2; L2=0.1; L3=0.22; L4=0.13;

for i=1:1000

co2=cos(2*5*i*L1/340);
H1(i)=(1-(1-k1)*co2)/(1-2*(1-k1)*co2+(1-k1)^2);

co1=cos(5*i*L2/340);
co2=cos(2*5*i*L2/340);
si1=sin(5*i*L2/340);
si2=sin(2*5*i*L2/340);
H2(i)=10*k1*(1+(1-k1)*(-co1+(1-k2)*co2))/(1+(1-k1)^2+(1-k1)^2*(1-k2)^2-2*(1-
k1)*co1+2*(1-k1)*(1-k2)*co2-2*(1-k1)^2*(1-k2)*co1*co2-2*(1-k1)^2*(1-k2)*si1*si2);

co1=cos(5*i*L3/340);
co2=cos(2*5*i*L3/340);
si1=sin(5*i*L3/340);
si2=sin(2*5*i*L3/340);
H3(i)=10*k2*(1+(1-k2)*(-co1+(1-k3)*co2))/(1+(1-k2)^2+(1-k2)^2*(1-k2)^2-2*(1-
k2)*co1+2*(1-k2)*(1-k3)*co2-2*(1-k2)^2*(1-k3)*co1*co2-2*(1-k2)^2*(1-k3)*si1*si2);

co1=cos(5*i*L4/340);
co2=cos(2*5*i*L4/340);
si1=sin(5*i*L4/340);
si2=sin(2*5*i*L4/340);
H4(i)=10*k3*(1+(1-k3)*(-co1+(1-k4)*co2))/(1+(1-k3)^2+(1-k4)^2*(1-k3)^2-2*(1-
k4)*co1+2*(1-k3)*(1-k4)*co2-2*(1-k3)^2*(1-k4)*co1*co2-2*(1-k3)^2*(1-k4)*si1*si2);

H(i)=(H1(i)*H2(i)*H3(i)*H4(i))^(1/4);
end;

figure(1);
plot(H);
title(Fonction de transfert globale);
figure(2);
plot([1:1:1000],H1,[1:1:1000],H2,[1:1:1000],H3,[1:1:1000],H4);
title(Fonction de transfert de chaque tube);
```

Figure 7 : Listing du programme Matlab pour le calcul des fonctions de transfert des tubes du bloc « Résonateur ».

En testant différents paramétrages, on observe que :

- la partie imaginaire des fonctions de transfert est négligeable devant la partie réelle.

- le terme dépendant de f (voir ci-dessus) permettant une transition continue entre les tubes n'a que peu d'influence.

On obtient finalement des courbes comme celles ci-dessous (Figure 8, 9).

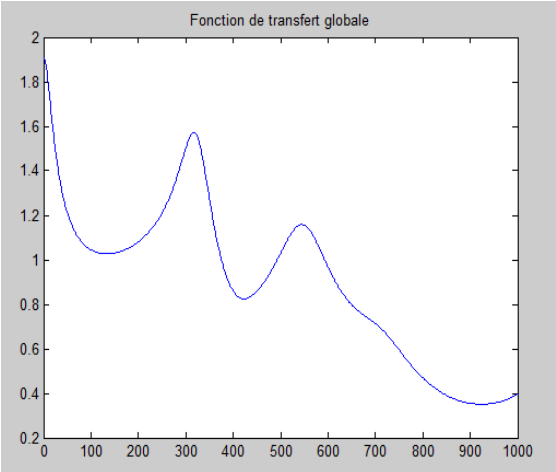


Figure 8 : Courbe représentant la fonction de transfert obtenu pour un résonateur modélisé. En X, les fréquences (Hz); en Y, l'amplitude.

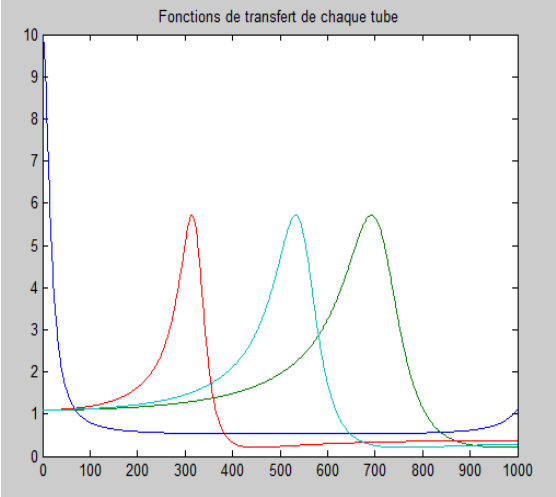


Figure 9 : Courbe représentant la fonction de transfert des différents tubes modélisés. On note que la taille des tubes va influencer la position centrale de pics des courbes de chaque tube. De même, on observe que les variations des coefficients de réflexion et de transmission retentissent sur la largeur des pics.

1.3.3. Bloc « Récepteur »

Écriture des équations

On notera f_c la fréquence centrale du filtre passe bande, f_b la largeur de bande et f_s la fréquence d'échantillonnage dans les équations. Le calcul des équations s'est effectué en plusieurs étapes lesquelles sont présentées comme suit.

a) Calcul de la transformée en Z pour un filtre

Dans un premier temps, on a effectué le calcul de la transformée en *Z* pour les filtres Gammatone. La transformée en *Z*[*] est un outil mathématique de traitement du signal qui est l'équivalent discret de la transformée de *Laplace*. On considère un signal temporel $x(t)$, à partir duquel on peut définir le signal échantillonné $x.(t)$ en ne prenant du signal $x(t)$ que des valeurs à des instants définis. Par le calcul et la simplification des équations en appliquant la transformée de *Laplace*, on obtient alors la transformée en *Z* du signal $x(t)$:

$$X(z) = \sum_{k=0}^{+\infty} x(k.T_e).z^{-k}$$

T_e est la période d'échantillonnage; $t = k . T_e$

b) Modélisation d'un filtre Gammatone

Dans un deuxième temps, nous avons modélisé le filtre Gammatone par la fonction de transfert simplifiée suivante :

$$K(z) = 1/(1 - \tilde{a} . z^{-1})^4$$

Les différents paramètres de cette fonction sont alors contenus dans le coefficient $\tilde{a} = \lambda e^{i\beta}$, où $\beta = 2\pi \frac{f_c}{f_s}$ et $\lambda = -\frac{p}{2} - \sqrt{\frac{p^2}{4} - 1}$ avec $p = \frac{2 - 2.10^{u/10} \cos\Phi}{1 - 10^{u/10}}$ et $u = \frac{3}{4}$.

c) Synthèse d'un filtre Gammatone

Ensuite, nous avons synthétisé un filtre à partir d'une fréquence centrale et d'une largeur de bande. L'approximation tout pôle que nous avons faite limite la synthèse d'un filtre au calcul de $\tilde{a}$. Tout d'abord, la fréquence centrale est donnée par la phase β de $\tilde{a}$. Pour une fréquence centrale et une fréquence d'échantillonnage donnée, le paramètre β prend la valeur :

$$\beta = 2\pi . (f_c / f_s)$$

[*] On l'utilise notamment pour la mise en place de filtres discrets. C'est notre cas ici puisque le modèle cherche à incrémenter sur le temps. Sa définition mathématique est la suivante : la transformée en z est une application qui transforme une suite *S* (définie sur les entiers) en une fonction *S* d'une variable complexe nommée *Z*, telle que : $S(z) = Z(s^*(n)) = \sum s(n) z^{-n}$. La variable *n* représente en général le temps échantillonné. La transformée en *Z* présente l'avantage d'avoir certaines propriétés comparables à la transformée de Laplace, laquelle s'utilise dans le cas continu.

Il représente alors directement l'angle correspondant à f_c dans le plan complexe en z. Pour le calcul du paramètre λ, on trouve une relation avec la largeur de bande fb en calculant le gain de la fonction de transfert à la moitié de la largeur de bande désirée.

Nous avons utilisé la méthode de mesure de la largeur des filtres appelée *largeur de bande rectangulaire équivalente* ou *Equivalent Rectangular Bandwidth* (ERB) proposé par Moore et Glasberg (1990)[*]. Pour chaque filtre auditif, la valeur d'*ERB* est définie comme la largeur d'un filtre passe-bande idéal de même fréquence centrale qui, alimenté par un bruit blanc, produirait la même énergie en sortie. La mesure des *ERB* illustre la résolution fréquentielle du système auditif. La formule suivante donne la relation moyenne pour plusieurs normo-entendants entre la fréquence centrale *f* et *l'ERB* en Hz :

$$ERB(f) = \quad 1 + \frac{f}{q}$$

avec l = 24,7 et q = 9,265

Pour mémoire, les filtres auditifs sont définis pour toute fréquence audible et la largeur des filtres augmente avec la fréquence, de telle sorte que cette largeur vaut approximativement 11% à 17% de la valeur de la fréquence centrale.

d) Synthèse du banc de filtres Gammatone complet

Enfin, nous avons construit le banc de filtres complet. Notre modèle a implémenté 24 filtres branchés en parallèle. Comme nous l'avons auparavant signalé, le nombre de filtres de 24 a été fixé afin d'effectuer une modélisation analogue à celle utilisée pour l'implant cochléaire. Il nous a fallut pour coder ces filtres, connaître le coefficient complexe $\tilde{a}$. Pour cela, nous avons décomposé ce dernier de la façon suivante :

$\tilde{a}$ = real ($\tilde{a}$) + i imag ($\tilde{a}$), où real ($\tilde{a}$) est la partie réelle de $\tilde{a}$ et imag ($\tilde{a}$), sa partie imaginaire.

[*] GLASBERG, BR., MOORE, BCJ., « Derivation of auditory filter shapes from notchednoise data », *Hearing Research*, Vol. 47, 1990, p. 103–138.

Calculer $\tilde{a}$ demandait de connaître les paramètres f_c et f_s. Une fois choisie la valeur de la fréquence d'échantillonnage f_s, il suffisait donc de calculer f_c ou la fréquence centrale de chacun des filtres Gammatone.

« Numéro ERB » et « pas ERB »

Pour la réalisation du banc de filtres complet, nous avons donc dû définir la largeur des filtres. Pour cela, on reprend la notion de largeur de bande rectangulaire équivalente (ERB) de Moore et Glasberg (1990)[217] vue précédemment afin de choisir les fréquences centrales et les largeurs de bande de nos filtres Gammatone. La génération de ces paramètres s'est effectuée de la façon suivante :

- on choisit une fréquence de départ (qui sera la fréquence centrale du premier filtre du banc), ou la fréquence la plus basse du banc de filtre ou f_{min},
- on choisit un *« pas en ERB »* ou l'inverse du nombre de filtres par ERB désiré, ou ERB_{scale},
- on choisit un numéro d'*ERB*, chaque numéro correspondant à une distance fréquentielle d'une bande rectangulaire équivalente :

$$n°ERB(f) = q.\log_{10}(1+\frac{f}{lq})$$

- on choisit le nombre de bandes du banc ou $N_{channels}$.

Cette largeur conditionne le degré de précision de l'analyse fréquentielle du signal (plus le filtre est étroit et plus l'analyse est fine). Nous avons d'abord choisi une fréquence minimale, une fréquence maximale et un *« pas de largeur de filtres »* : f_{min} = 50 Hz, f_{max} = 5000 Hz et un *pas* de 1,5. Le *pas* ERB_{scale} a été choisi de façon à répartir au mieux les N filtres dans l'intervalle de fréquences sélectionnée : différents tests de la fonction *paramètres* (cf. ci-dessous) ont conduit à prendre $ERB_{scale} = 1.5.\ N$, le nombre de filtres étant égal, comme expliqué précédemment, à 24. Enfin, la fréquence d'échantillonnage a été fixée à f_s = 30 000 Hz pour que le théorème de Shannon soit respecté et pour éviter des temps de calculs trop importants. On rappelle que, d'après le théorème de Shannon, pour pouvoir restaurer un signal continu dont est issu un signal échantillonné, la fréquence d'échantillonnage doit être au moins deux fois supérieure à la fréquence maximale contenue dans le signal échantillonné. Nous avons alors pu caractériser chaque filtre.

Le principe du calcul est comme suit : on multiplie le signal temporel du son émis par la fonction de transfert temporelle du banc de filtres complet, qui est en fait la somme des fonctions de transfert de chaque filtre. On récupère ainsi le signal perçu par l'individu, toujours en variable temporelle.

Etant donné que nous avons choisi d'effectuer l'implémentation numérique de notre modèle en langage C++, nous avons traité directement des signaux temporels. Pour cette raison, nous avons utilisée l'expression de la fonction de transfert d'un filtre cochléaire d'après le modèle de Patterson (1987) :

$$G(t) = a.t^{n-1} \cos(2\pi f_c t)\, e^{-2i\pi bt}$$

où :

n désigne l'ordre du filtre, nous avons choisi des filtre *d'ordre 4*,

a désigne la valeur du gain du filtre,

f_c est la fréquence centrale déterminée d'un filtre,

$b = 1.018 * ERB$,

$ERB(f_c) = 24{,}7(1 + 4{,}37\frac{f_c}{1000})$

Implémentation de la fonction « Paramètres »

On écrit une fonction, que l'on nomme « *paramètres* », sous forme de boucle récursive qui calcule pour chacun des filtres la fréquence centrale et le coefficient $\tilde{a}$ sous la forme exprimée auparavant par la fonction. « *Paramètres* » a comme arguments d'entrée f_{min}, ERB_{scale}, N et la fréquence d'échantillonnage f_s ; et comme arguments de sortie le coefficient $\tilde{a}$ de chacun des N filtres.

Reconstitution et retard du signal

Une fois le banc de filtres créé, le signal temporel a été recomposé après passage dans la cochlée. Ceci ne peut se faire par simple addition car le signal obtenu serait trop dégradé. En effet, la réponse impulsionnelle de chaque filtre a une structure fine et une phase différente, ce qui rend la réponse impulsionnelle du banc de filtres trop dispersive. On a donc dû, en sortie de chaque filtre, multiplier le signal par un retard dépendant du filtre considéré, de façon à faire coïncider les maxima des 24 filtres. On choisit de faire coïncider les maxima à $T = 80$

ms, cet instant étant légèrement postérieur à l'instant auquel arrive le maximum le plus tardif parmi les 24 maxima.

En reconstruisant le signal temporel, on peut obtenir alors le son tel qu'il est perçu par « l'oreille ». Puisque les calculs sont faits en temps réel, on peut utiliser le banc de filtres dans une boucle temporelle et le faire fonctionner dans un simulateur comme nous le souhaitions.

Pondération

Nous avons effectué trois types de pondération : 1) une « pondération avant », 2) une « pondération après », et 3) une « pondération extérieure ». La « pondération avant » permettait d'affecter un gain précédant pour sommer les participations de chaque filtre et ainsi réduire au maximum la distorsion du signal. Initialement, les maximas des signaux sont alignés. La « pondération après » permettait de compenser les défauts dus aux interférences constructives afin d'obtenir une réponse fréquentielle la plus plate possible en sortie. Les coefficients de pondération ont tous été donnés égaux à 100. Enfin, la « pondération extérieure » facilitait le paramétrage et la modification du bloc « Récepteur » de façon à adapter la modélisation aux défauts de perception désirés. Pour cela, on dispose à la suite de chaque bloc « pondération après », d'un bloc « pondération extérieure » auquel on affecte la valeur *pc (i)*, *i* étant le numéro du filtre correspondant, *i* varie de 1 à 24. On a donc 24 blocs « pondération extérieure », auxquels correspond dans le programme principal un vecteur *pc* de dimension 24. Les coefficients de ce vecteur sont des valeurs comprises entre *0* et *100*, *0* si le filtre correspond à un intervalle de fréquences non perçues, *100* si au contraire il s'agit d'un des intervalles de fréquences les mieux perçues. Ces valeurs peuvent être alors saisies directement. Nous avons effectivement implémenté le banc de filtres complet. La structure du modèle travaillé se présente de la façon suivante :

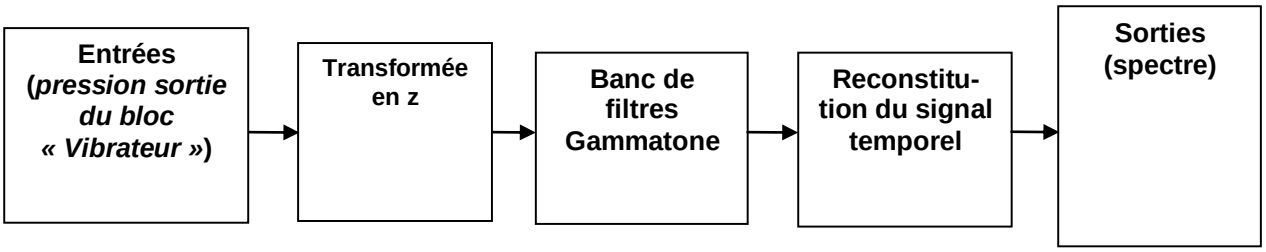


Figure 13 : Schéma de la structure du modèle du bloc « Récepteur » développé dans nos travaux.

1.3.3. Module de calcul du banc d'essai

a) Concernant le stockage de l'information du signal sonore, nous avons associé à chaque bloc du logiciel une manière de stocker les données. Le son est enregistré de deux manières différentes :

- sous forme d'un tableau de 2 000 valeurs représentant les variations de pressions en fonction du temps sur 0,5 s.
- sous forme de deux tableaux de 1 000 valeurs représentant les parties réelles et imaginaires de la transformée de Fourier du signal de pression (les valeurs oscillent entre 0 et 5 000 Hz).

A part, le bloc « Vibrateur », les blocs de calcul fonctionnent avec une transformée de Fourier, et il en est de même pour la création de graphiques. Néanmoins, l'import et l'export de fichier **.wav* se fait à partir des variations de pression en fonction du temps. Ainsi, pour une simulation, on a donc au minimum deux transformations de Fourier et une transformation de Fourier inverse. Ces opérations prennent un temps de calcul considérable, compte tenu de la taille des tableaux de données (Figure 14).

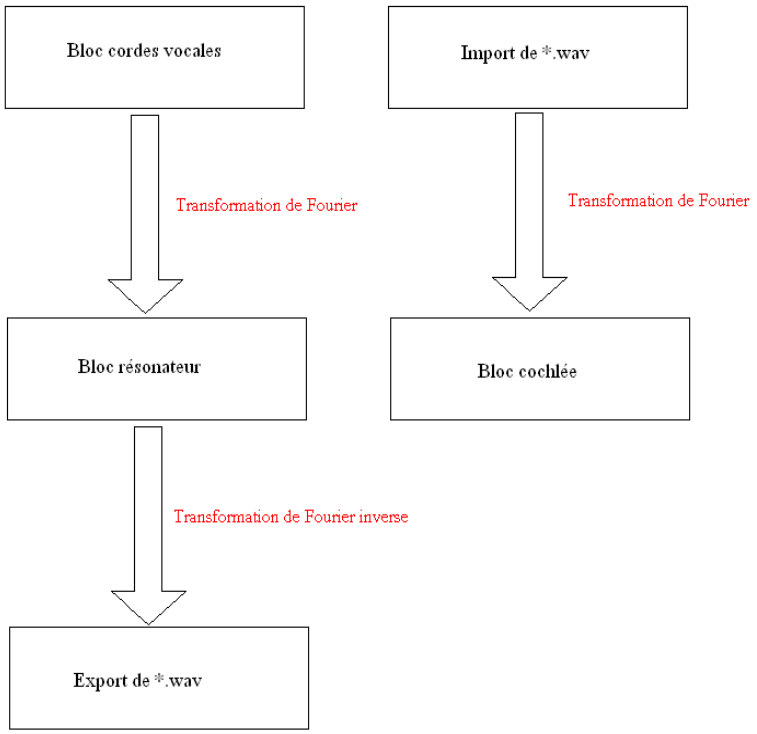


Figure 14 : Liaison entre les différents blocs du banc d'essaie pour le stockage de l'information

b) Quant à l'algorithme de la transformée de Fourier, il a été implanté à partir de la formule suivante :

$$x(n) = \sum_{k=0}^{N-1} x(k) e^{2i\pi kn/N}$$

On obtient alors le code suivant :

```
for(int i=0;i<n/2;i++)
{
        fftpressionR[i]=0;
        fftpressionI[i]=0;
        for (int j=0; j<n;j++)
        {
                fftpressionR[i]=fftpressionR[i]+pression[j]*cos(2*3.1415*i*j/n)/n;
                fftpressionI[i]=fftpressionI[i]-pression[j]*sin(2*3.1415*i*j/n)/n;
        }
}
```

Ici « fftpressionR » et « fftpressionI » désignent les tableaux contenant respectivement la partie réelle et la partie imaginaire de la transformée de Fourier de “pression”. Le temps de calcul de cet algorithme est proportionnel au carré de la taille des tableaux traités. Par conséquent travailler avec des tableaux de données plus grands aurait un énorme impact sur le temps de simulation du logiciel, c’est pour cela que nous nous sommes limités à des fréquences variant de 0 à 5000 Hz.

c) Concernant les blocs de calcul du banc d’essai, le *bloc Résonateur* et le *bloc Récepteur* fonctionnent tous les deux d'une manière similaire car ils correspondent tous deux à l'application d'un filtre sur le signal de pression. Dans un premier temps, on calcule la fonction de transfert du filtre. Ici, pour le *bloc Résonateur*, la fonction de transfert est le produit des fonctions de transfert des quatre tubes H1, H2, H3 et H4 (Figure 15).

```
for(int i=0;i<1000;i++)
{
    co2=cos(2*3.1415*2*5*i*L1R/340);
    H1=(1-(1-Q1R)*co2)/(1-2*(1-Q1R)*co2+pow((1-Q1R),2));

    co1=cos(2*3.1415*5*i*L2R/340);
    co2=cos(2*3.1415*2*5*i*L2R/340);
    si1=sin(2*3.1415*5*i*L2R/340);
    si2=sin(2*3.1415*2*5*i*L2R/340);
    H2=10*Q1R*(1+(1-Q1R)*(-co1+(1-Q2R)*co2))/(1+pow((1-Q1R),2)+pow((1-Q1R)*(1-Q2R),2)-
2*(1-Q1R)*co1+2*(1-Q1R)*(1-Q2R)*co2-2*pow((1-Q1R),2)*(1-Q2R)*co1*co2-2*pow((1-Q1R),2)*(1-
Q2R)*si1*si2);

    co1=cos(2*3.1415*5*i*L3R/340);
    co2=cos(2*3.1415*2*5*i*L3R/340);
    si1=sin(2*3.1415*5*i*L3R/340);
    si2=sin(2*3.1415*2*5*i*L3R/340);
    H3=10*Q2R*(1+(1-Q2R)*(-co1+(1-Q3R)*co2))/(1+pow((1-Q2R),2)+pow((1-Q2R)*(1-Q3R),2)-
2*(1-Q2R)*co1+2*(1-Q2R)*(1-Q3R)*co2-2*pow((1-Q2R),2)*(1-Q3R)*co1*co2-2*pow((1-Q2R),2)*(1-
Q2R)*si1*si2);

    co1=cos(2*3.1415*5*i*L4R/340);
    co2=cos(2*3.1415*2*5*i*L4R/340);
    si1=sin(2*3.1415*5*i*L4R/340);
    si2=sin(2*3.1415*2*5*i*L4R/340);
    H4=10*Q3R*(1+(1-Q3R)*(-co1+(1-Q4R)*co2))/(1+pow((1-Q3R),2)+pow((1-Q3R)*(1-Q4R),2)-
2*(1-Q3R)*co1+2*(1-Q3R)*(1-Q4R)*co2-2*pow((1-Q3R),2)*(1-Q4R)*co1*co2-2*pow((1-Q3R),2)*(1-
Q3R)*si1*si2);

    H=abs(H1*H2*H3*H4);
```

Figure 15 : Algorithme de calcul de la fonction de transfert des filtres pour le bloc Résonateur. La fonction de transfert est le produit des fonctions de transfert des quatre tubes : H1, H2, H3 et H4.

Puis on multiplie la transformée de Fourier du signal par cette fonction de transfert :

```
    fftpressionR[i]=fftpressionR[i]*H;
    fftpressionI[i]=fftpressionI[i]*H;
}
```

Pour le bloc Cochlée, le calcul est similaire et n’est pas détaillé ici.

d) A l'aide de l'application *iostream de C++*, que nous ne développerons pas ici, il est possible de récupérer dans un tableau d'entiers les valeurs des échantillons encodés dans un fichier son. C'est en travaillant directement sur ces valeurs que nous avons pu réaliser l'émission d'un son. Un fichier **.wav* se découpe en deux parties. La première partie est le fichier d'en-tête, il est composé de 56 octets et il contient diverses informations sur le fichier : sa nature (ici fichier **.wav*, *mono* ou *stéréo*…), sa taille globale, sa fréquence d'échantillonnage. Pour pouvoir créer un fichier son, il faut jouer sur l'en-tête du document[*]. La deuxième partie est la succession des valeurs des échantillons exprimés en entier compris entre 0 et 255 (chaque échantillon est codé sur un octet). Il faut savoir qu'on travaille souvent en hexadécimal car un octet (8 bits) est codé avec deux caractères hexadécimaux.

L'export des sons a été effectué à partir d'un tableau de nombres réels issus du calcul. Pour cela, on commence par fabriquer un nouveau tableau d'entiers, dans lequel on met le fichier d'en-tête sur les octets adéquats, et avec les caractéristiques qui nous conviennent, par exemple, la taille de l'échantillon (2000 points : 00 00 07 D0) et la fréquence d'échantillonnage (10 000Hz : 00 00 27 10). Par ailleurs, on doit transformer les valeurs réelles de l'échantillon en entiers. Pour cela, on applique juste une loi proportionnelle pour se ramener sur 0-255 puis une troncature de chaque valeur pour passer en entier. On perd ainsi l'information sur l'intensité du signal (volume) mais on conserve toutes les données fréquentielles. On réalise ensuite, grâce à la procédure *iostream* de C++, la génération d'un son au format **.wav* directement à partir du tableau d'entiers. Nous avons ainsi créé un son que l'on peut écouter avec un simple lecteur (type lecteur Windows Media).

De la même façon, nous avons voulu importer un son pour pouvoir disposer d'un *« son consigne »* que l'on comparera au « son émis ». Pour cela, nous avons récupéré avec la procédure *iostream* le tableau de valeurs entières du son. On lit dans l'en-tête quelle est la fréquence d'échantillonnage de ce fichier. On peut ainsi récupérer le nombre d'échantillons qui correspond à une fréquence de 10 000Hz (1 échantillon sur 4 par exemple si le fichier son en entrée est à 40kHz). Cette procédure a été automatisée pour n'importe quel fichier **.wav* d'entrée. Ensuite, on transfère les valeurs entières sur l'échelle des valeurs du son calculé (ou *« son émis »*) et, on perd l'information sur l'intensité du signal en conservant l'information fréquentielle. On a ainsi créé un tableau de valeurs réelles contenant l'information fréquentielle sur notre signal de consigne. On peut alors se servir de ce tableau pour la comparaison avec le *« son émis »*.

[*] http://col2000.free.fr/vocal/formawav.htm

Nous avons donc réussi à importer et à exporter des sons au format **.wav* dans le logiciel. Ceci était nécessaire au fonctionnement du logiciel ARTIVox v4.0, et permet, en plus, de simplifier la compréhension du résultat obtenu.

Annexe 15 : Résultats de l'observation clinique du comportement vocal et de la qualité de l'émission sonore d'un /a/ tenu

1. Objectif

L'évaluation du comportement vocal est un des éléments du bilan vocal orthophonique qui contribue au diagnostic de la dysphonie (Estienne, 1998[230]; Faure, 1987[231]; Amy de la Bretèque, 1997[232]; Heuillet-Martin, 1995[233]; Roubeau, 2002[234]). Puisque l'un des objectifs de notre travail était de confirmer si les voix des groupes ESA et ESIC étaient dysphoniques, nous l'avons inclus dans le protocole d'évaluation PEMVox (cf. Annexe 8 et 9). Malheureusement, nous n'avons pas effectué un traitement statistique de cette observation clinique pour affirmer ou infirmer l'existence d'une différence entre les trois groupes, et la corrélation avec les paramètres acoustiques. Etant donné que nous avons axé notre travail principalement sur les approches perceptive et acoustique, nous nous sommes vus contraints, pour des raisons de temps, de délaisser ces éléments d'analyse.

2. Technique d'évaluation

L'évaluation du geste vocal a été obtenue à partir de l'observation de différents aspects nécessaires à la production vocale tels que la posture, la tension musculaire et la respiration. L'enregistrement vidéo couplé à l'enregistrement audio est un excellent outil d'analyse du geste vocal. Il permet de visualiser le geste vocal autant de fois que possible et de le décomposer également en segments. Cependant, l'acquisition de ce type de matériel et sa mise en place demandent des disponibilités financières et temporelles importantes que nous n'avions pas à disposition.

C'est pourquoi, nous avons effectué une observation subjective (de face et de profil) comme en pratique clinique, de l'attitude posturale et du comportement respiratoire utilisés par l'enfant pour les différentes situations de phonation (voix conversationnelle, voix forte et d'appel) pendant l'enregistrement.

2.1. Posture

En ce qui concerne la posture, elle peut être définie selon la position des différents segments corporels à un moment donné. La « position debout » est considérée comme la posture de référence. Il existe plusieurs variantes dans cette posture de référence, la position debout demandant le moins de travail musculaire est dénommée position debout « commode » (Bouisset et Maton, 1995)[235]. Pour cette posture de référence, la ligne de gravité passe par le

sommet du crâne, le lobe de l'oreille, la pointe de l'épaule, l'articulation de la hanche, le centre de l'articulation du genou et légèrement en avant de l'articulation de la cheville (Figure 1) (Bouisset et Maton, 1995[235]; Kendall et cols. 2005[236]).

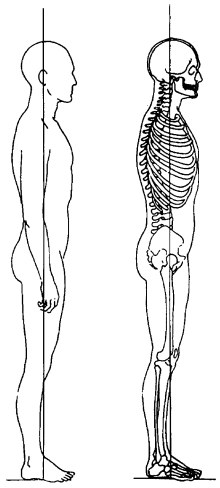


Figure 1: Vue de profil de la posture débout dite « commode ». On note la ligne de gravité passant par les différentes parties du corps humain.

a) Pour la voix conversationnelle, nous avons apprécié le geste vocal adopté en fonction de la posture et de la statique corporelle. L'axe de posture était évalué selon les critères suivants.

- *vertical* : si la position « commode » était observée,

- *vers l'avant* : s'il y avait une projection vers l'avant du visage et une flexion de la partie haute du thorax aboutissant à une perte de verticalité.

- *vers l'arrière* : s'il existait une flexion vers l'arrière du rachis cervico-thoracique aboutissant à une perte de verticalité.

- *instable* : s'il y avait constamment une perte de la verticalité.

La position des différents segments corporels lors de l'émission vocale a été évaluée selon les critères suivants.

- Concernant *la tête,* nous avons qualifié la position de *normal,* si la position du massif facial était sur la ligne virtuelle de l'axe de posture « commode »; *étirée,* s'il y avait une avancée du menton; *relâchée,* si la tête était abaissée et détendue.

- Concernant *les épaules,* nous avons qualifié la position de *normal,* si la position de la pointe était sur l'axe de posture « commode »; *étirée,* si la position de la pointe était trop en avant ou en arrière manifestant une lordose cervicale ou un redressement du corps important; *élevée,* si la position de la pointe des épaules était précédée d'un mouvement d'ascension.

- Concernant *le dos,* il était considéré *normal,* si la position était verticale ou sur l'axe de posture « commode »; *étirée,* si la position était trop en avant ou en arrière manifestant une lordose lombaire, un redressement du corps important ou une malformation telle qu'une cyphose dorsale.

- Concernant *les bras,* il était considéré *normal,* si les bras étaient spontanément placés le long du corps reposés et la paume des mains orientées vers les jambes; *croisés,* si les bras étaient placés en avant ou en arrière du corps et les mains étaient serrées l'une contre l'autre entraînant une légère tension musculaire.

- Concernant *les jambes,* il était considéré *normal,* si les jambes étaient spontanément séparées à une certaine distance permettant l'équilibre des points de stabilité et l'acquisition de la verticalité du corps; *très serrées,* s'il n'y avait pas de distance de séparation entre les jambes; *très écartées,* s'il y avait une distance de séparation assez importante entraînant un déséquilibre.

- Concernant *les talons,* il était considéré *normal,* si l'appui au sol était complet; *en pointe,* si le contact avec le sol était notamment antérieur; *en talon,* si l'appui au sol était notamment postérieur; *latéral,* si le contact avec le sol était notamment latéral.

- Concernant *les pieds,* il était considéré *normal,* si les pieds étaient spontanément placés séparés et tournés vers l'extérieur formant un angle de 30° permettant l'équilibre et l'axe de verticalité; *très serrés,* s'il n'y avait pas de distance de séparation entre les pieds; *très écartés,* s'il y avait une distance de séparation assez importante entraînant un déséquilibre.

b) Pour la voix forte, nous nous sommes également portés à observer la posture et la statique corporelle précédemment décrites.

c) Pour la voix d'appel, l'observation était seulement portée sur la posture et le mécanisme vocal utilisé. Concernant le mécanisme vocal, nous avons également déterminé perceptivement :

- *mécanisme 1* : si l'enfant utilisait une « voix de poitrine » ou aggravée (obtenue par l'accolement de toute l'épaisseur des cordes vocales).

- *mécanisme 2* : si l'enfant utilisait une « voix de tête » ou aiguë (obtenue par l'accolement fin du bord libre des cordes vocales).

2.2. Comportement respiratoire

En ce qui concerne l'évaluation du comportement respiratoire, il a été observé dans la phase pré-phonatoire juste avant l'émission de la voix conversationnelle. La respiration est physiologiquement de type diaphragmatique (sauf dans les cas d'insuffisance respiratoire d'étiologie traumatique, neurologique). Les variations de la « statique corporelle » et les « hyper » ou « hypo » tensions musculaires amènent une modification des ouvertures costales normalement nécessaires à l'ouverture thoracique et autorisant la meilleure expansion pulmonaire lors de l'abaissement du diaphragme à l'inspiration (chapitre II, section 1.3). Elles peuvent aussi modifier la dynamique expiratoire par contraction excessive de la sangle abdominale en provoquant un soulèvement anormal du haut de la cage thoracique. Cette observation peut se pratiquer au repos et pendant la production sonore. Nous avons seulement effectué l'observation du type respiratoire pendant la phase pré-phonatoire. Lorsqu'on analyse le type respiratoire et la qualité de la production sonore (attaque du son, tenue et sortie), on apprend sur l'efficacité du comportement vocal malgré les dysfonctionnements de l'individu (Guerrier et cols., 2004[237]). Selon la localisation de l'amplitude du mouvement inspiratoire pendant la respiration, nous avons déterminé le *type respiratoire* :

- *scapulaire* : s'il y avait une ascension et un enroulement des épaules,

- *costal supérieur* ou *thoracique supérieur* (Faure, 1987[231]) : s'il manifestait un ou plusieurs signes physiques tels qu'un tirage cervical, tension des sterno-cléïdo-mastoïdiens et des

scalènes, hyperlordose cervicale, tensions de la musculature du visage, rétractation de la paroi abdominale,

- *costo-abdominal* ou *costo-diphragmatique* (Guerrier et col., 2004[237]) *ou complet* (Faure, 1987[231]) : s'il y avait une mobilisation globale, facile et harmonieuse de la cage thoracique et du diaphragme,

- *abdominal* : s'il y avait une mise en tension diaphragmatique.

Selon la façon prédominante dont l'inspiration s'achevait durant le cycle respiratoire, nous avons déterminé le *mode respiratoire* :

- *prédominance nasale* : si la prise d'air s'effectuait notamment par la cavité nasale, un mouvement d'expansion des narines suivant la force d'aspiration de l'air inspiré était observé.

- *prédominance orale* : si la prise d'air s'effectuait notamment par la cavité orale, une ouverture orale prononcée suivant la force d'aspiration de l'air inspiré était observée.

De même, nous avons évalué la fréquence respiratoire, la perméabilité nasale et la tonicité musculaire mais cette fois-ci au repos et en position assise.

- *Fréquence respiratoire* : Nous avons demandé à l'enfant de respirer tranquillement les yeux fermés (ou en regardant ailleurs) afin d'éviter une accélération ou un ralentissement de la fréquence respiratoire. L'examinateur comptait le nombre de cycles respiratoires par minute en fonction des mouvements extérieurs de la cage thoracique.

- *Perméabilité nasale* : Nous avons demandé à l'enfant d'émettre un /a/ tenu à hauteur et intensité confortables pendant quelques secondes. L'examinateur plaçait le miroir de GLATZEL* sous les narines et légèrement posé sur la lèvre supérieure de l'enfant pendant

* Le miroir de GLATZEL est un miroir métallique permettant d'objectiver la déperdition nasale. Il est par ailleurs, fréquemment utilisé dans la rééducation.

l'émission du /a/ tenu. Il observait l'échappement d'air par une ou deux narines se produisait par la présence de buée sur le miroir.

- *Tonicité musculaire* : Nous avons observé pendant l'évaluation de la fréquence respiratoire si le tonus musculaire du visage et du cou était *« hypo »* (diminution du tonus, tête relâchée, ouverture de la cavité orale au repos), *« hyper »* (augmentation du tonus, tensions cervicales, avancée du visage, contraction des scalènes et des sterno-cléïdo-mastoïdiens) ou *« normal »*.

2.3. Qualité de l'émission vocale

Enfin, puisque l'analyse de la qualité de la production sonore (attaque du son, partie stable et sortie) nous apprend sur l'efficacité du comportement phonatoire malgré les dysfonctionnements chez l'individu (Guerrier et cols., 2004[237]), nous avons analysé la qualité vocale d'un /a/ tenue employé dans l'étude des paramètres acoustiques (cf. chapitre…., *section ….*).

- *Attaque* : On évaluait si le début de l'émission était *brusque*, s'il était perçu comme un démarrage abrupte par rapport à l'intensité et la mise en vibration des cordes vocales; *faible*, si le début était caractérisé par une impuissance vocale et une fatigabilité; *normal*, si le début était perçu comme étant équilibré et ne manifestait pas une des caractéristiques précédemment mentionnées.

- *Partie stable* : On évaluait si la tenue vocalique après l'attaque était *instable*, ou perçue avec des petites variations en intensité et en fréquence; *interrompue*, si la partie stable était coupée ou avait une tendance à l'interruption; *normale*, si l'on percevait une stabilisation du signal due à la synchronisation des cordes vocales.

- *Sortie* : On évaluait si la fin de l'émission était *brusque*, si elle était perçue comme une fin intempestive par rapport à l'intensité et la vibration des cordes vocales; *faible*, si la fin était caractérisée par une impuissance vocale et une fatigabilité ne permettant pas d'aller jusqu'au bout de l'émission; *normal*, si la fin était perçue comme étant équilibrée et ne manifestait pas l'une des caractéristiques précédemment citées.

3. Résultats :

Nous avons voulu présenter ici, à titre descriptif et en termes de « fréquence statistique », un « tour d'horizon » des caractéristiques du comportement vocal observées chez les enfants. D'abord, nous présentons les données concernant le geste vocal à partir de l'observation de la posture, la respiration et la tension musculaire. Ensuite, nous présenterons les résultats de l'analyse de la qualité de l'émission d'un /a/ tenu.

3.1. Posture

Les postures adoptées par les enfants de l'étude pour les différents types de voix sont présentées dans les tableaux ci-dessous (Tableau 1a,b,c). Afin de visualiser à quoi correspondaient les différentes postures observées, on peut se reporter aux dessins de la figure 2 et à la description de chacune des postures précédemment présentée (section *2*).

Posture voix conv	Instable	Vertical	Vers l'avant	Vers l'arrière
EE	42%	40%	13%	5%
ESA	50%	25%	20%	5%
ESIC	80%	10%	5%	5%

a) Posture de la voix conversationnelle.

Posture voix forte	Instable	Vertical	Vers l'avant	Vers l'arrière
EE	18%	40%	37%	5%
ESA	30%	10%	55%	5%
ESIC	45%	20%	30%	5%

b) Posture de la voix forte.

Posture voix d'appel	Instable	Vertical	Vers l'avant	Vers l'arrière
EE	13%	11%	74%	2%
ESA	10%	5%	85%	0%
ESIC	25%	5%	65%	5%

c) Posture de la voix d'appel.

Tableau 1 : Tableaux comparatifs des postures observées pour les différents types de voix. On note en bleu, les pourcentages qui signalent la posture majoritairement adoptée par chacun des groupes.

Les résultats montrent que la posture adoptée pour émettre la voix conversationnelle est majoritairement « instable » pour tous les enfants. Bien que l'exemple de la posture « commode » et la consigne de conserver cette posture au cours de la parole étaient donnés quelques instants avant l'enregistrement, les enfants ont utilisé différentes postures. Néanmoins, on peut remarquer qu'un pourcentage important du groupe EE a tendance à adopter la posture verticale ou « commode ».

a) Posture « vertical ou commode » *b) Posture « vers l'avant »* *c) Posture « vers l'arrière »*

Figure 2 : Axes posturaux.

L'échantillon de notre étude semblerait ne pas avoir « spontanément » acquis la posture « commode » pour la voix conversationnelle. La posture « instable » entraîne des tensions musculaires péri-laryngées qui perturbent le bon fonctionnement du mécanisme laryngé.

Concernant la posture adoptée pour émettre la voix forte, les tendances sont diverses pour chacun des groupes. D'une part, le groupe EE montre une tendance à l'utilisation d'une posture « verticale » et une posture plutôt « vers l'avant ». Le groupe EE démontre avoir compris qu'il est nécessaire d'adopter une nouvelle posture plus fonctionnelle pour l'émission de la voix forte. L'utilisation d'une posture « commode » pour la voix forte facilite la sortie du flux aérien et la mobilité laryngée indispensables pour la projection vocale. D'autre part, les groupe ESA et ESIC emploient une posture « vers l'avant » et parfois « instable ». Les groupes ESA et ESIC n'effectuent pas les modifications nécessaires pour adapter la posture au type de voix désiré. La posture « vers l'avant » engendre un geste respiratoire défectueux et bloque le larynx quand il s'agit d'une émission vocale à forte intensité et de longue durée.

Enfin, la posture de la voix d'appel est plutôt « vers l'avant » pour tous les groupes. Cette posture est tout à fait adaptée à la voix d'appel ou de détresse. Pour la production de ce type de voix, il semblerait que tous les groupes comprennent l'intention communicative sous-jacente de ce type de voix et les modifications posturales nécessaires pour accomplir l'action.
Nous nous apercevons que la notion de modification posturale nécessaire pour émettre les différents types de voix n'a pas encore été développée notamment chez les groupes ESA et ESIC pour la voix conversationnelle et la voix forte.

En résumé :

- La posture de la voix conversationnelle est majoritairement « instable » pour tous les enfants.
- La posture correcte pour la voix forte est notamment plus adoptée par le groupe EE que par les groupes ESA et ESIC.
- La posture correcte pour la voix d'appel est majoritairement adoptée par tous les groupes.

3.2. Statique corporelle

La position des différents segments corporels qui déterminaient la posture lors de l'émission de la voix conversationnelle est présentée dans les tableaux ci-dessous (Tableau 2a, b, c, d, e, f, g).

Position de la tête	Normale	Etirée	Relâchée
EE	29%	58%	13%
ESA	10%	85%	5%
ESIC	20%	65%	15%

a) Position de la tête.

Position des épaules	Normale	Etirée	Elevée
EE	18%	50%	32%
ESA	10%	75%	15%
ESIC	5%	95%	0%

b) Position des épaules.

Position du dos	Normale	Etirée
EE	53%	47%
ESA	45%	55%
ESIC	10%	90%

c) Position du dos.

Position des bras	Normale	Croisée
EE	29%	71%
ESA	25%	75%
ESIC	5%	95%

d) Position des bras.

Position des jambes	Normale	Très serrée	Très écartée
EE	45%	47%	8%
ESA	30%	65%	5%
ESIC	50%	45%	5%

e) Position des jambes.

Position des talons	Normale	En pointe	En talon	Latérale
EE	74%	3%	0%	23%
ESA	65%	0%	0%	35%
ESIC	60%	0%	5%	35%

f) Position des talons.

.

Position des pieds	Normale	Très serrée	Très écartée
EE	47%	42%	11%
ESA	40%	55%	5%
ESIC	50%	45%	5%

g) Position des pieds.

Tableau 2 : Tableaux comparatifs de la position des segments corporels observés pour la voix conversationnelle. On note en bleu, les pourcentages qui signalent la position majoritairement adoptée par chacun des groupes.

Nous avons observé 7 segments corporels qui déterminaient la posture globale adoptée par l'enfant afin de déterminer à peu près à quel niveau du corps la posture était compromise. Cette observation reste tout à fait subjective. Concernant la voix conversationnelle, d'une part, les segments corporels de la partie haute correspondant à la tête, aux épaules et au dos se trouvent plutôt en « position étirée » (cf. Figure 3) pour tous les groupes. L'enfant projette son visage vers l'avant; ainsi, le menton est avancé, la position de la pointe des épaules et le dos sont trop en avant. D'autre part, la partie basse du corps est plutôt instable pour tous les groupes. L'instabilité est dûe à un positionnement des bras en avant entraînant des tensions musculaires de la partie haute du corps et à une position serrée des extrémités inférieures. En revanche, tous les groupes utilisent un appui des pieds au sol qui permet d'assurer l'équilibre général du corps.

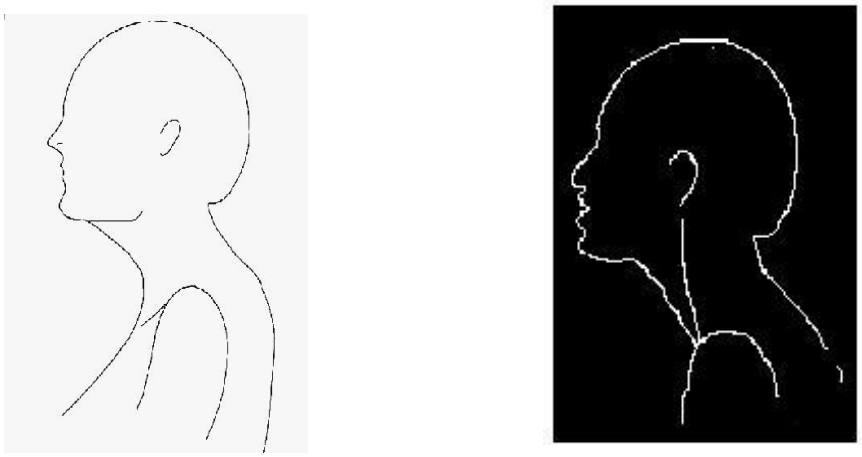


Figure 3 : Vues simplifiées montrant la position « étirée » des segments corporels de la tête, les épaules et la partie supérieure du dos.

La statique corporelle caractérisant la posture adoptée lors de l'émission de la voix forte est présentée dans les tableaux ci-dessous (Tableau 3a, b, c, d, e, f , g).

Position de la tête	Normale	Etirée	Relâchée
EE	18%	71%	11%
ESA	5%	95%	0%
ESIC	5%	70%	25%

a) Position de la tête.

Position des épaules	Normale	Etirée	Elevée
EE	24%	52%	24%
ESA	15%	70%	15%
ESIC	15%	80%	5%

b) Position des épaules.

Position du dos	Normale	Etirée
EE	53%	47%
ESA	40%	60%
ESIC	25%	75%

c) Position du dos.

Position des bras	Normale	Croisée
EE	42%	58%
ESA	45%	55%
ESIC	20%	80%

d) Position des bras.

Position des jambes	Normale	Très serrée	Très écartée
EE	63%	29%	8%
ESA	35%	60%	5%
ESIC	35%	55%	10%

e) Position des jambes.

Position des talons	Normale	En pointe	En talon	Latérale
EE	84%	0%	0%	16%
ESA	80%	0%	0%	20%
ESIC	70%	0%	5%	25%

f) Position des talons.

Position des pieds	Normale	Très serrée	Très écartée
EE	63%	26%	11%
ESA	45%	50%	5%
ESIC	40%	50%	10%

g) Position des pieds.

Tableau 3 : Tableaux comparatifs de la position des segments corporels observés pour la voix forte. On note en bleu, les pourcentages qui signalent la position majoritairement adoptée pour chacun des groupes.

Concernant la voix forte, les segments corporels de la tête, des épaules et du dos se trouvent plutôt en « position étirée » pour tous les groupes (Figure 3) comme pour la voix conversationnelle. En effet, l'enfant n'effectue pas de modification de la position des segments corporels afin d'adapter sa posture à un type de voix nécessitant un souffle phonatoire plus puissant. La partie basse du corps se manifeste plutôt instable chez les groupes ESA et ESIC alors que pour le groupe EE, elle apparait normale. Par ailleurs, tous les groupes utilisent un appui des pieds au sol qui permet d'assurer l'équilibre général du corps.

De manière générale, l'observation de la statique corporelle dans le groupe EE nous permet de dire que ce groupe est plus conscient que les groupes ESA et ESIC des modifications des segments corporels à effectuer pour l'émission de la voix forte, notamment au niveau de la stabilité de la partie basse du corps. Il semblerait que les enfants du groupe EE modifient leur posture car ils ont le ressenti que s'ils n'adoptent pas une nouvelle posture, ils ne seront pas entendus et par conséquent, ils ne seront pas efficaces dans leur intention communicative. En

revanche, cette attitude ne s'accomplit pas de manière spontanée chez les enfants des groupes ESA et ESIC.

En résumé :

- Pendant la voix conversationnelle, les segments corporels de la partie haute (tête, épaules et dos) se trouvent plutôt en « position étirée » et les segments de la partie basse (jambes) sont plutôt instables pour tous les groupes. Néanmoins, l'équilibre général du corps est compensé grâce à un appui des pieds correcte par tous les groupes.
- Pendant la voix forte, les segments corporels de la partie haute (tête, épaules et dos) se trouvent plutôt en « position étirée » pour tous les groupes alors que la partie basse est plutôt instable chez les groupes ESA et ESIC. Par ailleurs, tous les groupes utilisent un appui des pieds au sol qui permet d'assurer l'équilibre général du corps.

Quant à la statique corporelle caractérisant la posture adoptée lors de l'émission de la voix d'appel, nous ne l'avons pas observée. Cela car nous étions intéressés à observer plutôt les modifications des segments corporels pour la voix conversationnelle et la voix forte, deux types de voix émises à intensités différentes (Dalleas et col., 1987[238]) et les plus couramment employées dans la vie quotidienne de l'enfant. En revanche, nous avons évalué le mécanisme vocal utilisé lors de l'émission de la voix d'appel (Tableau 4).

Mécanisme vocal	(1) Poitrine	(2) Tête
EE	8%	92%
ESA	0%	100%
ESIC	25%	75%

Tableau 4 : Mécanisme vocal utilisé pour la voix d'appel.

Tous les groupes utilisent naturellement le mécanisme vocal 2 ou la « voix de tête » pour émettre la voix d'appel. La voix de tête est essentiellement produite par un contact des bords fins des cordes vocales et une faible masse vibrante qui facilite l'émission de fréquences aiguës (à partir de 415 Hz environ) et généralement à fortes intensités (à partir de 50 dB environ) de la voix d'appel.

3.3. Comportement respiratoire

De l'évaluation du comportement respiratoire, il ressort que :

- Tous les groupes utilisent principalement un type respiratoire costal supérieur caractérisé par un tirage cervical, des tensions des sterno-cléïdo-mastoïdiens et des scalènes, des tensions de la musculature du visage et une rétractation de la paroi abdominale (Tableau 5). La localisation du mouvement inspiratoire pendant la respiration pour la voix conversationnelle reste concentrée sur la partie supérieure de la cage thoracique.

Type respiratoire	Scapulaire	Costal supérieur
EE	8%	92%
ESA	20%	80%
ESIC	5%	95%

Tableau 5 : Type respiratoire.

- Les groupes ESA et ESIC utilisent un mode respiratoire de prédominance orale. En effet, la prise d'air pour ces groupes s'achevait durant le cycle respiratoire notamment par la cavité orale. Nous avons observé une ouverture orale prononcée précédant la force d'aspiration (Tableau 6).

Mode respiratoire	Prédominance nasale	Prédominance orale
EE	55%	45%
ESA	45%	55%
ESIC	30%	70%

Tableau 6 : Mode respiratoire.

- Tous les groupes ont une tendance à manifester une fréquence respiratoire qui oscille entre 16 et 21 cycles/minute. La fréquence semble correspondre à la norme pour les enfants de 5 à 13 ans et ne fourni pas de signes d'une insuffisance respiratoire (Tableau 7).

Fréquence respiratoire	10 – 15 cycles/min	16 – 21 cycles/min	22 – 28 cycles/min
EE	8%	74%	18%
ESA	20%	60%	20%
ESIC	30%	40%	30%

Tableau 7 : Fréquence respiratoire.

- Quant à l'évaluation de la perméabilité nasale, nous remarquons d'une manière générale que tous les groupes manifestent une maîtrise de la mobilité du voile du palais. Lorsque la nasalisation a lieu, elle est notamment bilatérale (Tableau 8).

Perméabilité nasale	Non	Unilatérale	Bilatérale
EE	76%	11%	13%
ESA	70%	10%	20%
ESIC	60%	10%	30%

Tableau 8 : Perméabilité nasale.

En résumé :

- Tous les enfants utilisent principalement un type respiratoire costal supérieur.
- Les enfants des groupes ESA et ESIC utilisent plutôt un mode respiratoire de prédominance orale.
- Tous les enfants manifestent plutôt une fréquence respiratoire qui oscille entre 16 et 21 cycles/minute ne révélant pas une insuffisance respiratoire.
- Tous les enfants manifestent plutôt une maîtrise de la mobilité du voile du palais.

3.4. Tension musculaire

Le tonus musculaire du visage des enfants est majoritairement normal. Quelques cas d'hypertonie musculaire ont été observés dans tous les groupes mais davantage chez l'ESA et l'ESIC. De même, une hypotonie a été observée plutôt chez l'EE et l'ESIC (Tableau 9).

Tonus musculaire visage	Hypo	Normal	Hyper
EE	11%	84%	5%
ESA	0%	75%	25%
ESIC	20%	45%	35%

Tableau 9 : Tonus musculaire du visage.

En revanche, le tonus musculaire du cou est majoritairement normal pour les groupes EE et ESA, cependant on note une tendance à une hypertonie chez le groupe ESIC (Tableau 10).

Tonus musculaire cou	Hypo	Normal	Hyper
EE	5%	53%	42%
ESA	0%	50%	50%
ESIC	20%	35%	45%

Tableau 10 : Tonus musculaire du cou.

3.5. Qualité de l'émission d'un /a/ tenu

L'analyse de la qualité de l'émission vocale révèle que :

- L'attaque ou le début de l'émission d'un /a/ tenu est majoritairement brusque pour tous les groupes. Le démarrage de la production sonore est abrupt par rapport à l'intensité et la mise en vibration des cordes vocales (Tableau 11).

Attaque	Normal	Brusque	Faible
EE	5%	92%	3%
ESA	10%	85%	5%
ESIC	10%	90%	0%

Tableau 11 : Attaque vocal d'un /a/ tenu.

- La partie stable ou la tenue vocalique après le démarrage est majoritairement instable pour les groupes ESA et ESIC, et normale pour le groupe EE. La tenue vocalique présente de petites variations en intensité et en fréquence (Tableau 12).

Partie stable	Normale	Instable	Interrompue
EE	68%	32%	0%
ESA	15%	80%	5%
ESIC	25%	75%	0%

Tableau 12 : Tenue vocalique d'un /a/ tenu.

- La sortie ou la fin de l'émission est brusque pour tous les groupes. La fin de l'émission est intempestive par rapport à l'intensité et à la mise en vibration des cordes vocales (Tableau 13).

Sortie	Normale	Brusque	Faible
EE	18%	66%	16%
ESA	15%	60%	25%
ESIC	0%	80%	20%

Tableau 13 : Sortie ou fin de l'émission vocale d'un /a/ tenu.

En résumé :

- L'attaque de l'émission vocale est majoritairement brusque pour tous les groupes.
- La tenue vocalique est majoritairement instable pour les groupes ESA et ESIC.
- La sortie de l'émission est plutôt brusque pour tous les groupes.

******

L'évaluation du comportement vocal repose sur des observations cliniques subjectives et qualitatives. La nature des informations recueillies permet en pratique clinique de mieux cibler l'approche thérapeutique et d'établir un programme de rééducation adapté au patient. Dans cette étude, ces informations nous ont donnée un aperçu en « mille feuilles » du comportement vocal de la population étudiée. Nos données actuelles ne nous permettent pas d'aller plus loin dans la réflexion et de trouver un corrélat avec d'autres variables comme les mesures acoustiques et l'analyse des paramètres aérodynamiques comme proposé par différents auteurs (Heuillet-Martin et col., 1995[233]).

---

229. UCHANSKI, Rosalie M. et GEERS, Ann E., "Acoustic characteristics of the speech of young cochlear implant users : a comparison with normal-hearing age-mates", *Ear and Hearing*, Vol. 24, n° 1S, 2003, p. 90-105.

230. ESTIENNE, F., *Voix parlée, voix chantée. Examen et thérapie*, Paris, Masson, 1998.

231. FAURE, MA., « Dynamique respiratoire et qualités acoustiques de la voix », *Rev laryngol (Bord)*, n° 108, 1987, p. 369-372.

232. AMY DE LA BRETEQUE, Benoît, *L'équilibre et le rayonnement de la voix*, Marseille, Solal, 1997.

233. HEUILLET-MARTIN, GH., GARSON-BAVARD, A. et LEGRÉ, A., *Une voix pour tous, vol. 1 et 2*, Marseille, Solal, 1995.

234. ROUBEAU, B., « Le bilan vocal », *Rééducation orthophonique*, n° 40, 2002, p. 54-59.

235. BOUISSET, S., et MATON, B., *Muscle, posture et mouvement : bases et applications de la méthode électromyographique*, Paris, Hermann, 1995.

236. KENDALL, F., MCCREARY, E., PROVANCE, P., RODGERS, M. et ROMAI, W., *Muscles testing and function with posture and pain*, V^e^ éd., PA, USA, Lippincott Williams & Wilkins, 2005.

237. GUERRIER, Bernard, GIOVANNI, Antoine et REMACLE, Marc, *Pathologie de la corde vocale chez l'adulte*, Editions de la Société Française d'Oto-rhino-laryngologie et de Chirurgie de la face et du cou, 2004.

238. DALLEAS, B., et DALLEAS, A., « Qualités acoustiques de la voix », *Revue de Laryngologie*, Suppl., 1987, p. 373-377.

Caractérisation de la voix de l'enfant sourd appareillé et implanté cochléaire : approches acoustique et perceptuelle et proposition de modélisation

Résumé

Cette thèse propose une analyse comparative acoustique et perceptive de la voix d'un effectif statistiquement fiable d'enfants sourds appareillés et implantés cochléaires. Peu de paramètres diffèrent de manière significative entre le groupe d'enfants sourds ayant été appareillés et implantés avant l'âge de trois ans, et le groupe d'enfants entendants. L'ensemble de résultats indiquent que la voix des enfants de notre étude ne présente pas les caractéristiques traditionnellement retenues pour déterminer la voix pathologique. Par ailleurs, les caractéristiques de la voix des enfants implantés cochléaires sont sensiblement comparables à celles des enfants entendants. Fort de ces résultats expérimentaux, nous avons proposé un modèle « vibro-acoustique » de la régulation de la voix des enfants sourds « oralisés », et développé un simulateur numérique de la boucle audio-phonatoire.

Characterising the voice of the fitted and cochlear implanted deaf child: acoustic and perceptive approaches with a view to modelling

Abstract

This dissertation presents an acoustic and perceptive comparative analysis of the voice in a reliable size group of fitted and cochlear implanted deaf children. There are very few significantly different parameters between fitted and implanted children before three years old and normal children. Results do not confirm that hearing-impaired children's voices of our study are pathological. Furthermore, characteristics of cochlear implanted voices are nearly comparable to normal children's voices. As a consequence of these results, we propose a « vibro-acoustic model » and a software of voice control mechanism in deaf children.

Mots-clés :
Voix
Enfants
Surdité
Prothèse auditive
Implant cochléaire
Paramètres acoustiques
Modèle
Boucle audio-phonatoire

Key Words :
Voice
Children
Deafness
Hearing aid
Cochlear implant
Acoustic parameters
Model
Audio-phonatory loop

PRAXILING UMR 5267 CNRS – Montpellier III

17 rue Abbé de l'Epée
34090 Montpellier - FRANCE