

Faculté des LETTRES
UR 1339 Linguistique, Langues et Parole - LILPA
ER Parole et Cognition & Institut de Phonétique de Strasbourg (IPS)

THESE

présentée pour l'obtention du doctorat de
Sciences du Langage
par

Camille FAUTH

Spécialité : Phonétique Générale et Expérimentale
Phonétique clinique

Perturbation de la production de la parole suite à une opération de la glande thyroïde

Jury

Mme Lise Crevier-Buchman, Chargée de Recherche au CNRS (Examinatrice)
LPL, CNRS - UMR 7018, Sorbonne Nouvelle Paris 3
M. le Professeur Bernard Harmegnies (Rapporteur)
Laboratoire de Phonétique, Université de Mons, Belgique
M. Yves Laprie, Directeur de Recherche au CNRS (Rapporteur)
LORIA/Groupe Parole, Nancy
M. le Professeur Jean-François Rodier (Examineur)
Centre Paul Strauss - Strasbourg
M. le Professeur Rudolph Sock (Examineur)
IPS & U.R. 1339 LiLPa, ER Parole et Cognition, Université de Strasbourg
Mme Béatrice Vaxelaire, Maître de Conférences - HDR (Directrice)
IPS & U.R. 1339 LiLPa, ER Parole et Cognition, Université de Strasbourg

Faculté des LETTRES
UR 1339 Linguistique, Langues et Parole - LILPA
ER Parole et Cognition & Institut de Phonétique de Strasbourg (IPS)

THESE
présentée pour l'obtention du doctorat de
Sciences du Langage
par

Camille FAUTH

Spécialité : Phonétique Générale et Expérimentale
Phonétique clinique

Perturbation de la production de la parole suite à une opération de la glande thyroïde

Jury

Mme Lise Crevier-Buchman, Chargée de Recherche au CNRS (Examinatrice)
LPL, CNRS - UMR 7018, Sorbonne Nouvelle Paris 3
M. le Professeur Bernard Harmegnies (Rapporteur)
Laboratoire de Phonétique, Université de Mons, Belgique
M. Yves Laprie, Directeur de Recherche au CNRS (Rapporteur)
LORIA/Groupe Parole, Nancy
M. le Professeur Jean-François Rodier (Examineur)
Centre Paul Strauss - Strasbourg
M. le Professeur Rudolph Sock (Examineur)
IPS & U.R. 1339 LiLPa, ER Parole et Cognition, Université de Strasbourg
Mme Béatrice Vaxelaire, Maître de Conférences - HDR (Directrice)
IPS & U.R. 1339 LiLPa, ER Parole et Cognition, Université de Strasbourg

Le pilote répondit : « Seigneur, ne vous effrayez de rien. On est ici aux confins de la mer de Glace, où, au début de l'hiver dernier, eut lieu une grande et cruelle bataille entre les Arismapiens et les Néphélibates. Alors gelèrent dans l'aire les paroles et les cris des hommes et des femmes, les chocs des masses d'armes, les heurts des armures, des caparaçons, les hennissements des chevaux et tout autre vacarme de combat. Maintenant la rigueur de l'hiver étant passée, le beau temps doux et serein étant arrivé, elles fondent et on les entend.

- Par Dieu, dit Panurge, je l'en crois. Mais pourrions-nous en voir quelqu'une ? Il me souvient d'avoir lu qu'au pied de la montagne où Moïse reçut la loi des Juifs, le peuple percevait les voix par la vue.

- Tenez, tenez, dit Pantagruel, voyez-en ici qui ne sont pas encore dégelées. »

Alors il nous jeta sur le tillac de pleines poignées de paroles gelées, et elles ressemblaient à des dragées perlées de diverses couleurs. Nous y vîmes des mots de gueule, des mots de sinople, des mots d'azur, des mots de sable, des mots dorés. Après avoir été échauffés entre nos mains, ils fondaient comme neige, et nous les entendions réellement, mais nous ne les comprenions pas car c'était un langage barbare. Un seul fit exception, assez gros, qui, Frère Jean l'ayant échauffé entre ses mains, produisit un son semblable à celui que font les châtaignes jetées dans la braise sans être entamées, lorsqu'elles éclatent, et nous fit tous tressaillir de peur. « C'était, dit Frère Jean, un coup de fauconneau, en son temps. »

Panurge demanda à Pantagruel de lui en donner encore. Pantagruel lui répondit que donner sa parole était acte d'amoureux.

« Vendez m'en donc, disait Panurge.

- Vendre des paroles, c'est ce que font les avocats, répondit Pantagruel. Je vous vendrais plutôt du silence, et plus cher, comme en vendit un jour Démosthène, contre des deniers qui le rendirent aphone. »

Néanmoins, il en jeta trois ou quatre poignées sur le tillac. Et j'y vis des paroles bien piquantes, des paroles sanglantes (...), des paroles horribles et d'autres assez désagréables à voir. Lorsqu'elles eurent fondu toutes ensemble, nous entendîmes hin, hin, hin, hin, his, tic, torche, lorgne, brededin, brededac, fr, frrr, frrr, bou, bou, bou, bou, bou, bou, bou, bou, bou, tracc, trac, trr, trr, trrr, trrrrr, on, on, on, on ououououon, goth, magoth, et je ne sais quels autres mots barbares : et il disait que c'était ce qu'on perçoit de la charge et du hennissement des chevaux au moment du corps-à-corps (...).

François Rabelais, Quart Livre (1552), Chapitre LVI Comment entre les paroles gélées, Pantagruel trouva des mots de gueule.

Remerciements

Par ces quelques lignes, je souhaite remercier toutes les personnes qui m'ont accompagnée ces dernières années et ainsi contribué à l'aboutissement de ce travail.

Je souhaite, tout d'abord, remercier Béatrice Vaxelaire, ma directrice de thèse, sans qui ce travail n'aurait pas pu voir le jour. Par son dynamisme et son énergie, elle a su me transmettre le goût pour la recherche dès mon entrée en master. A ses côtés, j'ai appris que la recherche est un travail de patience, que la progression ne peut se faire que par tâtonnements et rigueur. Sa disponibilité et sa bienveillance à mon égard ont grandement participé à la rédaction de ce travail. Je tiens aussi à la remercier pour l'autonomie et les responsabilités qui m'ont été confiées tout au long de ma thèse, elles ont permis à faire de ma thèse une maturation personnelle et scientifique. Cela été pour moi une grande fierté et un plaisir constant de travailler sous la direction de Béatrice Vaxelaire.

Je remercie également Rudolph Sock qui par sa présence et sa bonne humeur m'aura toujours poussé à faire mieux. Ses conseils, mais aussi ses plaisanteries, ont rendu mon travail de recherche plaisant et enrichissant. L'accompagner dans les diverses manifestations scientifiques a toujours été une belle opportunité et a été riche en enseignements.

Je remercie vivement les membres de mon jury, qui ont accepté de donner de leur temps pour lire et évaluer ce travail. Un grand merci aux rapporteurs Monsieur le Professeur Bernard Harmegnies et Monsieur le Directeur de Recherche Yves Laprie. Un grand merci également à Madame Lise Crevier-Buchman qui a montré un intérêt pour mon travail dès mes premières années de recherche et qui m'a accueillie dans son cabinet à l'HEGP pour compléter mes recherches.

Mes remerciements vont tout naturellement aux membres du Département de Chirurgie Oncologique du Centre Paul Strauss de Strasbourg qui m'ont accueillie et ainsi rendu possible l'acquisition de mes données. Au Professeur Jean-François Rodier,

Remerciements

directeur du département, plus particulièrement qui s'est toujours montré curieux pour mes recherches, sans lui ce travail n'aurait jamais pu voir le jour. Je le remercie également pour sa présence dans mon jury.

A toutes les orthophonistes, Mesdames Cécile Dauer-Seel, Fanny Delattre, Marie Delobbe, Valérie Dolisi, Christine Fronsacq, Laetitia Gobert, Charlotte Louis, Marie Masson, qui ont suivi mes locuteurs pendant leur rééducation et m'ont permis de compléter mes enregistrements chaque fois que cela était nécessaire.

Je remercie tous les patients qui ont accepté de prêter leurs voix pour mener à bien mon étude. Les contraintes hospitalières et leurs soucis ne les ont pas empêchés de se montrer enthousiastes et de répéter consciencieusement des séquences qui, pour eux, ne voulaient « rien dire ».

Parce que mener sa recherche doctorale c'est aussi avoir la chance de rencontrer des personnes enrichissantes, tant scientifiquement qu'humainement, je tiens à remercier très sincèrement Mesdames Catherine Schnedecker et Marie-Carmen Ramirez.

Je voudrais aussi remercier mes précieux relecteurs, pour leur application et leur disponibilité, Marion, Thomas, Maeva et Cyril qui auront souvent fait bien plus que simplement traquer les coquilles. Je remercie également Fabrice qui m'aura initiée aux subtilités informatiques en phonétique.

Ecrire une thèse est sûrement un exercice solitaire, mais le mien n'aurait pas été possible sans la compagnie et l'amitié de mes gnous de l'IPS, Marion, Lucie, Zach, Cyril, Fayssal, Xuelu, Thomas et Amel qui donnent vie au labo. J'ai une pensée particulière pour Marion qui est ma colocataire de séminaire préférée et pour Thomas qui partage avec patience mon bureau, mes colères, mes doutes et mes fou-rires.

Je remercie particulièrement ma famille dont le soutien a été primordial pour la réalisation de ce travail, de près ou de loin, et avec ses moyens, chacun a apporté sa

Remerciements

contribution. Je souhaite remercier mes parents du fond du cœur, qui ont su voir au-delà de mes doutes et me pousser à toujours tout entreprendre avec passion et conviction. A mon papa, pour ses « papamadit » toujours bénéfiques et visionnaires. A ma maman, pour son pragmatisme et son dévouement au quotidien. Leurs valeurs et leurs conseils rendent la vie facile à vivre. Sans concertation, ils ont toujours eu le même discours vrai et encourageant. Merci à mes grands-parents toujours fiers et disponibles pour pallier les tracas du quotidien.

Je remercie aussi ma tribu, Aline, Lilly, Hélo, Joanna, Julia, Claire, Céline, François, Maeva, Charlotte et Mathieu. Merci de m'avoir témoigné votre amitié, de m'avoir écoutée et encouragée. Votre présence à mes côtés a toujours été précieuse. Sans jamais y avoir rien compris, vous avez toujours mesuré l'importance de mes « lubies » en recherche. Vous êtes mon cocon en coton. Une pensée particulière pour Aline et Lilly, qui dans les derniers instants se seront appliquées à me vider la tête chaque fois que cela était nécessaire.

Un merci tout particulier va à Philippe qui s'est retrouvé « embarqué » dans cette aventure en cours de route. Sa bienveillance et son dévouement à mon égard ont été très précieux. Merci d'avoir toujours eu une oreille pour mes questionnements et mes doutes inopinés. Je le remercie de toujours croire en moi et lui témoigne tout mon amour.

Je remercie mes collègues de toujours, Marc, Martine, Joëlle, Michel et Jean qui m'ont témoigné leur amitié et ont su apporter leurs compétences informatiques pour régler certains détails en fin de parcours.

J'ai une pensée particulière pour Mamama, aujourd'hui disparue, mais qui m'accompagne toujours. Sa présence me manque.

Sommaire

Introduction.....	17
Organisation de la thèse.....	21
Etat de la Question.....	23
I. Théories en production de la parole.....	25
I.1. Planification et Exécution - Variation et Contraintes.....	26
I.2. Théories en production de la parole	27
II. Perturbation de la production et de la perception de la parole normale	45
II.1. Perturbation de la production de la parole	46
II.2. Perturbation de la perception de la parole.....	58
II.3. Perturber la perception pour perturber la production.....	61
III. Troubles de la parole et de la voix	66
III.1. Les dysarthries – troubles du mouvement ?	66
III.2. Dysphonies – Troubles laryngés.....	72
IV. Dysphonie et Immobilités laryngées	89
IV.1. Classification des paralysies laryngées	89
IV.2. Anatomie et physiologie du larynx.....	90
IV.3. Epidémiologie de la paralysie laryngée unilatérale.....	91
IV.4. Thyroïdectomie et paralysie laryngée unilatérale.....	93
IV.5. Immobilités laryngées post-thyroïdectomie	95
IV.6. Conséquences acoustiques de la thyroïdectomie.....	97
Protocole expérimental	109
V. Protocole Expérimental.....	111
V.1. Conditions d'enregistrement.....	111
V.2. Matériel.....	112
V.3. Corpus.....	113
V.4. Locuteurs	115

V.5. Enregistrements	118
V.6. Mesures – Analyses	120
V.7. Analyses Statistiques – Remarques générales	128
Résultats expérimentaux.....	130
VI. Expérience 1 : Etude acoustique de voyelles soutenues produites par des patients ne présentant pas de paralysie récurrentielle après thyroïdectomie	132
VI.1. Introduction	132
VI.2. Procédure expérimentale	133
VI.3. Hypothèses	134
VI.4. Résultats	135
VI.5. Discussion.....	154
VII. Expérience 2 : Etude acoustique de voyelles soutenues produites par des patients présentant une paralysie récurrentielle après thyroïdectomie.....	160
VII.1. Introduction	160
VII.2. Procédure expérimentale	161
VII.3. Hypothèses	162
VII.4. Résultats	164
VII.5. Discussion	185
VIII. Expérience 3 : Etude spatio-temporelle de la voix de patients présentant une paralysie récurrentielle après thyroïdectomie	191
VIII.1. Introduction.....	191
VIII.2. Procédure expérimentale.....	192
VIII.3. Résultats.....	198
VIII.4. Discussion	229
Synthèse Générale	236
Conclusions et perspectives	245
Quelques limites de cette recherche	248
Bibliographie	252
Tables et indexes.....	282
Index des auteurs.....	283
Index des figures	286
Index des tableaux.....	289

Table des matières.....	292
Annexes	299
I. Annexes 1 : Anatomie du larynx	300
I.1. Cartilages du larynx.....	300
I.2. Muscles du larynx	301
I.3. Le système membrano-ligamentaire	301
I.4. Les articulations	302
I.5. Les muscles	303
I.6. La vascularisation.....	307
I.7. L'innervation	307
II. Annexes 2 : Fiches de renseignements.....	309
II.1. Fiche Signalétique Patient.....	309
II.2. Questionnaire Post-Opératoire adapté du Voice Quality Related Life	311
III. Annexes 3 : Résultats.....	313
III.1. Valeurs formantiques Expérience 1	313
III.2. Valeurs formantiques Expérience 2.....	315
III.3. Durées du VTT – Expérience 3	317
III.4. Durées vocaliques vs durées consonantiques	318

INTRODUCTION

Nous proposons dans la présente étude d'apporter une contribution aux recherches menées dans le domaine de la phonétique clinique, et plus particulièrement dans celui de l'étude de la dysphonie, à partir de production de patients ayant été opérés de la glande thyroïde. Nous avons porté notre choix sur ce sujet en raison de notre intérêt pour la compréhension de la production de la parole et plus spécifiquement pour celle de la parole perturbée. Nos connaissances de l'organisation spatio-temporelle de la parole saine devraient permettre de définir convenablement les « cibles » articulatoires-acoustiques normales, afin d'aboutir à une meilleure évaluation des productions pathologiques déviantes.

Notre travail de recherche s'insère dans le cadre spécifique des problématiques liées aux chirurgies de la glande thyroïde. Cette glande est située à la base du cou. Sa fonction majeure est de sécréter des hormones indispensables à la régulation de fonctions importantes de l'organisme, telles que le rythme cardiaque, la motricité de l'intestin, la température du corps, la transformation des graisses et des sucres, l'appétit, l'humeur, le sommeil, l'équilibre du poids corporel et la croissance chez l'enfant¹.

A chaque dysfonctionnement thyroïdien correspond une intervention chirurgicale différente. En tout état de cause, le chirurgien prendra soin de préserver les nerfs récurrents laryngés, droit et gauche, qui participent à la phonation notamment, puisque ces nerfs sont responsables de l'activation de quatre des cinq muscles du larynx.

¹ (La glande thyroïde, Chirurgie, livret disponible sur le site du Centre Paul Strauss, http://www.centre-paul-strauss.fr/_mm/pdf/livret_chirurgie_glande_thy.pdf).

En cas d'atteinte de l'un des nerfs récurrents, il peut survenir une immobilité laryngée responsable de fuites laryngées. Ces fuites laryngées peuvent affecter le pli vocal membraneux et/ou la glotte postérieure. Toutefois une opération au niveau du larynx est rarement anodine, aussi la voix du patient peut se trouver modifiée alors même que les nerfs laryngés sont préservés.

Les immobilités laryngées affectent aussi bien la déglutition que la parole du patient. La voix du patient peut alors se trouver altérée avec pour conséquences : une voix soufflée ou rauque ; des vibrations laryngées irrégulières ou la diplophonie (perception simultanée de deux fréquences fondamentales) ; une modification de la périodicité de la voix ; une augmentation des mesures du jitter, du shimmer et du bruit par rapport aux harmoniques ; une intensité et une amplitude d'intensité diminuées ; une fatigue vocale ; une altération du débit aérien ; etc.

Il s'agira dans le cadre de ce travail, mené en partenariat avec le Département de Chirurgie du Centre Paul Strauss, Centre Régional de Lutte contre le cancer, d'évaluer les conséquences d'une chirurgie thyroïdienne sur la voix des patients, et de mettre au jour les différentes *perturbations* qu'entraîne ce type de chirurgie, ainsi que les éventuelles améliorations à apporter à ces interventions, grâce à la mise au jour objective de ces perturbations au niveau de la production. Notre étude sera longitudinale afin d'étudier les possibles stratégies de *compensation* ou de *réajustement* que le patient pourra mettre en place, seul ou à l'aide d'une rééducation orthophonique.

Il s'agit également d'étudier la flexibilité du système de production de la parole et de tenter de comprendre ce système à partir d'un dysfonctionnement. Nous pourrions ainsi peut-être déceler les limites des déviations physiques (articulatoires et acoustiques) imposées par les exigences linguistiques de clarté. Nous chercherons alors à repérer les patrons spatio-temporels résistants et à ainsi analyser les différentes stratégies utilisées par les locuteurs pour maintenir la distinctivité phonétique des catégories phonologiques, dans des conditions difficiles, en mettant notamment au jour divers phénomènes de compensations (Sock & Vaxelaire 2004 ; Vaxelaire 2007). A partir des différentes analyses que nous présentons dans ce travail, nous souhaitons rendre compte de la notion de *viabilité* (Aubin, 1991), en tentant de révéler les diverses *possibles*,

articulatoires et acoustiques, qui sous-tendent les unités phonétiques et phonologiques, malgré une variabilité omniprésente.

Les travaux en production de la parole sur la voix de patients thyroïdectomisés sont, à notre avis, très peu nombreux. Nous pensons ainsi contribuer à l'avancement des connaissances, non seulement sur la parole perturbée, mais également sur la caractérisation objective des dysfonctionnements, suite à l'intervention chirurgicale, pour une meilleure prise en charge de la qualité de vie des patients.

ORGANISATION DE LA THESE

Notre étude est constituée de trois parties :

1/La première partie correspond à l'état de la question. Il sera question dans le premier chapitre (Chapitre I) des grands courants théoriques en production de la parole. Puis (Chapitre II), nous expliquerons les différentes sources de perturbations qui peuvent provoquer une réorganisation du système de production-perception de la parole. Dans le troisième chapitre (Chapitre III), il sera question des différentes pathologies qui peuvent perturber le système de production de la parole. Dans le dernier chapitre, qui clôt cette partie (Chapitre IV), nous traiterons des divers travaux qui ont été effectués sur des patients ayant subi une thyroïdectomie.

2/La deuxième partie expose la méthodologie. Le chapitre V comprend les conditions d'enregistrement, les différents locuteurs retenus et les corpus élaborés pour nos études. Nous aborderons également les démarches expérimentales de nos différentes analyses, en précisant les mesures retenues pour chacune de nos expériences.

3/La troisième partie présente les résultats issus de nos trois expériences. La première expérience, exposée dans le chapitre VI, traite des caractéristiques spectrales de la voix de patients ayant subi une thyroïdectomie et ne présentant pas d'immobilité laryngée. La deuxième étude (chapitre VII) est menée à partir des caractéristiques spectrales de la voix de locuteurs présentant une immobilité laryngée post-thyroïdectomie. Enfin, notre dernière expérience (Chapitre VIII) a pour objet l'étude des caractéristiques spatio-temporelles de la voix de locuteurs qui ont subi une thyroïdectomie et présentant une immobilité laryngée.

Organisation

Nous proposons, tout au long de notre étude, des résumés et des bilans partiels afin de faciliter la lecture du travail.

PARTIE 1 :

ETAT DE LA QUESTION

Cette partie est consacrée à une revue de nos lectures des études relatives à notre sujet de recherche. Quatre chapitres composent cet état de la question :

Chapitre I : Théories en production de la parole

Chapitre II : Perturbations en production et en perception de la parole normale

Chapitre III : Troubles de la parole et de la voix

Chapitre IV : Dysphonie et immobilités laryngées

RESUME CHAPITRE I

Le début de ce chapitre rappellera l'importance du Système Nerveux Central (SNC) et des structures périphériques pour la planification et l'exécution des gestes articulatoires en production de la parole. L'importance du feedback, qui permet les réajustements articulatoires et donc d'atteindre les cibles, sera esquissée.

La suite du chapitre définira la production de la parole comme un système complexe régi par des contraintes de différents niveaux. Nous verrons que le message linguistique tolère une certaine variabilité, notion centrale de notre travail.

La fin du chapitre définira les notions de variabilité, d'invariance et de contraintes à l'aide des principaux courants théoriques en production de la parole : la Théorie de la Variabilité Adaptative de Lindblom, la Théorie Quantique de Stevens, le modèle DIVA de Guenther et Perkell, la notion du geste en Phonologie Articulatoire, le modèle de la Dynamique des Tâches, les notions d'Hyper et d'Hypo articulation développées par Lindblom, le principe de coordination et de coarticulation de Fowler et Saltzman, et enfin la Théorie de la Viabilité en production perception de la parole développée notamment au sein de notre laboratoire.

Ces principaux courants théoriques seront discutés afin de fournir le soubassement conceptuel qui pourrait servir pour l'analyse et la compréhension de nos données sur la production de la parole par nos sujets souffrant ou non de paralysie récurrentielle, après une opération de la glande thyroïde. Nous espérons pouvoir rationaliser nos données à la lumière du paradigme de la perturbation et du réajustement en production de la parole.

CHAPITRE I

I. Théories en production de la parole

Les premiers modèles portant sur la production de la parole posaient que le cerveau serait chargé de stocker les représentations mentales phonétiques issues des cibles articulatoires, acoustiques et perceptives. Une commande motrice provenant du Système Nerveux Central (SNC) serait responsable de l'activation des articulateurs. Leurs différents mouvements seraient alors orientés vers des cibles à atteindre pour l'émergence du son à produire. Cette description linéaire et unidirectionnelle (du cerveau vers les structures périphériques de la production) est maintenant nuancée par les récentes recherches en neurophysiologie. Les processus impliqués dans la production de la parole sont en réalité plus complexes, notamment si l'on considère le rôle des « *feedbacks* » ou des réafférences sensorielles : l'information provient également des régions périphériques en direction du SNC. Dès lors, la production de la parole ne peut plus être considérée comme un processus linéaire et doit être comprise comme le résultat d'interactions sensori-motrices complexes : l'activation des articulateurs est commandée par le Système Nerveux Central et constamment réajustée à l'aide d'informations sensorielles provenant des régions non centrales.

Nous proposons ici de mettre en évidence le rôle de quelques structures périphériques dans le système de production-perception de la parole, afin de mettre en exergue l'importance du contrôle des gestes articulatoires, ainsi que les conséquences des déviances éventuelles de ce contrôle.

I.1. Planification et Exécution - Variation et Contraintes

La communication reposant aussi sur la réception correcte du message linguistique, le locuteur est capable d'ajuster en permanence les caractéristiques propres à la production de la parole (vitesse d'élocution, qualité d'élocution, pauses, prosodie...) en fonction de son auditoire, de ses propres possibilités, ainsi qu'en fonction des capacités réceptives de l'auditeur (Lindblom, 1996). La variabilité est naturellement présente en production de la parole, et elle est liée aussi bien aux facultés perceptives de l'auditeur qu'aux stratégies individuelles suivant les interlocuteurs. Certaines études (Perkell et *al.* 1995, notamment) supposent que la programmation de la production de la parole serait liée à des *cibles* ou à des séquences de segments qui contiennent également des informations temporelles (appelé aussi *timing*). A cette étape de la planification serait associé un processus d'exécution qui vise à convertir les cibles discrètes en mouvements articulatoires. Ces cibles seraient alors une combinaison de mouvements des articulateurs et de paramètres acoustiques. Dès lors, les mouvements articulatoires contraindraient les variations acoustiques, tout en respectant les limites perceptives acceptables définissant des domaines plus ou moins étendues. Si cette description semble expliquer les processus de production de la parole, elle n'en est pas moins difficilement vérifiable. Il n'est pas, en effet, aisé d'identifier clairement les processus de programmation des cibles au niveau du SNC notamment.

Il est en revanche certain, surtout grâce à l'analyse de données articulatoires et acoustiques, que la production de la parole est une activité motrice volontaire régie par un système complexe de contraintes (Stetson, 1951). Ces contraintes peuvent être phonologiques (notion de traits distinctifs propres à chaque langue), bio-mécaniques (limites anatomiques et physiologiques) ou coarticulatoires (liées au contexte phonétique). Ces nombreuses contraintes, responsables de la variabilité en production de la parole, conduisent le système de production à s'adapter en adoptant diverses stratégies linguistiquement viables (*cf. infra*).

La cible serait atteinte si les contraintes articulatoires et acoustiques le permettent. Dans certains cas, lorsque ces contraintes sont plus importantes, la cible ne sera que partiellement ou pas du tout atteinte. Une certaine variabilité est affichée par les locuteurs et ne gêne pas la réception du message, même si les cibles ne sont pas

atteintes de façon optimales (Sock, 1998). Cela suppose que, dans une certaine mesure, un déplacement inhabituel des articulateurs peut être compensé de façon à atteindre la cible acoustique désirée, à défaut de mettre en place une cible articulatoire « canonique ». Naturellement, si aucune des cibles, articulatoire ou acoustique, n'est atteinte, le message linguistique sera détérioré et la compréhension deviendra alors problématique. Ces constats supposent donc l'existence de régions cibles correspondant à des zones de stabilité spatio-temporelle qui sous-tendraient la constance du message linguistique.

Les capacités d'adaptation des locuteurs dans des situations de communication perturbée, de façon naturelle (vitesse d'élocution, milieu bruité...) ou dans le cas de la parole pathologique ont été largement démontrées, grâce à des travaux de recherche expérimentale en phonétique. Il semble alors difficile de croire, sans nuancer ces propos, à l'existence d'invariances phonétique, articulatoire ou acoustique, que celles-ci soient absolues ou relatives (nous y reviendrons).

De nombreuses théories se sont attaché à définir la production de la parole en termes de processus articulatoire-acoustiques invariants ou de variabilité autorisée. Certaines de ces théories seront présentées ici. L'objectif est de voir dans quelle mesure ces approches théoriques peuvent nous fournir le soubassement conceptuel nécessaire pour rendre compte des capacités d'adaptation et de réajustement des locuteurs en général, et celles des locuteurs pathologiques plus particulièrement.

I.2. Théories en production de la parole

Les théories de production de la parole traitent des différents facteurs et contraintes nécessaires à l'émergence de la parole. Elles abordent ainsi les notions de cibles articulatoires et acoustiques, de contrôle moteur, de timing des articulateurs, mais aussi de contraintes articulatoires. Cette revue de la littérature doit beaucoup à la lecture de Sock (1998).

Diverses théories se proposent de définir la production de la parole. Deux types de courants opposés existeraient selon que ces théories soient orientées « sortie

acoustique » ou « système ». Ces deux manières de concevoir la production de la parole correspondent à deux courants de pensée majeurs en phonétique.

D'une part, les théories orientées « sortie acoustique » supposent l'existence d'une cible acoustique. Le locuteur chercherait alors à atteindre cette cible afin de garantir l'intégrité du message linguistique.

D'autre part, les théories orientées « système » s'orientent plutôt vers une cible articulatoire. L'émergence du message linguistique serait dans ce cas assurée si le locuteur atteint les cibles articulatoires adéquates.

1.2.1. L'existence d'une cible acoustique ?

Les théories orientées « sortie acoustique » supposent l'existence d'une cible acoustique nécessaire à l'émergence du message linguistique. Pour assurer l'intégrité du message linguistique, le locuteur doit s'approcher au maximum des cibles acoustiques et perceptives. Ces théories, dites aussi acoustiques, admettent peu de variabilité sur le plan acoustique.

1.2.1.1. Lindblom : Théorie de la Variabilité Adaptative en parole

En ce qui concerne les travaux sur la variabilité en parole, ceux de Lindblom sont parmi les plus exploités. Sa théorie de la Variabilité Adaptative en parole date de 1971 et a été mise à jour à plusieurs reprises (1983, 1990 et 1996).

Cette théorie orientée sortie articulatoire-acoustique est évidemment centrée sur la production acoustique du locuteur. Il ferait le choix de structures grammaticales et lexicales en fonction de ses intentions communicatives.

Selon Lindblom, l'intelligibilité de la parole dépendrait de deux facteurs :

- La qualité et le contenu du signal acoustique ;
- Les connaissances linguistiques et situationnelles de l'auditeur au moment du traitement perceptif.

Certains traits acoustiques peuvent être modifiés ou occultés, les expressions les moins claires seront alors dites « réduites ». L'auditeur n'a toutefois pas besoin d'un signal parfait pour comprendre le message du locuteur. Il est capable, dans une certaine mesure, de décoder le message à l'aide de ses connaissances sur le monde ou du contexte syntaxico-pragmatique. En outre, le locuteur agit en fonction des capacités de l'auditeur qu'il évalue de façon permanente durant les échanges.

Cette théorie repose donc sur deux éléments importants, à savoir le rôle actif du locuteur et de l'auditeur dans la situation de communication, mais aussi sur l'évaluation constante de cette situation de communication. Il s'agit de prendre en compte le milieu de la communication (par ex. bruité ou non) et les connaissances que le locuteur et l'interlocuteur ont de leur langue.

Lindblom fait intervenir la notion de cible qui est pour lui, à la fois un paramètre acoustique et/ou articulatoire et un corrélat de signes distinctifs qui permet la distinction des différents items lexicaux.

Etant donné que cette théorie admet nettement le principe de variabilité en production de la parole, nous proposons d'y revenir lorsque nous traiterons de la variabilité plus en détails.

I.2.1.2. Stevens : La Théorie Quantique

Dans le cadre des théories de production de la parole orientées « sortie acoustique », il est également important de parler de la Théorie Quantique de Stevens (1972).

La Théorie Quantique suppose l'existence d'une invariance acoustico-auditive. Les propriétés acoustiques sont donc présumées invariantes en fonction des locuteurs et des contextes phonétiques, et auraient un rôle perceptif démontré. Le locuteur apprend des patterns de coordination des articulateurs. L'acoustique serait l'espace de contrôle du locuteur qui agit pour doter le signal de parole des indices temporels et des propriétés spectrales invariantes, en association directe avec le code phonologique. Ces patterns seraient invariants quels que soient les locuteurs, les contextes phonétiques et les langues, afin de remplir les exigences acoustico-auditives du code phonologique.

Les travaux de Blumstein & Stevens (1979) montrent que la forme globale du spectre, calculée sur la détente de la consonne, possède une forme distinctive selon le lieu d'articulation. En 1991, leurs travaux vont plus loin et cherchent à montrer qu'il existe des propriétés acoustiques pour caractériser le lieu d'articulation de la consonne indépendamment du contexte vocalique.

Les travaux de Blumstein & Stevens (1979) concluent à une relation non linéaire entre l'articulation et les propriétés acoustiques des sons à la sortie du système de production. En d'autres termes, cela signifie qu'un paramètre articulatoire peut varier de façon linéaire et avoir des conséquences acoustiques non linéaires. Cela suppose que

des variations à l'intérieur de certaines zones articulatoires n'ont pas de conséquences acoustiques (résultat acoustique stable). Inversement, une petite variation articulatoire dans certaines zones articulatoires peut avoir des répercussions acoustiques importantes.

Les relations articulatoire-acoustiques sont d'ordre quantique, c'est-à-dire que les variations acoustiques se produisent par palier, tandis que les variations articulatoires se font de manière continue. Les relations entre les ondes acoustiques et les sons perçus sont aussi quantiques.

Les relations quantiques des paramètres de la production de la parole (paramètres acoustiques, perceptifs et articulatoires) sont des facteurs déterminants dans le choix des attributs articulatoires et acoustiques utilisés à des fins distinctives dans la parole. Ces relations permettent l'émergence des traits distinctifs.

Ladefoged (1983) n'adhère pas à la théorie quantique de Stevens, étant donné qu'il pose que les similitudes entre les langues du monde soient le fait de « facilités » articulatoires. Il pense que si les voyelles [i a u] sont présentes dans la plupart des langues du monde, c'est parce qu'elles correspondent aux réalisations les plus extrêmes du conduit vocal, et sont donc très distinctes les unes des autres.

Browman et Goldstein (1989) retiennent une partie de la théorie quantique de Stevens parce qu'elle permet d'expliquer non seulement les gestes que les langues peuvent choisir à des fins contrastives, mais aussi l'organisation de ces gestes.

Plus récemment, Clements & Ridouane (2006), dans leur revue de la Théorie Quantique, proposent deux grandes familles de définitions des traits quantiques :

(1) Les définitions contextuelles : les indices acoustiques ou auditifs seraient détectés uniquement lorsque les traits ont une position dans un contexte approprié.

(2) Les définitions intrinsèques : les indices se trouveraient alors dans le segment lui-même.

Considérer que la production de la parole, au regard de la théorie quantique, permet de mettre au jour la relation non linéaire qu'entretiennent les gestes articulatoires et leurs conséquences acoustiques. En effet, il a été observé que la variation de certains paramètres articulatoires peut conduire à une réorganisation non négligeable de l'espace acoustique, tandis que la modification d'autres paramètres n'aura pas ou peu d'impact sur le niveau acoustique. Il faut également noter qu'il n'existe pas de corrélation entre l'importance de la variation articulatoire et sa conséquence acoustique. Certaines

productions, comme la voyelle [a], acceptent une plus grande variabilité des gestes articulatoires. Pour cette voyelle, la zone cible se situe au niveau de la constriction pharyngale, les autres régions du conduit vocal pouvant adopter des configurations plus ou moins libres. Il est alors dit que l'espace de viabilité de ce type de son est plus large, hors de la région pharyngale critique, c'est-à-dire que sa réalisation suppose moins de contraintes articulatoires pour un résultat acoustique stable. C'est pour cette raison, entre autres, que nous avons retenu le [a] pour l'analyse de nos données intra locuteurs ; nous y reviendrons.

1.2.1.3. Le modèle Directions Into Velocities of Articulators (DIVA) de la production de la parole

Développé dans les années 90 et exposé pour la première fois par Guenther (1994), le modèle DIVA (Directions Into Velocities of Articulators) tient compte de phénomènes liés au contrôle moteur, à la variabilité du contexte, aux variations de la vitesse d'élocution, ainsi qu'aux faits coarticulatoires (Perkell & Nelson (1985), Guenther (1994, 1995a,b) Guenther & Barreca (1997) Guenther et *al.* (1998), Perkell et *al.* (2000)). En outre, DIVA n'est pas seulement un modèle de production de la parole dans la mesure où il apporte un regard novateur sur les théories portant sur l'acquisition de la parole (Guenther, 1994).

En effet, le babillage, phase d'apprentissage de la production de la parole chez l'enfant, est au centre du modèle. Durant cette période, l'enfant explore les différentes possibilités articulatoires et leurs conséquences auditives. Plus exactement, il réajuste sans cesse ses mouvements, durant cette phase d'acquisition de la parole, dans le but de s'approcher du son cible, à l'aide d'un réseau neuronal multidimensionnel (auditif, perceptif et articulatoire). Notons que le feedback a un rôle prépondérant durant le babillage. Selon Guenther & Perkell (2004), les cibles acoustiques apprises dès le plus jeune âge et tout au long de la vie demeurent acquises et se maintiennent dans le temps. Un argument en faveur de cette hypothèse est que si les personnes nées sourdes présentent des retards d'apprentissage du langage ainsi que des difficultés d'intelligibilité, les personnes devenues sourdes sont capables de continuer à parler de façon intelligible longtemps après l'apparition de leur handicap.

D'un point de vue articulatoire, le modèle DIVA suppose l'existence d'un contrôleur de haut niveau qui planifie et contrôle le mouvement des articulateurs de la

parole, à l'aide du feedback auditif. Le mouvement des articulateurs est motivé par l'atteinte de cibles qui, contrairement aux hypothèses de Browman et Goldstein (1989 et 1992) ou de Kelso et *al.* (1986), seraient non pas articulatoires mais acoustiques et seraient le résultat de paramètres aussi bien acoustiques qu'articulatoires.

Ces cibles (qui peuvent être les formants des voyelles ou les constriction du conduit vocal par exemple) sont des régions spatio-temporelles et auditives qui seraient stockées au niveau du SNC. Il est utile de préciser qu'une cible ne correspond pas à un point, mais plutôt à une zone, raison pour laquelle les chercheurs ayant participé au développement de DIVA considèrent que plusieurs cibles peuvent correspondre à un même son. Ainsi les cibles peuvent varier en fonction du contexte ou de la vitesse d'élocution par exemple. De même, la cible d'un phonème peut être modifiée par rapport au phonème qui précède (coarticulation régressive) ou en fonction du phonème qui suit (coarticulation progressive), ce qui se justifie par les notions d'économies articulatoires qui régissent également la production de la parole (Guenther & Perkell, 2004).

Comme il est possible de le constater, dès le début du développement du modèle DIVA, la variabilité en production de la parole est prise en compte. Plus précisément, si le contrôle moteur s'applique à produire les traits distinctifs, les traits secondaires sont davantage variables dans les domaines acoustiques (Perkell & Nelson, 1985).

L'étude de la parole pathologique est source d'informations sur la parole normale. Il s'agira donc de confronter ces théories, qui mettent l'accent sur la notion de cible acoustique, à notre étude sur des patients dysphoniques. En effet, la dysphonie étant une perturbation de la voix, il s'agira d'évaluer les conséquences de celle-ci sur le signal acoustique.

I.2.2. L'existence d'une cible articulatoire ?

I.2.2.1. Phonologie Articulatoire

Les recherches de Stetson (1951) proposaient déjà de considérer la production de la parole comme « une séquence de mouvements rendus audibles [plutôt que comme] une séquence de sons engendrés par des mouvements ». Dans cette perspective, c'est la

notion de cible articulatoire qui était mise au centre du processus de production de la parole, une attention moindre étant portée au niveau acoustique et perceptif.

Les travaux d'Ohman (1966) mettent en lumière la difficulté de définir une catégorie phonétique uniquement à partir de paramètres acoustiques. Le contexte vocalique peut, par exemple, modifier l'information formantique liée à une consonne donnée (Liberman et *al.* 1954). Dès lors, l'existence d'une cible acoustique ne peut être suffisante pour définir les processus responsables de la production de la parole.

Browman et Goldstein (1989 ; 1990a, b, c) proposent le geste comme unité de base de la structure phonologique. Les gestes sont des unités discrètes pré-linguistiques et inhérentes au développement de l'enfant. Ils contiennent en outre une spécification spatio-temporelle et jouent le rôle de structures coordinatives. Le geste est invariant (Liberman & Mattingly (1985), par exemple) et il est défini dans l'espace articulatoire (géométrique) selon deux variables : le lieu d'articulation et l'aire de la constriction. L'apparente variabilité des configurations spatio-temporelles pour un phonème donné est expliquée par le fait que certains articulateurs peuvent être impliqués dans la réalisation de plusieurs segments simultanément. Ces chevauchements permettent notamment d'expliquer le phénomène de la coarticulation.

Les gestes ont donc une double fonction : ils participent non seulement à la représentation des contrastes phonologiques, mais également à la représentation de l'action articulatoire et la description physique du processus de parole ; une même unité sert ainsi à deux niveaux de description.

Les travaux de Lindblom (1991) nuancent le principe d'invariance du geste. En effet, le choix d'un geste parmi un ensemble est régenté par des lois et répond donc à des adaptations et à des contraintes, à des mécanismes moteurs et perceptifs, indépendants du langage et non spécifiques à la parole. Pour lui, le geste articulatoire est malléable et témoigne donc d'une grande plasticité ; nous y reviendrons.

Enfin, il faut également considérer que la Phonologie Articulatoire est une théorie qui souffre de certaines limites, notamment parce qu'elle manque de contraintes suffisamment explicites et prédictibles, ce qui lui donne une puissance trop importante. Elle est parfois perçue comme un modèle trop abstrait pour être réaliste. D'autres critiques, enfin, reprochent à la Phonologie Articulatoire la prépondérance donnée à

l'articulatoire au détriment des caractéristiques acoustiques et perceptives (Fougeron, 2005).(cf. par ex. pour un état de la question)

I.2.2.2. Le modèle de la Dynamique des Tâches

Dès 1986, les travaux de Kelso et *al.*, souhaitent réconcilier la variabilité observée par les expériences acoustico-articulatoires et l'existence d'une invariance profonde en parole, soutenue par les analyses linguistiques traditionnelles. Le modèle de la Dynamique des Tâches est alors proposé afin de rendre compte de la cinématique des articulateurs et de mettre au jour le rôle des synergies entre les articulateurs mobilisés pour la production de la parole, mais également pour saisir les actions dirigées vers une cible.

Ce modèle, qui met en évidence l'existence d'une partition gestuelle dans laquelle chaque geste est défini par un ensemble d'équations, spécifie non seulement l'ensemble des valeurs des paramètres dynamiques mais également l'intervalle temporel durant lequel il est activé (Saltzman, 1986 ; Saltzman & Munhall, 1989 et Saltzman & Byrd, 2000). Kelso et *al.* (1986) attachent une importance toute particulière à la notion de variables dans le conduit vocal, c'est-à-dire que chaque variable est associée à un ensemble d'articulateurs et leurs mouvements permettent de déterminer la valeur de cette variable.

Actuellement, ce modèle spécifie les gestes selon une ou deux variables (lieu et degré de la constriction par exemple), ce qui permet de les définir à l'aide de descripteurs renseignant sur l'articulateur impliqué, mais également sur sa dynamique. Chaque geste est alors défini par un modèle descriptif qui lui est propre. Cette description présente un double avantage : elle décrit le mouvement articulatoire et propose une classification contrastive. Pour la description du français, les critères traditionnellement retenus sont le mode articulatoire, le lieu d'articulation, les traits d'oralité et de nasalité et la labialité.

Le modèle de la Dynamique des Tâches présente l'avantage de prendre en considération la variabilité inhérente à la production de la parole. La suite de ce chapitre se focalisera sur cette notion, afin d'observer comment les recherches en phonétique ont tenté de prendre en considération ce paramètre.

Nous avons essayé de définir la notion de cible articulatoire en production de la parole, puisque dans notre travail expérimental, nous tenterons d'observer comment la perturbation de la source laryngée peut éventuellement avoir des conséquences sur les cibles articulatoires. Il s'agira d'observer si nos locuteurs sont gênés pour atteindre les cibles articulatoires, en raison des conséquences de l'opération chirurgicale. Nous savons qu'une opération au niveau du larynx peut gêner son fonctionnement adéquat ; il s'agira alors d'observer si nos locuteurs atteignent tout de même les cibles et si l'intégrité du message linguistique est conservée.

I.2.3. Invariance et Variabilité

La section précédente a montré que la notion d'un invariant, qu'il soit acoustique ou articulatoire n'est pas toujours facile à déceler à l'apparente variabilité en production de la parole. En effet, la production de la parole est un processus qui autorise une certaine variabilité, elle-même toutefois soumise à de fortes contraintes. C'est cette tension entre variabilité et contrainte que nous proposons d'étudier dans la suite de la section.

I.2.3.1. La Théorie de la Variabilité Adaptative

L'Hyper- & Hypo-Speech Theory (Lindblom, 1987) admet la variation phonétique inter et intra locuteurs. Selon cette théorie, les performances du locuteur doivent s'accorder avec les exigences situationnelles et les connaissances linguistiques et phonétiques de l'auditeur. Le fonctionnement du système moteur reposerait donc sur deux facteurs :

- une exigence de l'auditeur de distinctivité perceptive ;
- une demande d'économie articulatoire du locuteur.

Ces deux notions seraient responsables de la variabilité en production de la parole. Celle-ci est perçue comme un continuum, et serait adaptative dans le sens où des phases d'Hypo- (phases de sous-articulation) et d'Hyper-articulation (phases de sur-articulation) peuvent alterner en fonction de la situation de communication. Le locuteur contrôle de façon permanente les facteurs de production de la parole en fonction du signal acoustique, le signal de sortie. L'unité phonétique ne reste pas stable, puisqu'elle subit les fluctuations en fonction de la situation de communication, permettant au locuteur d'adapter ses stratégies, lors de la production de la parole.

La variabilité en parole s'exprime par la dichotomie Hyper- et Hypo-articulation ; le locuteur est confronté à une double exigence :

- d'économie articulatoire dans la production de son message linguistique ;
- de contraste perceptif suffisant pour l'émergence de la compréhension.

Un compromis entre la demande de compréhension de l'auditeur et l'effort articulatoire du locuteur est ainsi recherché. La variabilité en parole est donc une conséquence de la stratégie d'adaptabilité du locuteur et de la capacité de récupération de l'auditeur.

La dichotomie hyper- et hypo-articulation explique également le fait que les paramètres de description de la parole peuvent être variables à tous les niveaux, pour un message linguistique invariant.

Cette théorie conçoit également que les segments phonétiques sont produits par une chaîne de commandes sur un temps très court, pouvant amener le locuteur à produire une parole déviante à cause du ratage des cibles acoustiques et/ou articulatoire. L'accélération de la vitesse d'élocution, ou une autre contrainte articulatoire (*bite block*, anomalie de conduit vocal), peuvent être responsables de la non-réponse ou d'une réponse incomplète à une première commande, alors que la deuxième commande est déjà planifiée, entraînant alors un ratage de cible.

L'hypo-articulation répond à une tendance du système moteur à l'économie d'effort dans tous les types de mouvements, qu'ils soient ou non consacrés à la parole. Toutefois, si le système vise uniquement l'économie d'effort, alors les phénomènes de ratage de cibles (*undershoot* et *overshoot*) seront fréquents. Outre les propriétés relatives à l'économie d'effort et intrinsèques au système de production de la parole, le locuteur met en place des stratégies de compensation pour réorganiser le système, afin d'éviter d'éventuels ratages de cibles.

L'hyper-articulation est également un phénomène qui révèle la plasticité adaptative de notre système de production de la parole. Le locuteur est capable de sur-articuler si la situation de communication le demande. Autrement dit, le locuteur peut compenser des situations de perturbations externes (lieu bruité) ou internes (*bite block*, anomalie du conduit vocal) en hyper-articulant le message, cela se faisant au détriment de la loi du moindre effort. La sortie acoustique sera équivalente à celle qu'il aurait pu

produire en situation normale de communication ; en revanche, les mouvements des articulateurs sont différents (Lindblom & Sunberg 1971). Nous retiendrons donc que la plasticité du système est capable de compenser les contraintes articulatoires.

En conclusion, les notions de plasticité, de flexibilité et de moindre effort sont responsables de la variabilité de la parole. De plus, ce modèle permet d'expliquer les stratégies mises en place pour compenser des perturbations naturelles (effets des articulateurs les uns sur les autres) et provoquées (expérimentateur).

La théorie de la Variabilité Adaptative en production de la parole renonce à l'idée de l'existence d'un Invariant Physique Absolu (1987). De ce fait, elle prône la prise en considération de deux aspects dans la production de la parole qui sont les notions d'Hyper-articulation et d'Hypo-articulation. Ces deux notions sont le fondement du cadre théorique de la *H & H Theory* de Lindblom.

Il pourra se révéler intéressant d'étudier les stratégies de nos locuteurs souffrant de paralysies récurrentielles à la lumière de la Théorie de la Variabilité Adaptative. Comment ces locuteurs vont-ils compenser leur dysphonie ? Auront-ils recours à des stratégies de sur-articulation pour rendre leur parole intelligible en insistant notamment sur certains contrastes phonologiques, ou au contraire, préféreront-ils adopter des stratégies d'économie articulatoire ? L'analyse de nos données tentera de répondre à ces questions.

I.2.3.2. Coordination et coarticulation en production de la parole

Nous savons que les sons ne sont pas produits les uns indépendamment des autres, mais que la production de la parole repose sur un principe fin de coordination qui donne lieu à des degrés de coarticulation variables.

Cette section devrait nous permettre d'identifier les notions relatives aux principes de coordination et de coarticulation. Fowler & Saltzman (1993) proposent d'étudier ces principes à l'aide du modèle de la Dynamique des Tâches (voir *supra*).

La coordination fait partie intégrante du geste phonétique, elle reflète l'établissement transitoire des relations contraintes par le mouvement des articulateurs afin de faire émerger des traits linguistiquement significatifs. Ces contraintes sont liées aux chevauchements articulatoires nécessaires en production de la parole. Elles sont

également liées à des composants bien définis, et réduisent significativement le degré de liberté du système de production de la parole. Il existe une dépendance naturelle entre les différents articulateurs, ce qui permet de créer des ensembles macroscopiques donnant lieu à des gestes phonétiques. Le geste de fermeture bilabiale, par exemple, nécessite un mouvement de la mandibule et des lèvres inférieure et supérieure.

La coarticulation est alors le résultat de l'activation temporelle des contraintes de coordination essentielles pour différents gestes phonétiques. C'est-à-dire que pour n'importe quel point t de l'énoncé, il est possible d'observer, grâce aux modèles acoustique et articulatoire, l'influence des gestes associés à plusieurs segments adjacents.

Selon Fowler et Saltzman (1993), la coordination et la coarticulation sont nécessaires pour l'émergence d'une organisation spécifique du conduit vocal. Un geste serait alors le résultat de mouvements articulatoires, même si tous les mouvements ne peuvent pas être considérés comme des gestes. Enfin, ces deux notions expliquent pourquoi dans certains contextes, un segment peut transformer les caractéristiques d'un autre, si le trait phonologique de ce dernier ne présente pas ou peu de résistance (Daniloff & Hammarberg, 1973).

Ces principes de coarticulation et de coordination permettent également de mettre au jour les contraintes articulatoires et d'en identifier la nature et l'importance. Recasens (1999, 2002), dans son modèle DAC (*Degree of Articulatory Constraints*), indique que le degré de contraintes articulatoires associé à un geste aura un impact sur l'étendue temporelle et la direction du phénomène de coarticulation. Les effets de coarticulation pourraient ainsi être, pour différents articulateurs, étroitement liés aux modèles de coordination inter-articulatoire.

Toutefois, certains auteurs (Fowler 2007 ; Hardcastle, 1985, 1999, 2006 ; Marchal, 1988) s'accordent à dire que ce phénomène de coarticulation reste limité, dans la mesure où l'essence linguistique de chaque segment d'une séquence reste conservée. Ils admettent néanmoins l'existence de chevauchements temporels ou coproductions des contraintes, de coordination dans les processus de production de la parole. Le point clé résiderait alors dans la dynamique des interactions gestuelles durant les chevauchements.

Les travaux de Fowler & Saltzman (1993) indiquent d'une part que l'invariance réside dans l'atteinte de la cible, d'autre part que la trajectoire des articulateurs peut varier, et enfin que les contraintes permettent que la cible soit obtenue.

I.2.3.3. La Théorie de la Viabilité

I.2.3.3.1. Définition générale

Cette théorie, développée au sein de l'Institut de Phonétique de Strasbourg à partir de données sur la parole normale et pathologique (Sock, 1998, 2001 ; Sock et Vaxelaire, 2004 ; Vaxelaire, 2007) conçoit la notion de cible comme un espace de réalisation de possibles articulatoire-acoustiques, perceptivement acceptables. Cette conception nous semble particulièrement intéressante dans la mesure où les locuteurs pathologiques réorganisent leurs productions selon leurs propres contraintes physiologiques et anatomiques provoquées par leur pathologie. Il s'agit pour eux aussi de respecter un espace dit viable pour l'émergence du phonème souhaité.

La Théorie de la Viabilité est une théorie mathématique qui fournit des métaphores mathématiques de l'évolution de macro-systèmes apparaissant dans divers domaines scientifiques, dont les sciences cognitives (*cf.* Aubin, 1991). L'objectif principal de cette approche est de rationaliser, dans le domaine des sciences cognitives, l'évolution d'un organisme régulé par des dynamiques non déterministes spécifiques, ainsi que par des contraintes de viabilité, afin de mettre au jour l'interaction des diverses composantes auto-organisées du système sous-jacentes à sa régulation.

En admettant que le système de production-perception de la parole soit un système auto-organisé, intégré dans cette organisation globale qui est le tout du vivant, lui-même auto-producteur et auto-organisé, la démarche épistémologique consistera à appréhender l'apparition, ainsi que le comportement des catégories phonétiques et phonologiques, comme étant des processus neurocognitifs émergeant d'un système soumis à des contraintes sensori-motrices de viabilité (voir, par exemple, Sock, 1998 et 2001).

La notion de perturbation est au centre de cette théorie, elle est considérée comme l'interaction structurelle à laquelle une unité est capable ou non de résister grâce à sa plasticité (Sock, 1998).

La Théorie de la Viabilité suppose que perturber un système permet d'en évaluer la robustesse. En ce qui concerne la production de la parole, la perturbation peut émaner de deux sources majeures (Vaxelaire & Sock, 1996 ; Vaxelaire, 2007) : (i) L'environnement externe constitue la première source de perturbation. Il s'agit d'événements non déterminés et indépendants du système de production de la parole. Un environnement bruité par exemple constitue une entité physique ayant sa propre clôture opérationnelle. (ii) La perturbation du système peut également lui être intrinsèque, suite aux modifications des interactions structurelles de ses propres composantes. Les modifications provoquées par une pathologie du langage ou de la voix entrent directement dans cette catégorie.

I.2.3.3.2. Perturbations et réajustements

Les perturbations, qu'elles soient internes ou externes, du système de production-perception de la parole peuvent engendrer des changements compensatoires.

Ces changements peuvent être conservateurs, c'est-à-dire qu'ils ne provoquent pas de modifications dans la stratégie de réalisation de l'opposition linguistique : les composantes conservent une stabilité relative dans leurs interactions. A titre d'exemple, notons que faire varier la vitesse d'élocution pourrait parfois provoquer la compression des durées absolues des classes phonétiques mais ne modifie pas les stratégies gestuelles utilisées pour maintenir les oppositions de quantité vocalique, par exemple (Sock, 1998).

Les changements peuvent également être innovateurs et, dans ce cas, les composantes du système subissent elles-mêmes des modifications. Les stratégies de réalisation de l'opposition linguistique sont modifiées. Il est probable que nos données en parole pathologique s'inscrivent directement dans cette catégorie.

La plasticité du système de production de la parole, mais également sa tendance à l'équilibre homéostatique permettent l'intervention de changements compensatoires, sans que l'organisation globale du système en soit modifiée. Le système est donc capable de maintenir, de manière relativement stable, la dynamique de ses états internes et la cohérence globale de son organisation. Toutefois, il arrive que le système de production de la parole ne soit plus capable de pallier les perturbations internes ou externes. La stabilité nécessaire à l'émergence du message linguistique est alors

menacée, ce qui peut conduire à la désintégration du système et à la perte de l'identité de ses propriétés émergentes. Ce type de scénario est notamment observable dans le cas de la parole pathologique en cas d'atteinte sévère du système de production-perception de la parole. L'étude de Béchet (2011) sur des locuteurs présentant une fente labio-palatine illustre les limites de la viabilité de certains sons. Dans certains cas, lorsque certaines contraintes articulatoire-acoustiques ne peuvent plus être respectées, l'intégrité du message linguistique est menacée. C'est ainsi que certains locuteurs, malgré une tentative de réaménagement de leur système de production de la parole, ne parviennent pas à produire le son souhaité. La perturbation anatomique du conduit vocal entraîne alors un amoindrissement de l'intelligibilité ; ces productions déviantes indiquent les limites du cadre de la viabilité en production-perception de la parole. Toutefois, elles permettent de mettre en exergue les zones de viabilité des sons étudiés. Chaque locuteur, sain et pathologique, est capable à partir de son propre système de production de la parole de proposer des ajustements qui lui sont spécifiques. Ces ajustements sont naturellement plus accentués chez les locuteurs pathologiques. De ce fait, leurs données attestent de l'existence d'une zone, différente chez chaque locuteur, mais à l'intérieur de laquelle les gestes articulatoires réalisés pour l'émergence d'un son doivent être viables, dans la mesure où l'intégrité du message linguistique doit être conservée.

Il nous paraît important de signaler ici que la notion de viabilité en production de la parole est directement liée non seulement à la notion de résistance des oppositions phonologiques, mais est également en relation avec les trajectoires adoptées par des catégories articulatoire-acoustiques, sur le plan temporel. Il est évident que la production d'un locuteur pathologique oscille entre perturbation et résistance, afin de conserver les réalisations à l'intérieur de zones dites viables, dans le but de garantir l'émergence d'un message linguistique compréhensible.

Pour conclure, la Théorie de la Viabilité indique qu'une multitude de solutions articulatoires peut mener à une même cible acoustique du moment que ces solutions soient situées à l'intérieur de la zone de viabilité qui sous-tend l'émergence d'un son. De la même façon, plusieurs solutions acoustiques peuvent permettre d'accéder à une cible perceptive équivalente. Il y aurait donc deux niveaux de variabilité possibles dans le système de production-perception de la parole, offrant une multitude de stratégies envisageables et réalisables par les locuteurs. Cela est particulièrement prometteur pour

capturer certaines données sur la parole pathologique ; des compensations à différents niveaux sont possibles pour les locuteurs pathologiques, afin d'atteindre des cibles acoustiques et perceptives comparables à celles de locuteurs sans trouble de la parole.

Nous considérons donc que l'étude de la parole pathologique peut fournir des informations sur la production de la parole normale, notamment grâce à la notion de variabilité. Il devrait donc être concevable d'étudier l'étendue de la variabilité, mais également d'en définir les limites.

CONCLUSION CHAPITRE I

Si deux courants théoriques s’opposent en production de la parole selon que la théorie soit orientée cible acoustique ou articulatoire, elles convergent toutefois vers l’idée que le geste ne peut être décrit de façon univoque. Le geste est non seulement étroitement lié au contexte de production, mais également aux notions de contrôle moteur et de timing des articulateurs. De plus, si un système de contraintes de plusieurs niveaux (phonologique, bio-mécanique et coarticulatoire) régit effectivement la production de la parole, une certaine variabilité est acceptée et n’entrave pas la réception correcte du message linguistique.

La notion de variabilité étant intrinsèque à la production de la parole, notre étude acoustique sera menée afin d’évaluer la plasticité de ce système, et d’identifier les possibles réajustements ou compensations qu’un locuteur peut mettre en place suite à une perturbation d’origine pathologique. Nous tenterons de mettre en lumière des zones de variabilité et de viabilité possibles et d’en dessiner les limites dans le contexte de la parole pathologique.

RÉSUMÉ CHAPITRE II

Ce chapitre sera consacré aux phénomènes de perturbations et de réajustements au niveau de la production et de la perception de la parole. De manière générale, il s'agit d'observer le degré de flexibilité du système de production et de perception de la parole à la suite d'une perturbation non pathologique, de comprendre le fonctionnement du système de production et de perception de la parole à partir d'un dysfonctionnement et de connaître les limites des déviations physiques (articulatoires et acoustiques) imposées par les exigences linguistiques de clarté.

La première partie de ce chapitre présentera les différentes méthodes de perturbations de la production de la parole, alors que la deuxième partie sera réservée aux méthodes de perturbations de la perception de la parole. Dans les deux parties, il sera question des stratégies de compensation ou des réajustements mis en place en réponses aux différentes perturbations non pathologiques.

CHAPITRE II

II. Perturbation de la production et de la perception de la parole normale

Si la production de la parole est centrale aux sciences phonétiques, la perturbation du système de production de la parole a également donné lieu à un grand nombre de recherches, et cela depuis plus d'un siècle. Ainsi, l'Abbé Rousselot témoignait déjà de l'intérêt de la notion de perturbation en phonétique expérimentale lorsqu'il écrivait : « ... on peut, en effet, très bien arriver à émettre toutes les voyelles labiales en les [lèvres] écartant avec les doigts » (Rousselot, 1897). L'Abbé Rousselot avait ainsi déjà intégré la notion de réajustement dans son cadre conceptuel de l'étude de la parole.

Perturber le système de production de la parole est de ce fait un terrain d'investigation à part entière en phonétique expérimentale. Plus précisément, il s'agit de comprendre le fonctionnement de la production de la parole à l'aide d'une perturbation qui vient gêner le déplacement des différents articulateurs ; cela permet d'évaluer la plasticité du système de production, étant donné qu'il existe des bornes dans l'étendue de la variabilité, spatiale et temporelle, des gestes responsables de la production de la parole, au-delà desquels la compréhension du message linguistique n'est plus assurée. En d'autres termes, apporter une contrainte lors de la production de la parole permet non seulement d'étudier les stratégies de réajustement déployées par le système, mais aussi d'évaluer le résultat de cette réadaptation.

En ce qui concerne la nature de cette perturbation, on en distingue généralement deux types, à savoir les perturbations de la parole dites « non pathologiques » et les perturbations « pathologiques » qui affectent directement ou indirectement l'un des organes nécessaires à la production de la parole.

II.1. Perturbation de la production de la parole

Cette partie présentera les méthodes de perturbation « non pathologiques » de production de la parole, les plus étudiées en phonétique expérimentale. Dans un premier temps, seront présentées quelques méthodes dites écologiques de perturbation de la production de la parole qui consistent, par exemple, à faire varier la vitesse d'élocution (en accélérant ou en ralentissant le débit de parole), ou l'intensité (voix criée ou voix chuchotée). Dans un second temps, les perturbations non-écologiques nécessitant la mise en place d'un dispositif expérimental (Bite-block, lip-tube, palais artificiel, perturbation dynamique de la production de la parole,...) seront abordées.

II.1.1. Perturbations de types écologiques

Cette section passera en revue les différentes méthodes de perturbations dites « écologiques » de production de la parole. Nous choisissons le terme « écologique » dans le sens où un locuteur, dans une situation de communication de la vie quotidienne, à cause de contraintes extérieures, peut être amené à faire varier sa vitesse d'élocution ainsi que l'intensité de sa voix. La perturbation est dite écologique puisqu'elle est produite par le locuteur lui-même en opposition avec une perturbation non écologique qui serait la conséquence de l'ajout d'un corps étranger (bite-block, pipe-speech, etc.) dans le système de production de la parole.

Si ces perturbations sont écologiques, elles n'en sont pas moins intéressantes à étudier de par leur caractère naturel, car le locuteur y joue un double rôle : il produit mais subit également la perturbation, dans la mesure où il crée lui-même à la fois une perturbation et un réajustement au niveau de la production de la parole.

II.1.1.1. Variation de la vitesse d'élocution

Signalons tout d'abord qu'il convient de différencier la vitesse d'élocution de la vitesse d'articulation, cette dernière correspondant uniquement au nombre de syllabes produites par seconde. La vitesse d'élocution, dont il est question dans cette partie, prend en considération à la fois la vitesse d'articulation (*i.e.* le nombre de syllabes produites par seconde) et la durée des pauses. Signalons que nous utilisons ici le terme de « débit » et « vitesse d'élocution » indifféremment.

La vitesse d'élocution est le perturbateur naturel par excellence de l'organisation spatio-temporelle des gestes de la production de la parole. Ainsi, Vaxelaire (2007) qualifie la vitesse d'élocution comme « une variable naturelle mais complexe en production de la parole ». Chaque locuteur est ainsi capable de faire varier son débit de parole en fonction des contraintes de la situation de communication dans laquelle il se trouve.

Les travaux de Gaitenby (1965) sont parmi les premiers à souligner la notion d'élasticité du signal de parole qui est décrite comme la capacité, pour un énoncé, à subir des expansions ou des compressions selon que la vitesse d'élocution est ralentie ou accélérée. L'élasticité des énoncés reste toutefois limitée, d'une part par les contraintes imposées par le système anatomo-squelettique de la production de la parole et, d'autre part, par le besoin de clarté qu'exige la situation de communication.

Selon Grosjean & Deschamps (1975), la vitesse d'articulation serait une donnée propre et constante à chaque locuteur. D'après cette étude, accélérer le débit de parole consisterait uniquement en une réorganisation de la durée des pauses. Plus le débit de parole est rapide, plus la durée des pauses se raccourcit, sans que cela ait un impact sur le nombre de syllabes produit par seconde.

Néanmoins, cette étude ne fait pas consensus dans la littérature puisque d'autres travaux, comme ceux de Vaxelaire (1993) ou Sock & Vaxelaire (2001), proposent une réorganisation plus globale du timing des gestes de la parole. Ces recherches montrent en effet qu'une augmentation de la vitesse d'élocution aurait des répercussions, non seulement sur la durée des pauses, mais aussi sur l'ensemble des gestes de la parole concernés par la réorganisation temporelle. Il convient, selon ces deux études, de distinguer les patterns, ou les coordinations entre gestes du conduit vocal, robustes, des patterns plus fragiles. Si, en fonction de la vitesse d'élocution, les patterns restent

stables en termes absolus et/ou relatifs, il sera possible d'évaluer leur robustesse ; *a contrario*, s'ils sont sujets à des changements, il faudra évaluer leur fragilité (Vaxelaire, 2007). Les réajustements moteurs qui touchent parfois aux structures spécifiques du conduit vocal seraient donc la conséquence de la résistance ou de la robustesse de l'organisation spatio-temporelle des gestes. Ils révéleraient ainsi les éventuelles stratégies de compensation motrice.

Outre les contraintes physiques universelles qui restreignent intrinsèquement la vitesse d'élocution, chaque langue obéit différemment à l'augmentation de la vitesse d'élocution, du fait de ses caractéristiques linguistiques (phonologiques, phonétiques et/ou prosodiques) propres. Notre thèse ne portant pas sur une variation de la vitesse d'élocution en tant que facteur expérimental, nous n'avons pas la prétention de dresser un état exhaustif de la littérature sur le sujet.

Néanmoins, pour illustrer ce dernier propos, nous pouvons citer, par exemple, les travaux de Peterson & Lehiste (1962) ou de Port (1980) sur l'anglais, qui est une langue à accent lexical. Ces études indiquent que les gestes des voyelles accentuées sont plus robustes et semblent donc maintenus avec l'accélération de la vitesse d'élocution, au détriment des gestes des voyelles inaccentuées. Cela peut probablement s'expliquer par le fait que cette stratégie permet de conserver l'information lexicale contenue dans le segment accentué.

Il en est de même pour l'opposition entre voyelles longues et brèves en wolof (Sock et al. 2006) lorsque la vitesse d'élocution est rapide, la durée des voyelles longues reste plus importante que celle des voyelles brèves, afin de préserver le contraste phonologique.

Par rapport au français, l'étude de Vaxelaire (2007) sur les effets d'un changement de débit sur les consonnes simples et les consonnes doubles (hétéromorphémiques et homorganiques) a montré une étendue de contact bilabial, linguo-palatal ou linguo-vélaire plus importante pour les doubles par rapport aux simples, en vitesse d'élocution normale. Lorsque le débit est accéléré, l'opposition entre consonnes simples et doubles est maintenue d'un point de vue articulatoire ; l'étendue de contact restant plus grande pour les consonnes contigües.

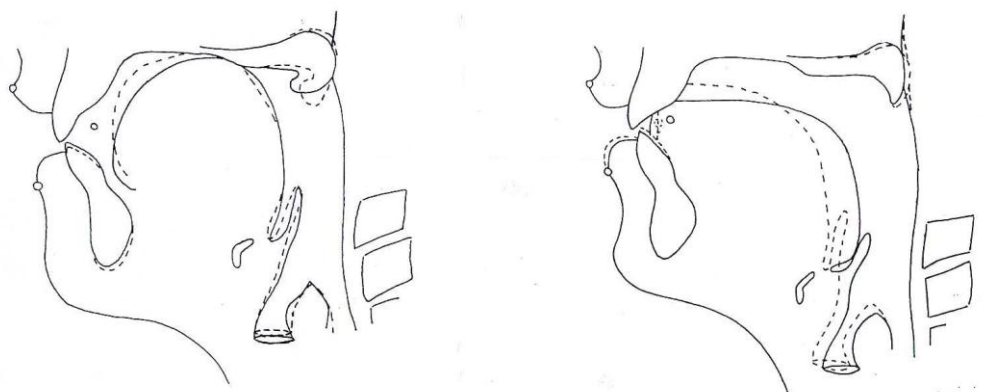


Figure 1 : Vues de profil montrant la production des séquences /aka/ et /akka/ (à gauche) et des séquences /ada/ et /adda/ (à droite), en vitesse d'élocution rapide. On observe des étendues de contact plus larges pour les consonnes doubles (pointillés) par rapport aux consonnes simples (lignes continues). Locuteur A.E. (Vaxelaire, 1995)

Si, jusqu'à présent, il n'a été question dans cette partie que de travaux portant sur l'augmentation de la vitesse d'élocution, d'autres recherches ont également été menées sur les conséquences d'un ralentissement du débit. C'est notamment le cas des travaux de (Zellner, 1998) qui a montré qu'un ralentissement de la vitesse d'élocution provoquait une augmentation des pauses, un allongement de ces mêmes pauses, une augmentation de la durée de l'ensemble des segments et une hausse du nombre de syllabes. En outre, les résultats qualitatifs présents dans la même étude (Zellner, 1998) révèlent qu'un ralentissement du débit de parole aurait tendance à altérer les liens interlexicaux. L'étude ne fait toutefois pas consensus dans la littérature et conclut à une multiplicité de stratégies lors du ralentissement du débit de parole.

Un changement de vitesse d'élocution peut aussi apparaître dans le contexte d'une parole spontanée ou lors de la variation du style de parole. Ainsi, Harmegnies & Poch-Olivé (1992) ont pu montrer que la variabilité était plus remarquable dans la production des locuteurs en parole spontanée, comparée à une parole de laboratoire. En effet, les distances entre les différents sons étaient réduites en parole spontanée, puisque les valeurs de F1 et de F2 avaient tendance à se rapprocher de celles du schwa. Par ailleurs, en examinant les variations des valeurs formantiques, suivant le style de parole, Harmegnies & Poch-Olivé (1994) ont trouvé différentes valeurs de F1 et de F2 selon le style de parole. Ils posent ainsi que produire la parole dans différentes situations de

communication peut engendrer des différences de fréquences formantiques, ce qui aboutirait à de nouvelles catégorisations.

La vitesse d'élocution peut donc être considérée comme un facteur écologique de perturbation de la production de la parole. En faisant varier la vitesse d'élocution, le locuteur est amené à adopter différentes stratégies de production de la parole pour s'adapter aux exigences de clarté linguistique qui sous-tendent toute situation de communication (Lindblom, 1990).

II.1.1.2. Le chuchotement

Le chuchotement peut également être considéré comme une source de perturbation de la production de la parole. En effet, il s'agit d'un mode de communication naturel et relativement intelligible, qui peut être imposé par la situation de communication, au même titre que la variation de la vitesse d'élocution. Il est à noter que ce type de perturbation peut présenter un certain intérêt dans le cadre d'une étude des voix de patients souffrant de paralysies récurrentielles, étant donné que certaines propriétés acoustiques du timbre de ce type de patients sont proches de celles du chuchotement (absence de fréquence fondamentale notamment).

Si la caractéristique principale de la voix chuchotée est d'être produite sans vibration des plis vocaux, d'autres phénomènes anatomico-fonctionnels ont également pu être observés. Ainsi, Tsunoda et *al.* (1994) ont montré une activité plus importante du muscle crico-aryténoïde postérieur (Tsunoda et *al.* 1994) dans cette condition d'élocution par rapport à de la parole normale. L'étude de la glotte, en parole chuchotée, a également révélé que celle-ci prenait la forme d'un Y inversé, les deux plis vocaux étant accolés et ne laissant passer l'air qu'à travers un espace réduit dans la partie postérieure, au niveau des cartilages aryténoïdes. Enfin, les travaux de Tsunoda, et *al.* (1997) indiquent que les structures supraglottiques ne sont pas uniquement resserrées, elles sont également déplacées vers le bas pour éviter complètement la vibration des plis vocaux pendant le chuchotement.

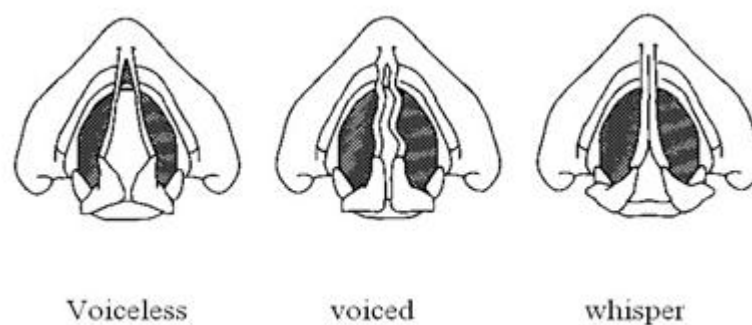


Figure 2 : La glotte en position de phonation à gauche (non voisée et voisée), de chuchotement à droite
(O'Grady, et al. 2004)

Au niveau aérodynamique, Stathopoulos, et *al.* (1991) ont mis en évidence le fait que le volume d'air dans les poumons serait moindre en parole chuchotée. Il en serait de même pour la pression d'air dans la trachée. On observe aussi une augmentation des débits d'air translaryngés et une diminution des résistances de la colonne d'air laryngée. En outre, cette recherche a aussi relevé un nombre moins élevé de syllabes par groupe respiratoire en parole chuchotée, ce qui pourrait être interprété comme un phénomène compensatoire lié à une dépense d'air plus importante en parole chuchotée. Notons encore que les réajustements constatés chez les locuteurs sains en parole chuchotée ne sont pas comparables aux patterns observés chez les locuteurs dysphoniques, caractérisés par des faibles charges de résistance au niveau musculaire (Stathopoulos et *al.* 1991).

Au niveau articulaire toujours, Rubin et *al.* (2006) ont cherché à savoir si la parole chuchotée pouvait provoquer des traumatismes sur les plis vocaux. Pour vérifier cette assertion, le larynx de 100 locuteurs a été observé en voix normale et en voix chuchotée à l'aide de la nasofibroscopie. Leurs résultats indiquent qu'en dehors de quelques cas d'hyperfonction laryngée, la parole chuchotée serait moins traumatique que la parole normale.

Comme cela a été mentionné plus haut, si la voix chuchotée est produite sans vibration des plis vocaux, elle n'en est pas pour autant inintelligible. C'est notamment ce qu'ont observé Higashikawa et *al.* (1996) dans leur recherche portant sur la perception des tons du japonais en parole chuchotée. En d'autres termes, il s'agissait de vérifier si des auditeurs étaient en mesure d'identifier une voyelle à ton montant ou descendant, en sachant que la production d'un ton s'effectue essentiellement grâce à des

variations de la fréquence fondamentale. Les résultats obtenus sur la voyelle /a/ montrent qu'il est possible de distinguer les tons, et cela même en voix chuchotée. De ce fait, les auteurs suggèrent que les modulations nécessaires à la distinction entre les tons ont lieu, en parole chuchotée, à d'autres niveaux que la glotte comme, par exemple, au niveau labial ou lingual.

D'un point de vue acoustique, les différences entre les voix normale et chuchotée ne se situent pas uniquement au niveau de la fréquence fondamentale, puisque la structure formantique des voyelles se trouve également modifiée en parole chuchotée. (Higashikawa *et al.* 1996) ont ainsi montré que les valeurs de F1 chez leurs sujets masculins se rapprochaient des valeurs attendues pour le « ton-cible » en parole chuchotée, ce qui laisse supposer que la hauteur tonale est exprimée par le premier formant lorsque la fréquence fondamentale n'est pas perceptible.

Cette modification des valeurs formantiques est également observée pour l'anglais puisque (Tartter, 1991) a noté une confusion significative entre le son /a/ et le son /o/ en parole chuchotée, confusion qui a été expliquée à la fois par une modification de l'espace vocalique et par un chevauchement des zones formantiques, lors de la réalisation de ces deux sons en voix chuchotée.

Comme nous venons de le voir, chuchoter est un acte perturbateur du système de production de la parole : le locuteur est en effet contraint d'utiliser des stratégies de production différentes que celles employées lors de la production de la parole dite « normale ».

L'étude de la variation de la vitesse d'élocution et de l'intensité a permis de montrer que même en situation écologique de perturbation de la parole, le locuteur est amené à réaliser diverses stratégies pour que le message linguistique puisse émerger correctement.

Outre les différentes perturbations écologiques que nous venons de voir, il est également possible de perturber la production de la parole de façon moins écologique, c'est-à-dire en ajoutant un dispositif externe qui vient gêner les articulateurs mobilisés pour la production de la parole.

II.1.2. Perturbations avec ajout d'un dispositif

Afin de perturber la production de la parole, l'ajout d'un dispositif externe qui vient empêcher (Bite-block), modifier (lip-tube ou palais artificiel) ou gêner (perturbation dynamique) le mouvement des articulateurs mobilisés pour la production de la parole est également envisageable.

Contrairement aux perturbations précédemment évoquées, ce type de perturbation est non écologique, dans le sens où le locuteur n'est pas directement responsable de la perturbation, puisqu'il ne fait cette fois, que la subir. Nous verrons toutefois qu'il peut s'adapter et mettre en place des stratégies de compensation face à ce type de perturbation.

Cette partie permettra d'étudier trois dispositifs externes de perturbations statiques, à savoir le Bite-block, le lip-tube et le palais artificiel, ainsi qu'un dispositif de perturbation dynamique qui vient modifier les déplacements de la mandibule.

II.1.2.1. Bite-block

La méthode du Bite-block a souvent été utilisée pour perturber la production de la parole. Plus invasive que la variation de la vitesse d'élocution, elle reste toutefois relativement naturelle puisqu'elle consiste à bloquer la mâchoire à l'aide d'un dispositif léger placé entre les molaires, qui permet, en mordant dessus, de bloquer les mouvements de la mâchoire inférieure. Cette méthode reproduit en quelque sorte la parole des fumeurs de pipe (*pipe-speech*) qui, lorsqu'ils ont la pipe dans la bouche, voient leurs mouvements mandibulaires remarquablement restreints.

De nombreuses recherches (Lindblom et *al.* (1977) ; Lindblom et *al.* (1979) ; Fowler & Turvey, (1980) ; Kelso & Tuller, (1983), etc.) ont été menées sur la production de la parole à partir du paradigme de bite-block. De manière générale, les études montrent que le locuteur porteur d'un bite-block est capable de réajuster rapidement les mouvements de la langue et des lèvres, permettant ainsi d'atteindre les configurations géométriques du conduit vocal nécessaires aux resserrements, aux constriction et aux occlusions propres à la réalisation des segments phonétiques. La compensation serait immédiate et sans phase d'adaptation longue (Lindblom & Sundberg, 1971). Notons que la position de la mâchoire influe sur les valeurs des trois premiers formants et notamment sur F1 qui renseigne plus ou moins sur l'aperture. Si la

mâchoire est bloquée, la langue ne peut plus s'abaisser avec la même amplitude que lorsque la mâchoire est libre. Ces études n'indiquent généralement pas de différences au niveau de la durée des voyelles et des consonnes par rapport à la perturbation.

Un bémol doit cependant être apporté à l'étude de Lindblom & Sundberg, (1971). En effet, si une compensation immédiate a bien lieu, celle-ci n'est qu'incomplète, dans la mesure où les stratégies de compensation se développent avec le temps. C'est ce que révèle l'étude de Baum & Mcfarland (1997) réalisée à partir de la production d'un corpus en français canadien comportant les voyelles /i, a, u/ et les consonnes /p, t, k, s/, produit deux fois, à 15 minutes d'intervalle. Leurs résultats, avec Bite-block, montrent notamment un abaissement des valeurs centrales pour les occlusives, un allongement de la durée du bruit de friction du /s/ et enfin, une augmentation des valeurs formantiques, liée à l'agrandissement de la cavité antérieure. Lors de la seconde répétition, qui a eu lieu 15 minutes après la première, les différences observées avaient tendance à s'estomper.

Si les études acoustiques menées sur de la parole produite avec ou sans bite-block ont pu révéler des différences minimales en termes de durée et de formant, il semblait intéressant de mener une étude articulatoire sur la parole perturbée à l'aide de cette technique. C'est ce que Lindblom, et *al.* (1979) ont entrepris en étudiant à partir d'images radiologiques des voyelles produites avec un bite-block. L'étude révèle que si les formants des voyelles produites avec bite-block sont relativement analogues à ceux des voyelles en production normale, la configuration du conduit vocal est largement modifiée par le blocage de la mandibule. (Gay, Lindblom, & Lubker, 1981) parlent également de compensation sélective, dans la mesure où ils ont pu constater que les écarts non compensés par rapport aux positions des articulateurs en parole normale ont lieu aux moments où l'aire du conduit vocal est la plus grande. En revanche, une compensation quasi complète a pu être observée aux points de constriction maximale. L'étude conclut à un mécanisme d'atteinte de cible perceptive grâce à la réalisation, même dans le cas de la parole perturbée, de la constriction maximale qui serait alors l'information sensorielle nécessaire.

Enfin, il est important de souligner que les stratégies employées lors de la production de la parole est perturbée par un bite-block sont très variables d'un locuteur à l'autre. Les travaux de Clairet, (2004) portant sur la réorganisation des gestes

articulatoires lors de la production d'occlusives en français avec bite-block ont, en effet, révélé une importante variabilité inter-locuteurs puisque, si des compensations ont été observées suite à l'introduction du Bite-block, chacun des locuteurs testés a développé des réajustements différents sur le plan qualitatif : les électrodes contactées étaient soit diminuées soit renforcées selon les locuteurs.

La méthode du bite-block perturbe comme attendu la production de la parole ; elle permet néanmoins de montrer les stratégies de compensation que le locuteur est capable de mettre en place, dès l'introduction de la perturbation, et d'étudier les réajustements opérés qui peuvent, dans certains cas, amener le locuteur à atteindre des cibles acoustiques comparables à celles de productions non-perturbées de parole. Outre le blocage de la mandibule, l'on peut aussi perturber de façon non naturelle le mouvement des lèvres à l'aide, par exemple, d'un tube labial ou lip-tube.

II.1.2.2. Lip-tube ou tube labial

La méthode du lip-tube se rapproche de celle du bite-block. Elle consiste à introduire un tube entre les lèvres du sujet. La perturbation ainsi introduite empêche les déplacements labiaux, notamment dans l'axe vertical. De plus, en fonction du diamètre du tube, l'ouverture labiale est plus grande que celle normalement requise pour la production de la voyelle choisie. (Savariaux, et *al.* (1995) ont utilisé cette méthode pour étudier la production de la voyelle /u/ chez 11 locuteurs. Pour cette étude, les sujets n'ont pas tous développé les mêmes stratégies de compensation, ce qui met en lumière la notion de variabilité interlocuteur lorsqu'il s'agit de compensation articulatoire. Cependant, leurs résultats vont dans le sens d'une réorganisation complète du timing des articulateurs, ce qui dépendrait de la capacité de chacun à ré-exploiter la relation articulatoire-acoustique déjà mémorisée et éprouvée. Selon ces auteurs, leurs résultats semblent donc indiquer l'existence d'une cible auditive dans la tâche du locuteur.

Placer un tube labial entre les lèvres du locuteur est donc bien une perturbation de la production de la parole qui met au jour des stratégies de compensations nouvelles. Ce corps étranger n'empêche pas le locuteur de mettre en place des stratégies de compensation qui permettent de maintenir la production de la parole, même pour des sons qui nécessitent la protrusion des lèvres, comme la voyelle /u/.

Dans ce qui suit, nous verrons comment gêner une zone de contact ou point d'articulation, en ajoutant un élément dans la cavité buccale. La région alvéolaire peut ainsi se voir modifiée par l'ajout d'un palais artificiel, par exemple.

II.1.2.3. Palais artificiel

Un autre moyen de perturber la production de la parole consiste à placer un palais artificiel dans la région alvéolaire de la cavité buccale, et de mesurer ainsi la capacité du locuteur à produire des sons réalisés dans cette zone. L'épaisseur du palais peut varier de façon à augmenter la perturbation dans le but de tester l'étendue des stratégies de compensation que le locuteur peut mettre en place, suite à l'ajout du palais artificiel. Cette méthode propose donc de réduire les mouvements linguaux pour les sons qui nécessitent une constriction ou une occlusion dans la région alvéolaire.

Les perturbations qu'introduit le palais artificiel sont nombreuses : il déplace le centre de gravité des consonnes, modifie le déplacement de la mandibule et enfin gêne la réalisation des fricatives alvéolaires, en réduisant le resserrement possible.

Les travaux de Baum & Mcfarland, (1997) et de Aasland et *al.* (2006) proposent de tester la production de la séquence /asa/ en introduisant un palais artificiel. Ces études cherchent à mettre en évidence et à quantifier l'adaptation du sujet à l'ajout du palais artificiel dans la cavité buccale. Les données acoustiques (Baum & Mcfarland, 1997) et électropalatographiques (Aasland et *al.* 2006), recueillies dans ces deux recherches, ont révélé que les sujets sont capables de compenser la présence d'un palais artificiel et, qu'après une heure, les mesures relevées avec le palais artificiel sont comparables à celles que les auteurs ont pu constater avant la mise en place du palais artificiel.

Cette même expérimentation a été reprise par Thibeault et *al.* (2011), en ajoutant l'occlusive [k] dans le corpus, consonne pour laquelle le palais artificiel ne devrait pas directement gêner la réalisation. L'étude relève toutefois des stratégies de réorganisation spatio-temporelle pour la réalisation de l'occlusive, et conclut que l'ajout d'un palais artificiel conduit à une re-calibration globale de l'organisation-temporelle des gestes de la parole.

Enfin, signalons que Brunner et *al.* (2006), qui ont repris le même protocole expérimental en insérant un palais artificiel dans la cavité buccale de 4 locuteurs pendant deux semaines, ont montré que les stratégies de compensation utilisées par

chacun d'entre eux se sont développées de façon progressive. Les auteurs concluent à l'existence de relations étroites entre la forme du conduit vocal, le contrôle moteur et la cible perceptive.

Comme les autres méthodes précédemment évoquées, le palais artificiel se révèle être une méthode de perturbation dans la mesure où elle permet, d'une part, d'étudier les stratégies de compensations qu'un locuteur est capable de mettre en place seul, et d'autre part, de quantifier ces stratégies de manière longitudinale. Le palais artificiel semble être bien « accepté » par le système de production de la parole, puisque les recherches réalisées sur ce type de perturbations ont montré que les séquences sonores, produites par des locuteurs munis d'un palais artificiel, sont similaires à celles observées avant l'insertion de ce même palais artificiel, et cela seulement après une heure de familiarisation, ce qui, selon la plupart des auteurs, va dans le sens d'un réajustement s'opérant à partir du feedback auditif. Il est toutefois judicieux de penser que le réajustement s'opère aussi bien par l'intermédiaire d'un feedback auditif que par l'utilisation d'informations d'origines proprioceptives.

Nous venons d'étudier les différentes perturbations statiques qu'il est possible de créer afin d'observer les différents réajustements moteurs que le locuteur est capable de déployer pour maintenir un message linguistique clair. Cette revue ne s'est pas voulue exhaustive. Outre le bite-block, le lip-tub ou le palais artificiel, il existe également d'autres moyens de perturber la parole de façon non naturelle comme, par exemple, l'ajout d'un poids sur la lèvre inférieure (Van Lieshout, 2011).

Notons qu'il est également possible de perturber la parole de façon dynamique. Notre thèse ne portant pas sur ce type de perturbation, nous nous proposons d'en étudier une seule. Il convient toutefois de mentionner l'existence de ce type de perturbation.

II.1.2.4. Perturbation dynamique

L'étude de Kelso et *al.* (1984) propose d'analyser les compensations articulatoires en introduisant une perturbation dynamique. Une masse est placée sur la mandibule et est activée de façon aléatoire par l'expérimentateur sans avertir le locuteur.

Leur étude conclut à des compensations dites fonctionnelles et flexibles, c'est-à-dire que les réponses compensatoires ne sont pas figées mais adaptées à l'acte moteur en

cours. Les conclusions de cette recherche vont dans le sens de l'existence d'une cible articulatoire.

Quant à l'étude de Laboissière et *al.* (2009), elle porte sur la raideur mécanique de la mâchoire, c'est-à-dire sur la capacité de la mandibule à subir ou non des déformations élastiques. Pour ce faire, un dispositif délivrant une perturbation mécanique est placé sur la mâchoire de 31 locuteurs, afin d'évaluer la raideur mandibulaire dans le plan sagittal. Leurs résultats indiquent que la raideur de la mâchoire est plus importante et la variabilité cinématique moindre, lors des phases d'élévation mandibulaire et *vice versa*. Ce pattern n'est d'ailleurs pas propre à la production de la parole, puisque des résultats similaires ont été observés pour des tâches non verbales.

Ces deux études ont permis d'illustrer le fait qu'il existe également des méthodes de perturbation dynamique de la production de la parole. Tout comme pour les perturbations statiques, le locuteur est capable de s'adapter à ce type de perturbation, plus aléatoire, en renforçant notamment son activité musculaire.

Pour résumer, l'introduction d'une perturbation non-pathologique lors de la production de la parole est un sujet ayant donné lieu à un grand nombre de travaux en phonétique. Cela s'explique par le fait que la perturbation permet d'en savoir davantage aussi bien sur les contraintes physiques propres à chaque locuteur, que sur les contraintes liées aux spécificités de la structure phonologique d'une langue. En outre, il ressort de ces recherches que les locuteurs sont en mesure de répondre immédiatement à une perturbation, et que cette réponse a tendance à s'améliorer avec de l'entraînement, pour atteindre des « cibles » articulatoire et/ou acoustique proches de celles pouvant être constatées sans perturbation ou de la parole dite « standard ».

II.2. Perturbation de la perception de la parole

La perturbation de la production de la parole peut entraîner des réajustements immédiats, afin de maintenir un message linguistique clair. Néanmoins, il est opportun d'examiner la dimension perceptive en relation avec ses facteurs perturbateurs. Ainsi,

nous proposons dans cette partie d'étudier les possibles réajustements et stratégies de compensation que le locuteur est capable de mettre en place lorsque c'est ce système de perception de la parole qui est perturbé. Il s'agit dans cette section, comme dans la précédente, de n'étudier que les procédés de perturbation non pathologique.

II.2.1. Le feedback auditif et son rôle

On parle de feedback auditif ou d'exteroception pour désigner le retour qu'un locuteur a de sa propre voix. Ce retour est notamment utilisé par chaque locuteur pour contrôler certains paramètres de la parole. Afin de mettre en avant son importance, de nombreuses recherches ont tenté de perturber le feedback auditif dans le but d'étudier les conséquences éventuelles de cette perturbation sur la production de la parole.

L'une des premières recherches menée sur le sujet était celle de Lombard (1911) qui a mis en avant la relation entre l'intensité sonore d'un environnement, où se trouve un locuteur, et l'intensité sonore de sa propre voix. Plus exactement, et comme nous allons le voir plus bas Lombard (1911) a montré qu'en milieu fortement bruyé, un locuteur va augmenter l'intensité de sa voix. Ce phénomène de régulation d'amplitude de la production vocale, en fonction du bruit, est depuis connu sous le nom de « l'effet Lombard ». Cette découverte, probablement considérée comme naïve aujourd'hui, atteste néanmoins de l'étroite relation qu'entretiennent production et perception de la parole. En effet, il ne s'agit pas uniquement pour un locuteur d'être intelligible, mais encore de pouvoir bénéficier de son propre feedback auditif pour vérifier *en ligne* la cohérence de ses propres productions.

Ces résultats s'inscrivent également dans le paradigme de la perturbation et des réajustements en production et en perception de la parole, la perturbation étant cette fois introduite au niveau de la perception. On a pu observer qu'une perturbation de la perception peut avoir des conséquences directes sur la production du locuteur.

D'autres travaux (Lee, 1950 ; ou Yates, 1963 par ex.) ont montré qu'un feedback retardé pouvait également avoir des répercussions sur la parole. Il semble donc important de traiter également, dans cette partie, les différentes formes de perturbations de la perception de la parole et les conséquences qu'elles peuvent éventuellement avoir sur la production de la parole.

II.2.2. Effet Lombard et milieu bruité

L'effet Lombard permet d'étudier l'importance de la situation de communication qui peut également se révéler être une source de perturbation et de réajustement en production de la parole. Cet effet part du constat que l'intensité phonatoire d'un locuteur augmente lorsque son feedback auditif est diminué, du fait d'un bruit ambiant ou du port de protections auditives par exemple. D'après Lombard (1911), « Le locuteur, qui n'entend pas mieux sa propre voix que les bruits extérieurs, s'efforce de remédier à cette audition défailante en augmentant l'intensité du son glottique » (p.108). Les causes de cet effet sont tantôt interprétées comme un simple réflexe physiologique, tantôt comme une accommodation aux exigences de clarté dans une situation de communication donnée.

Communiquer en milieu bruité aurait des conséquences diverses à la fois sur le plan acoustique et sur le plan articulatoire. Sur le plan acoustique, on retiendra un ralentissement du rythme, une élévation de l'intensité, mais également de la fréquence fondamentale et une modification des formants (Junqua, 1992). Une accentuation des gestes labiaux, indices importants en lecture visuelle Benoit et *al.* (1992), mais aussi une hyperarticulation globale (Bailly, 2010) sont notées sur le plan articulatoire.

Les effets compensatoires de la communication en milieu bruité sont également à rapprocher des effets que l'on peut relever en parole criée, même si l'effet Lombard ne conduit pas systématiquement le locuteur à hausser sa voix jusqu'à crier. En effet, les conséquences de la parole criée ne se résument pas à une amplification des mouvements produits en parole normale, mais plutôt en une réorganisation globale du timing des mouvements des articulateurs (Schulman, 1989).

Les réajustements provoqués par l'effet Lombard montrent qu'il est également possible de perturber la production de la parole de façon relativement naturelle en perturbant le feedback auditif, et donc la perception de la parole.

II.2.3. Feedback différé ou DAF (*Delayed Auditory Feedback*)

Outre l'altération du feedback auditif par le bruit, il est également possible d'agir sur le feedback en le retardant. Cette méthode, notamment utilisée dans le traitement de certaines pathologies comme le bégaiement et l'aphasie, aurait un impact sur la production de la parole chez des sujets sains. Cette technique consiste à faire parler les

participants munis d'un casque dans lequel ils perçoivent le retour de leur propre voix, en le retardant de 5ms à 200ms, par exemple.

Les conséquences du feedback retardé ont été étudiées dans un grand nombre de recherches, et cela depuis de nombreuses années. Ainsi, (Yates, 1963), par exemple, rapportait que les locuteurs, sur lesquels le DAF avait été testé, tendaient notamment à allonger la durée des voyelles, à répéter certaines consonnes et à augmenter l'intensité des répétitions. Il conclut en observant que les erreurs commises par les locuteurs sous ces conditions de parole sont proches des disfluences des personnes qui bégaiement et propose alors sa méthode comme un moyen de comprendre le bégaiement.

Cette recherche n'est pas isolée dans la mesure où, de manière générale, la littérature confirme que retarder le feedback chez un sujet non-pathologique aurait tendance à augmenter significativement les troubles de la fluidité verbale (Stuart, et *al.* 2002). Plus précisément, ces auteurs ont observé que plus le feedback était retardé, plus le nombre de disfluences était important. En outre, ils ont constaté que cet effet était également lié à la vitesse d'élocution, étant donné que les conséquences d'un DAF auraient tendance à être d'autant plus importantes que le locuteur est obligé d'augmenter sa vitesse d'élocution. Les auteurs concluent toutefois leur recherche en signalant bien que, même si les locuteurs sains sont significativement plus disfluents sous DAF, ils le sont toujours moins que les locuteurs bègues, dont le feedback n'a pas été retardé.

II.3. Perturber la perception pour perturber la production

D'autres techniques de perturbations de la perception de la parole utilisant notamment le feedback engendrent une perturbation de la production de la parole.

Shiller et *al.* (2009) ont ainsi mené une expérience sur la perception et la production de fricatives sourdes. Les auteurs ont utilisé un système permettant de modifier les fréquences en temps réel, ce qui rendait possible la modification du feedback, de telle manière que le son /s/ se rapprochait du son /ʃ/ lorsque les locuteurs entendaient le retour de leur propre voix. Les participants avaient alors tendance à accentuer leur production pour qu'elles ressemblent davantage à des /s/.

Shiller et *al.* (2009b) ont également conduit des tests de perception avant et après la tâche de production avec le feedback modifié. Les tests de perception visaient à identifier la frontière entre la catégorie du /s/ et la catégorie du /ʃ/. Leurs résultats perceptifs ont révélé que la frontière catégorielle avait tendance à glisser vers /ʃ/. Les locuteurs témoins, non soumis au feedback modifié, n'ont pas présenté ce type de stratégies. Ces résultats montrent que les stratégies d'adaptation, suite à la perturbation du feedback, ne sont pas limitées au domaine de la production de la parole mais ont également des répercussions sur les représentations auditives. Selon les auteurs, c'est la combinaison des réajustements moteur et perceptif qui permettrait de réduire l'impact de la perturbation.

Cette étude montre qu'il est non seulement possible de perturber la production de la parole en modifiant le feedback auditif mais qu'en plus, à terme, la perturbation du feedback a des répercussions sur l'identification des catégories des phonèmes. Il semble donc que les notions de perception et de production de la parole soient étroitement liées et fonctionnent de façon multimodale, les perturbations introduites peuvent agir sur plusieurs niveaux, provoquant à terme une « boule de neige ».

II.3.1. Effet McGurk

L'effet McGurk, mis en évidence en 1976 par McGurk & MacDonald (1976), est une illusion auditive créée par la combinaison d'un stimulus auditif et un stimulus visuel discordant. L'illusion dépend des stimuli choisis. Ainsi, si l'on présente à un auditeur le stimulus auditif /ba/ et le stimulus visuel /ga/, ce dernier aura tendance à percevoir la syllabe /da/. Cet effet renforce l'existence d'une interférence forte entre les informations visuelles et les informations auditives dans la perception de la parole. L'effet McGurk est considéré comme un indice de fusion des informations acoustiques et visuelles.

Cet effet est souvent qualifié de robuste, puisque l'illusion persiste souvent même après que le locuteur a été informé de l'effet qu'elle produit. Toutefois, une étude récente (Nohoma et *al.* 2010) indique que l'effet McGurk est relié au contexte et qu'il est possible de le supprimer en fonction du contexte créé. Pour ce faire, ils mettent en place un protocole expérimental visant à supprimer ou à modifier l'effet McGurk. Deux séquences de stimuli sont créées : une séquence dite « cohérente », en modalité audiovisuelle : le sujet voit le visage du locuteur qui prononce effectivement des

syllabes synchronisées avec des syllabes audio que le sujet entend ; une séquence dite « incohérente » pour laquelle les stimuli sont constitués du même matériel audio, superposé à la vision du même locuteur mais qui prononce de la parole quelconque. Dans les deux cas, les cibles auditives à identifier sont les mêmes. Leur étude montre que dans le cas du contexte incohérent, l'effet McGurk disparaît complètement et que l'identification de la cible est essentiellement effectuée à partir de l'information auditive.

L'effet McGurk indique que la réception correcte du message linguistique repose sur différents niveaux perceptifs. La perception de la parole est donc à considérer comme un processus multimodal ou comme un réseau d'interférences entre les niveaux visuel et auditif.

Pour résumer, l'étude de la dimension perceptive du système de production-perception de la parole en relation avec ses facteurs perturbateurs nous a permis de mettre en évidence l'importance du feedback. Les réajustements que le locuteur peut mettre en place face aux perturbations de son système perceptif renforcent l'idée que le niveau de la production et de la perception de la parole sont étroitement liés.

CONCLUSION CHAPITRE II

Etudier les différentes perturbations que le système de production-perception de la parole peut subir a permis de mettre en lumière les diverses stratégies de compensation ou réajustements que le locuteur est capable de développer.

Il est possible de perturber le système de production-perception de la parole à différents niveaux (production uniquement, production et perception...) et de différentes manières (écologiques ou non écologiques...). Dans la plupart des cas, le système est capable de s'adapter, de dépasser la perturbation et donc de garantir l'intégrité du message linguistique.

La plasticité du système de production perception de la parole peut ainsi être mise au jour, les stratégies employées par le locuteur sont diverses mais concourent toutes au maintien du message linguistique. Dans certains cas, même lorsque la perturbation est importante, le locuteur est capable d'atteindre des cibles acoustiques comparables à celles de productions non perturbées. Ces résultats vont dans le sens d'une cible qui serait en réalité une zone de viabilité. Il s'agira de vérifier éventuellement cette hypothèse à l'aide de nos données en parole pathologique.

RÉSUMÉ CHAPITRE III

La littérature propose généralement de distinguer les troubles dysarthriques des troubles dysphoniques : les premiers cités sont souvent définis comme des altérations de la planification, ainsi que de l'exécution motrice de la parole et portent sur la parole. En revanche, les seconds sont considérés comme des troubles de la phonation et touchent davantage à la voix. Ce chapitre aura pour objectifs, d'une part, de définir de manière plus précise ces troubles et, d'autre part, de montrer que cette dichotomie entre dysarthrie et dysphonie présente certaines limites, dans la mesure où elle ne permet pas de classer tous les dysfonctionnements liés à la voix ou à la parole.

Pour ce faire, une définition de la dysarthrie sera proposée dans un premier temps, en expliquant comment les différents types de dysarthries sont classifiés dans la littérature, et en soulignant les caractéristiques articulatoires des personnes souffrant de ce trouble de la parole. Afin d'illustrer notre propos, un cas particulier de dysarthrie, celui du sujet parkinsonien, sera davantage développé. Cet exemple permettra notamment d'illustrer les liens étroits qu'entretiennent dysarthries et dysphonie.

Dans un second temps, la notion de dysphonie, qui est au centre de notre sujet de recherches, sera approfondie. Une classification des différents types de dysphonie sera proposée, avant de mettre en avant les causes pouvant expliquer son apparition dans la voix d'un locuteur. Parallèlement à cela, le lecteur pourra constater que les conséquences acoustiques de la dysphonie diffèrent en fonction de son étiologie.

CHAPITRE III

III. Troubles de la parole et de la voix

III.1. Les dysarthries – troubles du mouvement ?

III.1.1. Définition et limites

D'ordinaire, la dysarthrie est définie dans la littérature comme un trouble de la parole d'origine neurologique, soit une maladie dégénérative (de type maladie de Parkinson, sclérose en plaques, ...), soit un Accident Vasculaire Cérébral (AVC), qui se manifeste par une déficience motrice (Fougeron et *al.* 2010) . Mais limiter la dysarthrie à un trouble neurologique est quelque peu réducteur. En effet, un patient ayant subi une chirurgie ablative oro-faciale (comme une exérèse de la langue, par exemple) peut présenter également un trouble de l'articulation alors qu'aucun trouble neurologique n'est identifié. De plus, certaines formes de dysphonie, comme l'hypophonie parkinsonienne, entrerait également dans la catégorie des dysarthries puisque leur origine est neurologique. La dysarthrie serait, en réalité, le signe clinique d'une symptomatologie plus large d'une pathologie dite « du mouvement » (Pinto & Ghio, 2008).

Cette définition et ses limites sont représentatives de la dualité qui existe entre les notions de dysarthrie et de dysphonie. C'est pourquoi certains auteurs (Ackermann et *al.* 1989 ; Klostermann et *al.* 2008) préfèrent parler de dysarthrophonie,

terme qui permet d'englober les troubles articulatoires mais aussi phonatoires liés à la dysarthrie.

Par ailleurs, si la frontière entre la dysarthrie et la dysphonie semble relativement poreuse, il est important de signaler qu'il n'existe pas qu'un type de dysarthrie mais plusieurs. Pinto et Ghio (2008) ont, en effet, mis en avant que, si la dysarthrie est provoquée par une lésion du système nerveux central ou périphérique, les conséquences motrices de ces lésions sur la production de la parole sont multiples et peuvent toucher aussi bien la respiration, la phonation, l'articulation, les résonances dans le conduit vocal et/ou la prosodie. C'est ce que nous proposons d'étudier dans la partie suivante.

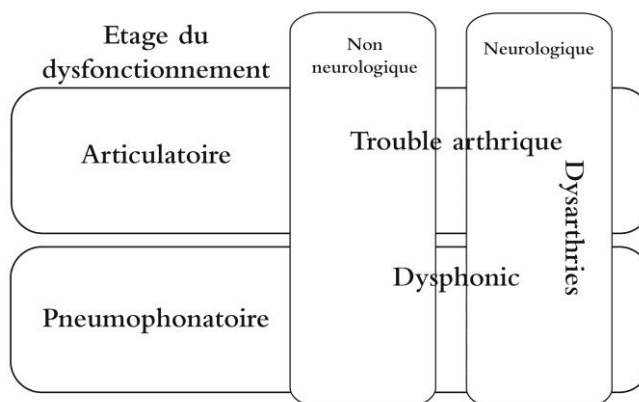


Figure 3 : Terminologie des troubles de la parole, d'après Pinto et Ghio, (2008)

III.1.2. Classification des dysarthries

La classification des dysarthries, même si elle ne fait pas consensus, doit beaucoup aux travaux de Darley *et al.* (1969a, 1969b et 1975) effectués à la Mayo Clinic (Etats-Unis). Elle repose essentiellement sur une évaluation perceptive d'anomalies relevées à partir de l'étude de 212 sujets pathologiques, regroupés en fonction de leurs symptômes : bulbaire, pseudo-bulbaire, cérébelleux, choréo-athétosique, dystonique, parkinsonien ou Sclérose Latérale Amyotrophique – SLA (Auzou, 2001).

La classification de la Mayo Clinic identifie et décrit six types de dysarthries. Sa particularité est qu'elle est la première à proposer d'associer une atteinte anatomique précise à un trouble moteur de la parole. Nous reprendrons ici l'adaptation proposée par Pinto et Ghio, (2008) qui est la suivante :

1) la dysarthrie flasque (faible tonicité, paralysie ou hypotonie) du groupe pathologique bulbaire (atteinte de la voie motrice finale commune qui comprend les nerfs périphériques, la jonction neuromusculaire, comme dans le cas de la myasthénie, et les muscles effecteurs de la parole) ;

2) la dysarthrie spastique (augmentation exagérée et permanente du tonus musculaire) du groupe pathologique pseudo-bulbaire (atteinte bilatérale du motoneurone central : atteintes pseudo-bulbaires, accidents vasculaires cérébraux) ;

3) la dysarthrie ataxique (défaut de coordination des mouvements volontaires) du groupe pathologique cérébelleux (atteinte du cervelet ou des voies cérébelleuses) ;

4) la dysarthrie hypokinétique (faible amplitude et lenteur des mouvements) du groupe pathologique parkinsonien (atteinte des noyaux gris centraux, avec prédominance akinétique, comme dans la maladie de Parkinson) ;

5) la dysarthrie hyperkinétique (grande amplitude de mouvements) des groupes pathologiques dystonique et choréo-athétosique (atteinte des noyaux gris centraux, avec prédominance choréo-athétosique, comme dans les dystonies ou la maladie de Huntington) ;

6) les dysarthries mixtes associant les troubles de la parole résultant de l'atteinte de plusieurs systèmes neurologiques (groupe pathologique : SLA, sclérose en plaques, traumatismes crâniens).

Les travaux de Duffy (2005), viendront compléter cette classification en proposant notamment deux types de dysarthries supplémentaires : les dysarthries par atteinte unilatérale du premier neurone moteur, généralement associées à une aphasie et une apraxie de la parole pour lesquels les accidents vasculaires cérébraux sont la principale cause, et les dysarthries d'étiologie indéterminée. L'ajout de ces deux catégories avait probablement pour but de rendre cette classification exhaustive notamment grâce à la dernière citée. Toutefois, il n'est pas toujours possible, à cause de la grande variabilité qui existe dans la parole dysarthrique, d'inscrire le trouble observé par le praticien dans une des cases prédéfinie par la classification de la Mayo clinic.

Cette classification, qui reste la plus utilisée dans le domaine clinique a, outre le problème de l'exhaustivité, certaines limites qui ont notamment été relevées par Ozsancak (2007), . Une de ses critiques provient du fait que l'imagerie cérébrale n'était pas aussi développée dans les années soixante-dix, période où la classification a été

mise au point. De ce fait, une interrogation sur la réelle homogénéité des groupes étiologiques étudiés subsiste, d'autant plus qu'aucune étude récente n'a été entreprise afin de valider cette classification sur un groupe plus important de sujets pathologiques. Par conséquent, aucun élément ne permet de conclure qu'une évaluation perceptive de la parole dysarthrique peut permettre d'identifier clairement une zone lésionnelle sous-jacente. Enfin, l'étude repose sur une analyse perceptive de critères déviants, mais la variabilité au sein de chaque type de dysarthrie rend difficile une analyse fondée sur des critères différentiels. Certaines productions déviantes sont particulièrement courantes au sein de toutes les catégories, les différences s'établissant parfois à l'aide de productions déviantes peu fréquentes mais pourtant différentielles.

Comme nous l'avons signalé précédemment, la classification de la Mayo Clinic ne fait pas consensus et n'est donc pas la seule disponible. Ainsi, Lecours & Lhermitte (1979), ou ont proposé des classifications différentes. La classification des dysarthries de Lecours & Lhermitte (1979), propose trois grands groupes de dysarthries : les dysarthries paralytiques, les dysarthries extrapyramidales et les dysarthries cérébelleuses. Celle de Le Huche & Allali (2010) est dérivée de la précédente, mais propose une classification plus fine à l'aide de six catégories, la dysarthrie est alors paralytique, akinétique, dyskinétique, ataxique, apraxique ou dystonique. Selon les auteurs, cette classification possède l'avantage de mieux prendre en compte l'étiologie tout en considérant la symptomatologie. Ces deux classifications sont toutefois moins connues que celle mise en place par les travaux de Darley *et al.* (1969a, 1969b et 1975) et complétée par Duffy (2005).

Cette section nous a permis de proposer une définition de la dysarthrie et d'en identifier ses limites. En résumé, la dysarthrie est une conséquence de l'atteinte du système nerveux central ou périphérique qui se traduit par un trouble de l'exécution motrice de la parole (Darley *et al.* 1975).

Prendre en charge la dysarthrie consiste à rechercher les signes cliniques associés, pointer les caractéristiques déviantes et observer leur évolution. C'est pourquoi la prise en charge et l'évaluation de la dysarthrie requiert une bonne connaissance de la production de la parole normale.

III.1.3. Parole Dysarthrique – Le cas du sujet parkinsonien

Comme nous l'avons vu précédemment, la dysarthrie est un trouble moteur de la production de la parole résultant d'une atteinte du système nerveux central ou périphérique. En fonction de la localisation et de l'étendue de la lésion au niveau cérébral mais aussi en fonction des particularités propres à chaque locuteur et à chaque maladie, les caractéristiques de la parole dysarthrique sont très variées (Fougeron et *al.* 2010).

De façon générale, la dysarthrie peut modifier, à des degrés variables, tous les niveaux de la parole : articulation, timing, phonation, prosodie, fluence, *etc.* Comme cela a été dit précédemment, la variabilité de la parole dysarthrique est également le résultat de la nature du trouble moteur. Ainsi, on sait par exemple que l'amplitude des mouvements et l'atteinte des cibles articulatoires seront plus affectées par la perte de la force musculaire (dysarthrie flasque, par ex.) alors que la modification de la durée ou des transitions seront davantage liées à un mauvais contrôle temporel (timing) des mouvements (dysarthrie ataxique, par ex.). Toutefois, la parole dysarthrique est globalement caractérisée par une perte de l'intelligibilité (Kain et *al.* 2007), c'est-à-dire par la capacité d'un auditeur à comprendre le discours d'un locuteur souffrant de dysarthrie.

Cette perte de l'intelligibilité chez les sujets dysarthriques est directement liée à l'altération d'un ou de plusieurs des niveaux de la parole précédemment cités. Les dysarthries et leurs conséquences étant nombreuses, nous nous attacherons ici à ne décrire que les causes de la perte d'intelligibilité chez le sujet parkinsonien, qui semble le plus étudié à ce jour.

Nous choisissons en effet de développer dans cette partie le cas du sujet parkinsonien car sa voix présente de nombreuses caractéristiques communes, à tous les niveaux de la production de la parole, avec celle du patient souffrant d'une paralysie récurrentielle (respiration moins efficace, mouvements de plis vocaux altérés, intensité diminuée...). En ce qui concerne le sujet atteint de la maladie de Parkinson et souffrant d'une dysarthrie de type hypokinétique, la lésion du système nerveux altère les trois niveaux de la production de la parole (Viallet & Gentil, 2001), à savoir :

- Le niveau subglottique ou la respiration : le sujet parkinsonien présente des difficultés lors des phases d'expiration ;

- Le niveau glottique : le sujet parkinsonien présente des difficultés pour contrôler la vibration de ses plis vocaux ;
- Le niveau supraglottique : le sujet parkinsonien peine à contrôler le mouvement de ses articulateurs. En outre, Ramig, et *al.* (2001) notent une diminution d'intensité dans la parole parkinsonienne.

Lehiste (1965) et Darley *et al.* (1975) notent par ailleurs un manque de coordination, chez le sujet parkinsonien, entre le mouvement des articulateurs et les vibrations des plis vocaux. Pour les deux études, ce résultat est la conséquence de l'akinésie propre à la maladie de Parkinson. Conséquemment, lors des tests de perception, la distinction entre les consonnes et les voyelles est rendue difficile.

En outre, les problèmes rencontrés par les sujets parkinsoniens lors de la production de la parole ont des répercussions directes sur les sons, et plus précisément sur les consonnes. Gentil, et *al.* (1995) ont ainsi montré que le locuteur souffrant de la maladie de Parkinson éprouvait des difficultés à produire une occlusion dans le conduit vocal ; la conséquence de ce dysfonctionnement est que l'air s'échappe de façon constante, ce qui crée un bruit de friction continu à la place de la tenue de l'occlusive. De même, les fricatives semblent également altérées chez ce type de locuteurs, du fait d'une déperdition constante d'air (Logemann, et *al.* 1978). Enfin, Ackermann & Ziegler (1991), ainsi que Lieberman et *al.* (1992) ont relevé une réduction des transitions formantiques, toujours chez les locuteurs parkinsoniens.

Par rapport au niveau suprasegmental, la parole parkinsonienne se caractérise par de la dysprosodie et de la monotonie (Ramig et *al.* 2001). Dès les premiers stades de la maladie, le sujet parkinsonien éprouve des difficultés pour effectuer des variations prosodiques, le contrôle de la durée des pauses, de la vitesse d'articulation, de l'intensité et de la hauteur étant rendu difficile (Ackermann, et *al.* 1997). En outre, le manque d'intensité déjà mentionné plus haut peut, dans les cas les plus graves, amener le locuteur dysarthrique à avoir une parole comparable au chuchotement.

Les résultats qui concernent l'étude de la fréquence fondamentale dans la dysarthrie parkinsonienne ne font pas consensus. En effet, certaines études notent une augmentation de la fréquence fondamentale (Ludlow & Bassich, 1984) alors que d'autres ont relevé une diminution de F0 (Kiebertz et *al.* 1994). Quant à Weismer (1984), il conclut son étude en soulignant que le sujet parkinsonien éprouve

des difficultés pour moduler sa fréquence fondamentale, ce qui constitue un indice supplémentaire pour expliquer la monotonie de la parole pour ce type de locuteurs.

Comme nous le verrons ultérieurement, la dysphonie du sujet parkinsonien présente de nombreuses similitudes avec celle du patient souffrant de paralysie récurrentielle, c'est pourquoi il nous a semblé intéressant de développer cet exemple.

Nous avons tenté dans cette section de définir les dysarthries et de mettre au jour les liens étroits qu'entretiennent les dysarthries et les dysphonies, la frontière entre les deux phénomènes n'étant pas toujours évidente à établir, comme en témoigne le terme de dysarthrophonie proposé par Ackermann et Ziegler (1989). Nous retiendrons toutefois que les dysarthries sont un trouble d'origine neurologique qui a notamment des conséquences sur la production de la parole. Cette parole déviante peut, dans certains cas, partager des caractéristiques avec la dysphonie (*cf.* le cas du sujet parkinsonien), qui peut également se révéler être une conséquence de l'absence de contrôle du mouvement des articulateurs.

III.2. Dysphonies – Troubles laryngés

Dès l'Antiquité, les notions de dysphonie et plus particulièrement d'aphonie sont connues, comme en atteste les différentes références présentes dans la littérature. Chariton, auteur grec du II^{ème} siècle ou Homère décrivent des personnages rendus aphones : « Elle se demanda ce qui s'était passé : il avait perdu l'usage de la parole... Il était comme paralysé et tout tremblant. » (Chariton, II^{ème} siècle). « A cette nouvelle, Antiloque, saisi d'horreur, demeure muet ; ses yeux se remplissent de larmes ; sa voix sonore est étouffée. » (Homère, Iliade chant XVII). L'aphonie est donc associée à un choc émotionnel fort qui vient traumatiser le personnage ; ces procédés littéraires ne sont pas sans rappeler toutefois la réalité d'une dysphonie dysfonctionnelle qui peut être le résultat d'une inhibition vocale.

Cette section aura pour objectif de définir la notion de dysphonie. Elle proposera également une classification des différents types de dysphonies avant d'en présenter les caractéristiques anatomiques, acoustiques et aérodynamiques.

Classer les différents types de dysphonie peut se révéler complexe tant parfois il n'est pas aisé de différencier la cause de la conséquence de ce trouble. En effet, la dysphonie peut être responsable de la modification des structures laryngées comme elle peut en être la suite.

III.2.1. Définition et classification

La dysphonie est un trouble de la voix pouvant porter sur différents paramètres comme le timbre, l'intensité ou la fréquence fondamentale (Dubreuil & Céruse, 2004), qui a été défini par Le Huche & Allali (2010) comme « *un trouble momentané ou durable de la fonction vocale ressenti comme tel par le sujet lui-même ou son entourage* » (p.47). Plus précisément, il s'agit d'une perturbation de l'émission du son laryngé perçu comme anormal par un locuteur et/ou un auditeur (Ghio & Pinto, 2008). Il est possible de classer les dysphonies en trois catégories qui seront davantage développées ultérieurement : 1) les dysphonies d'origine dysfonctionnelle ; 2) celles d'origine organique et enfin ; 3) celles d'origine neurologique. Cette section permettra également de montrer que cette classification n'est pas figée, étant donné qu'il n'est pas toujours aisé de déterminer si les atteintes organiques sont les causes ou les conséquences d'une dysphonie, c'est pourquoi le clivage organique/fonctionnel est largement discuté dans la communauté scientifique. En outre, il est utile d'ajouter que les classifications sont loin de faire consensus, et que les paralysies récurrentielles ne sont pas classées dans les mêmes catégories (organiques ou neurologiques) suivant les auteurs.

III.2.2. Les dysphonies dysfonctionnelles

Cette section s'attachera à la définition de la dysphonie dysfonctionnelle et permettra d'en étudier les différentes caractéristiques. Il en existe de nombreuses formes, nous nous limiterons aux formes qui affectent les adultes puisque notre population d'étude en est exclusivement constituée. Les formes particulières, comme la raucité infantile, les troubles de la mue ou les formes spécifiques comme la dysphonie du chanteur ne seront pas abordées.

Nous proposerons une définition de la dysphonie dysfonctionnelle avant d'étudier plus en détail les différents comportements qui peuvent conduire à la perturbation du geste vocal.

III.2.2.1. Définition de la dysphonie dysfonctionnelle

Les dysphonies dysfonctionnelles (ou fonctionnelles) sont le résultat d'une utilisation défaillante de l'appareil phonatoire. Elles sont caractérisées, à l'inverse des dysphonies organiques, par la non-atteinte permanente du système anatomique de production de la parole (Remacle, et *al.* 1999), c'est-à-dire que les lésions sont généralement non définitives. Dès lors que le patient a conscience de la mauvaise utilisation de son appareil phonatoire, la rééducation vocale, précédée d'un acte chirurgical dans les cas les plus graves, peut diminuer voire supprimer les lésions anatomiques. Le nodule et le polype, qui sont parmi les lésions les plus courantes, constituent évidemment une modification de la structure anatomique du larynx. Ces atteintes ont toutefois leur place parmi les dysphonies dysfonctionnelles, dans la mesure où leur apparition est la conséquence du malmenage et du forçage vocal (Remacle, et *al.*, 2003), comportement caractéristique des dysphonies dysfonctionnelles. Crevier-Buchman, et *al.* (2005) les définit comme « le résultat d'une inadéquation du comportement vocal responsable d'une diminution de l'efficacité du geste de la parole » (p.353). Les dysphonies dysfonctionnelles sont caractérisées par des phénomènes d'hypotonie et d'hypertonie pouvant conduire à l'apparition de lésions organiques. En effet, du fait d'un malmenage vocal, la muqueuse des plis vocaux risque de s'épaissir, ce qui diminue d'autant plus le rendement vocal et oblige le locuteur à toujours forcer davantage sur les mécanismes de phonation, avec comme conséquence, l'apparition de nodules, puis de polypes. Notons que ces dysphonies s'inscrivent dans le cercle vicieux du forçage vocal très présent chez les enseignants notamment. La prise en charge des dysphonies dysfonctionnelles passe d'abord par la prise de conscience de la mauvaise utilisation de l'appareil pneumo-respiratoire.

III.2.2.2. Etiologie et développement

De nombreux auteurs ont ainsi cherché à quantifier les changements vocaux qui peuvent intervenir au cours de la journée, lorsque le locuteur est amené à utiliser sa voix de façon continue, dans son travail notamment. Schoentgen & Grenez (2011) ont

développé un appareillage discret permettant d'enregistrer un locuteur en continu. L'étude qui en a résulté et qui portait sur un enseignant a montré une légère modification de la qualité vocale au cours du temps. Ce dispositif a été repris par Verduyck, et *al.* (2011) afin de quantifier l'inconfort vocal chez les enseignants au cours d'une journée de travail. Leurs observations préliminaires permettent de conclure à des effets de fatigue vocale en fin de journée. Il est à noter que ces conclusions sont à rapprocher de l'étude de Laukkanen, et *al.* (2008) qui a été conduite sur des enseignantes, et qui avait pour objectif d'évaluer leur voix avant et après une journée de travail en mesurant la fréquence fondamentale, le jitter et le shimmer notamment. L'augmentation de la fréquence fondamentale et la diminution des mesures de jitter et de shimmer, de façon non significative toutefois, ont permis de conclure à une adaptation du tonus musculaire à l'effort vocal. Ces résultats sont également corrélés à une augmentation du nombre de raclements de gorge et à l'impression de diminution d'efficacité vocale que le sujet ressent à la fin de la journée. Si les résultats de cette étude ne sont pas particulièrement concluants, ils permettent toutefois de souligner non seulement la difficulté d'évaluer l'effort et la fatigue vocale mais aussi de déceler de façon précoce les mauvais comportements phonatoires. Une étude de Remacle et *al.* (2012) indique que c'est essentiellement la durée, mais également, et de façon moindre, l'intensité, de la phonation qui ont un impact sur la fatigue vocale. Les auteurs notent également une augmentation de la fréquence fondamentale après un effort phonatoire.

Si les études précédemment évoquées montrent que les signes de la fatigue vocale peuvent apparaître (et à nouveau disparaître) sur un temps très court, il est alors facile de comprendre que les caractéristiques de la dysphonie dysfonctionnelle peuvent s'installer de façon insidieuse dans le comportement vocal de chacun. Certains auteurs, comme Heaver, (1959), pensent que l'apparition d'une dysphonie dysfonctionnelle peut être liée à des facteurs psychologiques divers. Ainsi, les nodules vocaux pourraient apparaître plus facilement chez des patients dont l'état psychologique est susceptible de créer un état de tension. Aronson & Bless (2009) précisent ces facteurs psychologiques comme des problèmes interpersonnels aigus et chroniques, des interactions sociales agressives, un état général tendu... Ces sujets, selon Hugh-Munier, et *al.* (1997) auraient des valeurs de fondamentaux plus variables et donc une phonation moins régulière

probablement responsable de l'apparition du comportement dysphonique. Ces études supposent l'existence d'un terrain psychologique propice au développement du malmenage vocal. Leurs résultats sont toutefois à nuancer car généralement très peu significatifs puisque souvent tirés d'échantillons de population relativement faibles. En outre, il est souvent difficile de circonscrire de manière objective la notion de « facteur psychologique ».

Enfin, et dans la mesure où cela concerne indirectement notre sujet de recherche, nous signalerons également que les désordres hormonaux, telle l'hyperthyroïdie, peuvent être responsables de difficultés vocales se manifestant par une perte de l'étendue vocale, une raucité, une fatigue vocale, ou encore une voix soufflée (Estienne, *et al.* 2006). Abitbol, *et al.* (1999) signalent qu'une réduction de l'efficacité vocale accompagnée d'une raucité, d'une plus grande fatigue vocale et d'un timbre plus sombre du fait d'une perte partielle des fréquences aiguës peut être constatée chez les femmes (33% de leur population test) au 14^{ème} et au 25^{ème} jour de leur cycle menstruel, ou pendant la grossesse.

Cette section nous a permis de définir la notion de dysphonie dysfonctionnelle et d'en aborder les différentes caractéristiques. Il semble important de retenir qu'une dysphonie dysfonctionnelle est une modification temporaire de la voix et généralement bénigne. Si le patient prend conscience rapidement du mauvais usage de son appareil phonatoire, les effets sont généralement réversibles. Il n'est d'ailleurs pas rare qu'un sujet rééduqué ait, en fin de rééducation, une utilisation optimale de ses structures laryngées.

La suite de ce chapitre sera consacrée aux pathologies vocales d'origine organique. Celles-ci sont généralement plus graves, puisqu'elles supposent une modification anatomique du larynx qui vient empêcher une production vocale efficace. Leur traitement passera souvent par une intervention chirurgicale.

III.2.3. Les pathologies d'origine organique

Cette catégorie reprend la terminologie de Le Huche *et al.* (2001). Les dysphonies d'origine organique concernent les altérations vocales (œdème de Reinke, kystes, sillons, tumeurs...) directement liées à une modification anatomique du larynx ou des plis vocaux. Elles regroupent les lésions dites « acquises » et les lésions

congénitales. Il n'est toutefois pas rare qu'une dysphonie dite organique s'accompagne des aspects d'une dysphonie dite dysfonctionnelle. En effet, le sujet a souvent tendance à compenser le déficit organique par un comportement de forçage ou, à l'inverse, de retenue (Le Huche et *al.* 2010), ce qui vient augmenter la sévérité de la dysphonie. Notons que dans de nombreux cas, l'intoxication tabagique peut venir compliquer et aggraver l'atteinte organique et donc la dysphonie.

Les pathologies d'origine organique sont nombreuses et ont des étiologies variées. La suite de cette section permettra de passer en revue les différentes atteintes organiques susceptibles de provoquer la dysphonie. Dans un premier temps, nous traiterons les différentes formes de laryngites, généralement déclenchées par un agent bactériologique infectieux. Les traumatismes (chirurgicaux et accidentels) laryngés seront traités dans un deuxième temps. Dans un troisième temps, nous aborderons les anomalies congénitales laryngées et extra laryngées responsables d'une altération vocale. Enfin, dans un quatrième temps, nous verrons qu'un dérèglement hormonal peut également avoir des conséquences sur la qualité vocale. Le degré de dysphonie est évidemment fonction de l'organe l'atteint et de la gravité de l'atteinte. Notre sujet de recherche ayant pour thème central la dysphonie, nous nous attacherons à ne décrire ici les pathologies que succinctement, pour se concentrer sur leurs conséquences sur la qualité vocale, afin de dégager éventuellement des caractéristiques communes avec la voix de nos locuteurs.

III.2.3.1. Les laryngites

Les laryngites sont des affections courantes et généralement bénignes du larynx. Elles sont dues à une inflammation des plis vocaux (Le Huche *et al.* 2010). En cas de laryngite, la dysphonie est souvent aggravée encore par une utilisation trop importante de la voix et peut aller, dans certains cas, jusqu'à l'aphonie. La forme la plus bénigne est la laryngite aiguë. On constate alors un timbre irrégulier ou rauque, voire étouffé en cas d'aphonie. Signalons également que ces symptômes sont limités dans le temps, puisqu'ils ne perdurent que rarement au-delà d'une semaine, et que la phonation ainsi que la déglutition sont rendues difficiles voire douloureuses.

Suite à la consommation excessive d'alcool et/ou de tabac ou suite à un malmenage vocal, les symptômes propres à la laryngite peuvent persister ou survenir fréquemment. On parlera alors de laryngite chronique. D'autres facteurs, comme le

reflux gastro-œsophagien, peuvent également être à l'origine de laryngites chroniques (Verhulst et *al.* (1998), par ex.).

A ce sujet, Pribuisiene et *al.* (2006) ont proposé d'étudier la qualité vocale de 108 patients souffrant de reflux gastro-œsophagien, en se servant de mesures subjectives comme l'auto-évaluation vocale du patient, le *Voice Handicap Index* (VHI) et l'échelle GBR (Hirano, 1981) et de mesures objectives (jitter, shimmer, Normalized Noise Energy et temps maximum de phonation). Leurs résultats indiquent une détérioration significative ($p < 0.001$) des paramètres subjectifs et objectifs entre le groupe de locuteurs souffrant de reflux gastro-œsophagien et le groupe de sujets contrôle. Les mesures de jitter, de shimmer et de Normalized Noise Energy sont significativement plus élevées, alors que le temps maximum de phonation est réduit pour les locuteurs pathologiques. Ces mesures peuvent être corrélées aux résultats subjectifs obtenus à l'aide des échelles GBR et VHI, ce qui signifie que les modifications des paramètres acoustiques sont directement liées non seulement à l'impression auditive de la voix du patient (mesurée à l'aide de l'échelle GRB) mais également à la détérioration de la qualité de vie du patient (échelle VHI). Les auteurs concluent que le traitement médical du syndrome de reflux gastro-œsophagien doit s'accompagner d'un suivi vocal, afin de prévenir la mauvaise utilisation de l'appareil phonatoire et donc l'installation des mécanismes de forçage.

D'autres formes de laryngites existent également. Ainsi, Babin et *al.* (2005), par exemple, ont regroupé sous le terme de laryngites spécifiques les atteintes organiques comme l'œdème chronique des plis vocaux (aussi appelé œdème de Reinke), la papillomatose laryngée et la tuberculose laryngée. Comme pour la laryngite chronique, la dysphonie constitue le symptôme d'appel des laryngites spécifiques : l'intensité vocale est diminuée, la fatigabilité vocale, notamment au réveil, est anormale et tend à s'aggraver au cours de la journée, la voix étant alors généralement décrite comme éraillée ou rauque (*roughness*). Le « simple enrrouement » du début de la maladie peut toutefois conduire à des périodes d'aphonies (Verhulst et *al.* 1998). En ce qui concerne l'œdème de Reinke, il se caractérise par un abaissement de la fréquence fondamentale, abaissement provoquant notamment la perte des caractéristiques sexuelles dans la voix des locuteurs souffrant de cette pathologie ; le patient a alors le sentiment de muer et

l'auditeur a du mal à différencier le sexe du locuteur à partir du seul critère vocal. La modification vocale entraînée par l'œdème de Reinke est donc une gêne manifeste pour les femmes notamment. Si la chirurgie n'est pas toujours préconisée en cas de dysphonie, elle est toutefois obligatoire si celle-ci s'accompagne d'une gêne respiratoire qui peut conduire à l'asphyxie (Remacle, 2003).

La papillomatose laryngée consiste en une prolifération de tumeurs bénignes, d'un épithélium malpighien, soit des cellules qui constituent le larynx (Lindeberg, 1991). Dans les cas les plus graves, elle peut conduire à une dyspnée laryngée, c'est-à-dire une difficulté respiratoire sévère, et donc nécessiter une trachéotomie (Nicollas et *al.* 2008). Niedzielska (2001) a étudié la qualité vocale de 120 enfants souffrant de papillomatose laryngée comparée à celle d'un groupe de sujets contrôles. La qualité vocale des sujets pathologiques a préalablement été évaluée à l'aide de l'échelle GBRAS d'Hirano (1981), en ne tenant compte que du paramètre R (*Rough* – raucité). Le groupe pathologique a été noté 4 sur une échelle où 5 est la plus mauvaise note. Cette étude visait à évaluer l'impact de la maladie sur les paramètres suivants : la fréquence fondamentale, le jitter, le shimmer, le coefficient de variation de la fréquence fondamentale aussi appelé « F0 tremor » et le Harmonics-to-Noise Ratio (HNR). L'ensemble des paramètres est significativement modifié pour le groupe de locuteurs pathologiques. Les mesures de jitter et de shimmer sont significativement plus élevées pour le groupe de locuteurs pathologiques (3.28% ; 13.86% respectivement) que pour le groupe de locuteurs contrôles (0.43% ; 4.07% respectivement). Comme attendu, les valeurs de HNR sont significativement plus basses (9.46 dB) pour les sujets souffrant de papillomatose laryngées que pour les locuteurs sains (23.70 dB), compte tenu du caractère rauque de leur voix. Enfin, signalons que les travaux de Jégoux (2002) indiquent que la sévérité de la dysphonie, due à la papillomatose laryngée, n'est pas proportionnelle à la taille de la lésion.

Il existe d'autres formes de laryngites spécifiques, comme la syphilis laryngée par exemple, qui peuvent également provoquer la dysphonie. Comme dans le cas de la tuberculose laryngée, la dysphonie fait alors partie des séquelles de l'infection et persiste souvent après la fin du traitement médicamenteux. La prise en charge orthophonique est alors largement recommandée, afin d'éviter les phénomènes de forçage vocal et donc l'aggravation de la dysphonie.

Notons que les dysphonies provoquées par des laryngites sont souvent réversibles. La mise en place d'un traitement médicamenteux adapté, éventuellement d'une intervention microchirurgicale et d'un suivi orthophonique précoce dans les cas les plus sévères, permettent une bonne récupération vocale. Babin *et al.* (2005) soulignent l'importance d'une prise en charge orthophonique à la fois dans les contextes de forçage vocal mais également pour proscrire la voix chuchotée. En effet, le patient utilisera parfois la voix chuchotée dans un but d'économie vocale ; or, comme nous l'avons vu précédemment (II.1.1.2.), la voix chuchotée induit un fonctionnement anormal de l'ensemble des structures laryngées.

III.2.4. Dysphonie et traumatismes laryngés

Précédemment, il a été avancé qu'une infection bactériologique, plus ou moins grave, pouvait être à l'origine de troubles vocaux parfois sévères. Dans cette section, nous souhaitons étudier les traumatismes auxquels le larynx peut se trouver exposé. Les structures laryngées peuvent être abimées par un accident (choc, brûlure...) ou par un acte chirurgical (intubation notamment) avec des conséquences parfois importantes sur la respiration et la phonation. Notons que, dans les cas traumatiques les plus graves, l'altération du larynx peut conduire à l'asphyxie, et donc à la mort. Dans la plupart des cas, des séquelles respiratoires et phonatoires plus ou moins définitives peuvent résulter des traumatismes laryngés.

Si les chocs ou les brûlures peuvent endommager les structures laryngées, et donc modifier parfois sévèrement leur fonctionnement, cette partie se focalisera plus particulièrement sur les caractéristiques de la voix après l'ablation totale ou partielle du larynx. Cette opération est le geste chirurgical dont les conséquences sur la qualité de vie et la qualité vocale du patient sont les plus lourdes puisque les actes de respiration, de phonation et de déglutition se voient modifiés.

III.2.4.1. Laryngectomies

Comme pour les pathologies précédemment évoquées, la dysphonie est le symptôme majeur et précoce du cancer du larynx. L'opération chirurgicale proposée dépendra naturellement du siège et de l'étendue de la lésion (Brasnu *et al.* 1988 et Laccourreye, 1993 par exemple).

L'ablation totale ou partielle du larynx est généralement recommandée dans les cas de cancers du larynx ou du pharynx. Ces cancers touchent majoritairement les hommes et sont favorisés par la consommation d'alcool et de tabac (OMS, 2004).

Lorsque cela est possible, la chirurgie partielle du larynx sera préférée à une laryngectomie totale, afin de conserver les fonctions laryngées de respiration, de phonation et de déglutition par les voies naturelles (voir, par ex. Brasnu, 1988 ou Laccourreye, *et al.* 1991).

III.2.4.2. Laryngectomies partielles

Les laryngectomies partielles supraglottiques ont généralement peu de conséquences sur la qualité vocale des patients, puisque ce type d'opération préserve les plis vocaux. Toutefois, des difficultés vocales ne sont pas à exclure comme lors de toute opération du larynx. Elles sont, dans ce cas, souvent légères et passagères (Le Huche *et al.* 2010).

Sur le plan glottique, les laryngectomies partielles ont, en revanche, toujours des conséquences sur la phonation, puisqu'elles consistent en l'ablation d'une partie ou de la totalité des plis vocaux en fonction du type de chirurgie requis. La rééducation vocale permettra au patient de mettre en place des stratégies de compensation phonatoire en utilisant les autres structures du larynx comme les plis vestibulaires, les aryténoïdes, la margelle postérieure, les replis muqueux ou l'épiglotte. L'altération vocale a alors tendance à s'estomper, le patient ne retrouvant toutefois que rarement son efficacité vocale pré-opératoire. Après une laryngectomie partielle, la voix est caractérisée, de façon générale, par une faible intensité, une déperdition du souffle et un timbre sourd. C'est notamment ce que révèle l'étude de Giovanni *et al.* (2002) qui ont travaillé sur les propriétés acoustiques et aérodynamiques de la voix de 21 patients ayant subi une laryngectomie frontale antérieure avec épiglotoplastie (méthode de Tucker), comparées à celles d'un groupe de 10 locuteurs contrôles. Comme attendues, les mesures de jitter et de shimmer se sont révélées plus élevées pour le groupe de locuteurs pathologiques. Le flux d'air oral, ainsi que la pression subglottique estimée (ESPG) étaient significativement plus importants pour les locuteurs pathologiques (400 cm³/s et 23 hPa respectivement) que pour le groupe de locuteurs contrôles (169 cm³/s et 6.2 hPa respectivement). Les résultats de cette étude laissent penser que les mesures

aérodynamiques sont un indice d'évaluation objective de la qualité vocale en phase post-opératoire.

Enfin, signalons encore que les laryngectomies supracricoiidiennes consistent en la résection d'une tranche horizontale complète du conduit laryngé qui comporte l'ablation du cartilage thyroïde, des deux plis vocaux, des deux plis vestibulaires et le plus souvent de l'épiglotte et d'un aryténoïde. Cette opération conserve toutefois l'os hyoïde, le cartilage cricoïde, un ou deux cartilages aryténoïdes mobiles et l'épiglotte suprahyoïdienne. La reconstruction consiste en une suture du cartilage cricoïde à l'os hyoïde avec la base de la langue et l'épiglotte suprahyoïdienne restante afin d'assurer la protection des voies aériennes et une source de vibration ; la neoglote ainsi recréée a alors une forme de « T » contrairement à la glotte anatomiquement normale en « V » (Crevier-Buchman, 1999).

Compte tenu de l'étendue de la résection que les laryngectomies supracricoiidiennes supposent, les conséquences sur le vibreur sont importantes. L'étude de Crevier-Buchman (1999), menée sur la voix de 20 patients en phases pré-opératoire et post-opératoires comparée à celle d'une population témoin, indique que les valeurs moyennes de F0 de la population de patients en phases post-opératoires 6 mois (111 Hz), 12 mois (109 Hz), 18 mois (111 Hz) et 24 mois (112 Hz) ne sont pas statistiquement différentes de celles de la population témoin (151 Hz). Les différences sont à remarquer en phase pré-opératoire, où la fréquence fondamentale moyenne des patients (151 Hz) est mesurée comme plus aigüe que celle des locuteurs témoins (122 Hz). Notons que cette observation confirme que la dysphonie est un symptôme du cancer du larynx comme nous l'avions évoqué *supra*.

L'étude indique toutefois que les différences entre les deux populations existent et se situent notamment dans l'impression auditive de la voix, notée comme grave et rauque, après l'opération, ce qui peut notamment s'expliquer par l'irrégularité du voisement mesuré à l'aide de la déviation standard de F0 (dsF0), de jitter, de shimmer et du NHR (Noise Harmonics Ratio). Les mesures sont alors significativement différentes entre les deux groupes.

La sdF0 moyenne pour le groupe de locuteurs témoins est de 1,65 tandis qu'elle est de 12,32 en phase pré-opératoire, de 22,21 en phase post-opératoire + 6 mois, de 19,89 en phase post-opératoire + 12 mois, 17,71 en phase post-opératoire + 18 mois et de

21,08 en phase post-opératoire +24mois pour le groupe de locuteurs pathologiques. Ces résultats peuvent être corrélés à ceux obtenus pour le jitter mesuré à 1,65% pour le groupe de locuteurs témoins, à 4,09% en phase pré-opératoire, à 10,67% en phase post-opératoire + 6mois, à 10,76% en phase post-opératoire +12mois, à 11,10% en phase post-opératoire +18mois et à 8,04% en phase post-opératoire +24mois pour le groupe de locuteurs pathologiques. L'étude révèle également une différence significative entre les témoins et les patients en post-opératoire pour le paramètre de shimmer et du NHR.

Cette étude présente un double intérêt, puisqu'elle permet de mesurer les conséquences de la tumeur avant l'opération sur la voix des patients. Les paramètres étudiés sont, en effet, tous significativement modifiés par rapport aux valeurs de référence mesurées dans le groupe témoin. Le deuxième intérêt de l'étude se situe dans son aspect longitudinal ; le temps (sur une période de deux ans) ne semble pas avoir un effet positif sur la récupération vocale. Certains paramètres acoustiques, comme le jitter par exemple, ont tendance à se normaliser à 24 mois sans toutefois atteindre des valeurs standards.

Makeieff et *al.* (2004) cherchent à évaluer les conséquences d'une laryngectomie supracricoidienne sur la qualité de vie de 26 patients actifs professionnellement. Ils proposent d'utiliser le *Voice Handicap Index* et le *Voice Related Quality of Life* afin de quantifier les effets de cet acte chirurgical dans la vie quotidienne des patients. Leurs résultats témoignent de la non-suffisance de la voix de substitution que les patients sont capables de produire après l'opération, puisque 57,6% d'entre eux sont obligés d'aménager ou de cesser leurs activités professionnelles.

III.2.4.3. Laryngectomies totales

Les possibilités de phonation de cette catégorie de patients laryngectomisés étant trop éloignées de notre propos d'étude, nous ne développerons pas ce sujet.

Les laryngectomies totales sont des interventions extrêmement lourdes autant sur le plan anatomique que psychologique. L'ablation complète du larynx suppose la dérivation du courant d'air pulmonaire, le laryngectomisé respirant alors par un trachéostome (un orifice) situé sur la paroi cervicale antérieure. Cette intervention comprend également l'ablation des plis vocaux dans leur totalité. Le patient perd donc toute possibilité de voix laryngée.

Il nous semble également important de noter que la notion de dysphonie peut ici rencontrer une limite importante. La pertinence de la notion de dysphonie peut être questionnée dès lors que l'on s'éloigne trop des mécanismes canoniques de production de la parole. Certains auteurs (Robbins, et *al.* 1984 ou Debruyne, et *al.* 1994) ont toutefois montré que la voix trachéo-oesophagienne comporte quelques caractéristiques acoustiques communes avec la voix laryngée. Il est notamment possible de détecter une « fréquence fondamentale » stable dans la voix trachéo-oesophagienne.

III.2.4.4. Intubation et conséquences sur la qualité vocale

La dysphonie après intubation est intéressante dans le cadre de cette recherche, puisque l'intubation est fréquemment utilisée lors des interventions chirurgicales et notamment lors des opérations de la thyroïde qui constituent notre sujet de recherche. Cette partie s'attachera ainsi à décrire les répercussions de l'intubation sur la qualité vocale sans paralysie des cordes vocales, les paralysies récurrentielles faisant l'objet du chapitre suivant.

Les conséquences d'une intubation pouvant s'étendre des lèvres à la trachée (Lacau Saint Guily et *al.* 2003), nous nous attacherons ici à ne décrire que les problèmes ayant un impact sur l'acte phonatoire lui-même.

Les complications laryngées ne sont pas exceptionnelles après une intubation, surtout lorsque celle-ci elle a été réalisée en urgence ou dans des conditions difficiles. La dysphonie post-intubation est un symptôme dont l'incidence varie en fonction des études entre 5 et 32% (Domino et *al.* 1999). La difficulté majeure de diagnostic réside dans le fait que les conséquences de l'intubation peuvent être précoces ou tardives et ne dépendent pas toujours de la durée de l'intubation, de la taille du tube utilisé lors de celle-ci (Pröschel & Eysholdt, 1993). Ces paramètres sont loin de faire consensus dans la communauté scientifique, puisque selon (Schwartz et *al.* 2009) par exemple, une intubation prolongée, supérieure à quatre jours, accroît de façon significative les risques de lésions laryngées et donc les complications vocales (Schwartz et *al.* 2009).

La dysphonie faisant suite à une intubation est généralement décrite comme un enrouement qui peut survenir alors même qu'aucune lésion n'est visible sur les plis vocaux (Yamanaka et *al.* 2009). En outre, sans lésions laryngées, la dysphonie régresse rapidement et de façon spontanée (Jones et *al.* 1992). Ces symptômes sont alors

comparables à ceux d'une laryngite bénigne décrite précédemment dans ce chapitre. Le timbre est rauque et l'intensité vocale diminuée.

Ces observations ne font pas consensus, puisque certaines études n'ont pas relevé de modification vocale après une intubation, et plus spécifiquement si cette intubation a été de très courte durée. C'est notamment le cas de l'étude de (Pröschel & Eysholdt, 1993) qui a montré que, pour une majorité des sujets testés, la qualité vocale ne variait pas avant et après l'intubation. En revanche, cette même étude a mis en avant que si des modifications sont observées au niveau de la voix d'un patient, celles-ci sont alors considérables. Cette recherche a ainsi révélé l'importance de prendre en compte la variabilité inter-locuteurs.

Une dysphonie plus marquée est fréquemment due à la présence d'un hématome sur l'un des plis vocaux, l'hématome et l'enrouement qui en résultent finissant par disparaître de façon spontanée en quelques jours. Malgré sa fréquence, la durée et la sévérité de l'enrouement sont toutefois peu prédictibles.

Si les dysphonies post-intubation sont souvent sans conséquence à long terme sur la voix du patient, le risque majeur réside toutefois dans l'installation du malmenage vocal. En effet, si les symptômes dysphoniques perdurent dans le temps, le patient peut malheureusement mettre en place un comportement de retenue ou, au contraire, de forçage vocal, ce qui peut alors avoir des conséquences à long terme sur l'utilisation de l'appareil laryngé, et conduire à des phénomènes proches de ceux observés précédemment dans la description des dysphonies dysfonctionnelles (Le Huche et *al.* 2010). Ce constat confirme encore que la frontière entre dysphonie d'origine organique et dysphonie dysfonctionnelle n'est pas clairement établie et qu'une zone mixte doit subsister.

Notons que si la dysphonie perdure dans le temps, c'est généralement synonyme de lésions laryngées plus importantes qui peuvent nécessiter une prise en charge orthophonique ou chirurgicale. Aucun de nos locuteurs ne présentant ce type de complications post-intubation, nous n'aborderons pas ce sujet plus en détails. Nous retiendrons uniquement que l'intubation peut être responsable d'un enrouement passager qu'il faudra prendre en compte pour l'analyse de nos données acoustiques, notamment dans les phases d'enregistrement post-opératoire précoces.

III.2.4.5. Dysphonie et radiothérapie

Certains de nos locuteurs ayant subi des séances de radiothérapie, il nous semble important de considérer que l'irradiation de la région laryngée provoque des réactions inflammatoires qui conduisent à une sécheresse de la bouche, des douleurs à la déglutition et une aggravation variable du timbre vocal.

III.2.5. Les dysphonies d'origine neurologique

Les dysphonies d'origine neurologique, appelées ainsi par Remacle (1999), peuvent être subdivisées en deux sous-catégories :

-Les dysphonies avec altérations neurologiques centrales, qui peuvent être considérées comme une conséquence d'un trouble dysarthrique. Elles concernent les pathologies qui affectent les voies nerveuses centrales des organes de la phonation et de la respiration, notamment dans le cas de maladies neurodégénératives. Notons que Ghio & Pinto (2008) utilisent le terme de « neurodysarthrophonie » pour désigner cette sous-catégorie étant donné qu'il s'agit d'un dysfonctionnement arthrique et phonatoire d'origine neurologique.

-Les dysphonies avec altérations neurologiques périphériques concernent les paralysies laryngées. Elles sont provoquées par un dysfonctionnement, à n'importe quel niveau (thorax, glotte...) de l'innervation laryngée d'origine tumorale ou traumatique. Cette classification ne fait toutefois pas consensus, puisque certaines recherches, comme celles menées par Le Huche *et al.* (2001) ou Teston (2001), classent les paralysies laryngées parmi les dysphonies d'origine organique. Cette sous-catégorie de dysphonies d'origine neurologique étant le sujet central de notre étude, nous y apporterons plus de détails dans le chapitre suivant.

CONCLUSION CHAPITRE III

Ce chapitre a permis de distinguer les notions de dysarthrie et de dysphonie. Si la première concerne les troubles de la planification et de l'exécution motrice de la parole, la seconde s'applique aux troubles de la phonation. Toutefois cette dichotomie est loin d'être figée, puisque de nombreux liens existent entre les deux notions comme cela a pu être montré à l'aide de l'exemple du sujet parkinsonien. Une pathologie dysarthrique conduit souvent à des troubles dysphoniques, raison pour laquelle le terme de neurodysarthrophonie semble particulièrement adapté pour décrire ce type de pathologie.

En outre, le chapitre III a mis en avant le fait que la dysphonie, thème central de notre recherche, regroupe toutes les atteintes qui modifient un ou plusieurs des paramètres vocaux (timbre, fréquence fondamentale, intensité...).

Le classement des dysphonies en fonction de leur étiologie offre la possibilité de quantifier plus précisément les changements vocaux qu'elles peuvent provoquer. Par la suite, nous tenterons d'analyser les résultats tirés de nos données pour les comparer aux autres pathologies mises en lumière dans ce chapitre.

RÉSUMÉ CHAPITRE IV

Le chapitre précédent a traité de la classification de la dysphonie. Deux types de dysphonies, les dysphonies d'origine organique et les dysphonies dysfonctionnelles, ont été mis en évidence en fonction de leur origine. Ce chapitre se focalisera sur la dysphonie provenant d'une paralysie laryngée unilatérale.

Dans un premier temps, nous reviendrons sur les problèmes que pose la classification de la paralysie laryngée dans les atteintes vocales. Il conviendra ensuite de faire un bref rappel anatomique afin d'identifier les origines de la paralysie laryngée unilatérale.

La suite du chapitre insistera sur les conséquences anatomiques, acoustiques et aérodynamiques d'une opération de la thyroïde. Nous verrons notamment que la qualité vocale des patients peut se trouver modifiée, même lorsque les nerfs récurrents sont préservés. A fortiori, dans les cas où les nerfs récurrents sont endommagés, la qualité vocale des patients pourra être grandement dégradée, rendant parfois son évaluation objective difficile.

CHAPITRE IV

IV. Dysphonie et Immobilités laryngées

IV.1. Classification des paralysies laryngées

La classification des paralysies laryngées fait l'objet de débats dans la communauté scientifique. En effet, certaines études, comme celle de Le Huche & Allali (2010b) proposent de classer les paralysies laryngées dans les dysphonies d'origine organique, ce qu'elles justifient dans la mesure où l'immobilité du pli vocal est provoquée par une lésion organique sur le trajet de l'un des nerfs récurrents.

Cette classification présente un inconvénient majeur : elle propose de traiter la paralysie du nerf laryngé supérieur avec les dysarthries, puisque sa principale conséquence est un défaut de tension des plis vocaux, alors que la paralysie du nerf laryngé inférieur est traitée avec les atteintes organiques. Dans la mesure où l'examen laryngoscopique simple ne permet pas toujours de les différencier clairement, il nous paraît plus cohérent de les traiter dans une même catégorie. En effet, l'examen ORL permet d'observer les anomalies des mouvements des plis vocaux et d'en supposer l'origine nerveuse sans toutefois pouvoir s'en assurer sans examen électromyographique (Yin et *al.* 1997). C'est pourquoi nous préférons classer les paralysies laryngées selon Remacle (1999), et les considérer comme la cause d'une dysphonie d'origine neurologique, avec comme acception du terme « neurologique », ce qui a endommagé le « nerf » que ce soit le nerf récurrent ou le nerf X. Il convient toutefois de retenir que l'immobilité laryngée peut également, dans de rares cas, être la conséquence d'une

atteinte de l'organe lui-même sans atteinte nerveuse, citons à titre d'exemple l'arthrite crico-aryténoïdienne (Le Huche & Allali, 2010a).

La catégorie des dysphonies neurologiques, proposée par Remacle (1999), comprend d'une part les altérations neurologiques centrales que nous avons regroupées précédemment sous le mot dysarthrie c'est-à-dire une pathologie du mouvement et, d'autre part, les altérations neurologiques périphériques (nerfs laryngés et nerf X), c'est-à-dire les différentes autres formes de paralysies laryngées. Pour les premières citées, Remacle (1999) parle de dysphonie dysarthrique, notion que l'on peut rapprocher de la « neurodysarthrophonie » proposée par Pinto & Ghio (2008), dans la mesure où l'altération du Système Nerveux Central touche également les mécanismes de la phonation.

IV.2. Anatomie et physiologie du larynx

Nous proposons ici de faire un bref rappel anatomique afin d'explicitier le lien qu'entretiennent les plis vocaux avec les nerfs laryngés.

L'innervation du larynx est assurée par le nerf laryngé supérieur et le nerf récurrent (également appelé nerf laryngé inférieur) qui sont tous deux des branches du nerf vague (dixième nerf crânien). Le nerf laryngé supérieur innerve uniquement le muscle crico-thyroïdien (muscle intrinsèque) tandis que tous les autres muscles intrinsèques du larynx (notamment le muscle vocal ou thyro-aryténoïdien inférieur) sont innervés par le nerf laryngé récurrent (Ekberg & Nylander, 2004).

Le muscle inter-aryténoïdien et le muscle crico-aryténoïdien latéral doivent travailler en synergie afin de rapprocher l'insertion postérieure des cordes vocales. Le crico-aryténoïdien postérieur est lui chargé de l'ouverture de la glotte.

Le muscle vocal est un muscle isométrique. Il est responsable de la tension de la corde vocale puisque sa contraction permet de la rigidifier. Il permet donc de modifier les propriétés vibratoires et la fréquence du son émis (Ormezzano, 2000).

Le nerf récurrent droit naît sous l'artère sous-clavière et remonte vers le larynx en suivant le côté droit de la trachée et de l'œsophage. Le nerf récurrent gauche naît dans le thorax sous la crosse de l'aorte. Son trajet est plus long, étant donné qu'il

poursuit son chemin vers le larynx en suivant le bord gauche de la trachée. Chaque nerf est responsable de l'innervation d'un côté du larynx. La longueur, ainsi que le trajet respectif de chacun de ces nerfs expliquent partiellement que les causes de paralysies récurrentielles peuvent être nombreuses (Schweizer & Dörfl, 1997).

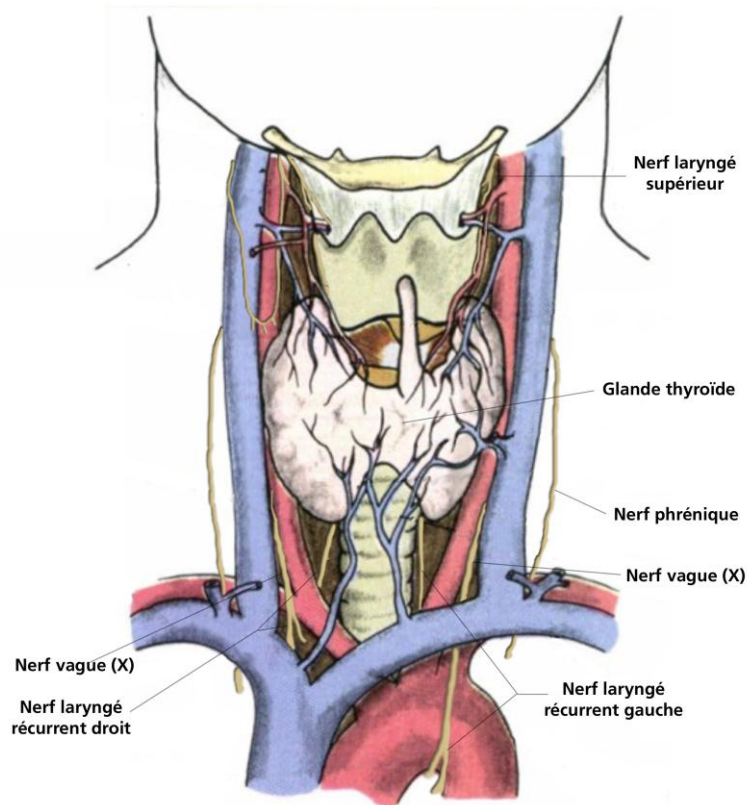


Figure 4 : Glande thyroïde, vue antérieure d'après Netter (2011)

IV.3. Epidémiologie de la paralysie laryngée unilatérale

La paralysie laryngée unilatérale est la conséquence d'une atteinte du Système Nerveux Central ou périphérique des motoneurones ou des axones du nerf récurrent (Hartl, 2004), ce qui signifie que toute atteinte sur le trajet de l'un des nerfs récurrents peut conduire à une paralysie laryngée unilatérale.

Les causes d'une paralysie laryngée unilatérale sont diverses, et son incidence dépend largement des populations étudiées mais également des moyens d'investigation (voir par exemple Benninger, 1998), raison pour laquelle nous proposons ici de lister

quelques-unes des étiologies possibles, sans toutefois en indiquer leur fréquence. Cette revue non exhaustive doit beaucoup aux travaux de Tucker, (1980), de Benninger (1998) et de Laccourreye et *al.* (2003). Selon ces recherches, la paralysie laryngée unilatérale peut être la conséquence d'un traumatisme chirurgical, telle qu'une opération du cou (glande thyroïde, artère carotide...) ou du tronc (poumon, œsophage, chirurgie cardiaque), mais également d'un traumatisme non chirurgical comme l'intubation, une plaie pénétrante ou un traumatisme non pénétrant (choc violent). Une tumeur ou un ganglion, aussi appelées causes compressives, qui auraient pour conséquence la compression du nerf récurrent peuvent également provoquer une paralysie laryngée unilatérale. Enfin, la paralysie peut aussi être d'origine toxique (chimiothérapie, radiothérapie, saturnisme ou diabète), idiopathique ou inflammatoire (névrite).

90% des atteintes neurologiques périphériques sont unilatérales, c'est-à-dire qu'elles se traduisent pas une paralysie (immobilité) ou parésie (paresse) de l'un des plis vocaux et, dans la majorité des cas, il s'agit d'une paralysie unilatérale gauche (Titcher, 1976). Toutefois, les paralysies bilatérales sont possibles, dans ce cas, la mobilité des deux plis vocaux est altérée. Notons que la paralysie bilatérale, notamment si la glotte est largement obstruée (paralysie bilatérale en fermeture), peut nécessiter une trachéotomie puisque le pronostic vital est engagé. Dans le cas d'une paralysie bilatérale en ouverture afin de protéger les voies aériennes d'une fausse-route alimentaire, une trachéotomie sera pratiquée avant une opération chirurgicale de compensation. Ces cas restent néanmoins rares.

Grâce aux travaux de Koufman et *al.* (1995) notamment, il est maintenant établi qu'il n'y a pas de lien direct entre la position du pli vocal paralysé et le niveau anatomique de l'atteinte. Le pli vocal peut être immobilisé dans quatre positions par rapport à la ligne médiane du larynx, cela ayant naturellement un impact sur les fonctions de respiration et de déglutition, mais également sur la phonation et donc sur la sévérité de la dysphonie. Plus l'écartement entre la ligne médiane du larynx et le pli vocal paralysé est grand, et plus la dysphonie sera sévère. A l'inverse, la gêne respiratoire sera moins importante si la glotte est ouverte, même si des troubles respiratoires pourront se manifester lors d'un effort. Les troubles de la déglutition n'y sont pas systématiquement associés, puisque les mouvements de l'épiglotte sont conservés ; toutefois la fermeture de la glotte dans le processus de déglutition constitue

un réflexe supplémentaire pour la protection des voies aériennes, puisque cette fermeture réflexe empêche l'entrée dans la trachée de la salive ou d'aliments qui peuvent s'introduire dans le vestibule laryngé. Aussi, des troubles de la déglutition légers sont possibles en cas de fermeture glottique incomplète (Remacle, 1999).

Selon Brasnu et *al.* (2008), le pli vocal peut être paralysé :

En abduction, plus de 2.5mm de la ligne médiane (la glotte est alors grande ouverte) ;

- En adduction (aussi appelé médiane, il s'agit de la position de phonation) ;
- Dans une position intermédiaire entre 1.5mm et 2.5mm (ou de repos) ;
- Dans une position paramédiane, entre 0 et 1.5mm de la ligne médiane.
- Le pli vocal est le plus souvent paralysé en position intermédiaire ou paramédiane. La paralysie en abduction est très rare.

IV.4. Thyroïdectomie et paralysie laryngée unilatérale

IV.4.1. Pronostic de la paralysie laryngée post-thyroïdectomie

Notons que si la cause la plus fréquente de la paralysie laryngée unilatérale demeure controversée (voir *supra*), l'opération de la glande thyroïde nécessite une attention particulière. Si l'agression directe du (des) nerf(s) laryngé(s) inférieur(s), que ce soit par section, dévascularisation, compression ou élongation, est classiquement incriminée dans la survenue d'une immobilité laryngée, ce mécanisme n'est pas le seul à générer une telle complication (Farrag et *al.* 2009 ; Dralle et *al.* 2004). La paralysie unilatérale post-thyroïdectomie peut effectivement être attribuée au geste chirurgical lui-même, mais également à l'intubation trachéale, même si cette cause est relativement rare (Friedrich et *al.*, 2000). Les travaux Baujat et *al.* (2001) indiquent que les avancées techniques permettent à présent de différencier une atteinte récurrentielle d'un traumatisme lié à l'intubation. Cette observation ne fait toutefois pas consensus.

Certains auteurs (Jamski et *al.* 2004 ; Randolph & Kamani, 2006) insistent également sur l'importance de l'examen clinique du larynx en pré et post-opératoire précoce au décours de la chirurgie de la glande thyroïde notamment afin de dépister un

éventuel trouble de la mobilité laryngée chez des patients a priori asymptomatiques (Liu et *al.* 1998).

La paralysie laryngée unilatérale post-thyroidectomie se justifie, entre autres, par une réalité anatomique, étant donné que les nerfs récurrents droit et gauche passent sur la face postérieure du lobe latéral de la glande thyroïde. Une opération de la glande thyroïde peut donc conduire à la lésion accidentelle ou nécessaire de l'un des nerfs récurrents, même si le chirurgien attache une importance particulière à la préservation de ces nerfs.

L'incidence de la paralysie laryngée unilatérale post-thyroidectomie reste heureusement relativement faible. Selon les études publiées ces dernières années (voir par exemple (Benninger et *al.* 1998 ou Benninger, 1998 ou Lo, et *al.* 2000) prenant en compte 500 patients minimum, le taux de paralysies laryngées en postopératoire immédiat (soit un mois après l'opération) après une opération de la glande thyroïde est compris entre 0,5% et 8,3%. Selon les séries, la paralysie laryngée inférieure initiale peut varier de 0,5% à 12%. De façon générale, le taux de paralysie laryngée inférieure définitive dépasse rarement 1% (voir par exemple Wittekindt et *al.* 2007). L'incidence définitive de la paralysie laryngée serait en réalité comprise entre 0 et 3,2% avec une médiane de 1% (Trésallet et *al.* 2006 et Laccourreye et *al.* 2008). La majorité des immobilités laryngées consécutives à une chirurgie thyroïdienne sont paucisymptomatiques et environ un tiers des patients récupère spontanément (Laccourreye et *al.* 2009). Si l'on considère uniquement les thyroidectomies, ce sont les interventions pour un goitre, un cancer ou une maladie de Basedow qui présentent plus de risque de paralysie laryngée postopératoire (Lo et *al.* 2000). Si certaines études insistent sur l'importance de repérer le tronc du nerf laryngé pour prévenir les paralysies laryngées définitives (Jatzko & Wette, 1994), ou sur l'utilisation d'un neurostimulateur afin de faciliter le repérage du trajet du nerf laryngé, ces recommandations semblent toutefois ne pas réduire statistiquement les risques de paralysie laryngée définitive (Koch et *al.* 1996).

IV.5. Immobilités laryngées post-thyroïdectomie

IV.5.1. Données chiffrées

La littérature ne s'accorde pas pour dire si la paralysie laryngée unilatérale, consécutive à une opération de la glande thyroïde, est passagère ou définitive, aussi est-il difficile d'avancer ici des chiffres. La plupart des études présente le biais méthodologique de ne pas différencier les atteintes unilatérales des atteintes bilatérales. Il est donc probablement plus prudent de parler en termes de tendance. Une étude ancienne de Hockauf & Sailer (1982) indiquait une récupération totale de la mobilité du pli vocal paralysé pour 43,4% des patients, une récupération partielle dans 3,7% des cas, et une absence de récupération dans 53,9% des cas. Ces chiffres sont corroborés par l'étude de Martensson & Terins (1985) qui indique que la paralysie était temporaire dans 50% des cas. L'étude plus récente de Wagner & Seiler (1994) avance, quant à elle, 60% de récupération suite à une paralysie laryngée unilatérale. En tout état de cause, une paralysie laryngée qui persiste au-delà d'un an sera considéré comme définitive.

Statistiquement, la récupération de la mobilité intervient durant les six premiers mois dans environ 75% des cas. De très rares récupérations ont pu être observées au-delà du douzième mois. Une paralysie laryngée qui persiste au-delà d'un an sera considérée comme définitive (Steinmüller et *al.* 2001 ; Thomusch et *al.* 2003). Compte tenu des possibilités de récupération spontanée, il convient généralement d'attendre jusqu'à un an avant d'envisager une nouvelle opération chirurgicale visant à compenser le déficit de mobilité laryngée (Hartl & Brasnu, 2000).

IV.5.2. Traitement des paralysies laryngées

IV.5.2.1. La rééducation orthophonique de la paralysie laryngée

Actuellement, le traitement orthophonique est le traitement le plus fréquemment proposé. Il est recommandé de commencer les séances le plus rapidement possible, afin de prévenir tous mécanismes de forçage néfastes pour les structures laryngées (voir chapitre précédent). Les patients ayant été inclus dans notre protocole expérimental ont tous bénéficié d'un traitement orthophonique, raison pour laquelle il nous semble important d'aborder cet aspect ici.

La rééducation vocale passe d'abord par l'information du patient ; il s'agit de lui proposer une description détaillée de l'organe vocal et de son fonctionnement. Le traitement orthophonique repose sur trois principes majeurs : 1) l'éviction de la déperdition du souffle ; 2) l'obtention de l'affrontement des plis vocaux et ; 3) la recherche de sons d'une tonalité précise. Le protocole rééducatif se compose alors d'exercices de respiration abdominale et thoracique intérieure, de relaxation et de détente, de mouvements de compensation par le pli vocal mobile et de répétitions sonores (Le Huche et al, 2001). Le protocole vise surtout à éviter la compensation supra-glottique aussi appelée voix des bandes ventriculaires. Un exemple de protocole rééducatif détaillé est disponible en annexe.

La place de la rééducation orthophonique dans le traitement de la paralysie laryngée ne fait pas consensus dans la communauté scientifique (Hoffman & McCulloch, 1996). Si une récupération spontanée est espérée et que la paralysie laryngée ne menace pas directement les fonctions de respiration et de déglutition, les séances d'orthophonie sont généralement le seul traitement proposé (Remacle, 1999). Le réel bénéfice de ces séances est toutefois discuté dans la mesure où il n'existe actuellement aucune étude pouvant apporter des renseignements sur le nombre de séances, la durée de la rééducation, ni même la méthode à suivre pour répondre au mieux à une paralysie laryngée (Laccourreye et al, 2003). Crumley (1994) va plus loin en indiquant que le rôle principal de la rééducation orthophonique est d'éduquer le patient sur son comportement vocal et de lui apporter un soutien psychologique. L'étude de Kelchner et al. (1999) nuance ce propos en indiquant que la rééducation orthophonique est d'autant plus bénéfique que la dysphonie est légère et qu'en cas de dysphonie et de symptômes associés plus importants, seule l'intervention chirurgicale améliorera l'état du patient.

IV.5.2.2. Les traitements chirurgicaux de la paralysie laryngée

Dans la mesure où aucun de nos locuteurs n'ait subi d'intervention chirurgicale suite au diagnostic de sa paralysie unilatérale laryngée, nous nous contenterons ici d'évoquer les possibilités chirurgicales sans entrer dans les détails.

Deux types d'intervention sont possibles : d'une part, les techniques de réinnervation et, d'autre part, les techniques de médialisation cordale par injection intracordale, adduction aryténoïdienne ou thyroplastie.

En ce qui concerne la réinnervation, il s'agit de redonner un volume et un tonus musculaires au muscle thyro-aryténoïdien. Cette technique permet théoriquement d'améliorer l'affrontement des plis vocaux et leurs qualités vibratoires. Cette technique est peu employée (Crumley, 1990).

L'injection intracordale consiste en l'injection par voie endoscopique d'une substance comme le téflon, la graisse autologue ou la fascia autologue (seuls produits autorisés en France actuellement). Cette technique apporte d'excellents résultats phonatoires mais une certaine résorption du produit injecté peut nécessiter une nouvelle intervention (Laccoureye, 1999).

Il est possible de reproduire l'action adductrice des muscles thyro-aryténoïdien et crico-aryténoïdien latéral en plaçant un fil de traction au niveau du processus musculaire de l'aryténoïde. Il s'agit de la technique de l'adduction aryténoïdienne. Cette technique associée à la thyroplastie permet un résultat vocal perceptif amélioré (McCulloch et *al.* 2000).

La thyroplastie est une technique chirurgicale par laquelle un implant est placé au travers d'une fenêtre confectionnée dans le cartilage thyroïde. Cet implant permet la médialisation du pli vocal et du processus vocal. La qualité vocale est généralement stabilisée environ trois mois après l'intervention (Netterville et *al.* 1993).

IV.6. Conséquences acoustiques de la thyroïdectomie

IV.6.1. Altérations vocales sans paralysies laryngées

Dans la littérature, la paralysie récurrentielle est l'une des complications post-thyroïdectomie ayant des conséquences sur la qualité vocale les plus étudiées. Il convient toutefois de noter qu'une opération au niveau du larynx est rarement anodine, dans la mesure où la voix d'un patient peut se trouver altérée et la déglutition gênée après une thyroïdectomie, alors même que les nerfs laryngés sont préservés (Lombardi et *al.* 2006). Cette section propose de présenter les modifications de la qualité vocale des patients pour lesquels le diagnostic ORL n'a pas révélé de paralysie laryngée.

IV.6.1.1. Fréquence fondamentale

Selon l'étude de Hong & Kim (1997), la fréquence fondamentale calculée à partir de la voyelle soutenue /i/ n'est significativement pas modifiée par la thyroïdectomie ($p>0.05$). En effet, la fréquence fondamentale est mesurée à 207 Hz avant l'opération puis à 196 Hz, 214 Hz et 204 Hz respectivement à un mois, trois mois et six mois après l'opération. Si ce paramètre n'est pas significativement modifié, les auteurs concluent toutefois à une grande variabilité inter et intra locuteur.

Ces résultats sont confirmés par l'étude de Stojadinovic et *al.* (2002), qui n'indique pas de changements significatifs pour la hauteur de la fréquence fondamentale entre les différentes phases d'enregistrement. En ce qui concerne les locuteurs féminins, la fréquence fondamentale est quantifiée à 183 Hz en pré-opératoire, à 186 Hz une semaine après l'opération et à 184 Hz trois mois après l'intervention.

Enfin l'étude de Netto et *al.* (2006) indique également le même type de schéma en ce qui concerne la fréquence fondamentale. Ce paramètre n'est pas significativement modifié par la thyroïdectomie.

IV.6.1.2. Temps maximum de phonation

L'étude de Hong & Kim (1997) a montré que le temps maximum de phonation n'était pas significativement modifié après une opération de la thyroïde pour la voyelle soutenue /i/. Le temps maximum de phonation moyen était de 21.7 secondes avant l'opération, puis de 18.8 secondes, 20.3 secondes et 19 secondes un mois, trois mois et six mois après l'opération. Il convient toutefois de remarquer que si ce paramètre n'a pas été significativement modifié ($p>0.05$), il est toutefois diminué après l'opération. Par la suite, le temps maximum de phonation tend à augmenter à nouveau et à se rapprocher des valeurs mesurées en pré-opératoire.

L'étude de Stojadinovic et *al.* (2002), conduite à partir de la voyelle soutenue /a/ notamment, indique la même tendance. Le temps maximum de phonation n'a pas été diminué par l'opération de la thyroïde. Il est de 17.62 secondes en pré-opératoire puis de 17.86 secondes une semaine après l'intervention et enfin de 18.38 secondes 3 mois après l'opération. Les écarts-types restent importants quelle que soit la phase d'enregistrement étudiée.

L'étude de Keilmann & Hülse (1992) indique que, malgré un examen laryngoscopique normal, un changement vocal était remarquable pour les patients ayant subi une thyroïdectomie alors même que les nerfs laryngés étaient préservés. Outre une diminution de la fréquence fondamentale et une intensité diminuée, les locuteurs se plaignent également de fatigue vocale. Cette fatigue vocale se traduit par un essoufflement en situation de lecture et une réduction du temps maximum de phonation. Ces changements ne sont toutefois pas significatifs ($p > 0.05$) et leurs résultats témoignent d'une certaine variabilité inter-locuteurs.

IV.6.1.3. Fréquence fondamentale de la voix parlée

L'étude de Debruyne *et al.* (1997) a porté sur la voix de 47 patientes ayant subi une thyroïdectomie et pour lesquelles l'examen laryngoscopique a révélé une préservation de la mobilité des cordes vocales et une absence de signe d'intubation traumatique. Les locutrices ont été enregistrées avant l'opération (phase pré-opératoire), quatre jours après l'opération (phase post-opératoire 1) et deux semaines après l'intervention (phase post-opératoire 2).

Leur étude a révélé que la fréquence fondamentale de la voix parlée était significativement plus basse lors des phases d'enregistrement post-opératoires (-11,59 Hz ; $p < 0.001$ et -4,97 Hz à $p < 0,05$ respectivement) que celle mesurée avant l'opération (210,4 Hz). L'étude indique également que ce sont les locutrices opérées pour une thyroïdectomie totale qui présentent une baisse de fréquence fondamentale en voix parlée plus importante (-13,48 Hz en post-opératoire 1 et -5,65 Hz en post-opératoire 2) que celles opérées pour une thyroïdectomie unilatérale (-7,67 Hz en post-opératoire 1 et -3,53 Hz en post-opératoire 2). Ce paramètre révèle également une grande variabilité inter-locutrice, puisque en post-opératoire 1 par exemple, 26 patientes présentent une diminution de la fréquence fondamentale en voix parlée, inférieure à un demi-ton, alors que 8 montrent une diminution de ce paramètre, supérieure à deux demi-tons. Aucune locutrice ne présentait toutefois d'augmentation de la fréquence fondamentale en voix parlée supérieure à deux demi-tons.

Une baisse de la fréquence fondamentale, suite à une opération de la thyroïde, a également été constatée par l'étude de Hong & Kim (1997). La voix de patients ayant subi une thyroïdectomie sans atteinte des nerfs laryngés était enregistrée avant l'opération puis un mois, trois mois et six mois après l'opération. L'étude remarque une

baisse de la fréquence fondamentale en voix parlée de façon significative un mois et trois mois après l'opération ; six mois après l'opération la fréquence fondamentale est toujours mesurée plus basse qu'avant l'opération mais de façon non significative. Elle était mesurée à 205 Hz avant l'opération, à 180 Hz un mois après l'opération, à 177 Hz trois mois après l'opération et à 191 Hz six mois après l'opération.

IV.6.1.4. Amplitude de la fréquence fondamentale

Debruyne et *al.* (1997) ont également porté attention à l'amplitude de la fréquence fondamentale en voix parlée (c'est-à-dire l'intervalle entre la fréquence fondamentale la plus basse et la fréquence fondamentale la plus haute dans la chaîne parlée) comme un indice intonatif. Leurs mesures indiquent que pour le groupe ayant subi une thyroïdectomie totale, la voix était plus monotone en phase post-opératoire 1. Ce paramètre ne s'est plus révélé significatif lors de la phase d'enregistrement post-opératoire 2. L'amplitude de la fréquence fondamentale n'a pas été modifiée pour le groupe de locutrice ayant subi une thyroïdectomie partielle. Ces résultats corroborent ceux de Hong & Kim (1997) qui observent une amplitude de la fréquence fondamentale réduite un mois, trois mois et six mois après l'opération. Notons que cette étude indique des mesures significativement réduites lors des deux premières phases d'enregistrement opératoires (un mois et trois mois), alors que l'étude de Debruyne et *al.* (1997) indique des différences non significatives dès quinze jours après l'intervention chirurgicale.

IV.6.1.5. Intensité de la voix parlée

En ce qui concerne l'intensité de la voix parlée, Hong & Kim (1997) notent que ce paramètre n'est pas modifié de manière significative après l'opération de la thyroïde pour leur cohorte de locuteurs. L'intensité moyenne était mesurée à 48.5 dB avant l'opération, 47.9, 50.2 et 49.2 dB un mois, trois mois et six mois après l'opération, respectivement.

IV.6.1.6. Jitter et Shimmer

L'étude de Debruyne et *al.* (1997) a également retenu le jitter comme indice de modification vocale. Il s'est également révélé être une mesure significativement modifiée en post-opératoire 1, pour le groupe de locutrices ayant subi une thyroïdectomie totale uniquement. Les valeurs sont à nouveau comparables à celles

mesurées en pré-opératoire dès la phase post-opératoire 2. Le même type de schéma a pu être observé pour la proéminence au niveau des harmoniques (harmonic prominence).

Ces résultats ne sont toutefois pas confirmés par l'étude de Stojadinovic et *al.* (2002). La voix de patients a été enregistrée en pré-opératoire, une semaine après l'intervention puis trois mois après l'opération. Si l'étude n'a pas permis de relever des différences significatives pour les mesures de jitter et de shimmer entre les différentes phases d'enregistrement, signalons toutefois que, d'un point de vue quantitatif, le jitter et le shimmer étaient respectivement mesurés à 0.63% et 1.80% en pré-opératoire, à 0.65% et 1.88% une semaine après l'intervention et à 0.60% et 1.90% trois mois après l'intervention. La phase post-opératoire 1 (c'est-à-dire une semaine après l'intervention) est caractérisée par des écarts-types plus importants et témoigne donc d'une plus grande variabilité inter et intra locuteur. Les résultats de cette étude sont à rapprocher de ceux de Hong & Kim (1997) qui n'indiquaient pas de différence significative pour la mesure de jitter entre les différentes phases d'enregistrement.

IV.6.1.7. Evaluation subjective

Nous avons jusqu'à présent détaillé les modificatifs objectives que peut entraîner une opération de la thyroïde sur la voix des patients, alors même que les nerfs laryngés sont préservés. Certaines études ont également cherché à évaluer les changements vocaux à l'aide d'échelles d'auto-évaluation comme le Voice Handicap Index ou VHI. Rappelons qu'une différence entre deux auto-évaluations vocales d'au moins 18 points au total ou de 8 points par catégorie suggèrent des modifications vocales importantes.

L'étude de Pedro Netto et *al.* (2006) a comparé les scores de VHI des patients avant et après l'opération (post-opératoire + 15 jours). L'étude révèle que, pour 2.2% des patients, les scores étaient diminués de plus de 18 points sur le score total de VHI. Par ailleurs, leurs résultats indiquent également que les scores dans les catégories intermédiaires étaient modifiés de plus de 8 points pour 3.4% des patients pour les paramètres relatifs aux aspects fonctionnels de la voix, pour 2.3% des patients pour les affirmations liées aux caractères physiques de la voix, et enfin pour 1.1% des patients pour les aspects émotionnels de la voix.

Signalons encore que ces données sont confirmées par les travaux de Inabnet et *al.* (2009). Selon cette étude, l'auto-évaluation vocale, effectuée à partir du

VHI, n'a pas permis de montrer de modifications vocales significatives suite à thyroïdectomie.

Enfin l'étude de Van Lierde et *al.* (2010) conduite à partir du Dysphonia Severity Index indique qu'après une thyroïdectomie, les critères objectifs et perceptifs d'évaluation de la qualité vocale des patients n'étaient pas significativement modifiés. De plus, les modifications vocales n'étaient que transitoires.

Les résultats de cette recherche sont toutefois à nuancer dans la mesure où l'évaluation vocale est effectuée entre deux et trois semaines après l'intervention. Il est probable que la qualité vocale ait été modifiée directement après l'opération, comme l'atteste les études objectives précédemment citées, mais qu'elle se soit déjà améliorée avec le temps.

Les différentes études présentées précédemment ont permis de montrer que la thyroïdectomie peut avoir des conséquences sur les caractéristiques acoustiques de la voix des patients alors même que les nerfs laryngés sont préservés. Les causes de ces modifications peuvent être nombreuses. L'intubation oro-trachéale a souvent été invoquée ; toutefois suite aux travaux de Pedro Netto et *al.* (2006), il est difficile d'établir un lien direct entre cette technique et les conséquences acoustiques. L'équipe de Pedro Netto a montré que si les complications vocales étaient plus souvent présentes dans le groupe de patients ayant subi une thyroïdectomie que dans le groupe de patients ayant subi une opération de la poitrine (leur groupe contrôle), l'intubation oro-trachéale n'était qu'un des multiples facteurs pouvant expliquer une altération vocale post-thyroïdectomie.

Ces études renforcent l'importance d'un diagnostic ORL pré et post-opératoire visant à évaluer les conséquences directes de la chirurgie thyroïdienne sur la qualité vocale des patients, dans la mesure où des troubles de la voix peuvent être observés après une thyroïdectomie, alors même que les nerfs laryngés sont préservés (voir par ex. Debruyne, et *al.* 1994 et Hong & Kim 1997). Cela s'explique d'ailleurs par le fait que la chirurgie thyroïdienne peut avoir des conséquences sur la musculature supra-glottique dans son ensemble, ce qui peut se traduire naturellement par des modifications acoustiques (Timon et *al.* 2010).

L'information du patient et les diagnostics pré et post-opératoires sont donc primordiaux, notamment pour éviter les problèmes médico-légaux (Laccourreye et *al.* 2009).

IV.6.2. Altérations vocales avec paralysies laryngées

La plupart des études acoustiques menées sur la voix de patients présentant une paralysie laryngée ont l'inconvénient, pour nos investigations, de traiter des améliorations vocales que les techniques chirurgicales peuvent apporter au patient. Dans la mesure où aucun des locuteurs analysés dans cette thèse n'a dû subir une opération chirurgicale visant à pallier la gêne fonctionnelle, il semble peu pertinent de les détailler ici. Aussi, nous nous focaliserons sur les quelques études qui évaluent la qualité vocale de patients présentant une paralysie laryngée sans ou avant une réintervention fonctionnelle. Signalons toutefois que quelle que soit la méthode chirurgicale employée, la qualité vocale du patient se trouve améliorée (Hartl, 2004).

Cette section détaillera les différentes altérations vocales liées à une paralysie laryngée. Rappelons qu'une atteinte de l'un des nerfs laryngés ne provoque pas nécessairement une dysphonie (Meek et *al.* 2008 ; Trinidad Ruiz et *al.* 2007).

Une immobilité laryngée peut modifier légèrement ou sévèrement les paramètres physiologiques, acoustiques et aérodynamiques de la voix. Stewart & Allen, (2006) signalent que les paramètres d'évaluation vocale sont parfois peu adaptés à l'évaluation de la dysphonie suite à une immobilité laryngée. En effet, quand la voix peut être trop dégradée, il convient de combiner les mesures objectives et les mesures perceptives pour tenter de quantifier la dysphonie.

IV.6.2.1. Mesures acoustiques

Conduite à partir d'échantillons de voix dégradées suite à une paralysie récurrentielle, l'étude de Inagi et *al.* (1997), confirme que les mesures de jitter et de shimmer sont à prendre avec précaution, notamment dans le cas de dysphonies sévères, où les mesures de perturbation du signal peuvent dépasser 10%. En effet, pour les voix très dégradées, ces mesures de la perturbation peuvent se révéler non pertinentes. En outre, ces paramètres ne permettent pas de distinguer clairement les degrés de dysphonie et n'aident pas au diagnostic laryngé.

L'étude de Hartl et *al.* (2001), menée sur deux locuteurs avant et après l'opération responsable de la paralysie récurrentielle, indique une modification de tous les paramètres acoustiques lors de la phase d'enregistrements post-opératoires. Les locuteurs présentent alors une immobilité laryngée gauche. Les mesures de jitter et de shimmer augmentent significativement (voir Tableau 2 ci-après). Les auteurs critiquent toutefois ces mesures de perturbation du signal dans le sens où elles manqueraient de sensibilité notamment pour identifier les phénomènes de diplophonie souvent présents chez les patients présentant une paralysie récurrentielle.

Tableau 1 : Données numériques pour le jitter et le shimmer d'après Hartl et *al.* (2001)

	Patient 1		Patient 2	
Paramètres	Pré-opératoire	Post-opératoire	Pré-opératoire	Post-opératoire
Jitter %	2,25	10,99	0,71	6,66
Shimmer %	6,49	21,78	2,42	11,26

En ce qui concerne les mesures d'Harmonics-to-Noise Ratio (HNR) et de Pics de Proéminence Cepstral (PPC), les valeurs diminuent pour les deux locuteurs, après le diagnostic en post-opératoire de la paralysie récurrentielle. Le PPC permet de quantifier les irrégularités de fréquence fondamentale. Il existerait une corrélation entre le critère « soufflé » d'une voix dysphonique et une diminution du CPP (voir Hillenbrand et *al.* 1994). Les résultats de Hartl et *al.* (2001) confirment cette hypothèse.

Tableau 2 : Données numériques pour le HNR et le CPP d'après Hartl et *al.* (2001)

	Patient 1		Patient 2	
Paramètres	Pré-opératoire	Post-opératoire	Pré-opératoire	Post-opératoire
HNR	5,67	1,23	7,34	3,29
CPP	4,39	1,23	5,21	2,02

Enfin, ces auteurs ont également cherché à quantifier les conséquences de l'immobilité laryngée à l'aide de mesures spectrales. Ils ont alors mesuré la différence de niveau entre le premier harmonique (H1 ou F0) et le second harmonique (H2) en décibels (H1-H2). Ils ont également mesuré la différence entre le niveau du premier

harmonique et le premier formant (H1 – L1), ainsi que la différence de niveau entre le premier harmonique et le troisième formant (H1 – L3). Notons que ces mesures renseignent sur le caractère « soufflé » de la voix. Au niveau des résultats, les deux locuteurs n'ont pas présenté le même type de schéma. Le patient 1 présentait une diminution de la différence entre H1-H2 et H1- L1 entre les deux phases d'enregistrement, alors que ces différences augmentent significativement pour le patient 2. Pour les deux patients en revanche, la différence de niveau pour H1 – L3 diminue. Les différences de schémas sont probablement un indice d'un comportement hypofonctionnel dans un cas (patient 1) et hyperfonctionnel dans l'autre (patient 2) (Hammarberg et *al.* 1986).

Tableau 3 : Données numériques pour les différences H1-H2, H1-L1 et H1-L3 d'après Hartl et al. (2001)

	Patient 1		Patient 2	
Paramètres	Pré-opératoire	Post-opératoire	Pré-opératoire	Post-opératoire
H1 – H2	15,25	8,07	2,91	19,37
H1 – L1	12,22	0,13	10,55	20,39
H1 – L3	38,29	12,60	37,19	33,13

Zhang et *al.* (2005) ont étudié la qualité vocale de patients présentant une paralysie récurrentielle. L'originalité de leur étude repose sur le fait qu'ils cherchent à comparer la pertinence des paramètres traditionnels permettant d'étudier la perturbation du signal, comme le jitter et le shimmer, et des méthodes dynamiques non linéaires comme la dimension de corrélation par exemple. Nous retiendrons de cette étude que les mesures de jitter et de shimmer se sont révélées significativement modifiées pour les patients par rapport aux sujets contrôles. Le jitter moyen est mesuré à 0,681% pour les sujets contrôles et à 1,339% pour les sujets pathologiques. Le shimmer moyen est mesuré à 2,188% pour les sujets contrôles et à 3,592% pour les sujets pathologiques. Signalons rapidement que les autres paramètres étudiés par ces auteurs se sont également révélés être significativement modifiés par la paralysie récurrentielle.

IV.6.2.2. Mesures aérodynamiques

Même si cette présente étude ne porte que sur des paramètres acoustiques, il semble important de signaler la paralysie récurrentielle modifie également les paramètres aérodynamiques de la voix. Hartl et *al.* (2001) indiquent que le débit d'air moyen et la pression intra-orale augmentent alors que la résistance glottique et l'efficacité vocale diminuent suite à la paralysie récurrentielle. Le processus vocal devient alors un effort phonatoire, ce qui se traduit naturellement par une efficacité phonatoire diminuée et des effets de fatigue vocale. Ceci peut notamment être évalué par le temps maximum de phonation. Pour un patient étudié, il était de 14,8 secondes avant l'opération et de 3,50 secondes après l'opération.

IV.6.2.3. Mesures perceptives

L'étude de Morsomme et *al.* (2001) a cherché à évaluer les conséquences de l'immobilité laryngée à l'aide de l'échelle GIRBAS (Grade – qualité générale, Instability - instabilité, Roughness - raucité, Breathness - souffle, Asthenia - asthénie et Strain - serrage). Leurs mesures indiquent que les notes pour les 6 paramètres étaient significativement plus hautes pour les patients que pour les sujets contrôles ($p < 0,001$). Leur étude conclut également qu'il existerait une corrélation entre les notes obtenues pour les paramètres G, B et A et les mesures objectives qui expriment l'apériodicité du signal, comme le coefficient de variabilité de l'intensité, le coefficient de variabilité de la fréquence fondamentale et le jitter.

Pour résumer, cette section a permis d'étudier les conséquences acoustiques, aérodynamiques et perceptives de la paralysie laryngée sur la voix des patients. La difficulté principale pour l'étude de la qualité vocale des patients présentant une paralysie récurrentielle repose sur le caractère dégradé de la voix. En effet, dans certains cas, la voix est rendue tellement apériodique qu'il devient alors difficile d'y appliquer les paramètres d'analyse traditionnels. C'est pourquoi de nombreux auteurs ont cherché à appliquer d'autres méthodes d'analyse (voir par exemple Giovanni et *al.* 1999 ou Zhang et *al.* 2005). Ils concluent toutefois à la non-redondance des différentes mesures et suggèrent d'ajouter les méthodes d'analyse pour rendre compte le plus fidèlement possible des modifications vocales.

CONCLUSION CHAPITRE IV

Dans le début de ce chapitre, nous avons indiqué que la littérature ne s'accordait pas pour classer les paralysies laryngées parmi les atteintes d'origine organique (Le Huche & Allali, 2010a) ou d'origine neurologique (Remacle, 2000). Nous avons préféré la classification de Remacle, parce qu'elle considère la paralysie laryngée dans une même catégorie que l'atteinte nerveuse se situant au niveau du nerf laryngé supérieur ou inférieur. L'acception du terme « neurologique » est alors à comprendre comme « ce qui a endommagé le nerf », que ce soit le nerf récurrent ou le nerf X.

La suite du chapitre s'est attachée à décrire anatomiquement les paralysies laryngées et les liens qu'entretiennent thyroïdectomie et immobilités laryngées. Il a également été question des différents traitements possibles (orthophonique ou chirurgical), suite au diagnostic d'une paralysie récurrentielle.

Enfin, la dernière partie de ce chapitre nous a permis d'étudier les conséquences d'une opération de la thyroïde sur la voix des patients. En tout état de cause, ni la technique ou le temps d'intubation, ni une anesthésie prolongée ou la taille importante d'un nodule ne permettent de prédire l'importance des altérations vocales ; c'est pour cette raison que l'information fournie au patient nous semble centrale. La qualité vocale peut se trouver modifiée, alors même que les nerfs récurrents sont préservés. En cas de paralysie récurrentielle ou « simplement » à cause de l'opération chirurgicale, les niveaux phonatoire, acoustique et perceptif peuvent se trouver dégradés.

PARTIE 2 :

PROTOCOLE

EXPERIMENTAL

Cette partie, composée d'un chapitre, détaille notre protocole expérimental.

Chapitre V : Protocole expérimental

RÉSUMÉ CHAPITRE V

Ce chapitre est consacré à notre démarche expérimentale. Il décrit les conditions d'acquisition des données, le corpus, les locuteurs, les paramètres et mesures retenus pour nos études.

Les données ont été obtenues grâce à la collaboration avec le département de Chirurgie Oncologique du Centre Paul Strauss de Strasbourg. Les enregistrements des locuteurs pathologiques ont eu lieu en milieu hospitalier dans les phases d'enregistrement précoces, puis en cabinet d'orthophonie, lorsqu'une rééducation vocale était nécessaire, pour les enregistrements plus tardifs. Les locuteurs contrôles ont été enregistrés en milieu hospitalier.

En ce qui concerne le corpus, les locuteurs devaient réaliser deux tâches : 1/ Prononcer à dix reprises au moins et soutenir environ cinq secondes les trois voyelles /i a u/. 2/ Répéter dix fois au moins une série de douze logatomes.

Notre démarche expérimentale se veut articulatoire-acoustique, l'objectif étant d'obtenir, autant que faire se peut, des informations articulatoires à partir de l'analyse du signal acoustique. Pour ce faire, nous décrirons sommairement les différents paramètres mesurés et donnerons une définition des événements articulatoire-acoustiques retenus.

CHAPITRE V

V. Protocole expérimental

Il sera ici question du protocole général mis en place dans notre travail expérimental. Les contraintes médicales et hospitalières ont rendu difficile la standardisation de nos méthodes d'acquisition de données, c'est pour cette raison que les procédures particulières adoptées seront précisées en début de chaque chapitre pour chacune de nos investigations.

V.1. Conditions d'enregistrement

Le présent travail a été mené en collaboration avec le Centre Paul Strauss. Nous bénéficions à Strasbourg d'une situation tout à fait exceptionnelle, puisque le Département de Chirurgie Oncologique du Centre Paul Strauss, Centre Régional de Lutte contre le cancer, reçoit entre 650 et 700 patients par an. Cela offre une possibilité unique de conduire des analyses systématiques sur les diverses conséquences d'une chirurgie thyroïdienne sur la voix des patients.

Les enregistrements acoustiques des patients et des locuteurs contrôles se sont déroulés au sein du Centre Paul Strauss.

Avant de procéder aux enregistrements, les patients ont tous signé une autorisation stipulant qu'ils étaient d'accord pour que leur voix soit analysée dans le cadre de nos recherches. Les patients ont été enregistrés dans une salle de consultation, au moment de leur admission hospitalière en préopératoire, et après la consultation du

médecin, en post-opératoire. Il est évident que les conditions d'enregistrement n'étaient pas optimales, à la fois pour des raisons logistiques mais aussi pour des raisons humaines. En effet, le contexte et les consultations peuvent être éprouvants pour le patient. Les locuteurs se sont toutefois investis totalement et se sont appliqués à réaliser les tâches demandées.

Les enregistrements plus tardifs (Post-opératoire 3, 4 et 5) ont eu lieu dans le cabinet d'orthophonie choisi par le patient. Ces enregistrements ont été rendus possibles grâce à la collaboration des orthophonistes, et ont toujours eu lieu après la séance de rééducation, afin de ne pas fatiguer le patient avant d'entamer son travail rééducatif. Cela implique naturellement que la voix des locuteurs ait été largement sollicitée juste avant notre séance d'enregistrement, ce qui introduit un certain taux de fatigue vocale relatif.

Ces conditions particulières expliquent le choix d'un corpus restreint. Dans certains cas, devant la pénibilité de la tâche de répétition, nous avons dû procéder à de légers aménagements, notamment en permettant au patient de faire des pauses lorsque nous jugions nécessaire.

L'enregistrement des locuteurs contrôles a également eu lieu au Centre Paul Strauss. Les locuteurs contrôles ont été appariés en genre, en âge et en langue maternelle. Notons qu'il semblait important d'apparier en langue maternelle nos locuteurs car, compte tenu de l'âge et de la provenance géographique des patients enregistrés, le français a souvent été appris avec la scolarisation, parallèlement à l'alsacien qui a souvent une place importante dans le système linguistique de nos locuteurs. Un diagnostic ORL nous a apporté la certitude que les locuteurs contrôles retenus ne présentaient aucune pathologie vocale et une mobilité laryngée préservée.

V.2. Matériel

Nous n'avons pas pu acquérir de données articulatoires ou aérodynamiques, ce qui nous aurait naturellement permis d'approfondir ou d'explorer certains de nos axes de recherche (*cf.* VIII.4.). Nous nous sommes donc limitée pour ce travail à des enregistrements acoustiques. Néanmoins, comme nous l'avons signalé plus haut, notre

approche du signal acoustique est à considérer comme une démarche articulatoire-acoustique. Il s'agit en effet d'extraire, autant que faire se peut, des informations articulatoires du signal acoustique (*cf.*, par ex., Sock, 1998).

La collecte de données acoustiques est réalisée sur un enregistreur numérique Marantz Professional© PMD661, relié à un microphone directif Sennheiser© e845S. Les enregistrements sont directement stockés sur une carte CompactFlash© dans des fichiers .wav lors de la prise de son. Afin de sécuriser au maximum les fichiers, un archivage de ces derniers a été effectué sur un disque dur Sohoaid© SR2 (2 terabyte) de la marque Stardom, ainsi que sur plusieurs DVD.

Le sujet était assis confortablement à environ 20 cm du microphone. Il avait comme consignes de parler normalement et de ne pas trop bouger, gardant ainsi une position plus ou moins stable par rapport au microphone.

V.3. Corpus

La taille de notre corpus est volontairement réduite afin de répondre aux contraintes hospitalières et ne pas solliciter trop longtemps la voix de patients qui sont souvent en souffrance vocale mais également psychologique (*cf.* *supra*).

Notre corpus comprend des voyelles soutenues, des logatomes ainsi qu'un texte lu.

V.3.1. Voyelles

Le corpus correspond aux voyelles soutenues, /i, a, u/ qui permettent d'explorer l'espace vocalique maximal de chaque locuteur. Il s'agit pour le locuteur de prononcer et de tenir environ 5 secondes la voyelle présentée. Chaque voyelle est répétée à 10 reprises, en tenant compte des éventuelles difficultés du locuteur, notamment dans les phases d'enregistrement post-opératoire précoces.

Les voyelles sont présentées sur des cartons au patient sous leurs formes graphiques : -i- -a- -ou- pour que le patient soit le plus à l'aise possible avec la lecture du corpus. Ce choix permet aussi d'éviter de passer par une phase de familiarisation

avec les symboles phonétiques, ce qui aurait pu perturber le patient et par conséquent certaines de ses productions.

Les cartons avec les représentations graphiques des voyelles sont présentés de façon aléatoire aux locuteurs afin d'éviter tout phénomène de liste ou d'ancrage. Il s'agit pour le patient de « simplement » prononcer les voyelles qui lui sont présentées, sans se soucier de l'item à venir.

V.3.2. Logatomes – Séquences VCV

Les logatomes ont été créés comme suit : si V1 est /i/ alors V2 est /a/ et vice versa.

V1 = [i] et V2 = [a]

Ou V1 = [a] et V2 = [i]

La consonne est l'une des 6 occlusives du français /p, t, k, b, d, g/.

C = [p t k b d g]

Notre corpus est alors composé de douze logatomes.

[iba] vs. [abi]

[ida] vs. [adi]

[iga] vs. [agi]

[ipa] vs. [api]

[ita] vs. [ati]

[ika] vs. [aki]

Les logatomes sont insérés dans une phrase porteuse qui permet de proposer un cadre semi-naturel, et de placer ce même logatome dans une position neutre, c'est-à-dire ni sous l'influence de l'accent rythmique (qui frappe habituellement la dernière syllabe d'un groupe de mots ou d'un syntagme) ni sur un accent d'insistance quelconque. Cela donne une séquence du type :

Cet VCV ça

Ex. : Cet abi ça

Les logatomes ont présentés aux sujets de la même manière que les voyelles tenues. Ils étaient inscrits sur des cartons que l'expérimentateur a montrés au patient de

manière aléatoire pour éviter tout artefact expérimental éventuel. Le patient est alors en situation de lecture.

V.4. Locuteurs

La base de données complète, patients présentant ou non une immobilité laryngée, comprend trente locuteurs. Toutefois, la totalité des enregistrements n'a pas pu être exploitée pour des raisons indépendantes de notre volonté. Nous avons tenté de définir un certain nombre de critères (détaillés ci-après). Les productions des locuteurs qui ne les remplissaient pas se voyaient alors exclues de nos recherches. Leurs données sont toutefois conservées, afin d'être éventuellement exploitées dans un travail futur. Au final, quatorze patients (sept locuteurs présentant une immobilité laryngée et sept locuteurs ne présentant pas d'immobilité laryngée) ont pu être retenus.

Pour obtenir une cohorte de locuteurs comparables, nous avons tenté de contrôler certains paramètres.

Comme cela a été précédemment dit, tous les locuteurs retenus étaient francophones. Toutefois, la situation linguistique de l'Alsace fait que tous nos locuteurs ont appris l'alsacien et le français conjointement.

Les locuteurs retenus ne devaient également présenter aucune pathologie vocale avant leur opération chirurgicale. Ce paramètre était facilement contrôlable puisque tous les patients bénéficient d'un bilan ORL pré-opératoire.

Certains locuteurs n'ont pas pu être retenus parce qu'ils n'arrivaient pas à effectuer les tâches demandées. En effet, si les logatomes permettent de neutraliser l'apport sémantique du mot, ils rendent parfois leur prononciation et leur répétition difficiles. Enfin, notre corpus suppose que les locuteurs soient suffisamment à l'aise avec la lecture. Dans certains cas, étant donné le niveau de scolarisation assez faible, et en conséquence une maîtrise problématique de la lecture, nous avons été obligée d'écarter la production de certains locuteurs.

Tous les locuteurs ont été informés de l'objet de notre étude dans les termes généraux suivants : « une étude de votre voix après une opération de la thyroïde ». Tous les sujets ont signé une autorisation d'enregistrement.

V.4.1. Locuteurs présentant une immobilité laryngée

Notre démarche expérimentale se veut longitudinale. En ce qui concerne les locuteurs présentant une immobilité laryngée, nous n'avons retenu que les patients qui ont suivi une rééducation orthophonique durant deux mois au minimum (ce qui correspond à la phase d'enregistrement post-opératoire 4 dans la suite de ce travail).

En résumé, les locuteurs présentant une immobilité laryngée et retenus dans ce travail remplissent les critères suivants :

- Ne présenter aucune pathologie vocale avant l'opération chirurgicale ;
- Subir une thyroïdectomie ;
- Être francophone ;
- Être capable de prononcer le corpus ;
- Présenter une immobilité laryngée post-thyroïdectomie ;
- Suivre une rééducation orthophonique.

Compte tenu de ces critères, nous n'avons pu retenir que sept locuteurs : deux hommes et cinq femmes. Si cette cohorte est relativement réduite, rappelons toutefois que l'incidence de la paralysie laryngée unilatérale post-thyroïdectomie reste heureusement relativement faible (voir chapitre précédent).

Tableau 4 : Locuteurs présentant une immobilité laryngée retenus pour notre étude

Locuteurs	Sexe	Années de naissance	Geste chirurgical	Examen ORL
UPPBAS	Féminin	1947	T Totale	C droite en abduction
UPPHEI	Féminin	1946	T Totale	C droite immobile
UPPHUB	Féminin	1953	T Totale	C gauche immobile
UPPWAL	Féminin	1977	T Totale	C gauche immobile paramédiane
UPPWAN	Féminin	1949	T Totale	C gauche immobile
UPPLAT	Masculin	1944	T Totale	C gauche adduction
UPPPAI	Masculin	1945	T Totale	C gauche adduction

UPP = Unilateral Paralysis Patient ou patient présentant une paralysie unilatérale (Fauth et *al.*, 2011) ; T = Thyroïdectomie ; C = Corde vocale

V.4.2. Locuteurs ne présentant pas d'immobilité laryngée post-thyroïdectomie

Afin de pouvoir comparer les résultats obtenus avec ceux des locuteurs présentant une immobilité laryngée post-thyroïdectomie, nous avons fait le choix de retenir le même nombre de patients ne présentant pas d'immobilité laryngée post-thyroïdectomie. Ce groupe est donc également composé de sept locuteurs, deux hommes et cinq femmes.

En résumé, les locuteurs ne présentant pas d'immobilité laryngée et retenus dans ce travail remplissent les critères suivants :

- Ne présenter aucune pathologie vocale avant l'opération chirurgicale ;
- Subir une thyroïdectomie ;
- Être francophone ;
- Être capable de prononcer le corpus ;
- Ne pas présenter d'immobilité laryngée post-thyroïdectomie.

Tableau 5 : Locuteurs retenus pour notre étude et qui ne présentaient pas d'immobilité laryngée

Locuteurs	Sexe	Années de naissance	Geste chirurgical
NPPGER	Féminin	1947	T Totale
NPPHOE	Féminin	1953	T Totale
NPPHOF	Féminin	1945	T Totale
NPPKRE	Féminin	1961	T Totale
NPPLEN	Féminin	1948	T Totale
NPPENS	Masculin	1958	T Totale
NPPKAU	Masculin	1954	T Totale

NPP = No Paralysis Patient ou patient ne présentant pas d'immobilité laryngée (Fauth et *al.* 2011) ; T = Thyroïdectomie

V.4.3. Locuteurs Contrôles

Nous choisissons ici le terme de « locuteur contrôle » en opposition à celui de « locuteur pathologique ». Ces locuteurs sont qualifiés comme sains dans la mesure où, au niveau de nos analyses acoustiques, ce sont des locuteurs qui satisfont aux tâches requises de l'expérimentateur, à savoir pouvoir réaliser aussi clairs que possible les voyelles soutenues, les logatomes, ainsi que la lecture du texte. Ces locuteurs ne

présentent aucun antécédent pathologique, que ce soit au niveau de la production ou de la perception de la parole. Le niveau de la production a été particulièrement contrôlé, puisque ces locuteurs bénéficient d'un diagnostic ORL vierge.

En résumé, les locuteurs contrôles retenus dans ce travail remplissent les critères suivants :

- Ne présenter aucune pathologie de production et de perception de la parole ;
- Être francophone ;
- Être capable de prononcer la totalité du corpus.

Tableau 6 : Locuteurs contrôles retenus pour notre étude

Locuteurs Contrôles	Locuteurs Pathologiques	Sexe	Années de naissance
LCBAS	UPPBAS	Féminin	1944
LCHEI	UPPHEI	Féminin	1960
LCHUB	UPPHUB	Féminin	1947
LCWAL	UPPWAL	Féminin	1961
LCWAN	UPPWAN	Féminin	1948
LCLAT	UPPLAT	Masculin	1959
LCPAI	UPPPAI	Masculin	1957

LC = Locuteur Contrôle ; UPP = Unilateral Paralysis Patient (Fauth et *al.* 2011)

V.5. Enregistrements

En raison de contraintes indépendantes de notre volonté, les différents groupes de locuteurs n'ont pas pu être enregistrés de la même manière. Ces contraintes sont notamment liées : 1) à la rareté de locuteurs présentant une immobilité laryngée post-thyroïdectomie ; aux patients qui décident de ne plus collaborer pour l'acquisition de nos données ; à l'éloignement géographique du patient et ses capacités de déplacement ; à une organisation du service d'oncologie qui ne coïncide pas toujours avec notre propre calendrier.

V.5.1. Locuteurs présentant une immobilité laryngée post-thyroïdectomie

Les enregistrements des locuteurs présentant une immobilité laryngée post-thyroïdectomie commencent au moment de la consultation de contrôle post-opératoire (Post-op 2), c'est-à-dire quinze jours après l'intervention. Le diagnostic ORL révèle une immobilité laryngée. La voix est alors de légèrement à sévèrement altérée avec paralysie laryngée. Le patient est ensuite enregistré une fois par mois (Post-op 3, Post-op 4, etc.) pendant toute la durée de sa thérapie vocale afin, de mesurer une possible récupération.

V.5.2. Locuteurs ne présentant pas d'immobilité laryngée post-thyroïdectomie

Les locuteurs ne présentant pas d'immobilité laryngée post-thyroïdectomie ont été enregistrés à trois reprises. Un premier enregistrement a lieu la veille de l'intervention. Il s'agit d'un enregistrement préopératoire (phase Préop) qui constitue la voix de référence du patient. Le patient est à nouveau enregistré le lendemain de l'intervention (Post-op1), où la voix est de légèrement à sévèrement altérée. Enfin, le patient est enregistré lors de la consultation de contrôle post-opératoire (phase Post-op2), c'est-à-dire quinze jours après l'intervention. Le diagnostic ORL révèle une mobilité laryngée normale. Cette phase permet de mesurer une possible récupération vocale.

V.5.3. Locuteurs Contrôles

Les locuteurs contrôles ont été enregistrés à une seule reprise. Leurs enregistrements constituent les valeurs de « référence » pour les locuteurs présentant une paralysie laryngée uniquement. Il est vrai que ce principe de retenir un locuteur contrôle comporte inéluctablement une part de variabilité inhérente aux productions de ce locuteur. Le terme de « référence » est, en conséquence, à relativiser lorsque nous procéderons aux comparaisons entre locuteurs contrôles et patients.

V.6. Mesures – Analyses

Notre investigation expérimentale est fondée sur une analyse acoustique. Nous analysons le signal acoustique temporel, ainsi que le spectrogramme afin de rechercher les différentes perturbations entraînées par une opération de la glande thyroïde. Diverses mesures sont effectuées afin de mettre au jour et d'évaluer les différences entre les signaux des locuteurs contrôles et les signaux de nos différents groupes de locuteurs pathologiques.

Toutes les mesures ont été obtenues à l'aide de logiciel PRAAT[®] (Boersma, 2001). Nous avons utilisé un script proposé par Gendrot (voir Figure 5) qui nous a permis de relever de façon automatique les mesures de durées, de fréquence fondamentale, de F1 et de F2 et de Harmonics-to-Noise Ratio (HNR). Les mesures automatiques ont été contrôlées manuellement et corrigées si nécessaire. De façon générale, les mesures effectuées sur les locuteurs sains ne posaient pas de problème particulier, alors que les mesures obtenues à partir des locuteurs pathologiques ont parfois dû être corrigées manuellement.

Figure 5 : Script de Gendrot utilisé pour nos mesures

```

outputfile$ = "path2$" + "\" + "resultats_consonnes.txt"
#fileappend "outputfile$" fic      label sex  tier2  hnr  cog
skewness    kurtosis      sdev  f0    place_p  manner_p
voice_p     place_fmanner_f  voice_f      duree  duree_norm
fo_norm     glissando      zcr   mot   mot_phon  place_syll
pause_prec1 pause_prec2  pause_suiv1  pause_suiv2  intensite      gramm
entour_prec voy_prec    entour_foll  voy_foll    position_in_syllable
position_phoneme_sur_total_phonemes 'newline$'
fileappend "outputfile$" fic label sex  tier2  hnr  cog  zcr
skewness    kurtosis      sdev  beg_f0 mid_f0 end_f0 place_p
manner_p    voice_p      place_fmanner_f  voice_f      duree
duree_norm  fo_norm      tier3  beg_intensite mid_intensite end_intensite
intensite_rms beg_F1      mid_F1      end_F1      beg_F2

```

```

mid_F2      end_F2      beg_F3      mid_F3      end_F3
beg_F4      mid_F4      end_F4 'newline$'
Create Strings as file list... liste 'path$\type$'
string = selected("Strings")
noliftt = Get number of strings
for x from 1 to noliftt
    select 'string'
    line$ = Get string... x
    line_depart$ = line$
    file$ = path$ + "\" + line$
    Open long sound file... 'file$'
    sound = selected("LongSound")
extension$ = right$(file$,3)
grille$ = path$ + "\" + line$ - extension$ + "TextGrid"
if fileReadable(grille$)
    Read from file... 'grille$'
    textgrid = selected("TextGrid")
call procedures
endif
endfor
#select all
#Remove
printline Vos résultats ont été enregistrés sous 'path2$'

```

V.6.1. Voyelles Soutenues

Les voyelles soutenues permettent de mesurer la qualité vocalique et de faire des mesures sur les formants de voyelles pour comparer les triangles vocaliques des sujets sains et des sujets pathologiques. Une attention particulière a été portée aux mesures de perturbations du signal comme le Harmonics-to-Noise Ratio, le jitter et le shimmer. Pour chaque voyelle, les valeurs formantiques ont été relevées au milieu de la voyelle, délimitée par une structure formantique stable. Pour ce qui concerne le jitter et le

shimmer, précisons que le calcul est effectué sur la totalité de la structure formantique stable.

Afin d'étudier plus en détail la variabilité inter et intra-individuelle des paramètres susmentionnés, nous observerons les moyennes calculées à partir des répétitions de chaque locuteur uniquement pour la voyelle [a]. Nous faisons ce choix en suivant les recommandations de la Société Européenne de Laryngologie (Dejonckere et *al.* 2001 et Dejonckere et *al.* 2003). Cette voyelle est souvent utilisée comme support pour les évaluer la dysphonie (Dejonckere et *al.*, 2003) notamment pour mesurer l'importance des perturbations vocales. En effet, les voyelles [i] et [a] sont plus difficiles à réaliser puisque, par rapport au [a], elles exigent toutes les deux une élévation linguale pour réaliser, respectivement, une constriction palatale et vélaire de taille réduite. L'ajout d'un tel contrôle spatial précis pourrait représenter une tâche supplémentaire pour le patient qui serait déjà face à un problème de gestion correcte du geste glottique.

V.6.1.1. Fréquence Fondamentale

La mesure de la fréquence fondamentale a été obtenue par détection automatique à l'aide du logiciel PRAAT[®]. Pour plus de finesse dans nos analyses, nous avons modifié les paramètres programmés par défaut dans le menu *Pitch Settings*. La fréquence fondamentale a été détectée entre 20 et 350 Hz. La méthode d'analyse retenue a été celle de la *cross-correlation* comme recommandée pour une analyse optimisée de la voix. Les paramètres du spectrogramme ont également été adaptés dans le menu *View Range* pour une détection de fréquence fondamentale entre 20 et 350 Hz, ce qui nous a permis de vérifier la réalité de la courbe.

V.6.1.2. Harmonics-to-Noise Ratio (HNR)

La mesure de HNR est considérée comme un indice de raucité, c'est-à-dire qu'elle renseigne sur le rapport bruit/harmoniques dans une voyelle. Cette mesure considère l'élément vocalique comme une entité qui serait composée d'un constituant périodique ou harmonique et d'un composant de bruit (Yumoto, et *al.* 1982). L'harmonicité (ou H) est l'énergie périodique du fondamental, alors que le bruit (ou N) est le composant aperiodique. Il s'agit alors de quantifier la proportion de chacun des éléments.

Cette valeur (quoique correspondant à un ratio) est habituellement exprimée en décibels. Traditionnellement, plus le ratio baisse, plus le signal est envahi par le bruit. Ce ratio est généralement mesuré autour de 20 dB pour /a/ et /i/ et de 30dB pour /u/, chez un locuteur non pathologique. Un ratio égal à 0 dB indique qu'il y a autant d'énergie dans les harmoniques que dans le bruit (Boersma, 1993).

La mesure de HNR a été obtenue à partir de l'équation logarithmique proposée par Boersma (1993).

$$\text{HNR (dB)} = 10 \cdot 10 \log \frac{r'_x(\tau_{max})}{1 - r'_x(\tau_{max})}$$

Équation 1 : équation logarithmique proposée par Boersma (1993)

V.6.1.3. Mesures formantiques - F1 et F2

De façon à obtenir des mesures fines, nous nous sommes appuyée sur les recommandations de Wood (1994) proposées sur son site internet, et avons donc modifié les différents paramètres programmés (dans le menu *Formant* et *Formant settings*) par défaut dans le logiciel PRAAT[®].

Le nombre de formants (*Number of formants*) a été monté à 5 pour un meilleur suivi, ce qui nous a également conduit à augmenter l'étendue du spectre (*Maximum formant*) à 5000 Hz, pour un locuteur masculin, et 5500 Hz pour un locuteur féminin. En effet, un formant est approximativement détecté dans chaque bande de 1000 Hz environ.

Enfin, la taille de la fenêtre (*Window length*), qui permet de spécifier la fenêtre de signal utilisée, et donc en quelque sorte la résolution temporelle pour mesurer chaque point, a été augmentée à 0,04 secondes. Une fenêtre trop petite pourrait en effet donner des valeurs formantiques fausses.

Les valeurs de F1 et de F2 ont été relevées automatiquement, puis manuellement au milieu de la structure formantique stable de la voyelle. Les valeurs ainsi acquises représentent les résonnances dans le conduit vocal, permettant d'inférer plus ou moins la configuration du conduit vocal, lors de la production des voyelles étudiées. La configuration du conduit vocale est alors inférée sur un graphique mettant en corrélation les valeurs de F2 en abscisse, et de F1 en ordonnée. Notons, toutefois, que cette relation entre l'espace acoustique et l'espace articulaire est loin d'être linéaire, étant donné la

nature mal-posée (*ill posed*) de cette relation, dans le sens mathématique du terme, entre les deux domaines, articulatoire et acoustique. En effet, nous savons, grâce aux recherches préoccupées par la notion d'inversion en parole, qu'il peut exister plusieurs solutions à un problème unique.

L'espace vocalique maximal (kHz²) a été calculé à partir de la formule de Héron :

$$\sqrt{p(p-a)(p-b)(p-c)}$$

$$\text{où: } p = \frac{1}{2}(a+b+c)$$

et a, b et c correspondent respectivement à la longueur des trois côtés du triangle.

$$a = \sqrt{(x_c - x_b)^2 + (y_c - y_b)^2}$$

$$b = \sqrt{(x_a - x_c)^2 + (y_a - y_c)^2}$$

$$c = \sqrt{(x_a - x_b)^2 + (y_a - y_b)^2}$$

Équation 2 : Formule de Héron

Cette valeur (en kHz²) fournit des informations sur l'espace vocalique qui pourrait permettre de distinguer les différentes configurations de cet espace suivant les locuteurs, pathologiques et sains.

V.6.1.4. Jitter

Le jitter est une mesure des perturbations à court terme de la fréquence fondamentale. Cette mesure, exprimée en pourcentage, pose de nombreux problèmes, puisque sa valeur dépend notamment de la technique de mesure de la fréquence fondamentale (Teston, 2001).

De nombreux auteurs ont travaillé sur le calcul de ce paramètre (voir Boek, et al.1997, Schoentgen, 2001, 2006 ou Schoentgen & Grenez, 2011)) qui dépend à la fois du logiciel utilisé mais également de la méthode de calcul retenue. Dans la mesure où nous travaillons à partir du logiciel PRAAT[®], nous avons choisi d'utiliser la méthode de calcul proposé par Boersma et al. (1993).

Le jitter local ou jitter ratio correspond à la moyenne de toutes les différences, en valeur absolue, entre les durées de deux périodes consécutives du signal. Cette moyenne est divisée par la durée moyenne d'une période du signal. Le manuel PRAAT[®] indique que le seuil normal/pathologique est de 1,04%.

Le jitter RAP (*Relative Average Perturbation*) propose de prendre en compte les variations volontaires de la fréquence fondamentale et de les atténuer en comparant la durée de chaque période (soit Tx) à la moyenne de trois périodes successives (soit Tx - 1, Tx et Tx + 1). Le seuil normal/pathologique est fixé à 0,68% pour ce critère.

V.6.1.5. Shimmer

Le shimmer est une mesure des perturbations à court terme de l'amplitude du signal sonore. Son calcul pose les mêmes problèmes que celui de jitter, c'est pourquoi nous adopterons la même méthode.

Pour obtenir la mesure de shimmer local, il s'agit de diviser la moyenne des différences, en valeur absolue, entre l'amplitude de deux périodes successives, par la moyenne des amplitudes maximales de chaque période. Le seuil normal/pathologique est de 3,81%.

Le shimmer APQ (*Amplitude Perturbation Quotient*) est obtenu sur le même principe que le jitter RAP. Il s'agit en effet de prendre en considération les modulations volontaires d'intensité. Pour ce faire, l'amplitude de chaque période (soit Tx) est comparée à l'amplitude des pics de onze périodes successives (soit Tx -5, Tx et Tx +5). Le seuil normal/pathologique est alors fixé à 3,07%.

V.6.2. Logatomes

V.6.2.1. Paramètres

Selon Abry et al. (1985), différents événements articulatoire-acoustiques peuvent être observés à partir d'un signal acoustique continu, et permettent de remonter aux configurations articulatoires. Ces événements, qui sont constitués en couples, ont été décrits par Sock (1998) :

V.6.2.1.1. Voice Onset Time (VOT)

Lisker & Ambramson (1964) ont défini le VOT comme l'intervalle allant du relâchement consonantique aux premières vibrations correspondant au début du voisement (VO). Klatt, (1975) délimite le VOT dans l'intervalle allant du relâchement consonantique jusqu'à l'apparition de la structure formantique stable de la voyelle (VVO).

Liberman et *al.* (1958) notent le VOT comme un indice majeur de la classification perceptive des occlusives voisées/non voisées. En d'autres termes, si ce délai est positif, l'occlusive sera perçue comme non voisée, en revanche si ce délai est négatif ou nul, l'occlusive sera perçue comme voisée.

Le VOT, comme paramètre de transition entre la consonne et la voyelle, serait un bon candidat pour l'étude de catégories phonétiques de lieux d'articulation consonantiques et du contexte vocalique post-consonantique.

V.6.2.1.2. Voice Termination Time (VTT)

Agnello (1975) définit le VTT comme l'intervalle allant de la dernière vibration glottique à la closure articulaire ou la pression intra-orale maximale. Il s'agit donc d'une période de transition qui conduit à la fermeture complète du canal buccal.

Certains auteurs (voir Sock, 1983, Abry et *al.* 1985 et Van den Berg, 1988, par ex.) considèrent cette mesure comme le temps nécessaire à l'arrêt de vibration des cordes vocales après la closure dans le conduit vocal. Par extrapolation, il serait alors possible de comprendre le VTT comme le temps nécessaire à la reconstruction de la pression d'air subglottique, indispensable à l'arrêt du voisement.

Le VTT aurait un lien étroit avec le mouvement d'abduction des cordes vocales et avec le niveau de la pression d'air subglottique. D'après certaines études, ce paramètre varierait selon le lieu d'articulation de la consonne et de la voyelle (voir Keating, 1984 ou Docherty, 1992, par ex.).

V.6.2.2. Mesures

Pour chaque logatome VCV, nous avons mesuré les durées de :

- La première voyelle (V1)
- La consonne intervocalique (C)

- La seconde voyelle (V2)

Les voyelles [i] vs [a] ont été volontairement choisies car leur réalisation permet d’observer, en contexte VCV, la traversée du larynx d’une position basse pour le [i] à une position haute pour le [a] et *vice versa*. Signalons que les données attestées dans la littérature (cf. par ex Vaxelaire et Sock, 1997 ; Vaxelaire, 2007) relatives à la position du larynx pour ces voyelles, dans ce contexte consonantique, ne corroborent pas le *Tongue-Pull Hypothesis* (Ladefoged, 1960 et Ohala, 1983 par ex.). Nous y reviendrons (cf. section XXX) lors de l’analyse de nos propres données.

Ces voyelles permettent en outre d’observer la dynamique maximale en termes de déplacements mandibulaire et lingual.

La durée absolue de la voyelle a été mesurée entre le VVO et le VVT c’est-à-dire entre le début et la fin de la structure formantique clairement définie.

Les valeurs obtenues ont été normalisées en calculant le pourcentage pris par la réalisation vocalique dans l’intervalle VC et CV. Il nous semble important de prendre en considération les valeurs normalisées en production de la parole, point qui sera développé *infra*.

Pour chaque consonne, les durées suivantes ont été quantifiées :

- le VTT (pour les consonnes non voisées) ;
- le silence acoustique (pour les consonnes non voisées) ;
- le VOT ;
- la tenue consonantique.

Les consonnes [p t k b d g] sont choisies car elles permettent d’observer l’effet du recul du lieu d’articulation (de l’avant vers l’arrière de la cavité buccale) sur la pression intra-orale en liaison avec l’activité laryngienne (mesure du VOT).

De même, ces consonnes permettent l’alternance de séquences entièrement sonores V1 [b d g] V2 et de séquences mixtes V1 [p t k] V2. Les plis vocaux ne sont donc pas sollicités de façon continue.

Toutes les séquences mesurées ont été analysées en valeurs absolues et en valeurs relatives. Il nous semble important de prendre en considération ces deux types de mesures. En effet, si les valeurs absolues permettent de quantifier l’étendue

temporelle d'un paramètre, les valeurs relatives nous renseignent sur le maintien ou non des rapports entre ces mêmes paramètres, dans des conditions différentes et, par conséquent, d'apprécier l'élasticité du signal de la parole (Gaitenby, 1965 ; Sock, 1998).

V.7. Analyses Statistiques – Remarques générales

Des analyses statistiques ont été conduites pour chacune de nos expériences.

Signalons qu'en déterminant la signification statistique des *effets principaux*, seuls les résultats significatifs avec une probabilité de moins de cinq pour cent d'avoir été obtenus par chance ($p < 0.05$) ont été retenus (pour plus de détails chiffrés sur les résultats obtenus de l'analyse des effets principaux, nous renvoyons aux différentes expériences). Des indications plus précises sur les différences significatives seront dérivées des oppositions *a posteriori* entre les moyennes de groupes, paire-par-paire.

CONCLUSION CHAPITRE V

Ce chapitre nous a permis de définir le protocole expérimental de cette thèse.

Tous les locuteurs pathologiques ont subi une thyroïdectomie au Centre Paul Strauss de Strasbourg. Une fiche signalétique permettant de contrôler les conditions d'inclusion dans le protocole expérimental a été remplie par chaque locuteur. Quatorze patients ont pu être retenus dans le cadre de cette recherche : sept patients ne présentant pas d'immobilité laryngée post-thyroïdectomie et sept patients présentant une immobilité laryngée post-thyroïdectomie.

Les sept locuteurs contrôles ne présentaient aucune pathologie vocale, ce qui a été vérifié par un examen ORL. Ils ont été appariés en genre et en âge avec les locuteurs du groupe présentant une immobilité laryngée post-thyroïdectomie.

Les conséquences opératoires sur la voix des patients ont été évaluées à l'aide d'un corpus composé de voyelles soutenues et de logatomes.

Les voyelles soutenues ont fait l'objet de la mesure de leur structure formantique (F1 et F2 et de l'aire de l'espace vocalique), de leur fréquence fondamentale, du Harmonics to Noise Ratio, du jitter et du shimmer.

Les logatomes ont permis les mesures temporelles suivantes : durées vocaliques et durées consonantiques. Une attention particulière sera accordée aux mesures du VOT et du VTT qui renseignent notamment sur l'activité laryngienne.

La fatigue vocale a pu être évaluée par la lecture de la *Chèvre de Monsieur Seguin*.

Des analyses statistiques ont été conduites pour chacune des expériences afin de vérifier la significativité des mesures prises.

PARTIE 3 :

RESULTATS

EXPERIMENTAUX

Cette partie expose les résultats de nos propres expériences dans le domaine acoustique. Trois chapitres composent cette partie.

Chapitre VI : Expérience 1 : Etude acoustique de voyelles soutenues produites par des patients ne présentant pas de paralysie récurrentielle après thyroïdectomie

Chapitre VII : Expérience 2 : Etude acoustique de voyelles soutenues produites par des patients présentant une paralysie récurrentielle après thyroïdectomie

Chapitre VIII : Expérience 3 : Etude spatio-temporelle de la voix de patients présentant une paralysie récurrentielle après thyroïdectomie

RÉSUMÉ CHAPITRE VI

Ce chapitre sera consacré à l'analyse des altérations vocales post-thyroïdectomie sans immobilité laryngée. Nous avons vu précédemment qu'une ablation de la glande thyroïde peut conduire à des modifications vocales alors même que les nerfs récurrents sont préservés.

Pour ce faire, nous avons enregistré à trois reprises (préopératoire, post-opératoire 1 et post-opératoire 2) sept patients (cinq femmes et deux hommes) ayant subi une thyroïdectomie et pour lesquels le diagnostic ORL n'a pas permis de déceler une immobilité laryngée consécutive à l'opération chirurgicale.

Notre analyse se fonde sur l'analyse d'un corpus composé de voyelles soutenues. Cette expérience se focalise sur les mesures spectrales telles que la fréquence fondamentale et les mesures formantiques. Une attention particulière sera accordée aux mesures de perturbations du signal acoustique, soit le Harmonics-to-Noise Ratio, le jitter et le shimmer.

Les résultats indiquent une modification de tous les paramètres étudiés lors de l'enregistrement post-opératoire 1, et un retour à des mesures comparables à la phase préopératoire en post-opératoire 2.

CHAPITRE VI

VI. Expérience 1 : Etude acoustique de voyelles soutenues produites par des patients ne présentant pas de paralysie récurrentielle après thyroïdectomie

VI.1. Introduction

Le présent travail est une étude acoustique de quelques caractéristiques spectrales de voyelles soutenues produites par des patients pour lesquels le diagnostic ORL n'a pas révélé d'immobilité laryngée mais dont la voix est altérée après une opération de la glande thyroïde.

Les conséquences de l'opération chirurgicale sont évaluées dans le but d'identifier les différentes perturbations que cette opération peut provoquer, mais également afin de mettre au jour les possibles stratégies de compensations et/ou de réajustements que le patient peut mettre en place seul, étant donné que ces patients ne suivent pas de rééducation orthophonique.

Notre étude se veut longitudinale puisque les patients sont enregistrés lors de différentes phases post-opératoires.

Rappelons qu'une opération au niveau du larynx est rarement anodine (*cf.* Chapitre IV.6.1), et que des modifications de la périodicité de la voix notamment peuvent intervenir après une opération de la thyroïde.

VI.2. Procédure expérimentale

VI.2.1. Locuteurs

Pour cette expérience, la voix de sept patients a été enregistrée au Centre Paul Strauss de Strasbourg. Il s'agit de cinq femmes (NPPGER, NPPHOE, NPPHOF, NPPKRE et NPPLN) et deux hommes (NPPENS et NPPKAU) ayant subi une thyroïdectomie totale et pour lesquels le diagnostic ORL post-opératoire n'a pas révélé d'immobilité laryngée. Nous (Fauth et *al.* 2011) avons appelé ce groupe de patients le groupe NPP (No Paralysis Patients ou patient ne présentant pas d'immobilité laryngée), dans la suite de ce travail.

Les locuteurs ont été enregistrés la veille de l'intervention (phase préopératoire), le lendemain de l'intervention (phase post-opératoire 1) et quinze jours après l'intervention (post-opératoire 2).

VI.2.2. Corpus

Notre corpus est constitué des trois voyelles /i a u/ qui ont été répétées chacune à dix reprises dans la mesure des possibilités des locuteurs. La consigne donnée aux différents patients était de produire et de soutenir durant 5 secondes environ les voyelles susmentionnées.

VI.2.3. Mesures

Le signal acoustique et le spectrogramme correspondant ont été segmentés conjointement. Pour chaque voyelle, les paramètres spectraux suivants ont été mesurés dans la structure formantique stable :

- la fréquence fondamentale ;
- les mesures formantiques (F1 et F2) ;
- les mesures de perturbations du signal, à savoir le Harmonics-to-Noise Ratio (HNR), le jitter et le shimmer.

VI.2.4. Analyses des données

Les résultats seront présentés sous forme de graphiques et de tableaux commentés, afin de faciliter leur lecture et de rendre plus aisée la comparaison entre les

différents locuteurs. Par ailleurs, ce type de présentation devrait également permettre de dégager des tendances.

Des analyses de variance (ANOVA) à mesures répétées ont été conduites pour toutes nos données (*cf. infra* pour plus de détails). Les ANOVA révèlent les différences significatives entre les groupes, mais ne stipulent pas ce sur quoi portent ces différences. Par conséquent, des tests *a posteriori* ou *post-hoc* ont été pratiqués. Ces tests indiquent entre quels groupes se situent la ou les différences significatives. Nous avons utilisé des tests *post-hoc* LSD (*“Least Significant Difference”*).

Les valeurs présentées sont toujours des moyennes calculées à partir des répétitions produites par nos locuteurs. Les valeurs chiffrées entre parenthèses renvoient aux écarts-types.

VI.3. Hypothèses

Etant donné qu’une opération au niveau de la glande thyroïde peut entraîner des modifications de la voix en phases post-opératoires, nous formulons les hypothèses suivantes :

1. La modification de la voix affecterait directement les valeurs de fréquence fondamentale, ou F0, dans la mesure où ce paramètre est lié à l’activité laryngienne ;
2. Dans le cas des locuteurs ayant subi une opération de la thyroïde n’entraînant pas de paralysie récurrentielle, la voix se trouve souvent altérée. Cette altération peut être due à une activité irrégulière du larynx qui pourrait avoir des conséquences sur les valeurs du Harmonics-to-Noise Ratio. Rappelons que ces valeurs sont généralement mesurées autour de 20 dB pour /i/ et /a/ et autour de 30 dB pour /u/, pour un locuteur non pathologique ;
3. Etant donné que les modifications de la fréquence fondamentale pourraient également se traduire par des modifications de la périodicité de la voix, les valeurs de jitter et de shimmer pourraient alors se trouver modifiées ;

4. Etant donné que la fréquence fondamentale risque d'être sensiblement perturbée par l'intervention chirurgicale, il est fort probable que les valeurs de F1 et de F2 soient déviantes des valeurs observées en phase préopératoire.
5. La perturbation de la source laryngée ayant des conséquences sur le couplage entre l'activité laryngienne et les cavités supraglottiques résonantes cela devrait, engendrer une modification de la taille et de la forme de l'espace vocalique.
6. La variabilité inter et intra locuteurs, traduite par des écarts-types plus importants, devrait être plus marquée dans les phases d'enregistrement post-opératoires et plus particulièrement en post-opératoire 1 à cause d'une coordination oro-laryngée problématique. Cette variabilité devrait se réduire dès la phase post-opératoire 2, c'est-à-dire dès le démarrage de la récupération des gestes glottiques.
7. Le temps devrait permettre une récupération vocale et une « standardisation » des paramètres précédemment évoqués ; les paramètres seraient alors comparables aux valeurs mesurées en phase préopératoire.

VI.4. Résultats

VI.4.1. Remarques Générales

Il semble utile de rappeler que les résultats présentés dans cette étude ont été obtenus à partir de :

- trois voyelles ;
- répétées à 30 reprises (soit 10 répétitions par voyelle) ;
- l'ensemble de ces répétitions a été prononcé par les 7 locuteurs ;
- lors de trois phases d'enregistrement (en préopératoire, en post-opératoire 1, puis en post-opératoire 2).

Soit 630 répétitions au total.

Des analyses de variance (ANOVA à mesures répétées) à deux facteurs (*temps* ou *phases d'enregistrement* et *voyelle* ou *qualité vocalique*) ont été effectuées pour les variables F0, F1, F2, HNR afin de savoir s'il existait des effets de *temps* ou de *voyelle*.

Par ailleurs, des analyses de variance (ANOVA à mesures répétées) à un facteur (*temps*) ont été faites pour les variables jitter, shimmer et l'aire de l'espace vocalique, afin de déterminer s'il existait des effets de *phases d'enregistrement (temps)* principalement. Par extrapolation, dans la mesure où la phase d'enregistrement préopératoire constitue les valeurs de référence attendues pour un locuteur contrôle, une différence significative entre la phase d'enregistrement préopératoire et les phases d'enregistrement post-opératoire pourrait être considérée comme un effet de groupe : locuteurs contrôles vs. patients. Cette démarche ne s'applique, bien entendu, que dans cette condition expérimentale de post-thyroïdectomie sans paralysie laryngée.

VI.4.1. Fréquence fondamentale (F0)

L'effet principal de *phases d'enregistrement* (soit *temps*) est significatif pour la variable F0 [$F(2,8) = 5,87$, $p < 0.028$] pour les locutrices du groupe NPP. Ces résultats révèlent que, tous contextes confondus, la fréquence fondamentale est modifiée en fonction de la phase d'enregistrement. De façon générale, ces résultats laissent supposer qu'une baisse de la fréquence fondamentale est attendue dans les phases post-opératoires. Notons que l'interaction des facteurs *voyelles* et *temps* ne s'est pas révélée significative ($p = ns$), ce qui indique que quelle que soit la voyelle étudiée, avec le temps les valeurs de fréquence fondamentale seront modifiées.

Les tests *post-hoc* LSD révèle que l'effet principal phase d'enregistrement repose sur une différence significative entre la phase d'enregistrement préopératoire (206 Hz) et les phases d'enregistrement post-opératoire 1 (181 Hz) et post-opératoire 2 (185 Hz).

Signalons que l'analyse fine de la fréquence fondamentale de nos locuteurs masculins a été, bien évidemment, effectuée indépendamment de celle de nos locuteurs féminins, afin de ne pas fausser nos mesures statistiques.

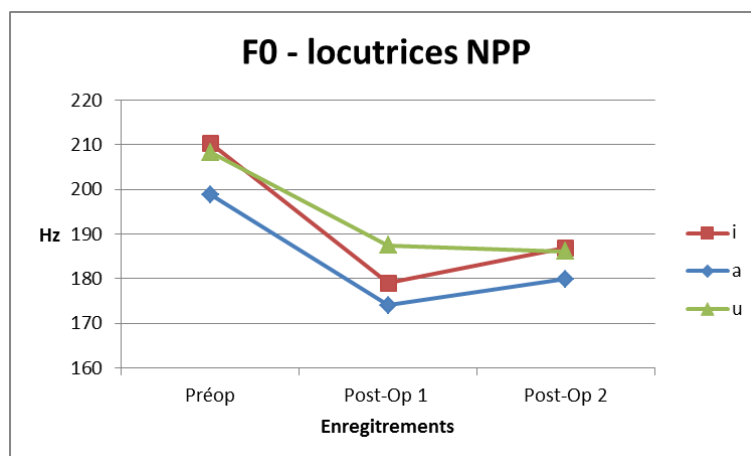


Figure 6 : Valeurs moyennes de la fréquence fondamentale (F0) en Hz pour [i a u] - locutrices du groupe NPP

Directement après l'opération, c'est-à-dire lors de la phase d'enregistrement post-opératoire 1, la tendance est à un abaissement de la fréquence fondamentale pour les locutrices du groupe NPP, comme l'illustre la

Figure 6. La fréquence fondamentale moyenne des locutrices est mesurée plus basse au lendemain de l'opération, et ce phénomène est observable quelle que soit la voyelle produite. Cette tendance est particulièrement marquée pour la voyelle [i], puisqu'elle était mesurée en phase préopératoire à 210 Hz (18,87 Hz) et qu'elle n'est plus qu'à 179 Hz (36,81) au stade post-opératoire 1, soit une diminution de 31 Hz, entre ces deux phases d'enregistrement. La baisse de la fréquence fondamentale est moins importante, d'environ 20 Hz, pour les voyelles [a] et [u] qui sont mesurées respectivement à 198 Hz (19,28 Hz) et à 208 Hz (23,96 Hz) en préopératoire et à 174 Hz (32,7 Hz) et à 187 Hz (32,07 Hz), respectivement en post-opératoire 1.

En ce qui concerne la phase d'enregistrement post-opératoire 2, les fréquences fondamentales moyennes, calculées à partir des trois voyelles, sont toujours plus basses que celles observées en préopératoire. Toutefois, on peut noter que pour les voyelles [a] et [i], la fréquence fondamentale moyenne augmente entre les deux phases d'enregistrement post-opératoires, alors mesurées à 180 Hz et 187 Hz environ, respectivement. Pour la voyelle [u] en revanche, la moyenne est comparable entre ces deux phases, soit environ 187 Hz.

Les phases d'enregistrement post-opératoires sont également marquées par des écarts-types plus importants (voir Tableau 7), raison pour laquelle il semble important

d'étudier les locuteurs individuellement afin de comprendre si cette plus grande variabilité est le résultat de stratégies inter ou intra individuelles.

Tableau 7 : Valeurs moyennes de la fréquence fondamentales (F0) en Hz pour [i a u] pour les locutrices du groupe NPP. Les écarts-types sont donnés entre parenthèses

NPPfem	Préop	Post-Op 1	Post-Op 2
i	210,42 (18,87)	179 (36,81)	186,96 (28,24)
a	198,9 (19,28)	174,1 (32,7)	179,98 (25,36)
u	208,34 (23,96)	187,48 (32,07)	186,1 (31,53)

Les tendances observées pour les locutrices se confirment également pour les sujets masculins, comme en atteste la figure 2.

Pour ces sujets également, la phase d'enregistrement post-opératoire 1 est caractérisée par des valeurs moyennes d'une fréquence fondamentale plus basses, quelle que soit la voyelle étudiée par rapport au stade préopératoire. La voyelle dont la fréquence fondamentale diminue le plus après l'opération est le [u], étant donné que le F0 a été mesuré à 144 Hz (2,6 Hz) en préopératoire et à 108 Hz (3,1 Hz) en post-opératoire 1. La fréquence fondamentale du [a] baisse également en passant de 127 Hz (2,17 Hz) en préopératoire à 103 Hz (2,35 Hz) en post-opératoire 1. Cette tendance est également observée pour le [i], voyelle pour laquelle le F0 était à 140 Hz (1,87 Hz) avant l'opération, puis à 112 Hz (3,74 Hz) le jour suivant cette même opération.

Tout comme pour les locutrices, les valeurs de F0 des trois voyelles produites par les sujets masculins augmentent dans la phase d'enregistrement post-opératoire 2. Cette augmentation n'atteint toutefois pas les valeurs standards mesurées en préopératoire pour le [i] et le [u], puisque le F0 est de 116 Hz (3,62 Hz) en postopératoire 2 pour la voyelle antérieure (140 Hz en préopératoire), et de 120 Hz (2,83 Hz) pour la voyelle postérieure (144 Hz en préopératoire) dans cette même phase d'enregistrement. Signalons que le [u] est la voyelle qui présente la baisse de F0 la plus remarquable en post-opératoire 1, mais dont l'augmentation de la fréquence fondamentale était la plus modérée en post-opératoire 2, parmi les trois voyelles testées.

La fréquence fondamentale moyenne de la voyelle [a] est mesurée plus haute en post-opératoire 2, où elle a été quantifiée à 135 Hz (26,75 Hz), par rapport à la phase

préopératoire, où elle était de 127 Hz (2,17 Hz). Il convient toutefois de signaler que la variabilité est relativement importante pour le [a] (écart-type : 26,75 Hz) (voir Tableau 8). Par conséquent, il serait intéressant d'étudier ces mesures de façon individuelle afin de comprendre si cette variabilité est imputable à des manœuvres inter ou intra individuelles.

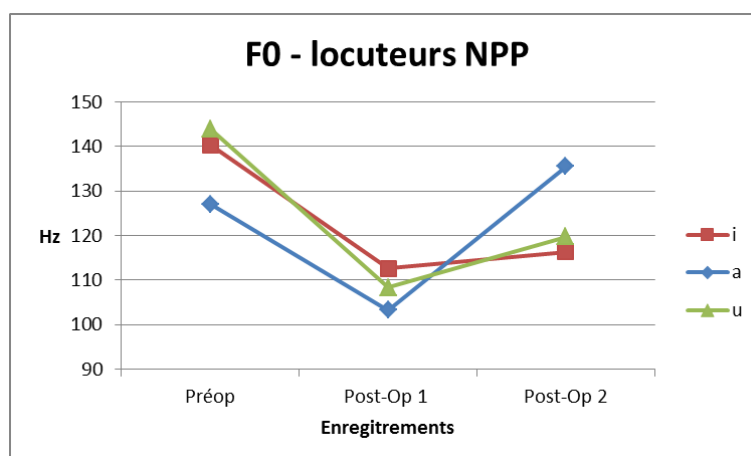


Figure 7 : Valeurs moyennes de la fréquence fondamentale (F0) en Hz pour [i a u] - locuteurs du groupe NPP

Tableau 8 : Valeurs moyennes de la fréquence fondamentales (F0) en Hz pour [i a u] pour les locuteurs du groupe NPP. Les écarts-types sont donnés entre parenthèses

NPPHom	Préop	Post-Op 1	Post-Op 2
i	140,35 (1,87)	112,7 (3,74)	116,4 (3,62)
a	127,1 (2,17)	103,3 (2,35)	135,5 (26,75)
u	144 (2,66)	108,4 (3,20)	119,75 (2,83)

Afin d'étudier plus en détail la variabilité inter et intra-individuelle de la fréquence fondamentale, déjà esquissée précédemment, nous observerons ci-après les moyennes calculées à partir des répétitions de chaque locuteur uniquement pour la voyelle [a]. Nous faisons ce choix (voir également V.6.1. en suivant les recommandations de la Société Européenne de Laryngologie (Dejonckere et *al.* 2001 et 2003).

La fréquence fondamentale, mesurée pour la voyelle [a], diminue pour tous les locuteurs du groupe NPP (excepté pour la locutrice NPPKRE en orange), entre la phase d'enregistrement préopératoire et la phase d'enregistrement post-opératoire 1 (voir Figure 8). Pour trois locuteurs en particulier, NPPHOF (en violet), NPPHOE (en vert) et NPPENS (en bleu foncé), la fréquence fondamentale mesurée en post-opératoire 1 est particulièrement dégradée. Il convient donc déjà de signaler qu'une certaine variabilité inter-locuteurs est à prendre en considération. En effet, si la fréquence fondamentale de six de nos sept locuteurs diminue entre les phases d'enregistrement préopératoire et post-opératoire 1, l'importance de la diminution dépend largement du locuteur étudié. Pour exemple, nous pouvons considérer le cas de la locutrice NPPHOE (en vert) pour laquelle la fréquence fondamentale est mesurée à environ 50 Hz plus basse en post-opératoire 1, alors que pour le locuteur NPPKAU (en bleu ciel) la baisse de la fréquence fondamentale n'est que de 10 Hz entre la phase préopératoire et la phase post-opératoire 2 (soit une baisse non significative).

La phase post-opératoire 2 renforce l'idée de variabilité inter-locuteurs déjà entamée à partir des mesures observées en post-opératoire 1. En effet, lors de cette dernière phase d'enregistrement, des comportements différents sont observables en fonction des locuteurs. Pour certains locuteurs (NPPHOF en violet, NPPKRE en orange ou NPPKAU en turquoise), la fréquence fondamentale mesurée en post-opératoire 2 est comparable à celle mesurée en préopératoire. Pour NPPLN (en bleu clair), NPPGER (en rouge), NPPENS (en bleu foncé) ou NPPHOE (en vert), les mesures prises en post-opératoire 2 n'indiquent pas d'amélioration significative, par rapport aux mesures relevées en post-opératoire 1. La fréquence fondamentale de ces deux locutrices est toujours mesurée plus basse que celle mesurée en préopératoire.

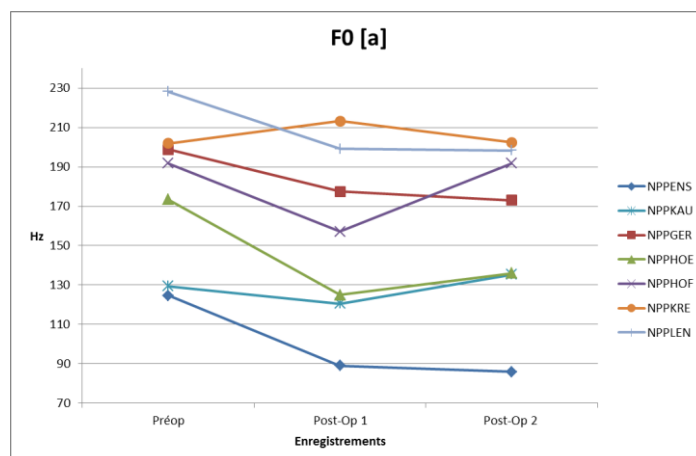


Figure 8 : Valeurs moyennes de fréquence fondamentale (F0) en fonction du locuteur pour la voyelle [a]

Le Tableau 9 est intéressant pour comprendre la variabilité intra-individuelle. La phase d'enregistrement post-opératoire 1 témoigne d'écarts-types plus importants qui se réduisent ensuite parfois en post-opératoire 2 (sauf pour NPPENS, NPPKAU et NPPLEN) mais restent toutefois plus importants que ceux mesurés en préopératoire pour la plupart des locuteurs (NPPHOF fait figure d'exception ici).

Tableau 9 : Valeurs moyennes de fréquence fondamentale (F0) en fonction du locuteur pour la voyelle [a] – Groupe NPP. Les écarts-types sont donnés entre parenthèses

[a]	Préop	Post-Op 1	Post-Op 2
NPPENS	124,6 (4,94)	80 (3,24)	86 (5,18)
NPPGER	198,8 (2,82)	177,4 (6,55)	173 (4,24)
NPPHOE	173,4 (3,37)	125 (7)	136 (4,9)
NPPHOF	220,68 (15,7)	157,23 (18,41)	193,17 (8,13)
NPPKAU	129,33 (2,17)	120,61 (2,36)	135,36 (3,21)
NPPKRE	201,92 (4,33)	213,18 (7,97)	202,44 (5,2)
NPPLEN	228,1 (1,91)	199,09 (2,63)	198,33 (3,14)

Pour résumer :

La fréquence fondamentale moyenne calculée à partir des réalisations des locuteurs masculins et féminins est mesurée plus basse en post-opératoire 1 pour les trois voyelles. Elle augmente à nouveau en post-opératoire 2, sans toutefois atteindre les valeurs observées en phase préopératoire.

Les phases post-opératoires sont également caractérisées par des écarts-types plus importants.

L'observation plus précise des mesures en fonction des locuteurs révèle une certaine variabilité inter-individuelle. Si celle-ci a pu être observée en post-opératoire 1, c'est surtout dans la phase d'enregistrement post-opératoire 2 que l'on observe le plus de différences entre les locuteurs, puisque des comportements très différents ont pu être mis au jour (amélioration des mesures, stagnation des mesures ou détérioration des mesures), en fonction des locuteurs étudiés.

La variabilité intra-individuelle est également plus présente dans les phases d'enregistrement post-opératoires.

VI.4.2. Harmonics-to-Noise Ratio (HNR)

La mesure de HNR est considérée comme un indice de raucité, c'est-à-dire qu'elle renseigne sur le rapport bruit/harmoniques dans le spectre d'une voyelle. Habituellement, plus le ratio baisse, plus le signal est envahi par le bruit. Rappelons que les valeurs standards attendues sont de 30 dB pour [u] et de 20 dB pour [i] et [a].

L'effet principal de *phases d'enregistrement* n'est pas significatif pour la variable HNR ($p=ns$).

En préopératoire, les valeurs moyenne de HNR sont conformes aux valeurs attendues pour des locuteurs non pathologiques : la voyelle [a] est ainsi mesurée à 23 dB, la voyelle [i] à 25 dB et la voyelle [u] à 30dB. Les écarts-types (voir Tableau 10) sont par ailleurs faibles et témoignent de la régularité des productions des différents locuteurs.

Quelle que soit la voyelle étudiée, les valeurs de HNR sont inférieures en post-opératoire 1 ([i] = 22,64 dB, [a] = 17,31 dB, et [u] = 24 dB), par rapport à celles mesurées en préopératoire et post-opératoire 2. Toutefois, quelle que soit la phase d'enregistrement, les mesures restent très proches des valeurs attendues pour chacune des voyelles (

Figure 9).

Lors de la dernière phase d'enregistrement, les valeurs moyennes de HNR tendent à se rapprocher des valeurs mesurées en préopératoire. Le HNR de la voyelle [i] est alors mesuré à 23,94 dB (8,16 dB), la voyelle [a] à 19,68 dB (8,89 dB) et la voyelle [u] à 25,65 dB (10,28 dB).

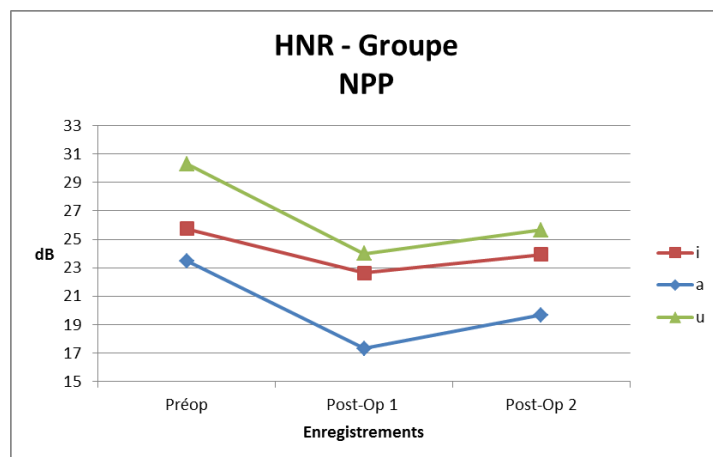


Figure 9 : Valeurs moyennes de HNR (dB) pour [i a u] – Groupe NPP

Les phases post-opératoires sont caractérisées par des écarts-types plus importants et témoignent d'une certaine variabilité inter et intra locuteur (Tableau 10).

Tableau 10 : Valeurs moyennes de HNR (dB) pour [i a u] – Groupe NPP. Les écarts-types sont donnés entre parenthèses

NPP	Préop	Post-Op 1	Post-Op 2
i	25,72 (3,98)	22,64 (9,37)	23,94 (8,16)
a	23,47 (3,55)	17,31 (8,35)	19,68 (8,89)
u	30,28 (3,60)	24 (9,97)	25,65 (10,28)

Compte tenu de l'importance des écarts-types dans les phases d'enregistrement post-opératoires, il nous semble nécessaire d'étudier les productions des locuteurs individuellement afin de mettre au jour les possibles stratégies propres à chacun. Nous nous focaliserons sur la voyelle [a].

La Figure 10 confirme la grande variabilité inter-locuteurs comme semblent le montrer les écarts-types importants (voir Tableau 10) mentionnés précédemment. Si les valeurs moyennes de chaque locuteur sont relativement similaires et conformes aux

valeurs attendues lors de la phase d'enregistrement préopératoire, les phases d'enregistrement post-opératoires indiquent des « stratégies individuelles » très différentes en fonction des locuteurs.

Pour cinq locuteurs (NPPENS en bleu foncé, NPPGER en rouge, NPPHOE en vert, NPPKAU en turquoise et NPPLLEN en bleu clair), les valeurs de HNR sont mesurées plus basses en post-opératoire 1 qu'en préopératoire. Toutefois, pour deux locuteurs, NPLEN (en bleu clair) et NPPGER (en rouge), les valeurs ne sont que faiblement diminuées, (-4 dB et à -5 dB, respectivement), alors que pour NPPENS (en bleu foncé), NPPHOE (en vert) et NPPKAU (en turquoise), le ratio baisse significativement, soit -17 dB, -13 dB et -11 dB, respectivement. En ce qui concerne les deux autres locuteurs, NPPKRE (en orange) et NPPHOF (en violet), les valeurs moyennes de HNR sont légèrement supérieures en postopératoire 1 (26,72 dB et 21,30 dB, respectivement) que celles mesurées en préopératoire (24,16 dB et 21,30 dB, respectivement).

Les valeurs deviennent « normales » pour cinq locuteurs (NPPGER en rouge, NPPHOE en vert, NPPHOF en violet, NPPKAU en turquoise et NPPKRE en orange) dès la phase d'enregistrement post-opératoire 2 et sont alors comparables à celles mesurées en préopératoire. Seuls deux locuteurs, NPPENS (en bleu foncé) et NPPLLEN (en bleu clair) indiquent des valeurs de HNR toujours inférieures en post-opératoire 2, par rapport à celles mesurées en préopératoire.

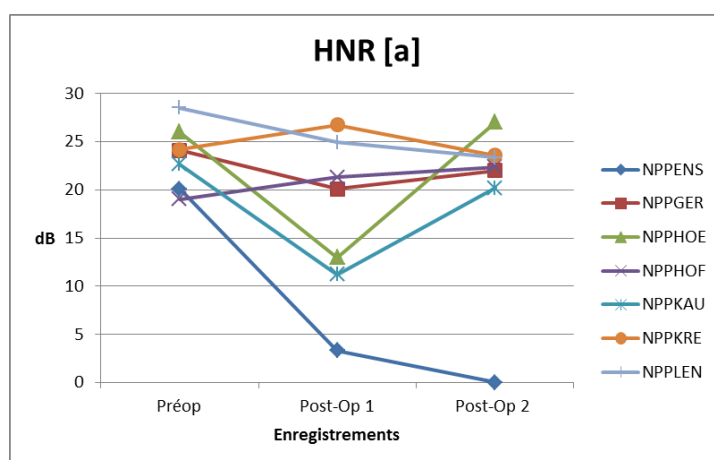


Figure 10 : Valeurs moyennes de HNR (dB) en fonction du locuteur pour la voyelle [a]

Comme le montre le Tableau 11, les écarts-types intra-locuteurs sont faibles, quelle que soit la phase d'enregistrement considérée, ce qui semble révéler une certaine régularité dans les productions vocaliques des locuteurs, même lorsque la qualité vocale est diminuée.

Tableau 11 : Valeurs moyennes de HNR (dB) en fonction du locuteur pour la voyelle [a] – Groupe NPP.

Les écarts-types sont donnés entre parenthèses

[a]	Préop	Post-Op 1	Post-Op 2
NPPENS	20,1 (1,59)	3,3 (2,21)	0 (0,7)
NPPGER	24,1 (1,37)	20,1 (2,02)	22 (1,63)
NPPHOE	26 (2,95)	13 (7,16)	27 (2,44)
NPPHOF	19 (1,89)	21,30 (1,71)	22,38 (2,47)
NPPKAU	22,66 (1,35)	11,23 (2,13)	20,18 (4,53)
NPPKRE	24,16 (2,83)	26,72 (1,96)	23,55 (1,17)
NPPLEN	28,5 (2,82)	24,90 (2,88)	23,33 (2,35)

Pour résumer :

De façon générale, les valeurs de HNR sont plus basses lors de la phase d'enregistrement post-opératoire 1, avant de ré-augmenter lors de la phase d'enregistrement post-opératoire 2, et ce quelle que soit la voyelle étudiée. Toutes phases d'enregistrement confondues, les valeurs de HNR sont très proches des valeurs attendues pour un locuteur non pathologique.

Les phases d'enregistrement post-opératoires présentent des écarts-types plus importants, ce qui nous a amenée à considérer la mesure de HNR en fonction des locuteurs. Cette démarche a permis de mieux mettre au jour une certaine variabilité inter-locuteurs (les écarts-types intra-locuteurs restant faibles).

Le temps permet donc une récupération vocale, puisque pour six locuteurs sur sept, les valeurs du HNR sont proches des valeurs mesurées en préopératoire, et cela dès la phase post-opératoire 2.

VI.4.3. Jitter et Shimmer

Les analyses de variance n'ont pas permis d'identifier un effet principal *temps* pour la variable jitter ($p=ns$). Cela indique que ce paramètre a été calculé autour des mêmes valeurs, quelle que soit la phase d'enregistrement étudiée.

Suivant les recommandations habituelles (*cf.* par ex., celles du manuel PRAAT[®]), le seuil normal/pathologique de la mesure de jitter est fixé à 1,04%.

L'étude de jitter révèle une augmentation de ce paramètre entre la phase d'enregistrement préopératoire (0,42%) et les phases d'enregistrement post-opératoires 1 et 2, où il est alors de 0,84%, sans toutefois dépasser le seuil normal/pathologique. Les écarts-types sont plus importants pour les enregistrements post-opératoires, puisqu'ils sont respectivement de 0,72 et 0,97, ce qui indique que certaines productions de certains locuteurs pourraient dépasser le seuil de 1,04%. Le Tableau 12 illustre ces observations.

Tableau 12 : Valeurs moyennes de jitter en % – Groupe NPP. Les écarts-types sont donnés entre parenthèses

% jitter	Préop	Post-Op 1	Post-Op 2
[a]	0,42 (0,17)	0,84 (0,72)	0,84 (0,97)

En étudiant la mesure de jitter en fonction du locuteur (Tableau 13), on remarque que si les mesures sont plus importantes lors de la phase d'enregistrement post-opératoire 1 pour tous les locuteurs, un seul locuteur dépasse le seuil normal/pathologique. Il s'agit du locuteur NPPENS pour lequel le jitter est mesuré à 1,83% en post-opératoire 1. Ce locuteur présente également des écarts-types plus importants que tous les autres locuteurs.

Lors de la phase d'enregistrement post-opératoire 2, les valeurs de jitter diminuent pour tous les locuteurs par rapport à la phase d'enregistrement précédente.

Les écarts-types intra-locuteurs restent faibles toute phase d'enregistrement confondue.

Tableau 13 : Valeurs moyennes de jitter en % en fonction du locuteur – Groupe NPP. Les écarts-types sont donnés entre parenthèses

% jitter	Pré-Op	Post-Op 1	Post-Op 2
NPPENS	0,69 (0,13)	1,83 (1,42)	3,07 (0,74)
NPPGER	0,36 (0,10)	0,97 (0,22)	0,69 (0,22)
NPPHOE	0,34 (0,15)	0,82 (0,30)	0,29 (0,12)
NPPHOF	0,36 (0,07)	0,39 (0,09)	0,27 (0,06)
NPPKAU	0,44 (0,12)	0,96 (0,23)	0,64 (0,17)
NPPKRE	0,50 (0,14)	0,57 (0,22)	0,47 (0,07)
NPPLN	0,27 (0,08)	0,35 (0,13)	0,48 (0,14)

Les ANOVA conduites à partir de la mesure de shimmer ont donné des résultats similaires à ceux observés pour le jitter, à savoir qu'un effet principal *temps* n'a pas pu être dégagé ($p=ns$).

En ce qui concerne le shimmer, le seuil normal/pathologique est généralement fixé à 3,81% (*cf.* le manuel PRAAT[®]). La mesure de shimmer, détaillée dans le Tableau 14, suit la même tendance que le jitter, à savoir que cette mesure augmente en post-opératoire 1, où elle est de 5,45% (contre 3,08% en préopératoire). Le shimmer diminue ensuite lors de la dernière phase d'enregistrement, étant donné qu'il est alors mesuré à 4,28%. Les mesures moyennes des deux phases d'enregistrement dépassent ainsi le seuil normal/pathologique.

Les phases post-opératoires (voir Tableau 14) sont caractérisées par des écarts-types importants (4,03 et 3,74 respectivement), ce qui témoigne d'une certaine variabilité probablement individuelle.

Tableau 14 : Valeurs moyennes de shimmer en % – Groupe NPP. Les écarts-types sont donnés entre parenthèses

% shimmer	Préop	Post-Op 1	Post-Op 2
[a]	3,08 (1,11)	5,45 (4,03)	4,28 (3,74)

Le Tableau 15 propose d'étudier la mesure de shimmer en fonction du locuteur. Il semble alors que la variabilité mise au jour par les écarts-types importants (voir Tableau 14) soit le résultat de stratégies inter-individuelles. En effet, si la mesure de shimmer augmente pour tous les locuteurs (sauf pour NPPLN) en phase post-opératoire 1, l'amplitude n'est pas la même pour tous les locuteurs. La mesure de shimmer est particulièrement dégradée pour le locuteur NPPNS, alors qu'elle augmente de façon non significative pour NPPHOE, par exemple. De plus, le seuil normal/pathologique n'est dépassé que pour NPPNS (12,82%), NPPHOE (4,46%) et NPPKAU (9,21%), soient trois locuteurs sur sept.

La phase d'enregistrement post-opératoire 2 renforce également l'idée d'une variabilité inter-locuteurs. Le shimmer diminue à nouveau pour NPPHOE (1,73%) et NPPKAU (3,22%) qui sont alors mesurés en dessous du seuil normal/pathologique. Le shimmer reste particulièrement important pour NPPNS (12,82%). Pour trois locuteurs, NPPGER, NPPLN et NPPKRE, le shimmer augmente de façon non significative entre les phases d'enregistrements post-opératoires, sans pour autant atteindre le seuil normal/pathologique. Enfin, pour NPPHOF, le shimmer diminue en post-opératoire 2.

Les écarts-types témoignent également d'une certaine variabilité intra-individuelle (voir Tableau 14).

Tableau 15 : Valeurs moyennes de shimmer en fonction du locuteur – Groupe NPP. Les écarts-types sont donnés entre parenthèses

% shimmer	Pré-Op	Post-Op 1	Post-Op 2
NPPNS	5,01 (1,46)	12,82 (2,83)	12,82 (2,82)
NPPGER	2,67 (0,58)	3,61 (0,70)	3,86 (0,77)
NPPHOE	2,57 (0,55)	4,46 (1,03)	1,73 (0,28)
NPPHOF	3,16 (0,23)	3,22 (1,20)	2,46 (0,40)
NPPKAU	2,65 (0,65)	9,21 (2,27)	3,22 (0,88)
NPPKRE	3,19 (0,55)	2,68 (0,88)	3,00 (0,37)
NPPLN	2,31 (0,65)	2,11 (0,68)	2,91 (0,69)

Pour résumer :

Le jitter augmente entre les phases d'enregistrement préopératoire et post-opératoires, sans toutefois dépasser le seuil normal/pathologique usuellement attesté. Les phases post-opératoires sont caractérisées par des écarts-types plus importants.

Pour tous les locuteurs, le jitter augmente lors de la phase d'enregistrement post-opératoire 1 avant de diminuer en post-opératoire 2. La variabilité intra-locuteurs est faible. Un seul locuteur présente un schéma différent et attendu. En effet, la mesure du jitter dépasse le seuil normal/pathologique lors des enregistrements post-opératoires. Ce locuteur est également caractérisé par des écarts-types plus importants.

Le shimmer, quant à lui, augmente également entre les phases d'enregistrement préopératoire et post-opératoire 1 avant de diminuer, sans toutefois atteindre les valeurs mesurées en préopératoire et en post-opératoire 2. Le seuil normal/pathologique est dépassé lors des deux phases d'enregistrement post-opératoires. Des écarts-types plus importants sont également à observer dans ces deux phases.

Le shimmer atteste d'une certaine variabilité inter et intra locuteurs, puisque si la tendance est la même, quel que soit le locuteur étudié, l'augmentation ou la diminution de cette mesure dépend étroitement du locuteur.

VI.4.4. Valeurs formantiques

Pour ce qui concerne les analyses de variance conduites à partir des valeurs formantiques des différentes voyelles, l'effet principal *temps* ne s'est pas révélé significatif ($p=ns$). Les modifications de ces valeurs (voir *infra*) dépendent largement de la voyelle étudiée. Ces résultats indiquent que quelle que soit la voyelle étudiée, le temps ne modifie pas les variables formantiques, soit F1 et F2, de façon significative. Il conviendra toutefois de vérifier en section VI.4.5. si ces modifications non significatives statistiquement n'altèrent pas tout de même l'aire de l'espace vocalique.

Enfin, signalons que le facteur *voyelle* pour la variable F1 s'est révélé significatif [$F(2,12) = 68,64, p<0.000000$] et également pour la variable F2 [$F(2,12) = 212,44, p<0.000000$], ce qui laisse supposer que si l'espace vocalique subit de légères modifications il ne devrait pas moins rester géométriquement conventionnel.

L'étude des valeurs formantiques, soit F1 et F2, révèle des tendances différentes en fonction de la voyelle étudiée (voir Tableau 16). Dans la mesure où la plupart de nos locuteurs développent des comportements similaires aux tendances que nous allons décrire, nous ne proposerons pas d'analyse individuelle, les valeurs sont toutefois disponibles en annexe.

En ce qui concerne la voyelle [i], les valeurs de F1 et de F2 ne varient pas en fonction de la phase d'enregistrement, et sont toujours proches des valeurs attendues. F1 est mesuré à 321 Hz en préopératoire, 320 Hz en post-opératoire 1 et 306 Hz en post-opératoire 2. Les écarts-types sont légèrement plus importants lors de la phase d'enregistrement post-opératoire 1, puisqu'ils sont alors de 61,90 Hz. Les phases d'enregistrement préopératoire (écart-type 43,27 Hz) et post-opératoire 2 (écart-type 44,60 Hz) indiquent une variabilité réduite. Les valeurs de F2 sont les suivantes : 2423 Hz en préopératoire, 2435 Hz en post-opératoire 1 et 2447 Hz en post-opératoire 2. Ces valeurs sont non seulement régulières mais également proches des valeurs attendues pour un locuteur non pathologique. Les écarts-types mesurés à partir des valeurs de F2 de la voyelle [i] sont relativement constants quelle que soit la phase d'enregistrement étudiée (environ 240 Hz en préopératoire et 210 Hz en post-opératoires).

Les valeurs formantiques des voyelles [a] et [u] subissent de légères modifications en fonction des phases d'enregistrement préopératoires et post-opératoire 1. Le premier formant du [a] est mesuré à 731 Hz en préopératoire, puis à 666 Hz en post-opératoire 1 et enfin à 719 Hz en post-opératoire 2. Ces résultats vont dans le sens d'une centralisation de l'espace vocalique sur l'axe vertical (voir VI.4.5. pour des informations relatives à l'aire de l'espace vocalique). Les modifications des valeurs de F2 suivent le même schéma, puisque les changements les plus importants sont à observer en post-opératoire 1, F2 étant alors mesuré à 1315 Hz. Les valeurs en préopératoire (1297 Hz) et post-opératoire 2 (1288 Hz) sont relativement similaires. En ce qui concerne la voyelle [u], les valeurs de F1 sont les suivantes : 340 Hz en préopératoire, 381 Hz en post-opératoire 1 et enfin 391 Hz en post-opératoire 2. Ces valeurs vont également dans le sens d'une centralisation de l'espace vocalique sur l'axe vertical. Les valeurs de F2 subissent également de légères modifications, puisque F2 est mesuré à 762 Hz en préopératoire, 878 Hz en post-opératoire 1 et 853 Hz en post-opératoire 2.

Signalons que les écarts-types relevés à partir des valeurs de F1 sont toujours plus importants en post-opératoire 1, quelle que soit la voyelle étudiée - soit 61,90 Hz pour [i], 214,12 Hz pour [a] et 104,02 Hz pour [u] - que ceux mesurés en préopératoire - soit 43,27 Hz pour [i], 114,12 Hz pour [a] et 44,09 Hz pour [u]. La phase post-opératoire 2 indique des écarts-types à nouveau comparables avec ceux de la phase préopératoire, sauf pour la voyelle [u] qui affiche toujours une grande variabilité - soit 44,60 Hz pour [i], 168,65 Hz pour [a] et 112,13 Hz pour [u].

L'étude des écarts-types de F2 dépend largement de la voyelle considérée. En effet, la voyelle [i] présente des écarts-types plus importants en préopératoire 243,51 Hz que ceux mesurés en post-opératoire 1 et 2, soit 210,4 Hz et 209,05 Hz respectivement. Les écarts-types pour la voyelle [a] suivent la tendance attendue, à savoir plus de variabilité (165,33 Hz) en post-opératoire 1 que dans les phases préopératoire (99,99 Hz) et post-opératoire 2 (149,81 Hz). Enfin, les écarts-types sont plus importants pour le [u], par rapport aux deux autres voyelles investiguées, lors des phases post-opératoires (soit 218,49 Hz et 302,43 Hz respectivement), alors que les écarts-types relevés en préopératoire étaient relativement faibles soit 86,78 Hz.

Tableau 16 : Valeurs formantiques moyennes – Groupe NPP. Les écarts-types sont donnés entre parenthèses

Voyelles	Formants	Préop	Postop1	Postop2
i	F1	321 (43,27)	320 (61,90)	306 (44,60)
	F2	2423 (243,51)	2435 (210,4)	2447 (209,05)
a	F1	731 (114,12)	666 (214,12)	719 (168,65)
	F2	1297 (99,99)	1315 (165,33)	1288 (149,81)
u	F1	340 (44,09)	381 (104,02)	391 (112,13)
	F2	762 (86,78)	878 (218,49)	853 (302,43)

Pour résumer :

Pour le groupe NPP (Patients ne présentant pas d'immobilité laryngée), les valeurs de F1 et F2 ne varient pas en fonction de la phase d'enregistrement pour la voyelle [i]. En revanche, F1 et/ou F2 subissent de légères modifications pour les voyelles [a] et [u] en phase d'enregistrement post-opératoire 1, avant de retrouver des

valeurs attendues en post-opératoire 2. Ces tendances ($p=ns$) pourraient tout de même influencer sur les tailles de l'espace vocalique maximal (cf. infra).

L'étude des écarts-types a révélé une variabilité plus importante, caractérisée par des écarts-types plus grands, dans la phase post-opératoire 1. Cela est particulièrement vérifié pour les voyelles [a] et [u].

VI.4.5. Aire de l'espace vocalique

Comme nous l'avions supposé *supra*, les faibles modifications des valeurs formantiques, extraites à partir des voyelles, ont des conséquences sur l'aire de l'espace vocalique. Les ANOVA indiquent un effet principal *phase d'enregistrement* pour la variable aire de l'espace vocalique $F(2,12) = 4,15$ $p < 0,043$. Les tests *post-hoc* révèlent que l'effet principal repose sur une différence significative entre la phase d'enregistrement préopératoire ($0,33 \text{ kHz}^2$) et la phase d'enregistrement post-opératoire 1 ($0,22 \text{ kHz}^2$). La phase post-opératoire 2 ($0,28 \text{ kHz}^2$), en revanche, n'est pas significativement différente des phases d'enregistrement préopératoire et post-opératoire 1.

L'aire de l'espace vocalique se réduit dans les phases post-opératoires. Elle est alors de $0,33 \text{ kHz}^2$ en préopératoire, de $0,23 \text{ kHz}^2$ en post-opératoire 1 et de $0,28 \text{ kHz}^2$ en post-opératoire 2. Ces mesures indiquent donc une tendance à la centralisation des voyelles [u] et [a] ; la Figure 11 illustre cette tendance.

Notons que si l'espace vocalique est réduit, il n'en reste pas moins géométriquement conventionnel (voir Figure 11). Les voyelles sont toujours réalisées à l'intérieur des régions habituellement observées.

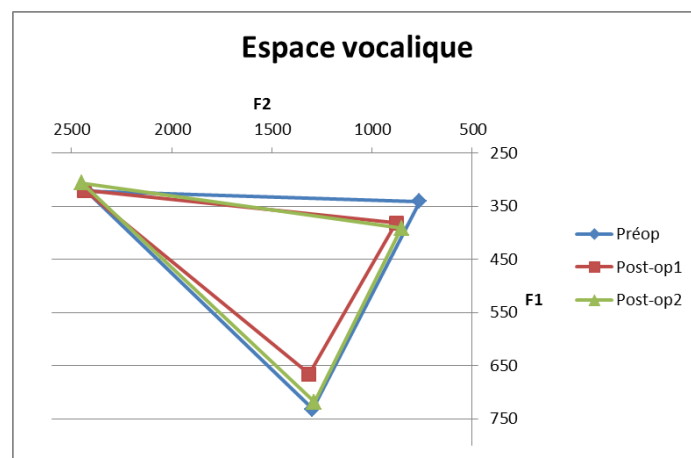


Figure 11 : Espace vocalique – Groupe NPP

En ce qui concerne l'étude des espaces vocaliques en fonction du locuteur (voir Tableau 17), la principale tendance est naturellement, compte tenu des remarques précédentes, à une diminution de cet espace en post-opératoire 1, puis à une augmentation en post-opératoire 2, sans atteindre toutefois les valeurs relevées en préopératoire. Néanmoins, ce paramètre atteste également d'une grande variabilité interlocuteurs, puisque l'espace vocalique de certains locuteurs se réduit de façon plus importante en post-opératoire 1 que celui d'autres locuteurs.

Les espaces vocaliques de NPPGER, NPPHOE et NPPKAU se réduisent de façon remarquable entre les phases d'enregistrement préopératoire, soit 0,46 kHz², 0,22 kHz², 0,42 kHz² et 0,20 kHz² respectivement, et post-opératoire 1, où ils sont alors plus que de 0,15 kHz², de 0,03 kHz² et de 0,07 kHz². Les espaces vocaliques de NPPHOF et NPPKRE diminuent également entre les deux phases préopératoire et post-opératoire 1, mais cela de façon moins importante ; ils sont mesurés à 0,42 kHz² et 0,33 kHz² en préopératoire et à 0,34 kHz² et 0,22 kHz² en post-opératoire 1. Enfin, les locuteurs NPPENS et NPPLEN présentent la tendance inverse, à savoir que leur espace vocalique en post-opératoire 1 (soit 0,31 kHz² et 0,40 kHz², respectivement) est légèrement plus important que celui mesuré en phase préopératoire (soit 0,29 kHz² et 0,38 kHz²).

Lors de la phase post-opératoire 2, l'espace vocalique de tous les locuteurs augmente à nouveau, sauf pour NPPENS pour qui l'espace vocalique est mesuré le plus petit dans cette phase (0,17 kHz²). Pour tous les autres locuteurs, l'aire de l'espace

vocalique en post-opératoire 2 est à nouveau comparable à celui quantifié en phase préopératoire. Elle reste toutefois légèrement plus petite pour NPPGER (0,34 kHz²), NPPHOE (0,16 kHz²), NPPHOF (0,38 kHz²), et NPPLEN (0,33 kHz²). Elle est mesurée plus grande pour NPPKRE à 0,37 kHz².

Tableau 17 : Espace vocalique (kHz²) en fonction des locuteurs – Groupe NPP

kHz ²	Préop	Post-Op 1	Post-Op 2
NPPENS	0,29	0,31	0,17
NPPGER	0,46	0,15	0,34
NPPHOE	0,22	0,03	0,16
NPPHOF	0,42	0,34	0,38
NPPKAU	0,20	0,07	0,17
NPPKRE	0,33	0,22	0,37
NPPLEN	0,38	0,40	0,33

Pour résumer :

L'aire de l'espace vocalique moyen est significativement réduite ($p < 0.05$) pour le groupe NPP entre les phases d'enregistrement préopératoire et postopératoire 1, avant de ré-augmenter ($p = ns$) en phase postopératoire 2, sans toutefois atteindre une valeur de référence de la phase préopératoire.

L'étude de l'aire de l'espace vocalique en fonction des locuteurs révèle une tendance similaire, à savoir que la plupart des locuteurs ont une aire de l'espace vocalique plus petite en post-opératoire 1 que celle mesurée en phase préopératoire, avant d'augmenter en post-opératoire 2. Les différences individuelles sont à relever dans l'importance de cette réduction.

VI.5. Discussion

Il convient à présent de vérifier si nos hypothèses initiales ont été confirmées ou infirmées.

1. Les valeurs de la fréquence fondamentale sont modifiées pour tous les locuteurs. La principale tendance semble être un abaissement de la fréquence fondamentale dans les phases d'enregistrement précoces. Cela est particulièrement remarquable lors de la première phase d'enregistrement post-opératoire. La fréquence fondamentale de la plupart des locuteurs est alors mesurée plus basse que celle relevée en préopératoire. La phase post-opératoire 2 montre une amélioration des mesures de la fréquence fondamentale, même si les données obtenues ne sont pas encore comparables aux valeurs préopératoires. La voix des locuteurs est donc plus grave (le fondamental plus bas) dans les phases post-opératoires. Il semble donc que le lien que nous avons prédit entre une perturbation de l'activité laryngienne et les valeurs de la fréquence fondamentale soit établi (*Hypothèse 1*).
2. L'activité irrégulière du larynx a également un impact sur les valeurs de Harmonics-to-Noise Ratio (*Hypothèse 2*). Cette mesure suit la même tendance que les mesures de la fréquence fondamentale, à savoir que les valeurs de HNR sont plus basses lors de la phase post-opératoire 1. Les valeurs augmentent à nouveau dès la phase post-opératoire 2 pour se rapprocher, sans toutefois les atteindre, les valeurs mesurées en préopératoire. Cette mesure indique qu'il y a plus de bruit dans la voix des locuteurs en post-opératoire ; leur voix est alors plus rauque.
3. Les valeurs de jitter et de shimmer sont modifiées en post-opératoire. Cette observation est vraisemblablement liée à une modification de la périodicité de la voix (*Hypothèse 3*). Les valeurs de jitter ne dépassent toutefois pas le seuil normal/pathologique. Le shimmer est modifié de façon plus importante, sans que cela soit statistiquement significatif. Ces mesures indiquent que le signal est perturbé (fréquence fondamentale et intensité) dans les phases d'enregistrement post-opératoires. Ce constat est à souligner plutôt dans la première phase d'enregistrement post-opératoire. La voix des patients serait alors plus irrégulière après l'opération chirurgicale. Notons que cette irrégularité s'estompe avec le temps, puisque les valeurs mesurées en post-opératoire 2 sont comparables à celles relevées en préopératoire.

4. Les valeurs formantiques, extraites à partir des voyelles soutenues, ont également été perturbées (sauf pour la voyelle [i]) mais de façon non significative. Cette perturbation des valeurs formantiques proviendrait de celle de l'activité laryngienne, en amont (*Hypothèse 4*).
5. Ces modifications ont naturellement des conséquences sur l'aire de l'espace vocalique qui peut subir une réduction mais pas de réorganisation géométrique. L'espace vocalique est plus petit en post-opératoire que celui mesuré en préopératoire. Dès la phase post-opératoire 2, il augmente pour se rapprocher des valeurs standards observées avant l'opération. L'espace vocalique reste conventionnel, malgré la centralisation des voyelles [u] et [a]. Notons que ces modifications formantiques sont probablement liées aux perturbations de la source laryngée plutôt qu'à des stratégies articulatoires différentes. Cette observation reste à vérifier (*Hypothèse 5*), puisqu'il est possible qu'en phase post-opératoire précoce, le locuteur soit gêné dans les réalisations postérieures [u] à cause de l'œdème chirurgical. Il pourrait également utiliser des stratégies d'économie articulatoire, en termes de réduction de l'ampleur de l'abaissement mandibulaire et lingual (avec une translation moindre de la masse linguale vers la paroi pharyngale), ce qui pourrait expliquer la centralisation de la voyelle [a].
6. Nous avons vu que quel que soit le paramètre étudié, la variabilité inter et/ou intra-locuteurs est toujours plus marquée dans les phases d'enregistrement post-opératoires. La phase post-opératoire 1, notamment, indique des écarts-types souvent plus importants que ceux mesurés en préopératoire. La phase post-opératoire 2 est généralement caractérisée par des écarts-types plus importants que ceux mesurés en préopératoire mais plus faibles que ceux mesurés en post-opératoire 1. Il semblerait donc que cette phase soit une phase de transition. On peut supposer que les écarts-types, notamment intra-individuels, devraient continuer à se réduire avec le temps, témoignant ainsi d'une régularité retrouvée dans les productions des locuteurs. Il s'agirait là de la récupération des coordinations oro-laryngées initiales (*Hypothèse 6*).
7. Enfin, de façon générale, le temps a un impact positif sur tous les paramètres précédemment mentionnés (*Hypothèse 7*). Les mesures se rapprochent des

valeurs préopératoires dès la phase d'enregistrement post-opératoire 2, et cela pour la plupart des patients. Ce constat indique que si la voix des locuteurs est modifiée directement après l'intervention, à court terme (post-opératoire 1), la qualité vocale s'améliore par la suite. Dans la plupart des cas, la voix est à nouveau comparable à celle d'avant l'opération, en post-opératoire 2.

CONCLUSION CHAPITRE VI

Dans ce chapitre, nous avons étudié la voix de sept locuteurs qui ont été opérés de la glande thyroïde et pour lesquels le diagnostic ORL n'a pas révélé d'immobilité laryngée. Notre étude se voulait longitudinale, puisque leur voix a été enregistrée à trois reprises : en préopératoire (la veille de l'intervention), en post-opératoire 1 (le lendemain de l'intervention) et en post-opératoire 2 (quinze jours après l'intervention).

Grâce à un corpus composé de voyelles soutenues, nous avons pu étudier les modifications que l'opération chirurgicale entraîne sur les paramètres acoustiques et spectraux suivants : la fréquence fondamentale, les valeurs formantiques, le Harmonics-to-Noise Ratio (HNR), le jitter, le shimmer et l'aire de l'espace vocalique.

Nous supposons que l'opération chirurgicale au niveau du larynx pourrait modifier l'activité laryngienne, ce qui aurait naturellement des conséquences sur les paramètres susmentionnés, et notamment sur les mesures qui permettent de quantifier les perturbations du signal acoustique, tels que le HNR, le jitter et le shimmer. Nous émettions également l'hypothèse que le temps devait avoir un effet positif sur les paramètres étudiés, ce qui se traduirait par des valeurs semblables en post-opératoire 2 et en phase préopératoire.

Les analyses de variance (ANOVA) à trois facteurs indiquent que l'effet principal *phase d'enregistrement* n'a été significatif ($p < 0,05$) que pour les variables *fréquence fondamentale* et *aire de l'espace vocalique*.

Plus précisément, la fréquence fondamentale diminue en post-opératoire 1 ($p < 0,05$), tout comme il est possible d'observer une réduction du triangle vocalique à cette même phase d'enregistrement. Des tendances ont également pu être relevées pour les autres paramètres : le Harmonics-to-Noise Ratio diminue en post-opératoire 1, alors que les valeurs du jitter et du shimmer sont plus importantes.

Le temps a un effet positif sur tous les paramètres puisque, lors de la phase d'enregistrement post-opératoire 2, tous les paramètres retrouvent des valeurs « standards » proches de celles mesurées en préopératoire.

Les phases post-opératoires sont également caractérisées par une variabilité inter et intra locuteur qui a toutefois tendance à se réduire en post-opératoire 2.

RESUME CHAPITRE VII

Précédemment, il a été montré qu'une modification de la voix après une ablation de la glande thyroïde pouvait intervenir alors même que les nerfs laryngés étaient préservés. Il s'agira dans ce chapitre d'évaluer les altérations vocales post-thyroïdectomies, liées à une immobilité laryngée unilatérale. Celle-ci est, en effet, une des complications post-thyroïdectomie la plus fréquente même si elle reste heureusement relativement rare (voir chapitre IV). La littérature indique que tous les paramètres (acoustiques, aérodynamiques et spectraux) de la voix se trouvent alors modifiés. Notre étude se veut longitudinale, il s'agit également d'évaluer la possible récupération vocale des patients ; tous les patients suivent une rééducation orthophonique.

Pour répondre à cet objectif, nous avons enregistré la voix de sept locuteurs (cinq femmes et deux hommes) présentant une immobilité laryngée. Ils ont été enregistrés à trois reprises : en post-opératoire 2 (quinze jours après l'intervention), en post-opératoire 3 (un mois après l'intervention) et en post-opératoire 4 (deux mois après l'intervention).

Afin d'obtenir des valeurs de référence, nous avons également enregistré la voix de sept locuteurs contrôles, appariés en genre et en âge avec les patients.

Notre corpus est composé de voyelles soutenues à partir desquelles nous avons mesuré la fréquence fondamentale, les valeurs formantiques (permettant ainsi de quantifier l'aire de l'espace vocalique). Des mesures de perturbations du signal telles que le Harmonics-to-Noise Ratio, le jitter et le shimmer ont également été effectuées.

Les résultats vont dans le sens d'une altération de tous les paramètres étudiés en post-opératoire 2 et d'une amélioration progressive, à partir de la phase post-opératoire 3, qui se poursuit jusqu'en post-opératoire 4.

CHAPITRE VII

VII. Expérience 2 : Etude acoustique de voyelles soutenues produites par des patients présentant une paralysie récurrentielle après thyroïdectomie

VII.1. Introduction

La présente investigation repose sur l'analyse de quelques caractéristiques spectrales de voyelles soutenues produites par des patients ayant subi une thyroïdectomie à la suite de laquelle une immobilité laryngée est apparue. Il s'agit ainsi d'une étude acoustique.

La voix des patients est étudiée afin d'identifier et d'évaluer les différentes perturbations que l'immobilité laryngée post-thyroïdectomie peut provoquer au niveau acoustique.

Il s'agit d'une étude longitudinale, puisque les patients sont enregistrés lors de différentes phases post-opératoires, ce qui devrait nous permettre de mettre au jour les possibles stratégies de compensation ou de réajustements que le patient peut initier à l'aide de sa rééducation orthophonique.

Rappelons que les immobilités laryngées post-thyroïdectomie demeurent heureusement relativement rares (voir Benninger et *al.*, 1998 par ex.) et que leur origine n'est pas toujours facilement identifiable : geste chirurgical ou intubation trachéal, etc (Friedrich et *al.*, 2000). Dans 75% des cas (Steinmüller et *al.*, 2001), la récupération de

la mobilité intervient dans les six premiers mois, et le traitement orthophonique reste la solution la plus fréquemment proposée (voir chapitre IV).

VII.2. Procédure expérimentale

VII.2.1. Locuteurs

Pour cette expérience, la voix de sept locuteurs a été enregistrée au Centre Paul Strauss, puis dans le cabinet de différentes orthophonistes. Il s'agit de cinq femmes (UPPBAS, UPPHEI, UPPHUB, UPPWAL et UPPWAN) et de deux hommes (UPPLAT et UPPPAI). Tous ces patients ont subi une thyroïdectomie totale, à la suite de laquelle une immobilité laryngée unilatérale est apparue. L'examen laryngé et le diagnostic ont été effectués par l'ORL choisi par le patient. Il est donc probable que les méthodes d'investigation (miroir, laryngoscopie...) soient différentes en fonction des praticiens retenus.

Nous (Fauth et *al.*, 2011) avons appelé ce groupe de patients le groupe UPP (Unilateral Paralysis Patients ou patient présentant une immobilité laryngée), dans la suite de ce travail.

Les patients ont été enregistrés à plusieurs reprises, tout au long de leur rééducation orthophonique. Afin d'avoir une cohorte comparable et de pouvoir conduire des analyses de variance (ANOVA à mesures répétées), nous n'avons pu retenir que trois enregistrements post-opératoires, soit post-opératoire 2 (quinze jours après l'intervention), post-opératoire 3 (environ un mois après l'intervention) et post-opératoire 4 (environ deux mois après l'intervention). Dans certains cas, la rééducation orthophonique n'était donc pas terminée.

Dans la mesure où nous n'avons pas pu acquérir de données en préopératoire (voir chapitre V), la voix de référence a été obtenue auprès d'un locuteur contrôle (dorénavant LC) apparié en genre et en âge avec le patient correspondant.

VII.2.2. Corpus

Notre corpus est constitué des trois voyelles /i a u/ qui ont été répétées chacune à dix reprises, dans la mesure des possibilités des locuteurs. La consigne donnée aux

différents patients était de produire et de soutenir, durant 5 secondes environ, les voyelles susmentionnées. En cas de nécessité, notamment dans les phases d'enregistrement post-opératoires précoces, une pause après cinq répétitions de chaque voyelle était proposée au patient.

VII.2.3. Mesures

Le signal acoustique et son spectrogramme correspondant ont été segmentés conjointement. Pour chaque voyelle, les paramètres spectraux suivants ont été mesurés dans la structure formantique stable :

- la fréquence fondamentale ;
- les mesures formantiques (F1 et F2) ;
- les mesures de perturbations du signal, à savoir le Harmonics-to-Noise Ratio (HNR), le jitter et le shimmer.

VII.2.4. Analyse des données

Les résultats seront présentés sous forme de graphiques et de tableaux commentés, afin de faciliter leur lecture et de rendre plus aisée la comparaison entre les différents locuteurs. Par ailleurs, ce type de présentation devrait également permettre de dégager aisément des tendances.

Des analyses de variance (ANOVA à mesures répétées) ont été conduites pour toutes nos données (*cf. infra* pour plus de détails). Les valeurs présentées sont toujours des moyennes, calculées à partir des répétitions produites par nos locuteurs. Les valeurs apparaissant entre parenthèses renvoient aux écarts-types.

VII.3. Hypothèses

Les hypothèses émises ici sont, à peu de choses près, les mêmes que celles de l'investigation précédentes. Néanmoins, nous les rappelons encore afin de faciliter la lecture de ce chapitre.

Etant donné qu'une immobilité laryngée entraîne des modifications de la voix, nous formulons les hypothèses suivantes :

1. La modification de la voix affecterait directement les valeurs de la fréquence fondamentale, ou F0, dans la mesure où ce paramètre est lié à l'activité laryngienne ;
2. L'altération vocale peut être due à une activité irrégulière du larynx qui pourrait avoir des conséquences sur les valeurs du Harmonics-to-Noise Ratio, généralement mesurées autour de 20 dB pour /i/ et /a/ et autour de 30 dB pour /u/, pour un locuteur non pathologique ;
3. Etant donné que les modifications de la fréquence fondamentale pourraient également se traduire par des modifications de la périodicité de la voix, les valeurs de jitter et de shimmer pourraient alors se trouver modifiées ;
4. La perturbation de la source laryngée aurait également des conséquences sur le couplage entre l'activité laryngienne et les cavités supraglottiques résonantes, et donc sur la taille et la forme de l'espace vocalique. Cela induirait une perturbation des résonances supra-glottiques *i.e.* des valeurs de F1 et de F2 ;
5. La variabilité inter et intra locuteurs, traduite par des écarts-types plus importants, devrait être plus marquée dans les phases d'enregistrement précoces, et plus particulièrement en post-opératoire 2, à cause d'une coordination oro-laryngée problématique. Cette variabilité devrait se réduire dès la phase post-opératoire 3, c'est-à-dire dès le démarrage de la rééducation orthophonique, révélant ainsi la récupération progressive des gestes glottiques ;
6. Le temps et la rééducation vocale devraient permettre une récupération vocale et une « standardisation » des paramètres précédemment évoqués ; les paramètres seraient alors comparables, toutes proportions gardées, aux valeurs acquises auprès des locuteurs contrôles. Cette récupération devrait être progressive et continue.

VII.4. Résultats

VII.4.1. Remarques générales

Il semble utile de rappeler que les résultats présentés dans cette étude ont été obtenus à partir de :

- trois voyelles ;
- répétées à 30 reprises (soit 10 répétitions par voyelle) ;
- l'ensemble de ces répétitions a été prononcé par les sept patients et par le locuteur contrôle apparié par patient ;
- lors de trois phases d'enregistrement (en post-opératoire 2, en post-opératoire 3 et en post-opératoire 4) et une seule fois pas les contrôles, soit quatre « sessions » par locuteur au total.

Nous aboutissons ainsi à 840 répétitions au total.

Des ANOVA à mesures répétées à deux facteurs (*temps* et *voyelle*) ont été effectuées pour les *variables* F0, F1, F2 et HNR, et à un facteur (*temps*) pour les variables jitter, shimmer et aire de l'espace vocalique, afin de déterminer s'il existait des effets de *phases d'enregistrement (temps)* et de *voyelle* pour les *variables* F0, F1, F2 et HNR, et de savoir s'il y avait un effet de phase temps pour les variable jitter, shimmer et aire de l'espace vocalique

Les enregistrements des locuteurs contrôles ont été considérés comme une *phase d'enregistrement* ou T0. Ainsi par extrapolation, une différence significative entre la phase d'enregistrement des locuteurs contrôles et celles des patients reviendrait à considérer qu'il existe un effet principal locuteur contrôles vs. patients.

VII.4.2. Fréquence fondamentale (F0)

Les analyses de variance (ANOVA à mesures répétées) n'ont pas permis de révéler une modification significative du facteur *phase d'enregistrement* ($p=ns$).

Afin de ne pas fausser nos analyses statistiques, il est entendu que l'étude fine de la fréquence fondamentale de nos locuteurs féminins a été effectuée, bien évidemment, indépendamment de celle de nos locuteurs masculins.

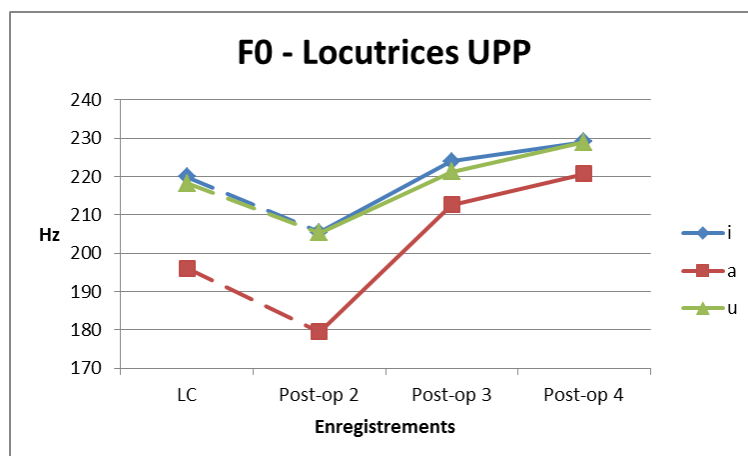


Figure 12 : Valeurs moyennes de la fréquence fondamentale (F0) en Hz pour [i a u] - locutrices du groupe UPP

Après l'opération, en post-opératoire 2, la fréquence fondamentale moyenne, mesurée à partir des trois voyelles [i a u] produites par les locutrices, est plus basse, soit 205,48 Hz, 179,52 Hz et 205,30 Hz respectivement, que celle obtenue à partir des voyelles produites par les locuteurs contrôles, soit 219,9 Hz, 196,04 Hz, et 218,14 Hz, respectivement (voir Figure 12). Dès la phase post-opératoire 3, la fréquence fondamentale augmente pour les trois voyelles ; elle est alors de 223,48 Hz pour [i], 212,54 Hz pour [a] et 221,14 Hz pour [u]. Cette augmentation se poursuit jusqu'en post-opératoire 4, où elle est alors de 229,02 Hz pour [i], 220,68 Hz pour [a] et 228,86 Hz pour [u]. Dès la phase post-opératoire 3, la fréquence fondamentale des trois voyelles est plus élevée que celle relevée à partir des productions des locuteurs contrôles. Ce phénomène est observable pour toutes les voyelles étudiées, mais il est particulièrement remarquable pour la voyelle [a].

Tableau 18 : Valeurs moyennes de la fréquence fondamentale (F0) en Hz pour [i a u] pour les locutrices du groupe UPP. Les écarts-types sont donnés entre parenthèses

UPPFEM	LC	Post-op 2	Post-op 3	Post-op 4
[i]	219,9 (11,1)	205,48 (51,85)	223,92 (64,68)	229,02 (43,56)
[a]	196,04 (21,33)	179,52 (63,63)	212,54 (61,19)	220,68 (40,34)
[u]	218,14 (14,06)	205,3 (66,12)	221,14 (65,58)	228,86 (42,29)

Les écarts-types, calculés à partir des productions des locuteurs contrôles, sont faibles comme attendus, 11,10 Hz pour [i], 21,33 Hz pour [a] et 14,06 Hz pour [u]. En ce qui concerne les patients, leurs écarts-types sont toujours plus importants quelles que soient la voyelle et la phase d'enregistrement considérées (voir Tableau 18), et témoignent ainsi d'une certaine variabilité. Cette variabilité est plus importante dans les phases d'enregistrement post-opératoire 2 et 3, puisque les écarts-types sont alors de 51,85 Hz et 64,68 Hz pour [i], 63,63 Hz et 61,19 Hz pour [a] et 66,12 Hz et 65,58 Hz pour [u]. Ces écarts-types se réduisent lors de la phase post-opératoire 4, mais restent toutefois importants (43,56 Hz pour [i], 40,34 Hz pour [a] et 42,29 Hz pour [u]). Afin de comprendre si cette variabilité est le fait de phénomènes inter ou intra individus, il nous semble important d'étudier *infra* les locuteurs de façon séparée.

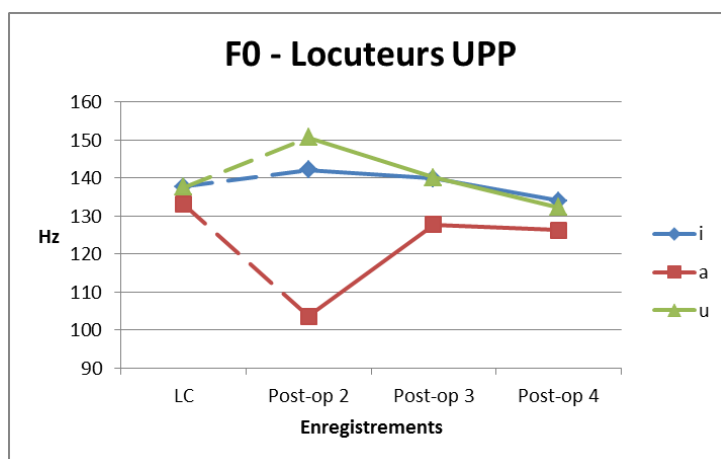


Figure 13 : Valeurs moyennes de la fréquence fondamentale (F0) en Hz pour [i a u] - locuteurs du groupe UPP

En ce qui concerne les locuteurs masculins, le scénario est différent en fonction de la voyelle étudiée (voir Figure 13). La fréquence fondamentale de la voyelle [a] suit les mêmes tendances que celle observée pour les locutrices, à savoir qu'elle est mesurée plus basse (103,60 Hz) en post-opératoire que celle obtenue à partir des productions des locuteurs contrôles (133,70 Hz). A partir de la phase post-opératoire 3, la fréquence fondamentale des patients augmente pour atteindre 127,70 Hz ; cette valeur est maintenue en post-opératoire 4 (126,30 Hz). La fréquence fondamentale des voyelles [i] et [u], en revanche, subit très peu de modifications, en fonction de la phase d'enregistrement considérée, et reste très proche des valeurs mesurées à partir des

productions des locuteurs contrôles (137 Hz pour [i] et [u]). La fréquence fondamentale de la voyelle [i] est de 142,12 Hz en post-opératoire 2, 139,90 Hz en post-opératoire 3 et 133,95 Hz en post-opératoire 4. La fréquence fondamentale de la voyelle [u] est de 150,65 Hz en post-opératoire 2, 140,05 Hz en post-opératoire 3 et 132,20 Hz en post-opératoire 4.

Tableau 19 : Valeurs moyennes de la fréquence fondamentale (F0) en Hz pour [i a u] pour les locuteurs du groupe NPP. Les écarts-types sont donnés entre parenthèses

UPPHOM	LC	Post-op 2	Post-op 3	Post-op 4
[i]	137,7 (1,52)	142,15 (25,94)	139,9 (24,18)	133,95 (29,19)
[a]	133 (1,29)	103,6 (18,18)	127,7 (18,46)	126,3 (19,73)
[u]	137,6 (1,23)	150,65 (22,22)	140,05 (20,16)	132,2 (25,67)

Les enregistrements des patients sont également marqués par une variabilité plus grande, traduite par des écarts-types importants (voir Tableau 19) ; les écarts-types des locuteurs contrôles restent faibles. Signalons que cette variabilité ne se réduit pas comme attendu, lors de la phase post-opératoire 4 ; elle est même plus importante pour les trois voyelles [i], [a] et [u] (29,19 Hz, 19,73 Hz et 25,67 Hz, respectivement) que celle observée lors des phases précédentes ([i] : 25,94 Hz en post-opératoire 3 et 24,18 Hz en post-opératoire 3 ; [a] : 18,18 Hz en post-opératoire 3 et 18,46 Hz en post-opératoire 3 ; [u] : 22,22 Hz en post-opératoire 3 et 20,16 Hz en post-opératoire 3).

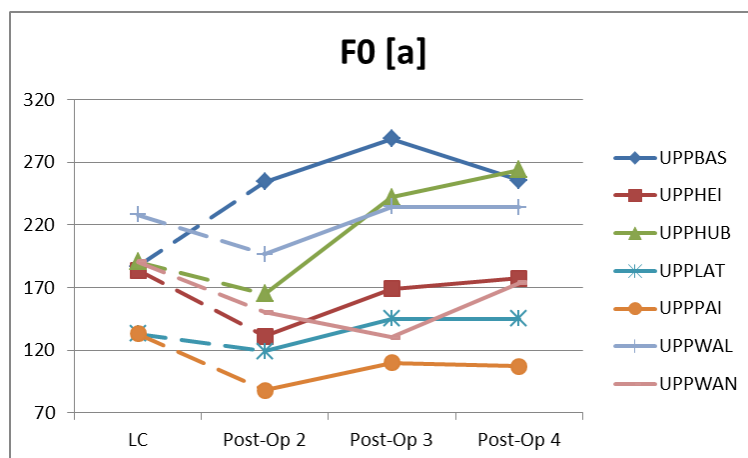


Figure 14 : Valeurs moyennes de la fréquence fondamentale (F0) en fonction du locuteur pour la voyelle [a]

Afin de comprendre les phénomènes de variabilité déjà mis en lumière par l'étude des écarts-types *supra*, nous étudierons la fréquence fondamentale de la voyelle [a] de chaque locuteur individuellement. Nous faisons ce choix (voir également V.6.1.) en suivant les recommandations de la Société Européenne de Laryngologie (Dejonckere et *al.* 2001 et 2003).

Pour tous les locuteurs, sauf UPPBAS (en bleu foncé), la fréquence fondamentale est mesurée plus basse en post-opératoire 2 que celle mesurée à partir des productions des locuteurs contrôles. L'importance de la baisse est toutefois à considérer en fonction du locuteur. Ainsi, pour UPPLAT (en turquoise) ou UPPHUB (en vert), la fréquence fondamentale n'est mesurée que légèrement plus basse en post-opératoire 2 (119,20 Hz et 165,30 Hz respectivement) que celle de son locuteur contrôle (133 Hz et 190 Hz respectivement). Pour les autres locuteurs, UPPHEI (en rouge), UPPPAI (en orange), UPPWAL (en bleu ciel) et UPPWAN (en rose), la phase post-opératoire 2 est marquée par des mesures de la fréquence fondamentale nettement plus basses (131 Hz, 88 Hz, 196,60 Hz et 150 Hz, respectivement) que celle de leur locuteur contrôle (183 Hz, 133 Hz, 228,10 Hz et 198,70 Hz, respectivement).

Signalons que pour deux locutrices, UPPHEI (en rouge) et UPPWAN (en rose), la fréquence fondamentale en post-opératoire 2 est mesurée (131 Hz et 150,30 Hz,

respectivement) en dessous du seuil de 160 Hz, qui est supposé différencier la fréquence fondamentale d'une femme de celle d'un homme (Vaissière, 2006).

À partir de la phase post-opératoire 3, les moyennes des fréquences fondamentales augmentent pour tous les locuteurs, sauf pour UPPWAN (en rose) mesurée à 130,27 Hz. Pour UPPHEI (en rouge), UPPLAT (en turquoise) UPPPAI (en orange) et UPPWAL (en bleu ciel), les valeurs de la fréquence fondamentale commencent à se rapprocher des valeurs des locuteurs contrôles. Elle est alors mesurée à 169 Hz, 145,20 Hz, 110,20 Hz et 234,70 Hz, respectivement. Pour UPPBAS (en bleu foncé), la fréquence fondamentale continue à augmenter et à être quantifiée au-dessus de celle de son locuteur contrôle, soit 288,72 Hz.

La fréquence fondamentale de UPPHUB (en vert) est mesurée plus élevée (242 Hz), dès la phase post-opératoire 3, que celle de son locuteur contrôle (190 Hz). Cette tendance se poursuit en post-opératoire 4, où sa fréquence fondamentale est alors mesurée à 264 Hz.

La phase post-opératoire 4 est marquée par la relative stagnation de la plupart des données de la fréquence fondamentale, qui sont alors comparables à celles observées en post-opératoire 3, et se rapprochent encore plus des valeurs des locuteurs contrôles.

Tableau 20 : Valeurs moyennes de la fréquence fondamentale (F0) en fonction du locuteur pour la voyelle [a] – Groupe NPP. Les écarts-types sont donnés entre parenthèses

UPP	LC	Post-Op 2	Post-Op 3	Post-Op 4
UPPBAS	187 (3,47)	254,4 (82,48)	288,72 (5,33)	255,63 (8,76)
UPPHEI	183,7 (9,17)	131 (37,10)	169 (5,89)	177,4 (7,85)
UPPHUB	190,7 (4,63)	165,3 (15,34)	242 (18,58)	264 (13,74)
UPPLAT	133 (1,33)	119,2 (4,57)	145,2 (3,77)	145,2 (3,77)
UPPPAI	133 (1,33)	88 (11,67)	110,2 (5,03)	107,4 (3,78)
UPPWAL	228,1 (1,91)	196,6 (45,25)	234,7 (2,21)	234,3 (4,11)
UPPWAN	190,7 (17,37)	150,3 (35,38)	130,27 (39,56)	173,73 (16,17)

Comme détaillé dans le Tableau 20, les écarts-types des locuteurs contrôles sont faibles, ce qui correspond à nos attentes. En ce qui concerne les patients, les écarts-types sont particulièrement importants, lors de la phase d'enregistrement post-opératoire 2 et notamment pour UPPBAS (82,48 Hz), UPPHEI (37,10 Hz), UPPWAL (45,25 Hz) et UPPWAN (35,38 Hz). Les écarts-types se réduisent à partir de la phase post-opératoire 3, sauf pour UPPWAN (39,56 Hz). La phase post-opératoire 4 est marquée par des écarts-types faibles, quel que soit le locuteur étudié. La variabilité esquissée *supra* est donc le résultat de différences de productions inter-individuelles mais également celui de réalisations intra-individuelles.

Pour résumer :

La fréquence fondamentale, calculée pour des locuteurs féminins du groupe UPP, est mesurée plus basse, en post-opératoire 2, que celle mesurée à partir des productions des locuteurs contrôles, et ce quelle que soit la voyelle étudiée. Dès la phase post-opératoire 3, la fréquence fondamentale moyenne des trois voyelles augmentent ; ce phénomène se poursuit en post-opératoire 4. La fréquence fondamentale de ces voyelles est alors mesurée plus haute que celle des locuteurs contrôles.

Les locuteurs masculins présentent la même tendance pour la voyelle [a] uniquement. La fréquence fondamentale des voyelles [i] et [u] est peu modifiée par l'immobilité laryngée.

Les phases post-opératoires précoces 2 et 3 sont marquées par des écarts-types plus importants ; leur réduction s'amorce lors de la phase post-opératoire 3, et se poursuit jusqu'en post-opératoire 4.

L'étude de la fréquence fondamentale moyenne individuelle de la voyelle [a] nous a permis de mettre au jour la grande variabilité inter et intra individuelle dans les productions des différents locuteurs. Une locutrice présente un schéma inverse, puisque pour celle-ci, la fréquence fondamentale est mesurée plus haute, dès la phase post-opératoire 2, que celle de son locuteur contrôle. Les autres locuteurs présentent des scénarios comparables, à savoir que la fréquence fondamentale est mesurée plus basse que celle du locuteur contrôle, en post-opératoire 2, avant d'augmenter à partir de la phase post-opératoire 3.

Les phases d'enregistrement post-opératoire 2 et 3 sont marquées par des écarts-types plus importants, ce qui témoigne d'une certaine variabilité intra-locuteur durant ces phases.

VII.4.2.1. Harmonics-to-Noise Ratio (HNR)

Nous avons considéré le Harmonics-to-Noise Ratio comme un indice de raucité. Les valeurs habituellement attendues pour un locuteur non pathologique sont de 20 dB pour [i] et [a] et de 30 dB pour [u] (cf., par ex. Mathieson, 2001).

Les analyses de variance ont permis de mettre au jour un effet principal *phase d'enregistrement* pour les voyelles [a] $F(2,12) = 4,15$ $p < 0,0069$ et [u] $F(2,12) = 3,54$ $p < 0,036$. L'effet principal repose, pour les deux voyelles, sur une différence significative entre la phase d'enregistrement post-opératoire 2 et toutes les autres phases d'enregistrement.

Un effet principal *voyelle* [$F(2,12) = 36,15$, $p < 0,000008$] a également pu être identifié, ce qui suppose que même si le spectre des voyelles est modifié, elles sont tout de même réalisées distinctement les unes par rapport aux autres.

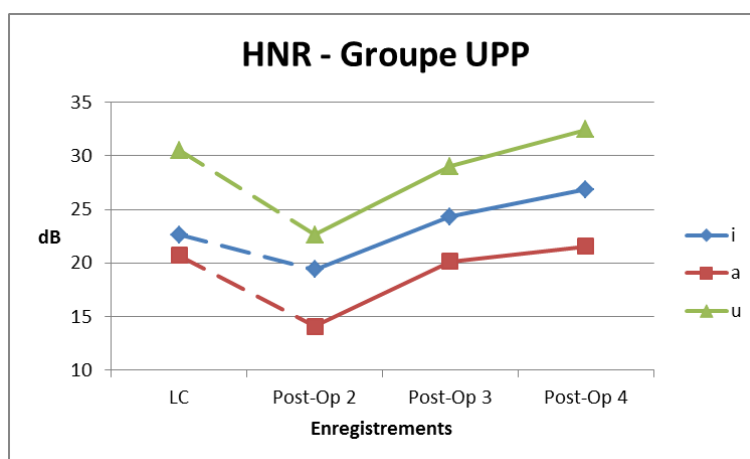


Figure 15 : Valeurs moyennes de HNR (dB) pour [i a u] – Groupe UPP

Les valeurs de HNR, calculées à partir des productions des locuteurs contrôles, sont conformes aux mesures attendues pour un locuteur non pathologique, soit 22,61 dB pour [i], 20,72dB pour [a] et 30,54 pour [u]. Les écarts-types sont par ailleurs faibles, soit 5,82 dB, 3,08 dB et 2,29 dB, respectivement.

Les valeurs de HNR, mesurées en post-opératoire 2, sont inférieures à celles mesurées à partir des productions des locuteurs contrôles pour [i a u]. L'écart est particulièrement notable pour les voyelles [a] et [u], pour lesquelles le HNR est calculé à 14,06 dB et 22,64 dB, respectivement, en post-opératoire 2. La voyelle [i] est également mesurée plus basse à 19,37 dB ; l'écart avec les productions des locuteurs contrôles et sa mesure reste proche des valeurs attendues pour un locuteur non pathologique.

Dès la phase post-opératoire 3, les valeurs de HNR deviennent conformes, et sont alors similaires aux valeurs de références, soit 24,30 dB pour [i], 20,13 dB pour [a] et 29,03 dB pour [u]. L'augmentation des valeurs de HNR se poursuit jusqu'en post-opératoire 4, où le ratio est alors de 26,87 dB pour [i], 21,56 dB pour [a] et 32,46 dB pour [u].

Tableau 21 : Valeurs moyennes de HNR (dB) pour [i a u] – Groupe NPP. Les écarts-types sont donnés entre parenthèses

UPP	LC	Post-Op 2	Post-Op 3	Post-Op 4
[i]	22,61 (5,82)	19,37 (6,37)	24,3 (8,18)	26,87 (5,88)
[a]	20,72 (3,08)	14,06 (7,76)	20,13 (6,77)	21,56 (4,11)
[u]	30,54 (2,29)	22,64 (10,03)	29,03 (8,05)	32,46 (5,29)

Les phases d'enregistrement des patients sont caractérisées par des écarts-types plus importants, quelle que soit la voyelle étudiée. Le Tableau 21 illustre ce phénomène. La variabilité est particulièrement importante, lors de la première phase d'enregistrement post-opératoire, et cela notamment pour la voyelle [u] (10,03 dB). En post-opératoire 3, la variabilité se réduit, reste importante pour [i] (8,18 dB) et [u] (8,05), mais augmente pour [a] (6,77 dB). La phase post-opératoire 4 témoigne d'une variabilité réduite, puisque les écarts-types sont les plus faibles mesurés à partir des productions des patients, quelle que soit la phase considérée, soit 5,88 dB pour [i], 4,11 dB pour [a] et 5,29 dB pour [u].

La voyelle [a] sera maintenant étudiée par individu, afin de comprendre si la variabilité relevée précédemment est le fait de différences inter ou intra individuelles.

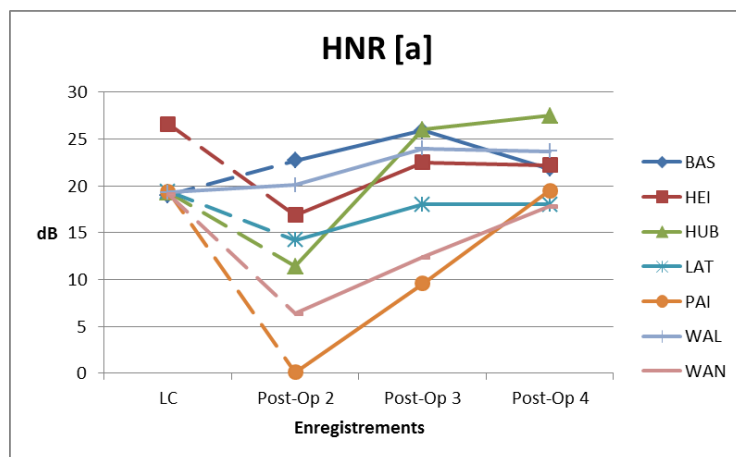


Figure 16 : Valeurs moyennes de HNR (dB) en fonction du locuteur pour la voyelle [a]

La Figure 16 témoigne d'une certaine variabilité inter-locuteurs. Ainsi, si pour tous les locuteurs, sauf UPPBAS (en bleu foncé), les valeurs de HNR sont mesurées plus basses, en post-opératoire 2, que celles des locuteurs contrôles, l'amplitude de la différence dépend étroitement du locuteur étudié. Ici aussi, la standardisation des valeurs, qui commence dès la phase post-opératoire 3, ne s'effectue pas au même rythme selon le locuteur observé.

Comme pour l'étude de la fréquence fondamentale, la locutrice UPPBAS ne suit pas la tendance générale. Son HNR moyen est mesuré plus élevé en post-opératoire 2 (22,70 dB) que celui de son locuteur contrôle (19 dB). Cette observation s'accroît en post-opératoire 3 (25,90 dB). La phase post-opératoire 4 semble marquer une inversion dans la tendance, puisque le HNR baisse de façon relativement importante, pour s'approcher des valeurs standards attendues (21,80 dB).

En post-opératoire 2, les valeurs moyennes de HNR des trois locuteurs HPPHEI (en rouge), UPPLAT (en turquoise) et UPPWAL (en rose) ne sont pas si inférieures aux valeurs attendues pour un locuteur non pathologique soit 16,87 dB, 14,20 dB et 20,10 dB. Les autres locuteurs présentent des valeurs de HNR plus basses et plus éloignées des valeurs attendues, c'est le cas de UPPHUB (11,40 dB), UPPWAN (6,40 dB). La voix du locuteur UPPAI est tellement dégradée que son HNR est alors mesuré à 0,10 dB, ce qui indique qu'il y a autant de bruits que d'harmoniques dans sa voix.

Les valeurs de HNR augmentent pour tous les locuteurs à partir de post-opératoire 3. Pour certains locuteurs, les valeurs sont, dès cette phase, standards ; c'est le cas de UPPHEI (en rouge), dont le HNR est de 22,5 dB et UPPLAT (en turquoise),

dont le HNR est de 18 dB. Les valeurs de UPPBAS (en bleu foncé), UPPHUB (en vert) et UPPWAL (en bleu clair) sont même mesurées au-dessus des valeurs attendues, soit 25,9 dB, 22,5 dB, 26 dB et 24 dB respectivement. UPPPAI (en orange) et UPPWAN (en rose) conservent, malgré une hausse importante par rapport à la phase post-opératoire 3, des valeurs de HNR inférieures à celles attendues, et de leur locuteur contrôle, soit 9,60 dB et 12,40 dB.

En post-opératoire 4, les valeurs de HNR sont proches des valeurs standards attendues pour tous les locuteurs étudiés. Notons que le HNR moyen de UPPBAS (en bleu foncé) et, particulièrement pour UPPHUB (en vert), sont supérieurs à celui de leur locuteur contrôle, soit 21,8 dB et 27,5 dB. Les valeurs de UPPHEI (en rouge), UPPLAT (en turquoise) et UPPWAL (en bleu clair), qui étaient déjà proches des valeurs standards lors de la phase d'enregistrement précédente, se maintiennent à 22,5 dB, 18 dB et 23,70 dB, respectivement. En ce qui concerne UPPPAI (en orange) et UPPWAN (en rose), dont les HNR étaient mesurés plus bas en post-opératoire 3, ils indiquent des mesures en hausse en post-opératoire 4, soit 19,5 dB et 17,80 dB, respectivement. Le HNR moyen de UPPPAI est alors proche des valeurs attendues pour un locuteur pathologique. La constante augmentation des valeurs de UPPWAN laisse supposer que la tendance se poursuivra. Ces valeurs sont d'ores et déjà satisfaisantes.

Tableau 22 : Valeurs moyennes de HNR (dB) en fonction du locuteur pour la voyelle [a] – Groupe NPP.

Les écarts-types sont donnés entre parenthèses

UPP	LC	Post-Op 2	Post-Op 3	Post-Op 4
UPPBAS	19 (1,76)	22,7 (4)	25,9 (2,54)	21,8 (4,14)
UPPHEI	26,6 (1,35)	16,87 (4,02)	22,5 (1,52)	22,2 (0,92)
UPPHUB	19,3 (1,89)	11,4 (2,55)	26 (1,64)	27,5 (3,14)
UPPLAT	19,4 (1,71)	14,2 (2,62)	18 (3,68)	18 (3,68)
UPPPAI	19,4 (1,71)	0,1 (0,57)	9,6 (2,67)	19,5 (0,71)
UPPWAL	19,3 (1,88)	20,1 (1,45)	24 (1,05)	23,7 (1,56)
UPPWAN	19,1 (1,79)	6,4 (2,91)	12,4 (4,33)	17,8 (2,62)

Le Tableau 22 nous renseigne notamment sur les écarts-types, donnés entre parenthèses. Il semble que la variabilité remarquée *infra* soit le résultat de différences

inter-individuelles puisque, quels que soient la phase d'enregistrement post-opératoire et le locuteur étudiés, les écarts-types restent faibles (entre 0,57 dB pour l'écart-type le plus faible et 4,14 dB pour l'écart-type le plus important). Les locuteurs sont donc relativement constants dans leurs productions ; la variabilité est alors à comprendre plutôt comme le résultat de comportements et de stratégies individuelles.

Pour résumer :

Le HNR moyen des trois voyelles [i a u] est mesuré plus bas, en post-opératoire 2, que celui des locuteurs contrôles. L'augmentation et la standardisation des valeurs moyennes commencent lors de la phase post-opératoire 3, et se poursuivent jusqu'en post-opératoire 4.

Les valeurs du HNR des locuteurs pathologiques sont inférieures à celles des locuteurs contrôles, uniquement lors de la première phase d'enregistrement, c'est-à-dire en post-opératoire 2. Dès la phase post-opératoire 3, les valeurs augmentent pour tous les locuteurs. Les mesures de la plupart des locuteurs atteignent déjà des valeurs standards. Enfin, en post-opératoire 4, les valeurs du HNR des patients sont comparables à celles des locuteurs contrôles.

Les forts écarts-types constatés lors de l'étude des valeurs moyennes du HNR sont le résultat d'une variabilité inter-individuelle ; les écart-types intra-locuteurs restent faibles.

VII.4.3. Jitter et Shimmer

Les analyses de variance ont permis d'identifier un effet principal *phase d'enregistrement* pour la variable jitter [$F(3,18) = 7,24$, $p < 0,0022$]. Les tests *post-hoc* révèlent que les productions des patients en post-opératoire 2 (1,28%) sont significativement différentes des mesures effectuées à partir de toutes les autres phases d'enregistrement (LC = ,43% ; Post-op 3 = 0,68% ; Post-op 4 = 0,42%).

Tableau 23 : Valeurs moyennes de jitter en % – Groupe UPP. Les écarts-types sont donnés entre parenthèses

% jitter	LC	Post-op 2	Post-op 3	Post-op 4
[a]	0,43 (0,14)	1,28 (0,84)	0,68 (0,49)	0,42 (0,23)

Le jitter est considéré comme un indice de perturbation du signal, le seuil normal/pathologique étant habituellement fixé à 1,04% (*cf.* par ex. manuel PRAAT[®]).

Le jitter des locuteurs contrôles est conforme aux valeurs attendues, il est faible et bien évidemment en dessous du seuil normal/pathologique, soit 0,43%.

En ce qui concerne les patients en phases post-opératoires, leur jitter est mesuré au-dessus du seuil normal/pathologique, en post-opératoire 2 à 1,28%, puis à nouveau en dessous de ce seuil à partir de post-opératoire 3 (0,68%), même s'il reste toutefois supérieur aux valeurs de référence des contrôles. La baisse se poursuit jusqu'en post-opératoire 4, où le jitter est alors calculé à 0,42%, soit une valeur comparable à celle des locuteurs contrôles. Le Tableau 23 illustre ces tendances.

Les phases post-opératoires sont caractérisées par des écarts-types plus importants (voir Tableau 23), soit 0,84%, 0,49% et 0,68% respectivement, que la phase des locuteurs contrôles, soit 0,14%. Notons que si les écarts-types sont plus importants pour les patients, ils diminuent toutefois au cours du temps, parallèlement à la diminution de jitter.

Tableau 24 : Valeurs moyennes de jitter en % en fonction du locuteur – Groupe UPP. Les écarts-types sont donnés entre parenthèses

% jitter	LC	Post-op 2	Post-op 3	Post-op 4
UPPBAS	0,36 (0,07)	0,56 (0,16)	0,41 (0,22)	0,50 (0,19)
UPPHEI	0,52 (0,23)	1,59 (0,52)	0,51 (0,24)	0,40 (0,05)
UPPHUB	0,36 (0,07)	2,39 (1,02)	0,45 (0,13)	0,32 (0,09)
UPPLAT	0,45 (0,12)	1,02 (0,20)	0,37 (0,10)	0,37 (0,10)
UPPPAI	0,36 (0,10)	0,96 (0,62)	1,17 (0,18)	0,33 (0,03)
UPPWAL	0,50 (0,14)	0,48 (0,18)	0,30 (0,10)	0,22 (0,04)
UPPWAN	0,44 (0,06)	1,97 (0,56)	1,48 (0,57)	0,82 (0,35)

Il semble, suite au Tableau 24, que les écarts-types importants relevés précédemment, lors de la phase post-opératoire 2, soient le résultat d'une variabilité inter et intra individuelle. En effet, pour certains locuteurs (voir par exemple UPPBAS dont le jitter est mesuré à 0,56% (0,16) ou UPPWAL dont le jitter est mesuré à 0,48% (0,18)), le jitter et les écarts-types sont particulièrement bas et comparables à ceux d'un

locuteur non pathologique, ce qui témoigne non seulement d'une absence de perturbation de leur fréquence fondamentale à court terme, mais aussi d'une régularité dans leurs productions. Pour d'autres locuteurs en revanche, comme UPPHEI, dont le jitter est mesuré à 1,59% (0,52) et UPPHUB, dont le jitter est mesuré à 2,39% (1,02), ce paramètre est non seulement situé au-dessus du seuil normal/pathologique, mais les écarts-types sont également plus importants, ce qui indique une certaine variabilité dans leurs productions. Le locuteur UPPLAT, dont le jitter est mesuré à 1,02% (0,20), se situe à la frontière de ces deux tendances.

Le jitter et les écarts-types se réduisent pour tous les locuteurs, à partir de la phase post-opératoire 3. Seul UPPPAI indique des résultats différents : le jitter de UPPPAI augmente (mesuré à 1,17%, écart-type 0,18) et dépasse alors le seuil normal/pathologique, alors même que ce ne fut pas le cas en post-opératoire 2. Celui de UPPWAN diminue (1,58% écart-type 0,57), mais reste supérieur à celui attendu pour un locuteur non pathologique. Pour tous les autres locuteurs, le jitter est mesuré plus bas qu'en post-opératoire 2, et en dessous du seuil normal/pathologique, soit à 0,41% (UPPBAS), 0,51% (UPPHEI), 0,45% (UPPHUB), 0,37% (UPPLAT) et 0,30% (UPPWAL).

La tendance se poursuit en post-opératoire 4 ; le jitter est alors non seulement mesuré en dessous du seuil normal/pathologique pour tous les patients, mais il est également plus bas que celui relevé en post-opératoire 3, à savoir à 0,50% (UPPBAS), 0,40% (UPPHEI), 0,32% (UPPHUB), 0,37% (UPPLAT), 0,33% (UPPPAI) 0,22% (UPPWAL) et 0,82% (UPPWAN).

Les phases post-opératoires 3 et 4 sont également caractérisées par des écarts-types plus faibles (et comparables à ceux des locuteurs contrôles) que ceux relevés en post-opératoire 2 (voir Tableau 24 pour les détails de ces valeurs).

Les mêmes tendances sont observables pour la mesure de shimmer. Cette variable est significativement modifiée par le facteur *phase d'enregistrement* [$F(3,18) = 8,33$, $p < 0,0012$]. Ces résultats vont dans le sens de modifications importantes au cours du temps. En effet, les tests *post-hoc* indiquent que les productions des patients en post-opératoire 2 sont significativement différentes des productions enregistrées lors de toutes les autres phases.

Tableau 25 : Valeurs moyennes de shimmer en % – Groupe UPP. Les écarts-types sont donnés entre parenthèses

% shimmer	LC	Post-op 2	Post-op 3	Post-op 4
[a]	2,80 (0,60)	8,80 (5,51)	5,23 (3,75)	3,39 (1,51)

En ce qui concerne le shimmer, le seuil normal/pathologique est généralement fixé à 3,81% (cf. le manuel PRAAT[®]). La mesure de shimmer, détaillée dans le Tableau 25, suit la même tendance globale que le jitter, à savoir que ces valeurs sont plus importantes (5,51%) que celles mesurées à partir des locuteurs contrôles (2,80%) en post-opératoire 2. Les valeurs diminuent à partir de la phase post-opératoire 3 (5,23%), ce qui se poursuit également en post-opératoire 4 (1,51%). Toutefois, l'analyse fine des valeurs de shimmer montre que les mesures sont supérieures au seuil normal/pathologique, jusqu'en post-opératoire 3, et en restent proche de ce seuil en post-opératoire 4.

Les écarts-types sont également plus importants dans les phases post-opératoires, soit 5,51%, 3,75% et 1,51%. Le temps et la rééducation orthophonique permettent toutefois de les réduire.

Tableau 26 : Valeurs moyennes de shimmer en fonction du locuteur – Groupe UPP. Les écarts-types sont donnés entre parenthèses

% shimmer	LC	Post-op 2	Post-op 3	Post-op 4
UPPBAS	3,16 (0,23)	6,07 (2,80)	4,61 (2,42)	3,90 (1,15)
UPPHEI	2,37 (0,61)	4,47 (1,27)	2,74 (0,33)	2,56 (0,28)
UPPHUB	3,16 (0,23)	9,82 (3,62)	3,49 (0,99)	2,27 (0,42)
UPPLAT	2,66 (0,65)	6,01 (1,22)	2,93 (1,98)	2,94 (1,98)
UPPPAI	2,67 (0,58)	16,67 (5,77)	10,94 (1,95)	3,96 (0,04)
UPPWAL	3,20 (0,55)	3,80 (0,49)	2,31 (0,86)	2,4 (0,36)
UPPWAN	2,37 (0,61)	14,64 (2,41)	9,56 (3,33)	5,71 (1,35)

Le Tableau 26 permet d'étudier les patients individuellement et de comprendre que la variabilité est le résultat de différences inter et intra individuelles, notamment lors de la phase post-opératoire 2.

En post-opératoire 2, tous les locuteurs indiquent des valeurs de shimmer supérieures au seuil normal/pathologique, soit 6,07% (UPPBAS), 4,47% (UPPHEI), 9,82% (UPPHUB), 6,01% (UPPLAT), 16,67% (UPPPAI) et 3,80% (UPPWAL) et 14,64% (UPPWAN). Le shimmer de UPPPAI et de UPPWAL est particulièrement important ; ces locuteurs avaient également des mesures de jitter plus importantes que celles des autres patients.

Lors de la phase post-opératoire 3, seuls UPPBAS (4,61%), UPPPAI (10,94%) et UPPWAN (9,56%) présentent des valeurs supérieures au seuil normal/pathologique. Tous les locuteurs vont toutefois dans le sens d'une diminution des valeurs de shimmer, entre la phase post-opératoire 2 et post-opératoire 3 (UPPHEI = 2,74% UPPHUB = 3,49% UPPLAT = 2,96 % UPPWAL = 2,31%).

Cette tendance se poursuit en post-opératoire 4, les valeurs de shimmer étant alors les plus basses pour les patients, toutes phases d'enregistrement confondues, soit 3,90% (UPPBAS), 2,56% (UPPHEI), 2,27% (UPPHUB), 2,94% (UPPLAT), 3,96% (UPPPAI) et 2,40% (UPPWAL) et 5,71% (UPPWAN). Signalons que UPPBAS, UPPAI et UPPWAN restent au-dessus du seuil normal/pathologique.

En ce qui concerne les écarts-types, ils sont plus importants lors de la phase post-opératoire 2 (entre 0,49% pour le plus faible et 5,77% pour le plus élevé), et se réduisent à partir de post-opératoire 3 (entre 0,33% pour le plus faible et 3,33% pour le plus élevé). Cette tendance se poursuit en post-opératoire 4 (entre 0,28% et 1,98%).

Pour résumer :

Les valeurs de jitter et de shimmer sont plus importantes, en post-opératoire 2, que celles calculées à partir des productions des locuteurs contrôles. Le temps permet toutefois une standardisation de ces deux mesures, qui sont alors quantifiées en-dessous du seuil normal/pathologique, en post-opératoire 4.

La variabilité est plus marquée lors des enregistrements en post-opératoire 2. Cela est le résultat de différences intra et inter-individuelles.

Signalons enfin que l'étude des locuteurs pris individuellement a permis de montrer une certaine variabilité inter-individuelle, puisque si le temps et la rééducation orthophonique ont un impact positif sur ces mesures, la récupération peut se révéler plus lente, en fonction du locuteur étudié.

Ces deux mesures, de jitter et de shimmer, peuvent être étudiées en parallèle puisque ce sont les locuteurs qui présentaient une plus forte présence de jitter qui indiquent également une plus forte présence de shimmer.

VII.4.4. Valeurs formantiques

Les analyses de variance n'ont pas permis de déterminer un effet principal *phase d'enregistrement* pour la variable *F1* ($p=ns$).

En ce qui concerne la variable *F2*, les valeurs de *F2* sont significativement modifiées par le facteur *phase d'enregistrement* [$F(3,18) = 4,22$ $p < 0,021$]. Il existe une différence significative entre la phase d'enregistrement post-opératoire 2 et toutes les autres phases d'enregistrement.

Enfin, signalons que le facteur *voyelle* pour la variable *F1* s'est révélé significatif [$F(2,12) = 302,40$, $p < 10^{-6}$] et également pour la variable *F2* [$F(2,12) = 220,48$, $p < 10^{-6}$], ce qui semble indiquer que si l'espace vocalique subit de légères modifications, il reste néanmoins géométriquement conventionnel.

En ce qui concerne la voyelle [i], les valeurs formantiques (soit *F1* et *F2*) des patients ne sont pas significativement différentes ($p=ns$) de celles des locuteurs contrôles (*F1* 315 Hz et *F2* 2456Hz). Cette observation est juste, quelle que soit la phase d'enregistrement post-opératoire étudiée (voir Tableau 27). Les valeurs formantiques des patients sont les suivantes : 317 Hz (*F1*) et 2446 Hz (*F2*) en post-opératoire 2 ; 318 Hz (*F1*) et 2413 Hz (*F2*) en post-opératoire 3 ; 332 Hz (*F1*) et 3419 Hz (*F2*) en post-opératoire 4.

Les valeurs formantiques des voyelles [a] et [u] subissent de légères modifications en fonction des phases d'enregistrement, par rapport à celles des locuteurs contrôles. Le premier formant du [a] est mesuré à 752 Hz pour les locuteurs contrôles, à 787 Hz en post-opératoire 2, puis à 733 Hz en post-opératoire 3, et enfin à 752 Hz en post-opératoire 4. Ces résultats vont dans le sens d'une expansion de l'espace vocalique sur l'axe vertical (voir VI.4.5. pour des informations relatives à l'aire de l'espace vocalique). Les modifications des valeurs de *F2* suivent le même schéma, puisque les modifications les plus importantes sont à observer en post-opératoire 2, *F2* étant alors mesuré à 1379 Hz. Les valeurs des locuteurs contrôles (1328 Hz), en post-opératoire 3

(1295 Hz) et post-opératoire 4 (1301 Hz) sont relativement similaires. En ce qui concerne la voyelle [u], les valeurs de F1 sont les suivantes : 349 Hz pour les locuteurs contrôles, 394 Hz en post-opératoire 2, puis 361 Hz en post-opératoire 3 et enfin 353 Hz en post-opératoire 4. Ces valeurs vont également dans le sens d'une centralisation de l'espace vocalique sur l'axe vertical. Les valeurs de F2 subissent des modifications plus importantes, puisque F2 est mesuré à 745 Hz pour les locuteurs contrôles, à 1057 Hz en post-opératoire 2, à 799 Hz en post-opératoire 3 et à 844 Hz en post-opératoire 4.

Les écarts-types, calculés à partir des valeurs de F1, sont toujours plus importants pour les patients que ceux des locuteurs contrôles, quelles que soient la phase d'enregistrement post-opératoire et la voyelle observées, soit pour [i] : 32,89 Hz (LC), 52,19 Hz (post-opératoire 2), 57,15 Hz (post-opératoire 3) et 83,54 Hz (post-opératoire 4) ; pour [a] : 106,77 Hz (LC), 162,41 Hz (post-opératoire 2), 150,79 Hz (post-opératoire 3) et 168,16 Hz (post-opératoire 4) ; pour [u] : 52,03 Hz (LC), 118,33 Hz (post-opératoire 2), 63,36 Hz (post-opératoire 3) et 68,74 Hz (post-opératoire 4). Notons que la variabilité a tendance à se réduire avec le temps et la rééducation orthophonique.

Les écarts-types, calculés à partir des valeurs de F2, affichent le même scénario, à savoir qu'ils sont plus importants pour les patients, et notamment en post-opératoire 2 – soit 349 Hz pour [i], 267,75 Hz pour [a] et 450 Hz pour [u]. Les écarts-types restent très importants dans les phases suivantes, soit 342,72 Hz pour [i], 143,83 Hz pour [a] et 168,70 Hz pour [u] en post-opératoire 3 ; et 358,86 Hz pour [i], 168,16 Hz pour [a] et 213,83 Hz pour [u] en post-opératoire 4.

Dans la mesure où la plupart de nos locuteurs développent des comportements similaires aux tendances que nous venons de décrire, nous ne proposons pas d'analyse individuelle, les valeurs étant toutefois disponibles en annexes.

Tableau 27 : Valeurs formantiques moyennes – Groupe UPP. Les écarts-types sont donnés entre parenthèses

Voyelles	Formants	LC	Post-op 2	Post-op 3	Post-op 4
[i]	F1	315 (32,89)	317 (52,19)	318 (57,15)	332 (83,54)
	F2	2456 (268,75)	2446 (349)	2413 (352,72)	2419 (358,96)
[a]	F1	752 (106,77)	787 (162,41)	733 (150,79)	752 (120,11)

	F2	1328 (154,13)	1379 (267,75)	1295 (143,83)	1301 (168,16)
[u]	F1	349 (52,03)	394 (118,33)	361 (63,36)	353 (68,74)
	F2	745 (189,99)	1057 (450,26)	799 (168,7)	844 (213,83)

Pour résumer :

Pour le groupe UPP, les valeurs de F1 et de F2 ne varient pas en fonction de la phase d'enregistrement pour la voyelle [i]. En revanche, F1 et plus particulièrement F2 (cf. analyses de variance) subissent des modifications pour les voyelles [a] et [u], en phases d'enregistrement post-opératoires, avant de retrouver des valeurs attendues en post-opératoire 4. Ces tendances pourraient tout de même influencer sur les tailles de l'espace vocalique maximal (cf. *infra*).

L'étude des écarts-types a révélé une variabilité plus importante, caractérisée par des écarts-types plus grands, pour les productions des patients. Cela est particulièrement vrai pour les voyelles [a] et [u].

VII.4.5. Espace Vocalique

Pour le groupe UPP, l'aire de l'espace vocalique en post-opératoire 2 (0,28kHz²) est modifiée ($p < ns$) par rapport à celle des contrôles (0,35 kHz²). Notons qu'en post-opératoire 2, l'aire de l'espace vocalique du groupe UPP (0,28 kHz²) est alors comparable à celle du groupe NPP pour lequel l'aire est de 0,28 kHz² (Voir chapitre VII.4.6). A partir de la phase post-opératoire 3, l'aire augmente (0,31 kHz²) pour atteindre 0,32 kHz² en post-opératoire 4. Ce résultat confirme ce qui était supposé *supra*, à savoir que les modifications des valeurs formantiques auraient un impact sur l'aire de l'espace vocalique.

L'aire de l'espace vocalique des patients est donc réduite en post-opératoire 2. L'espace vocalique reste géométriquement conventionnel, alors même qu'il est réduit et que les voyelles [u] et [a] sont déplacées (voir Figure 17).

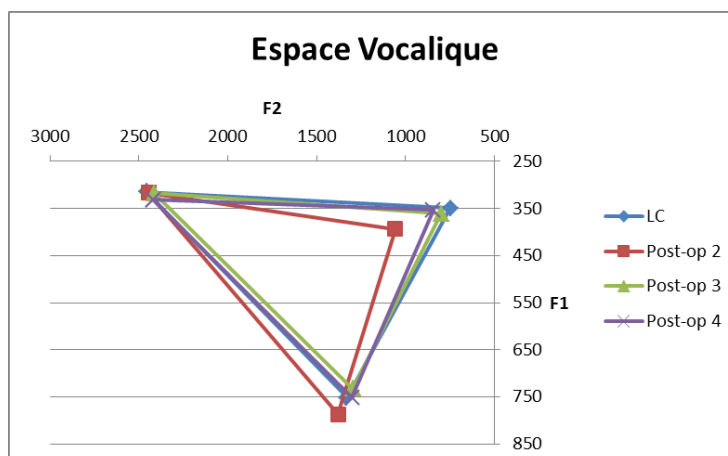


Figure 17 : Espace vocalique – Groupe UPP

L'étude de l'aire des espaces vocaliques en fonction des locuteurs indique une certaine variabilité inter-locuteurs. Ainsi, l'aire est plus petite en post-opératoire 2 pour UPPBAS (0,40 kHz²), UPPHEI (0,11 kHz²), UPPWAL (0,31 kHz²) et UPPWAN (0,19 kHz²) que celle de leur locuteur contrôle pour qui elle est respectivement de 0,42 kHz², 0,36 kHz², 0,41 kHz² et 0,49 kHz². En ce qui concerne les autres locuteurs, l'aire est plus grande en post-opératoire 2 que celle de leur locuteur contrôle : elle est de 0,54 kHz² pour UPPHUB (0,42 kHz² pour son contrôle), de 0,3 kHz² pour UPPPAI (0,18 kHz² pour son contrôle) et de 0,23 kHz² pour UPPPAI (0,21 kHz² pour son contrôle).

La phase post-opératoire 3 renforce cette variabilité inter-locuteurs. L'aire de l'espace vocalique se réduit pour UPPBAS (0,37 kHz²), UPPHUB (0,35 kHz²) et UPPPAI (0,18 kHz²), par rapport à la phase d'enregistrement précédente et à leur locuteur contrôle. UPPHEI, UPPWAL et UPPWAN suivent la tendance attendue, à savoir que l'aire de leur espace vocalique augmente à cette phase et est alors mesurée respectivement à 0,33 kHz², 0,39 kHz² et 0,44 kHz².

En post-opératoire 4, l'aire de l'espace vocalique est mesurée la plus petite pour UPPBAS (0,19 kHz²) et pour UPPHUB (0,29 kHz²). Les autres locuteurs suivent la tendance attendue, leur espace vocalique étant alors plus important et proche des valeurs des locuteurs contrôles, soit 0,34 kHz² pour UPPHEI, 0,23 kHz² pour UPPPAI, 0,34 kHz² pour UPPWAL et 0,49 kHz² pour UPPWAN. Enfin, il est intéressant de signaler que l'aire de l'espace vocalique de UPPLAT ne subit aucune modification quelle que soit la phase d'enregistrement post-opératoire étudiée, et elle est toujours de 0,30 kHz².

Tableau 28 : Espace vocalique (kHz²) en fonction des locuteurs – Groupe UPP

kHz ²	LC	Post-Op 2	Post-Op 3	Post-Op 4
BAS	0,42	0,40	0,37	0,19
HEI	0,36	0,11	0,33	0,34
HUB	0,42	0,54	0,35	0,29
LAT	0,18	0,30	0,30	0,30
PAI	0,21	0,23	0,18	0,23
WAL	0,41	0,31	0,39	0,34
WAN	0,49	0,19	0,44	0,49

Pour résumer :

L'aire de l'espace vocalique moyen est significativement réduite ($p < 0.05$) pour le groupe UPP entre la phase d'enregistrement des locuteurs contrôles et celle des patients en post-opératoire 2, avant de ré-augmenter en phase post-opératoire 3. Cette tendance se poursuit jusqu'en post-opératoire 4 ; les valeurs de référence des locuteurs contrôles ne sont toutefois pas atteintes lors de cette dernière phase d'enregistrement.

L'étude de l'aire de l'espace vocalique, en fonction des locuteurs, révèle une tendance similaire, à savoir que la plupart des locuteurs ont une aire de l'espace vocalique plus petite en post-opératoire 2 que celle mesurée pour les locuteurs contrôles, avant d'augmenter en post-opératoire 3. Ce phénomène se poursuit jusqu'en post-opératoire 4. Des différences individuelles subsistent quelle que soit la phase d'enregistrement étudiée.

VII.5. Discussion

1. Conformément à notre *hypothèse 1*, les valeurs de fréquence fondamentale des patients sont plus basses dans la phase d'enregistrement précoce que celles des locuteurs contrôles. Cela est particulièrement remarquable lors de la phase d'enregistrement post-opératoire 2 (soit quinze jours après l'intervention). Dès la phase post-opératoire 3, la fréquence fondamentale augmente, et les données obtenues sont alors comparables à celles des locuteurs contrôles ; cette tendance se poursuit en post-opératoire 4. De façon générale, la voix des patients présentant une immobilité laryngée est donc plus grave (le fondamental est mesuré plus bas) que celle des contrôles. Il y aurait donc un lien entre une perturbation de l'activité laryngienne et les mesures de fréquence fondamentale.
2. Les mesures de Harmonics-to-Noise Ratio (HNR) semblent être liées à l'activité irrégulière du larynx (*Hypothèse 2*). Cette mesure suit la même tendance que celle observée pour la fréquence fondamentale, à savoir que le ratio est mesuré plus bas pour les patients en post-opératoire 2 que pour les locuteurs contrôles. Dès la phase post-opératoire 3, le ratio augmente pour les patients ; il est alors

plus élevé pour l'ensemble des voyelles. Cette tendance se poursuit jusqu'en post-opératoire 4. Dans la mesure où cette mesure renseigne sur le rapport bruit/harmoniques dans une voyelle, il semble que la voix des patients soit plus bruitée, donc plus rauque en post-opératoire 2. Les phases d'enregistrement suivantes témoignent d'une harmonicité retrouvée.

3. Les valeurs de jitter et de shimmer sont modifiées en post-opératoire (*Hypothèse 3*). Selon notre hypothèse 3, il semble que ces modifications soient liées à un changement dans la périodicité de la voix. La voix des patients en phase post-opératoire serait alors plus aperiodique que celle des contrôles. Le signal des patients serait alors particulièrement perturbé, notamment dans la phase d'enregistrement post-opératoire 2. Cette observation est toujours avérée mais dans une moindre mesure pour les phases d'enregistrement post-opératoires suivantes. La voix des patients serait alors plus irrégulière. Cette irrégularité diminue avec le temps, puisque les valeurs de jitter et de shimmer sont quantifiées en dessous du seuil normal/pathologique en post-opératoire 4.
4. Les valeurs formantiques, extraites à partir des voyelles soutenues, ont également été perturbées (sauf pour la voyelle [i]). Les valeurs de F1 le sont de manière non significative ($p=ns$) tandis que les valeurs de F2 le sont de façon significative. Il est fort possible que la perturbation des valeurs formantiques soit liée à celle de l'activité laryngienne précédemment évoquée (*Hypothèse 4*).
5. Ces modifications ont toutefois des conséquences sur l'aire de l'espace vocalique (*Hypothèse 5*) qui subit une réduction en post-opératoire 2, par rapport à celle des locuteurs contrôles. Dès la phase post-opératoire 3, l'aire de l'espace vocalique des patients augmente pour se rapprocher des valeurs des locuteurs contrôles, sans toutefois les atteindre. Cette tendance se poursuit en post-opératoire 4. L'espace vocalique reste géométriquement conventionnel, malgré le déplacement des voyelles [u] et [a] et ne subit donc pas de réorganisation remarquable. Ces modifications seraient probablement liées aux perturbations de la source laryngée plutôt qu'à des stratégies articulatoires différentes. Cela reste toutefois à vérifier.

6. Nous avons vu que, quel que soit le paramètre étudié, la variabilité inter et/ou intra-locuteurs est toujours plus marquée pour tous les paramètres chez les patients (*Hypothèse 6*). La phase post-opératoire 2, notamment, indique des écarts-types souvent plus importants que ceux mesurés chez les locuteurs contrôles, et cela dans les phases d'enregistrement post-opératoires suivantes. Les écarts-types inter et intra individuels se réduisent dès la phase post-opératoire 3. Cette observation, qui se confirme en post-opératoire 4, témoigne d'une régularité retrouvée dans les productions des patients, et pourrait être un signe de récupération des coordinations oro-laryngées « standards ».
7. Enfin, de façon générale, le temps et la rééducation vocale ont un impact positif sur tous les paramètres précédemment mentionnés (*Hypothèse 7*). Les mesures se rapprochent des valeurs des locuteurs contrôles, dès la phase d'enregistrement post-opératoire 3, et cela pour la plupart des patients. Ce constat indique que, si la voix des locuteurs est modifiée par l'immobilité laryngée post-thyroidectomie en post-opératoire 2, la qualité vocale s'améliore par la suite (post-opératoire 4). Dans la plupart des cas, la voix est à nouveau comparable à celle des locuteurs contrôles dès la phase post-opératoire 4.

CONCLUSION CHAPITRE VII

Dans ce chapitre, nous avons étudié la voix de sept locuteurs qui ont été opérés de la glande thyroïde et pour lesquels une immobilité laryngée unilatérale a été diagnostiquée, suite à l'opération chirurgicale. Ne disposant pas d'enregistrement préopératoire, sept locuteurs contrôles appariés en genre et en âge ont été enregistrés. Cette étude se veut longitudinale, puisque la voix des patients a été enregistrée à trois reprises : en post-opératoire 2 (quinze jours après l'intervention), en post-opératoire 3 (un mois après l'intervention) et en post-opératoire 4 (deux mois après l'intervention).

Notre corpus était composé de voyelles soutenues, qui ont permis de quantifier les perturbations que l'immobilité laryngée entraîne sur les paramètres acoustiques et spectraux suivants : la fréquence fondamentale, les valeurs formantiques, le Harmonics-to-Noise Ratio (HNR), le jitter, le shimmer et l'aire de l'espace vocalique.

Il était supposé que l'activité laryngienne, perturbée par l'immobilité laryngée, aurait des conséquences sur les paramètres susmentionnés, et notamment sur les mesures qui permettent de quantifier les perturbations du signal acoustique, tels que le HNR, le jitter et le shimmer. Nous émettions également l'hypothèse que le temps et la rééducation vocale devaient avoir un effet positif sur les paramètres étudiés, ce qui se refléterait une amélioration de la qualité vocale des patients. Leurs mesures seraient alors comparables à celles obtenues auprès des locuteurs contrôles.

Les analyses de variance (ANOVA) à deux facteurs indiquent que l'effet principal *phase d'enregistrement* a été significatif ($p < 0,05$) pour les variables *Harmonics-to-Noise Ratio*, *jitter*, *shimmer*, *F2* et *aire de l'espace vocalique*.

Il semble intéressant de remarquer que ce sont les mesures de perturbations du signal, tel que le HNR, le jitter et le shimmer, qui sont significativement modifiées suite à l'immobilité laryngée. Les mesures du HNR sont alors plus basses en post-opératoire 2 que celles des locuteurs contrôles, tandis que les valeurs du jitter et du shimmer sont plus importantes lors de cette même phase pour les patients, par rapport aux locuteurs contrôles. Les valeurs commencent à se « standardiser » dès la phase post-opératoire 3, ce qui se poursuit jusqu'en post-opératoire 4 ; il semble donc que le temps et la rééducation vocale aient un impact positif sur tous les paramètres étudiés. Enfin, les

mêmes tendances ont également pu être observées pour les autres paramètres (fréquence fondamentale et F1).

Les enregistrements des patients présentant une immobilité laryngée sont également caractérisés par une variabilité inter et intra locuteur qui a toutefois tendance à se réduire en post-opératoire 3 ; cette tendance se poursuit et s'accroît en post-opératoire 4.

RÉSUMÉ CHAPITRE VIII

Ce chapitre entend étudier les caractéristiques temporelles de la voix de patients présentant une immobilité laryngée post-thyroïdectomie à l'aide de différents indices spatio-temporels. Il s'agit en réalité d'analyser les conséquences d'une paralysie récurrentielle sur le timing oro-laryngé des patients.

Notre démarche se veut articulatoire-acoustique, puisqu'il s'agit d'observer à partir du signal acoustique continu, les différents indices articulatoire-acoustiques qui peuvent nous permettre de remonter aux configurations articulatoires (Abry et *al.*, 1985). Il s'agira d'étudier les différentes réorganisations temporelles qu'une immobilité laryngée peut provoquer sur la voix des patients, et c'est pour cette raison que les perturbations temporelles éventuelles observées seront au centre de ce chapitre.

Notre corpus est composé de douze logatomes de type VCV, à partir desquelles ont été mesurés : les durées vocaliques, le VOT (Klatt, 1975), le VTT (Agnello, 1975), l'occlusion et la tenue consonantique.

Les résultats vont dans le sens de modifications tendanciennes de toutes les durées suite à l'immobilité laryngée. Une modification statistiquement significative de la durée du VOT apparaît pour les consonnes voisées uniquement dans les phases d'enregistrement post-opératoires proches de l'opération.

Le temps et la rééducation vocale permettent un retour à des valeurs proches de celles obtenues pour les locuteurs contrôles.

CHAPITRE VIII

VIII. Expérience 3 : Etude spatio-temporelle de la voix de patients présentant une paralysie récurrentielle après thyroïdectomie

VIII.1. Introduction

La présente étude propose d'analyser les caractéristiques spatio-temporelles de la voix de patients présentant une immobilité laryngée suite à une opération de la glande thyroïde. Il s'agit ici d'évaluer les conséquences de l'immobilité laryngée sur le *timing* des articulateurs, à l'aide d'un corpus de logatomes (ou séquences VCV).

Notre investigation, bien qu'acoustique, entend analyser les différents indices articulatoire-acoustiques qui peuvent être observés à partir d'un signal acoustique continu, et ainsi remonter aux configurations articulatoires (Abry et *al.*, 1985). Une attention particulière sera donc accordée aux différentes perturbations temporelles qu'une immobilité laryngée peut provoquer sur la voix des patients, grâce à l'étude de la réorganisation temporelle des paramètres retenus.

Notre étude se veut longitudinale puisque la voix des patients est enregistrée lors de différentes phases post-opératoires (post-opératoire 2, 3 et 4), ce qui devrait nous permettre de mettre au jour non seulement les *perturbations* provoquées par l'immobilité laryngée mais également les *réajustements* que les patients pourraient mettre en place au cours du temps et de leur rééducation orthophonique.

Rappelons qu'une immobilité laryngée a des conséquences acoustiques et aérodynamiques sur la voix des patients (voir Hartl, et al. 2001 par ex.). Dans la plupart des cas, l'immobilité laryngée post-thyroïdectomie est transitoire, une amélioration de la qualité vocale est donc mesurable au cours du temps (Laccourreye et *al.*, 2008). Il

s'agit ici de savoir si cette amélioration de la qualité vocale est manifeste dans les aspects temporels du signal de parole.

VIII.2. Procédure expérimentale

VIII.2.1. Locuteurs

Cette étude est conduite à partir de sept patients dont la voix a été enregistrée au Centre Paul Strauss, puis dans le cabinet de différentes orthophonistes. Il s'agit des mêmes locuteurs que ceux retenus pour l'expérience 2 (voir Chapitre VII), soit cinq femmes (UPPBAS, UPPHEI, UPPHUB, UPPWAL et UPPWAN) et deux hommes (UPPLAT et UPPPAI). Tous ces patients présentent une immobilité laryngée unilatérale consécutive à une thyroïdectomie totale. Différents ORL ont pratiqué l'examen laryngé et le diagnostic, ce qui s'explique par le fait que le choix du praticien était laissé au patient. Des méthodes d'investigation laryngée différentes (miroir, laryngoscopie...) ont donc pu être utilisées en fonction du praticien retenu.

Ce groupe sera appelé UPP (Unilateral Paralysis Patients ou patient présentant une immobilité laryngée), dans la suite de ce travail.

Les patients ont été enregistrés une fois par mois tout au long de leur rééducation orthophonique. Afin d'obtenir une cohorte comparable et ainsi de pouvoir conduire des analyses de variance (ANOVA à mesures répétées), seules trois phases post-opératoires ont pu être retenues : en post-opératoire 2 (soit quinze jours après l'intervention), en post-opératoire 3 (soit un mois après l'intervention) et en post-opératoire 4 (c'est-à-dire deux mois après l'intervention). Pour certains locuteurs, la rééducation orthophonique n'était alors pas terminée.

Comme les données en préopératoire n'étaient pas disponibles (voir V.4.), la voix des patients a été comparée avec celle de locuteurs contrôle (dorénavant LC) appariés en sexe et en âge avec le patient correspondant. Les locuteurs contrôles constituent alors la voix de référence.

VIII.2.2. Corpus

Notre corpus est constitué de douze logatomes (ou séquence VCV) insérés dans une phrase porteuse.

Ex : « Cet VCV ça. »

Les logatomes sont créés comme suit : si V1 est /i/ alors V2 est /a/ et vice versa. La consonne est l'une des 6 occlusives /p, t, k, b, d, g/. Ces consonnes ont été choisies car elles permettent l'alternance de séquences entièrement voisées V1 [b d g] V2 et de séquences mixtes V1 [p t k] V2, où les plis vocaux ne sont pas sollicités de façon continue.

Les douze logatomes sont donc les suivants : [abi], [adi], [agi], [api], [ati], [aki], [iba], [ida], [iga], [ipa], [ita] et [ika].

Il s'agit pour le locuteur de lire chaque logatome qui lui est présentés dix fois et de façon aléatoire.

VIII.2.3. Mesures

Le signal acoustique et le spectrogramme correspondant ont été segmentés conjointement. Pour chaque logatomes, les durées suivantes ont été mesurées :

- La première voyelle (V1)
- La consonne intervocalique (C)
- La seconde voyelle (V2)

Une attention particulière a été portée à la consonne intervocalique pour laquelle les paramètres temporels suivants ont été quantifiés :

- le Voice Termination Time (pour les consonnes non voisées) ;
- le silence acoustique (pour les consonnes non voisées) ;
- le Voice Onset Time ;
- la tenue consonantique.

Le Voice Onset Time (VOT) a été mesuré selon Klatt (1975) soit comme l'intervalle allant du relâchement consonantique jusqu'à l'apparition de la structure formantique stable de la voyelle subséquente. Rappelons que Liberman et al., (1958) notent le VOT comme un indice majeur de la classification perceptive des occlusives voisées/non voisées.

Le Voice Termination Time (VTT) a été décrit par Agnello (1975) comme l'intervalle allant de la closure articulaire ou la pression intra-orale maximale jusqu'à la dernière vibration glottique. Pour certains auteurs, le VTT coïnciderait avec le temps nécessaire à l'arrêt de vibration des plis vocaux après la closure dans le conduit vocal.

Le silence acoustique a été mesuré comme l'intervalle entre la fin du VTT et le début du VOT pour les consonnes non voisées. En ce qui concerne les consonnes voisées, la tenue consonantique a été mesurée entre la fin de la structure formantique stable de V1 et le début de la structure formantique stable de V2.

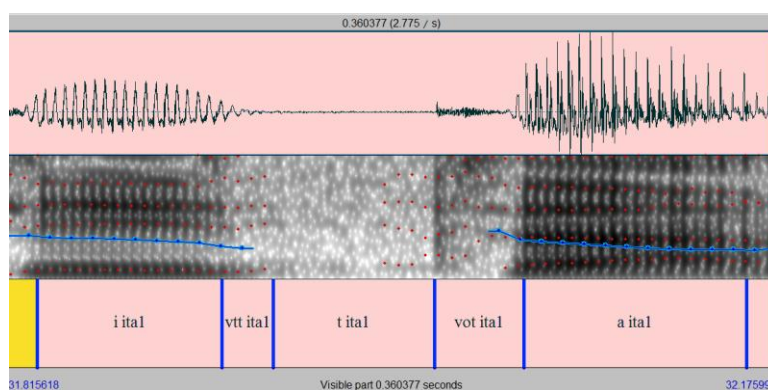


Figure 18 : Mesures temporelles pour une séquence VCV dans laquelle C est une consonne non voisée

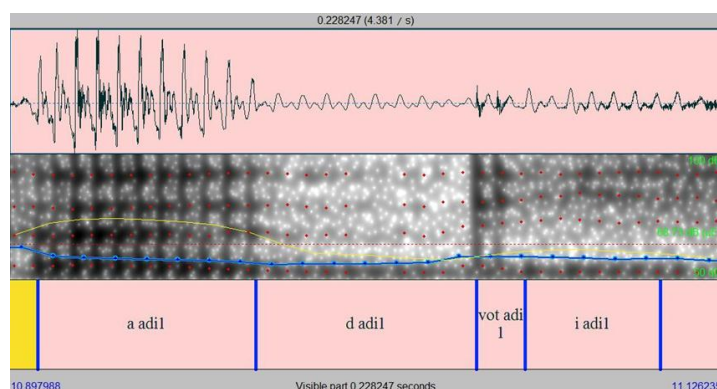


Figure 19 : Mesures temporelles pour une séquence VCV dans laquelle C est une consonne voisée

Les durées du VOT obtenues ont également été normalisées en calculant le pourcentage pris par cet intervalle dans les différentes réalisations de la tenue consonantique entière. La tenue consonantique est considérée comme l'intervalle allant

de la fin de la structure formantique stable de V1 au début de la structure formantique stable de V2. Le résultat ainsi obtenu est une valeur exprimée en pourcentage.

La durée relative du VOT a été calculée comme suit :

$$\frac{\text{Dures absolues du VOT} \times 100}{\text{Tenue consonantique}}$$

Équation 3 : Durée relative du VOT exprimée en %

Il nous semble important de prendre en considération les valeurs normalisées en production de la parole afin de pouvoir évaluer correctement les effets de l'élasticité du signal temporel de la parole (Sock, 1998). En effet, une telle normalisation permet d'apprécier, à sa juste valeur, l'effet des variations de vitesses d'élocution entre nos différents locuteurs. Nous pourrions ainsi évacuer toute variabilité de durées du VOT, liée simplement à des variations de vitesses d'élocution pour pouvoir déceler les véritables différences qui restent robustes malgré le changement de vitesse d'élocution.

VIII.2.4. Analyse des données

Afin de faciliter la lecture des résultats, ces derniers seront présentés sous forme de graphiques et de tableaux commentés.

Des analyses de variance (ANOVA) à mesures répétées ont été conduites pour toutes nos données (*cf. infra* pour plus de détails). Les valeurs présentées renvoient toujours aux moyennes calculées à partir des répétitions produites par nos locuteurs. Les écarts-types sont disponibles entre parenthèses.

VIII.2.5. Hypothèses

Etant donné qu'une immobilité laryngée entraîne des modifications de la périodicité de la voix et des difficultés dans le contrôle du mouvement des plis vocaux et donc du voisement, nous formulons les hypothèses suivantes :

1. Les difficultés de voisement rencontrées par les patients pourraient avoir un impact sur la durée du VOT. Celle-ci considérée comme un indice de vibration des plis vocaux et pouvant permettre ainsi la distinction des consonnes voisées des consonnes non voisées, serait modifiée. Traditionnellement, le VOT de Klatt (1975), allant du relâchement supraglottique de la consonne à l'apparition d'une

structure vocalique formantiquement stable, est plus court pour les occlusives voisées que pour les occlusives non voisées. Pour les occlusives voisées, la transition CV implique uniquement un changement d'un état obstrué du conduit vocal à un état suffisamment dégagé de celui-ci pour permettre l'apparition d'une structure formantique claire. En revanche, pour les occlusives non voisées, cette transition CV exige, outre le changement d'état du conduit vocal, une modification de la configuration de la glotte d'une position ouverte pour la consonne non voisée à une position fermée pour la voyelle. En conséquence, en raison de difficultés de maintien de vibrations laryngées, le VOT devrait alors être plus long pour les occlusives voisées (plus ou moins désonorisées, et produites avec difficulté par les patients) par rapport à celles produites par les locuteurs contrôles. Ainsi, la durée du VOT de Klatt (1975) des occlusives voisées des patients pourrait alors être comparable à celle mesurée pour leurs occlusives non voisées, rendant ainsi la distinction entre ces deux catégories potentiellement difficile dans une perspective quantitative.

2. Le VTT (Agnello, 1975), reflet du timing oro-laryngé, pourrait être modifié par l'immobilité laryngée. Rappelons qu'il s'agit en effet d'un paramètre qui correspond, dans une transition VC, à l'intervalle allant de la disparition de la structure formantique clairement définie de la voyelle (disparition provoquée par la closure supraglottique de la consonne) à l'arrêt du voisement dans la phase consonantique. Dans la mesure où ce paramètre est lié à l'arrêt des vibrations périodiques, donc par extrapolation à l'amortissement des vibrations des plis vocaux (Sack, 1998), il devrait être modifié par l'immobilité laryngée, avec un délai d'arrêt du voisement plus court pour les patients par rapport aux locuteurs contrôles.
3. La durée de l'occlusion pourrait également se voir modifiée par l'immobilité laryngée. En effet, suivant les travaux de Meynadier & Gaydina (2012) par exemple, il a été montré que cet indice participe à la distinction de voisement des consonnes en parole modale et en parole chuchotée. Ainsi, en parole chuchotée, la tenue consonantique serait significativement plus longue que celle observée en parole modale. Compte tenu des difficultés de voisement

rencontrées par nos patients, nous pouvons supposer que la tenue consonantique des occlusives voisées pourrait se voir allongée, ce qui impliquerait également que la tenue consonantique des occlusives non voisées le soit également pour maintenir le contraste de sonorité nécessaire à la distinction des sons. Dans tous les cas, les tenues consonantiques de nos patients devraient être plus longues que celles des locuteurs contrôles.

4. La tenue consonantique est l'intervalle qui sépare V1 de V2, dans une séquence VCV. Il se peut cependant qu'une phase, plus ou moins silencieuse, apparaisse dans la tenue consonantique des occlusives voisées. Cette apparition intrusive serait alors toujours imputable à cette difficulté de contrôle et de maintien de vibrations laryngées périodiques tout au long de la réalisation de l'occlusive voisée rendant alors cet intervalle plus long.
5. Il s'agira également d'observer la robustesse des différences potentielles de durées observées en valeurs absolues dans l'hypothèse 1 face à l'élasticité du signal. Une attention particulière sera alors accordée aux durées en valeurs relatives, afin de comprendre si l'observation précédente est significativement pertinente ou simplement liée à des stratégies différentes en matière de vitesse d'élocution. Il est en effet possible que l'immobilité laryngée contraigne les locuteurs à réajuster leurs vitesses d'élocution, ce qui se traduirait par des compressions/expansions du signal de parole. Cela pourrait être, *in fine*, un indice de modifications aérodynamiques.
6. En raison de stratégies compensatoires différentes selon les individus, la variabilité inter locuteurs devrait être plus marquée dans les phases d'enregistrement post-opératoires précoces (post-opératoire 2, notamment). La variabilité intra locuteur devrait également être plus importante lors de cette phase. Il est en effet possible que, lors de certaines productions, le locuteur se rapproche des cibles attendues, alors que pour d'autres, il en serait particulièrement éloigné. Cela devrait se traduire par des écarts-types plus importants pour ces phases d'enregistrements. Toute variabilité remarquable devrait se réduire avec le temps.

7. Le temps et la rééducation vocale devraient avoir un impact positif sur les paramètres modifiés par l'immobilité laryngée. C'est-à-dire que si la qualité vocale du patient s'améliore (voir Chapitre VII) au cours du temps, les paramètres temporels devraient également ressembler à ceux des locuteurs contrôles dès la phase d'enregistrement post-opératoire 4 (c'est-à-dire environ deux mois après l'intervention).

VIII.3. Résultats

VIII.3.1. Remarques générales

Les résultats de ce chapitre sont présentés sous forme de graphiques commentés. Les valeurs, qui y sont représentées, sont toujours des moyennes calculées à partir des productions des différents locuteurs. Nous proposons toujours dans un premier temps d'observer les moyennes calculées à partir des différentes productions, tous locuteurs confondus. Cette étape devrait permettre de dégager certaines tendances. Compte tenu de l'importance des écarts-types (donnés entre parenthèses) et pour offrir une analyse plus fine de nos résultats, les productions seront ensuite étudiées en fonction des locuteurs lorsque cela peut apporter des explications complémentaires pour comprendre mieux nos données. Cela devrait permettre de mettre au jour les différentes stratégies de compensations individuelles que les locuteurs peuvent mettre en place.

Des Analyses de Variance à trois facteurs (*voisement*, *lieu d'articulation* et *temps* ou *phases d'enregistrement*) ont été conduites pour toutes les variables étudiées. Les effets principaux, lorsqu'ils étaient significatifs, ont ensuite été analysés à l'aide de tests post-hoc (*LSD* ou *Least Significant Difference - LSD*) afin de déterminer entre quels groupes se situent la ou les différences significatives.

VIII.3.1. Durées du VOT

Des ANOVA à trois facteurs, *voisement*, *lieu d'articulation* et *phases d'enregistrement*, ont été conduites pour la variable VOT dans les deux contextes [a-i] et [i-a].

En contexte [a-i], nous avons observé un effet principal du *voisement* qui repose sur une différence significative entre les consonnes voisées et non voisées, $[F(1,6) = 430,27 \text{ } p < 0,00003]$, toutes phases et tous lieux d'articulation confondus. Il y a un effet principal du *lieu d'articulation* reposant sur une différence significative entre les consonnes bilabiales [p b] et les quatre autres consonnes [t d] et [k g], $[F(2,12) = 70,08 \text{ } p < 10^{-6}]$. Enfin, nous avons aussi constaté un effet principal de la *phase d'enregistrement* $[F(3,18) = 216,14 \text{ } p < 0,02]$, qui met en lumière une différence significative entre la phase d'enregistrement post-opératoire 2 et toutes les autres phases d'enregistrement.

En contexte [a-i], l'interaction entre les trois facteurs, *voisement*, *lieu d'articulation* et *phases d'enregistrement*, est statistiquement significative $[F(6,36) = 51,41 \text{ } p < 0,032]$. L'interaction repose notamment sur le fait que, quelle que soit la consonne voisée considérée, les productions des locuteurs contrôles sont significativement différentes de celles des patients lors de la phase d'enregistrement post-opératoire 2. A partir de la phase d'enregistrement post-opératoire 4, les tests *post-hoc* indiquent que le facteur de *voisement* ne permet plus de distinguer les productions des locuteurs contrôles de celles des patients, pour les consonnes voisées. En revanche, les consonnes non voisées des locuteurs contrôles et des patients en phases post-opératoires 2, 3 et 4 ne sont pas significativement différentes. En ce qui concerne le facteur *lieu d'articulation*, le test *post-hoc* révèle qu'en post-opératoire 2, la variable *VOT* ne permet pas toujours de distinguer les consonnes les unes des autres. Il n'existe alors plus de différence significative entre [d] et [g]. Enfin, signalons que le facteur *lieu d'articulation* pour la variable *VOT* indique que leurs productions sont significativement différentes en fonction de la consonne étudiée.

Nous avons observé un effet principal du *voisement* $[F(1,6) = 105,08 \text{ } p < 0,0002]$. Suite aux tests *LSD*, nous pouvons affirmer que cet effet principal de *voisement* repose, comme attendu, sur une différence significative entre les consonnes voisées et les consonnes non voisées. Une différence significative a également pu être mise au jour pour le facteur *lieu d'articulation* $[F(2,12) = 77,03 \text{ } p < 10^{-5}]$. Les tests *post-hoc* révèlent que la durée du *VOT* est significativement différente en fonction du lieu d'articulation des diverses occlusives. Ainsi, la durée du *VOT* de [p b] est différente de celle de [t d], qui est elle-même différente de celle de [k g]. Un effet principal de la

phase a pu être mis au jour pour la variable VOT en contexte [a-i] [$F(3,18) = 188,02$ $p < 0,013$]. La phase d'enregistrement post-opératoire 2 est significativement différente de toutes les autres phases d'enregistrement.

En contexte [i-a], on n'observe pas d'interaction significative entre les trois facteurs, *voisement*, *lieu d'articulation* et *phases d'enregistrement* ($p=ns$).

VIII.3.1.1. Occlusives voisées

Nous proposons dans cette section d'étudier la durée du VOT (Klatt, 1975) pour les occlusives voisées, soient [b d g]. Nous présenterons, pour chaque occlusive, les valeurs moyennes calculées à partir des productions des locuteurs contrôles et des patients. Compte tenu de l'importante variabilité dans les productions des patients, ces locuteurs seront ensuite analysés de façon individuelle afin de comprendre si cette variabilité est le reflet de différences inter et/ou intra individuelles. Une conclusion partielle sera donnée pour les deux catégories de consonnes (voisées et non voisées) en fin de partie.

❖ Contexte [b]

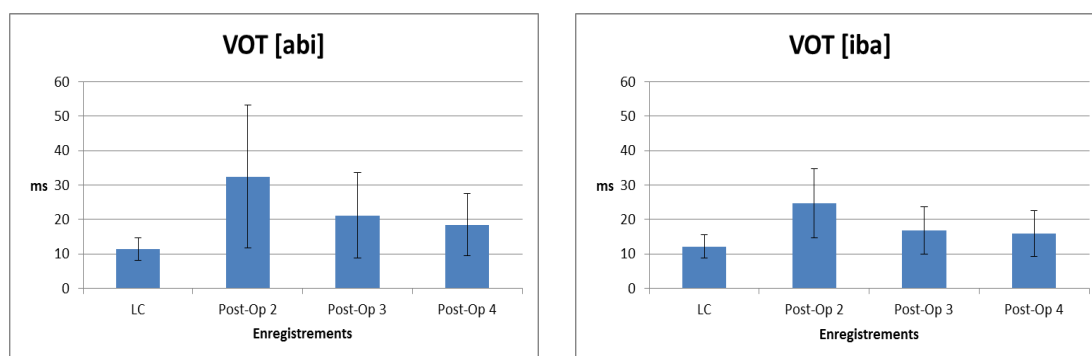


Figure 20 : Valeurs absolues de la durée du VOT (Klatt, 1975) en contexte [b] pour les logatomes [abi] (à gauche) et [iba] (à droite)

En ce qui concerne la consonne occlusive voisée [b], la durée du VOT des locuteurs contrôles est de 11,45 ms (3,34 ms) pour [abi] et de 12,17 ms (3,39ms) pour [iba] ; ces valeurs sont conformes à la littérature (*cf.* par ex. Sock & Benoit, 1986) ; de plus, les écarts-types réduits témoignent d'une certaine régularité dans les productions de ces locuteurs.

En contextes [a-i] et [i-a], le VOT des patients est toujours plus long que celui des locuteurs contrôles (voir Figure 20). Cette observation est particulièrement remarquable lors de la phase d'enregistrement post-opératoire 2, où le VOT des patients est de 32,51 ms (20,72 ms) pour [abi] et 24,77 ms (10,04 ms) pour [iba]. La durée du VOT se réduit dès la phase post-opératoire 3, où il est alors de 21,17 ms (12,41ms) pour [abi] et de 16,90 ms (6,88 ms) pour [iba]. Cette tendance se poursuit jusqu'en post-opératoire 4, où le VOT est alors mesuré à 18,48 ms (6,67ms) pour [abi] et à 16,01ms (9,09ms).

Si dans les deux contextes la durée du VOT est plus importante pour les patients, notamment en post-opératoire 2, cela est plus notable pour le contexte [a-i] que pour le contexte [i-a]. Le VOT mesuré pour [a-i] indique également plus de variabilité que le second contexte. Nous posons que la transition [C=> i] exige un contrôle plus précis et plus problématique que la transition [C=> a]. En effet, la voyelle [i] requérant une constriction réduite dans la zone palatale, sans qu'il y ait de contact entre le dos de la langue et la voûte palatine, les patients semblent avoir plus de difficulté à réaliser cette transition de la consonne vers la voyelle cible [i]. Cela ne semble pas être le cas pour le [a], dont l'abaissement mandibulaire, accompagné de celui de la masse de la langue avec qui la mandibule est fortement couplée, crée sans difficulté notoire la constriction radico-pharyngale nécessaire pour l'émergence acoustique de cette voyelle.

Les phases d'enregistrement des patients sont caractérisées par des écarts-types (voir valeurs entre parenthèses) plus importants que ceux des locuteurs contrôles, ce qui peut être le résultat d'une variabilité inter et/ou intra individuelle.

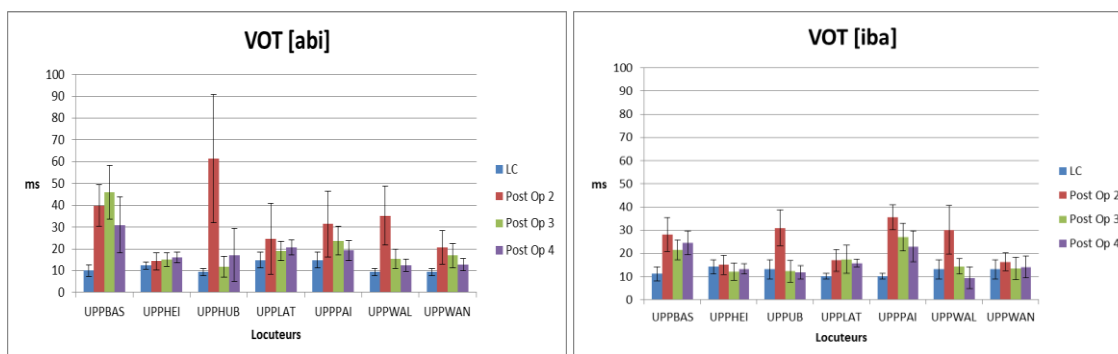


Figure 21 : Valeurs absolues de la durée du VOT (Klatt, 1975) en contexte [b] pour les logatomes [abi] (à gauche) et [iba] (à droite) en fonction du patient

La durée du VOT des locuteurs contrôles (phase en bleu) est conforme aux valeurs attendues pour la consonne [b] et est caractérisée par de faibles écarts-types, ce qui témoigne, encore une fois, de la régularité dans les productions des locuteurs contrôles.

Comme le laissaient supposer les écarts-types relevés à partir de la Figure 20, si le VOT est plus long pour les patients en post-opératoire 2 que pour les locuteurs contrôles, l'importance de ce phénomène est différente en fonction du patient étudié. En effet, la durée du VOT de quatre locuteurs (de UPPBAS, de UPPHUB, de UPPPAI et de UPPWAL) est particulièrement plus élevée, alors que celle de trois locuteurs (de UPPHEI et de UPPWAL) est moins modifiée lors de cette phase d'enregistrement. Cette observation confirme le fait que la variabilité précédemment évoquée (voir Figure 20) semble être en priorité inter-individuelle. Pour exemple, le VOT de UPPHUB, la patiente qui produit la durée de VOT la plus longue en post-opératoire 2 (phase en rouge), est de 31,50 ms (15,14 ms) pour [abi].

De plus, le contexte [a-i] (Figure 21, à gauche) ou [i-a] (Figure 21

Figure 21, à droite) n'a pas d'effet sur ce phénomène. Le VOT mesuré à partir de [iba] est alors également plus long pour les patients que pour les locuteurs contrôles. Il est alors de 31,00 ms (7,69 ms) chez UPPHUB.

La phase post-opératoire 2 est également marquée par des écarts-types importants (plus de 15 ms en moyenne), notamment pour les productions de UPPHUB, de UPPLAT, de UPPPAI et de UPPWAL, que ce soit en contexte [a-i] ou en contexte [i-a]. Il est à noter que la variabilité en contexte [i-a] est plus réduite que celle observée en contexte [a-i] ; elle reste toutefois plus importante que celle mesurée pour les locuteurs contrôles. La variabilité précédemment esquissée (voir Figure 20) serait donc liée à des productions différentes à un double niveau : elle serait le résultat de différences inter mais également intra-individuelles.

Dès la phase d'enregistrement post-opératoire 3, la durée du VOT se réduit pour tous les locuteurs dans les deux contextes (sauf UPPBAS en contexte [a-i] pour laquelle la durée du VOT est de 45,90 ms). UPPHUB est la patiente pour laquelle la diminution de la durée du VOT est la plus importante : elle est alors de 11,80 ms (4,83 ms) pour

[abi] et 12,30 ms (4,69 ms) pour [iba], rappelons que son VOT était de 61,30 ms pour [abi] et 31,00 ms pour [iba] en post-opératoire 2. La diminution de la durée du VOT est moins importante pour les autres locuteurs. La durée du VOT est alors de 19,10 ms (4,36 ms) chez UPPPAI, par exemple. Cette phase d'enregistrement est également marquée par des écarts-types qui commencent à se réduire, ce qui indique que les productions des différents locuteurs deviennent plus régulières.

Chez UPPBAS, la durée du VOT est plus courte en post-opératoire 4 que celles mesurées lors des phases d'enregistrement précédentes, soit 31,00 ms (12,84ms) pour [abi] et 24,50 ms (5,02 ms) pour [iba], sans toutefois atteindre les valeurs de référence de son locuteur homologue de contrôle. Rappelons que pour ce locuteur, nous n'avions pas observé de diminution du VOT, lors de la phase post-opératoire 3. Pour tous les autres locuteurs, la tendance amorcée en post-opératoire 3 se poursuit, les durées de VOT sont alors non seulement plus courtes mais également comparables à celles des locuteurs contrôles. Citons UPPWAL pour qui la durée du VOT est alors de 12,80 ms (2,70 ms), la durée du VOT de son locuteur contrôle est de 9,40 ms (1,65 ms). Cette phase d'enregistrement est également marquée par des écarts-types réduits, preuve que les productions des patients sont devenues plus régulières.

Un seul patient se démarque des observations précédentes (voir Figure 21). En effet, la durée du VOT, en contexte [a-i] ou [i-a], n'est pas modifiée en fonction des différentes phases d'enregistrement, et reste proche de celle du locuteur contrôle pour la locutrice UPPHEI : elle est alors de 14,30 ms (3,86ms) en post-opératoire 2, de 15,20 ms (3,08 ms) en post-opératoire 3 et de 16,10 ms (2,60 ms) pour [abi] et de 15,10 ms (4,09 ms) en post-opératoire 2, de 12,10 ms (3,63ms) en post-opératoire 3 et de 13,30 ms (2,26ms) en post-opératoire 4 pour [iba]. De plus, les productions de cette locutrice sont également régulières puisque les écarts-types sont faibles, quelle que soit la phase d'enregistrement étudiée.

❖ Contexte [d]

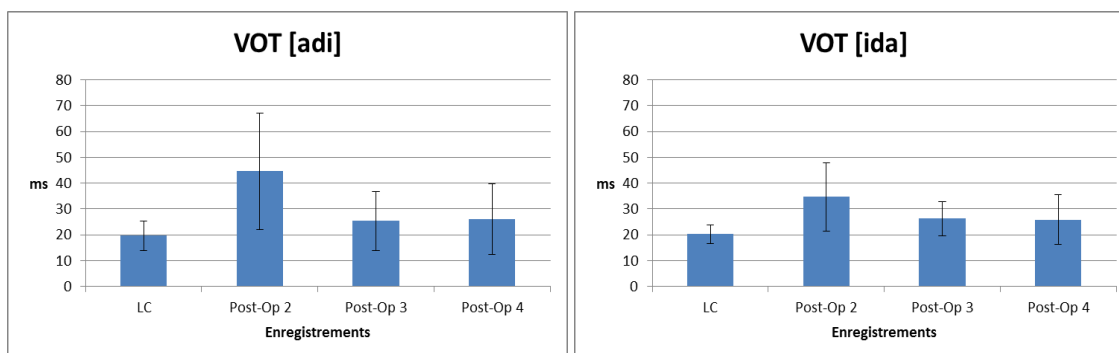


Figure 22 : Valeurs absolues de la durée du VOT (Klatt, 1975) en contexte [d] pour les logatomes [adi] (à gauche) et [ida] (à droite)

La Figure 22 illustre les valeurs de VOT pour la consonne [d]. Les durées de ce paramètre chez les locuteurs contrôles (LC) sont conformes aux valeurs attendues à savoir 19,67 ms pour le logatome [adi] et 20,25 ms pour le logatome [ida]. De plus, les faibles écarts-types, soit 5,74 ms et 3,61 ms respectivement, témoignent de la régularité de ce groupe de locuteurs.

La première phase d'enregistrement des patients, ou la phase post-opératoire 2, est caractérisée par des durées de VOT plus longues et des écarts-types plus importants que ceux relevés chez les locuteurs contrôles. La durée du VOT est alors de 44,61 ms (22,47 ms) pour [adi] et de 34,67 ms (13,3 ms) pour [ida]. Ces durées diminuent dès la phase post-opératoire 3, le VOT de [adi] étant alors 25,31 ms (11,54 ms) tandis que celui de [ida] est de 26,22 ms (6,63 ms). Notons que les écarts-types se réduisent également. Cette tendance se maintient jusqu'en post-opératoire 4, où le VOT de [ida] est mesuré à 26,05 ms (13,53 ms) et celui de [adi] à 25,78 ms (9,64 ms). La durée du VOT reste alors plus importante pour les patients que pour les locuteurs contrôles. Il convient toutefois de remarquer qu'elle s'est significativement réduite depuis la phase post-opératoire 2 et que les écarts-types ont diminué également, ce qui témoigne d'une amélioration du contrôle des vibrations laryngées et de la régularité des productions.

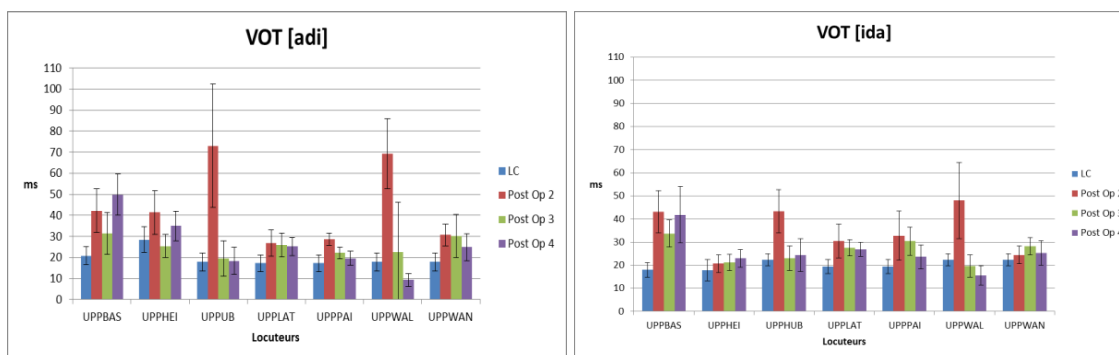


Figure 23 : Valeurs absolues de la durée du VOT (Klatt, 1975) en contexte [d] pour les logatomes [adi](à gauche) et [ida](à droite) en fonction du patient

La Figure 23 permet d'étudier les durées de VOT pour la consonne [d] en fonction du contexte et du locuteur, afin de comprendre l'origine des écarts-types importants relevés notamment lors de la phase d'enregistrement post-opératoire 2 (voir Figure 22). Comme attendu, les valeurs des locuteurs contrôles (phase en bleu) sont proches des durées moyennes attestées dans la littérature pour un locuteur sain, puisqu'elles sont à environ 20 ms pour chaque locuteur contrôle, et régulières comme en témoignent les écarts-types réduits, qui s'élèvent à environ 3 ms.

En ce qui concerne les productions des patients, la durée du VOT en post-opératoire 2 (phase en rouge) est toujours plus importante que celle relevée pour le locuteur contrôle associé. L'importance de ce phénomène est toutefois à considérer en fonction du locuteur étudié. Ainsi, chez certains patients, la durée du VOT en post-opératoire 2 est beaucoup plus importante que celle relevée pour leur locuteur contrôle associé. C'est le cas notamment chez UPPBAS, UPPHUB et UPPWAL pour les deux contextes [a-i] et [i-a] ; la durée du VOT est alors de 42,20 ms (10,30 ms), de 73,10 ms (29,27 ms), de 31,50 ms (6,53 ms), respectivement pour [adi] et de 43,10 ms (9,16 ms), de 43,30 ms (9,38ms) et de 48,00 ms (16,51 ms). Pour ces locuteurs, la durée du VOT est multipliée au moins par deux, comparée à celle obtenue auprès des locuteurs contrôles. En ce qui concerne les autres patients (UPPHEI, UPPPAI, UPPLAT et UPPWAN), la durée du VOT est également plus importante en post-opératoire 2 que celle mesurée auprès des locuteurs contrôles mais dans une moindre mesure. Le VOT de [adi] est alors de 41,40 ms (10,30 ms) chez UPPHEI, par exemple. La même tendance est observable pour le contexte [i-a], puisque le VOT est de 20,70 ms (3,89 ms) chez la même patiente. Signalons toutefois que, pour [d], en contexte [i-a], l'augmentation de la

durée du VOT en post-opératoire 2 est moins importante que celle relevée en contexte [a-i]. La phase d'enregistrement post-opératoire 2 est également caractérisée par des écarts-types plus importants. La variabilité précédemment évoquée (voir Figure 22) est donc le résultat de différences inter et intra individuelle, lors de cette phase d'enregistrement.

La phase d'enregistrement post-opératoire 3 (phase en vert) indique un retour vers des durées de VOT des patients comparables à celles mesurées à partir des productions des locuteurs contrôles, dans les deux contextes [a-i] et [i-a]. Une diminution de la durée de ce paramètre est donc observée pour tous les locuteurs. Il est intéressant de noter que, même les patients pour lesquels la durée du VOT était deux fois plus importante que celle relevée pour leur locuteur contrôle en post-opératoire 2, présentent des valeurs de VOT attendues pour un locuteur non pathologique dès cette phase d'enregistrement. Il semble donc qu'un meilleur contrôle des vibrations laryngées est à nouveau possible pour tous les locuteurs. Citons comme exemple, le cas de UPPBAS en contexte [a-i] : la durée du VOT est de 31,30 ms (9,93 ms). La même tendance se confirme pour le contexte [i-a] où la durée du VOT est de 33,70 ms (5,91 ms), toujours chez ce même locuteur. Quels que soient le locuteur et le contexte considérés, les écarts-types demeurent faibles.

Enfin, en post-opératoire 4 (phase en violet), les durées de VOT pour les deux contextes sont comparables à celles obtenues en post-opératoire 3, pour la plupart des locuteurs, ce qui traduit le fait que le contrôle des vibrations laryngées est maintenu. Notons que les écarts-types restent également faibles lors de cette phase d'enregistrement.

Concernant la patiente UPPBAS, les résultats indiquent une augmentation de la durée du VOT pour [adi] (VOT : 49,80 ms // écart-type : 9,80 ms) et [ida] (VOT : 59,60 ms // écart-type : 12,19 ms).

❖ Contexte [g]

Signalons que les moyennes pour le logatome [iga] sont calculées à partir des productions de six de nos sept locuteurs. En effet, UPPPAI n'a pas été en mesure de prononcer correctement ce logatome. Ses productions ont donc dû être écartées.

Les durées de VOT pour la consonne [g] sont illustrées par la Figure 24. Les productions des locuteurs contrôles sont conformes aux valeurs attendues pour cette consonne, étant donné que l'intervalle en question mesure 30 ms environ en moyenne. Les productions sont également régulières, puisque les écarts-types restent faibles, à 10 ms. Ces observations sont valables pour les deux contextes étudiés.

Lors de la première phase d'enregistrement des patients, soit en post-opératoire 2, la durée du VOT est plus importante, à savoir 41,31 ms et 43,65 ms pour [agi] et [iga] respectivement, que celle relevée lors de la phase d'enregistrement des locuteurs contrôles. La durée du VOT pour les autres phases d'enregistrement se réduit sans toutefois atteindre les valeurs de référence des locuteurs contrôles. En contexte [a-i], la durée diminue en post-opératoire 3, mesurée à 28,52 ms, avant de ré-augmenter légèrement en post-opératoire 4, où elle est mesurée à 33,25 ms. Le contexte [i-a] suit la tendance attendue, à savoir que le VOT commence à diminuer en post-opératoire 3, soit à 30,44 ms, ce qui se poursuit en post-opératoire 4, soit à 28,44 ms.

Pour la consonne [g] la variabilité, remarquable grâce à des écarts-types élevés, reste importante quels que soient la phase d'enregistrement et le contexte considérés. En contexte [a-i], la variabilité perdure puisque les écarts-types sont quasiment identiques lors des trois phases d'enregistrement, soit à 17,25 ms en post-opératoire 2, à 17,50 ms en post-opératoire 3 et à 19,73 ms en post-opératoire 4. En contexte [i-a], le scénario est légèrement différent, puisque la variabilité observée durant la phase d'enregistrement post-opératoire 2, qui est à 20,22 ms, se réduit en post-opératoire 3, à 9,29 ms, avant de ré-augmenter en post-opératoire 4 à 14,57 ms. Ces phénomènes peuvent être attribués aussi bien à des différences inter qu'intra-individuelle dans les productions des patients.

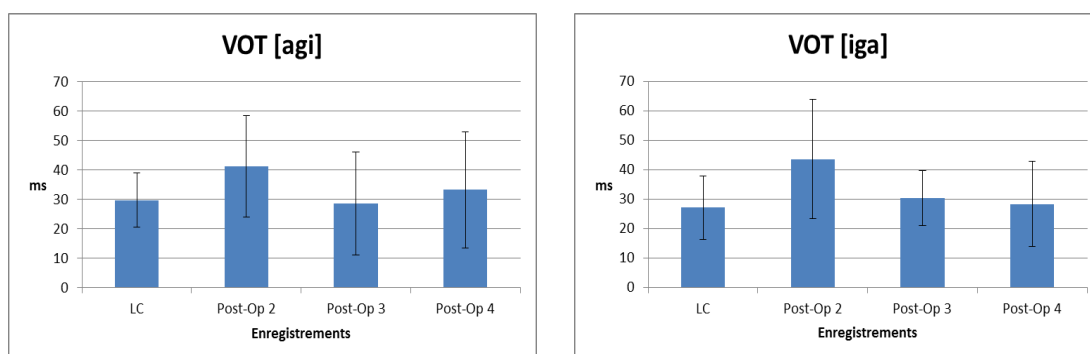


Figure 24 : Valeurs absolues de la durée du VOT (Klatt, 1975) en contexte [g] pour les logatomes [agi] (à gauche) et [iga] (à droite)

En ce qui concerne le logatome [iga], la durée du VOT est plus longue pour les patients en post-opératoire 2 que celle obtenue auprès de leurs locuteurs contrôles. Cette observation est valable pour tous les locuteurs. Le phénomène est particulièrement important pour UPPWAL, chez qui la durée du VOT est alors de 61,80 ms et chez UPPHUB pour qui le VOT est de 50,50 ms, en post-opératoire 2.

En contexte [i-a], la durée du VOT est plus importante pour les patients que pour les locuteurs contrôles (toujours mesurée à 30 ms environ). Ce phénomène est toutefois à nuancer en fonction du locuteur étudié. En effet, si l'on considère UPPBAS (57,50 ms), UPPHUB (64,40 ms), UPPLAT (42,40 ms), UPPPAI (30,40 ms) et UPPWAL (59,90 ms), la durée du VOT est significativement plus importante en post-opératoire 2 que celle de leurs locuteurs contrôles. La durée du VOT est alors au moins doublée pour ces locuteurs, comparée à celle obtenue pour leurs contrôles associés. En revanche, le VOT de UPPHEI (24,20 ms) et de UPPWAN (26,80 ms) n'est pas significativement plus long que celui de leur locuteur contrôle.

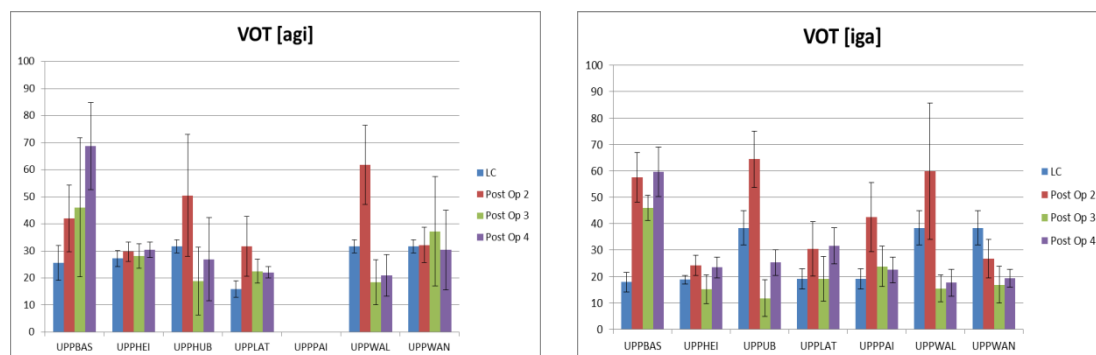


Figure 25 : Valeurs absolues de la durée du VOT (Klatt, 1975) en contexte [g] pour les logatomes [agi] (à gauche) et [iga] (à droite) en fonction du patient

En ce qui concerne la séquence [iga], la durée du VOT se réduit pour tous les patients dès la phase post-opératoire 3. Elle est alors de 45,20 ms chez UPPBAS et de 24,80 ms chez UPPWAL. Quant à la phase post-opératoire 4, elle ne correspond pas au phénomène attendu. Contrairement aux autres contextes, la durée du VOT présente une forte variabilité en fonction des locuteurs étudiés pour la consonne [g]. En contexte [a-i], par exemple, la durée du VOT est plus longue que celle relevée lors des phases d'enregistrement précédentes et auprès du locuteur contrôle pour UPPBAS (68,70 ms). Il s'agit là du VOT le plus long mesuré chez ce patient. Pour les autres locuteurs, la durée du VOT en post-opératoire 4 peut légèrement augmenter (c'est le cas pour UPPHUB : 26,90 ms) ou rester proches des mesures précédemment obtenues en post-opératoire 3 (c'est le cas pour UPPHEI : 30,50 ms). Pour ces derniers, la durée du VOT reste plus importante que celle obtenue pour les locuteurs contrôles, sans l'être significativement toutefois. Le contexte [i-a] suit le même scénario, à savoir qu'en post-opératoire 4, la durée du VOT est supérieure à celle mesurée en post-opératoire 3. Les mesures restent toutefois proches des valeurs de référence des locuteurs contrôles.

Il semble donc que la récupération du contrôle des vibrations laryngées soit plus problématique en contexte [g] que pour les deux autres occlusives voisées. Cette consonne témoigne également d'une importante variabilité quelle que soit la phase d'enregistrement étudiée. Comme attendu, la variabilité la plus importante est relevée lors de la phase d'enregistrement post-opératoire 2 (22,47 ms chez UPPHUB pour [agi] par exemple). Elle se réduit ensuite lors des phases d'enregistrement suivantes (12,48 ms en post-opératoire 3 et 15,48 ms en post-opératoire 4 chez UPPHUB pour [agi]) mais reste plus importante que celle des locuteurs contrôles. Le contrôle du VOT serait

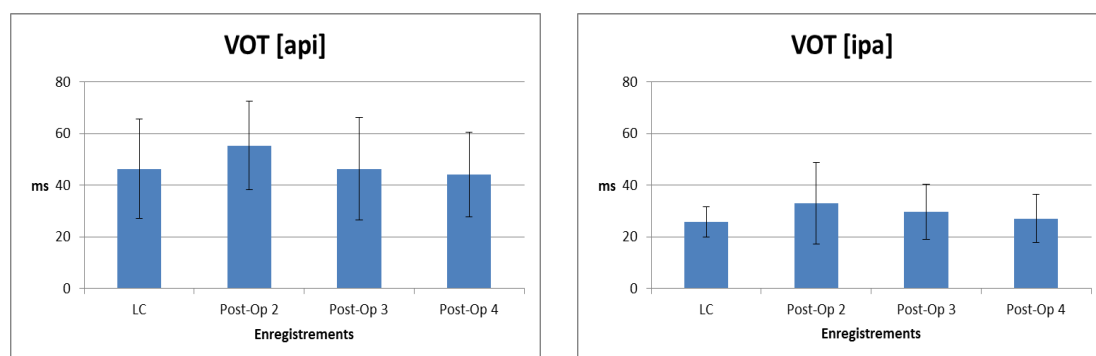
alors plus difficile pour cette consonne rendant les productions des patients plus variables. De manière générale, nous savons (*cf.* par ex., Abry, et al., 1990 in Sock 1998) que le relâchement dorso-vélaire pour réaliser le [g] peut présenter plusieurs pics (trois à quatre) sur le signal acoustique, témoignant ainsi des propriétés inertielles liées à la masse importante de cette structure. Etant donné que ce relâchement ne se fait pas toujours de manière nette (comme pour celui du [p]), le VOT des occlusives vélaires présente souvent une variabilité remarquable.

Nous proposons maintenant d'examiner le VOT des occlusives non voisées, avant de présenter une conclusion partielle pour les deux catégories de consonnes.

VIII.3.1.2. Occlusives non voisées

Nous présenterons dans cette section les résultats relatifs aux occlusives non voisées. Dans la mesure où le paramètre temporel du VOT n'a pas permis de dégager de différences significatives entre les différentes phases d'enregistrement et que la variabilité reste faible nous ne proposerons pas d'analyse par locuteur.

❖ Contexte [p]



Figures 26 : Valeurs absolues de la durée du VOT (Klatt, 1975) en contexte [p] pour les logatomes [api] (à gauche) et [ipa] (à droite)

En ce qui concerne la durée du VOT pour la consonne [p], il apparaît qu'elle n'est pas modifiée suite à l'immobilité laryngée. En effet, la durée de VOT des patients est plus longue (soit 55,32 ms pour [api] et 33,05 ms pour [ipa]) que celle mesurée à partir des productions des locuteurs contrôles (soit 46,31 ms pour [api] et 25,75 ms pour [ipa]) mais sans l'être significativement, que ce soit en contexte [a-i] ou en contexte [i-a]. De plus, quelle que soit la phase d'enregistrement post-opératoire considérée, les

durées obtenues à partir des productions des patients restent proches de celles des locuteurs contrôles pour les deux contextes, soit 46,27 ms et 44,04 ms pour [api], 29,71 ms et 27,10 ms pour [ipa], respectivement en phases post-opératoires 3 et 4.

Notons, comme constaté et expliqué précédemment, que la variabilité est plus marquée en contexte [a-i] qu'en contexte [i-a] pour toutes les phases d'enregistrement, et cela autant pour les locuteurs contrôles que pour les patients.

❖ Contexte [t]

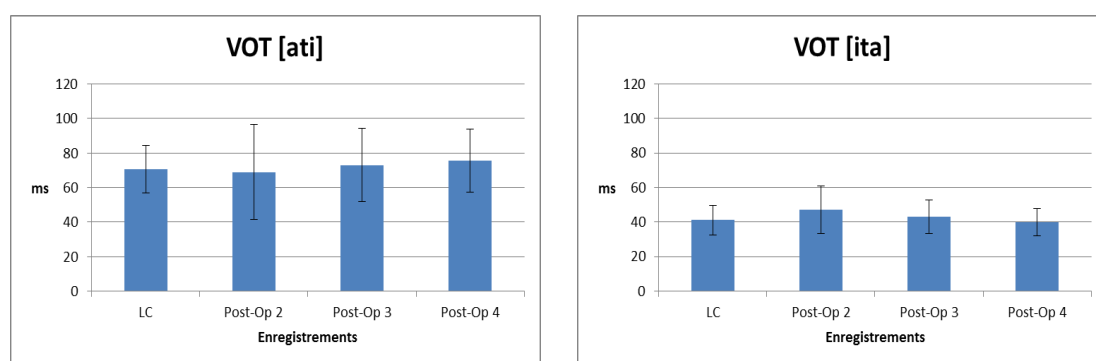


Figure 27 : Valeurs absolues de la durée du VOT (Klatt, 1975) en contexte [t] pour les logatomes [ati] (à gauche) et [ita] (à droite)

Le scénario est le même pour l'occlusive [t], à savoir que la durée du VOT n'est pas modifiée par l'immobilité laryngée pour les patients par rapport aux locuteurs contrôles, même en phase d'enregistrement post-opératoire 2. Les valeurs des patients restent donc proches des valeurs de référence des locuteurs contrôles (70,75 ms pour [ati] et 41,18 ms pour [ita]) en contexte [a-i], soit 69,00 ms en post-opératoire 2, 73,10 ms en post-opératoire 3 et 75,57 ms en post-opératoire 4 ; soit 47,21 ms en post-opératoire 2, 43,24 ms en post-opératoire 3 et 40,01 ms en post-opératoire 4.

Si l'on considère le contexte [a-i], la variabilité, traduite par les valeurs des écarts-types, est marquée à la fois pour les locuteurs de contrôle mais aussi pour les patients, et cela dans toutes les phases d'enregistrement. Elle est la plus importante pour la phase post-opératoire 2, à 27,47 ms, puis se réduit à partir de post-opératoire 3 et 4, à 21,31 ms et à 18,26 ms, respectivement.

En revanche, les écarts-types sont faibles pour le contexte [i-a] : ils sont de 8,44 ms pour les productions des locuteurs contrôles, de 13,71 ms en post-opératoire 2, de

9,67 ms en post-opératoire 3 et de 7,90 ms en post-opératoire 4. De plus, la variabilité suit la tendance attendue, à savoir qu'elle est la plus faible pour les productions des locuteurs contrôles, la plus importante en post-opératoire 2, avant de se réduire pour les phases d'enregistrement post-opératoires 3 et 4.

❖ Contexte [k]

L'occlusive [k] répond au même schéma que les deux occlusives non voisées décrites précédemment, à savoir que les valeurs des patients restent proches des valeurs des locuteurs contrôles pour les deux contextes, et cela quelle que soit la phase d'enregistrement post-opératoire étudiée. La valeur de référence des locuteurs contrôles pour le contexte [a-i] est de 68,00 ms (11,00 ms) ; en post-opératoire 2, la durée du VOT est de 71,42 ms (22,49 ms), alors qu'elle est de 77,37 ms (20,87 ms) en post-opératoire 3, et de 71,81 ms (18,98 ms) en post-opératoire 4. La durée de cet intervalle est donc plus importante, sans l'être significativement, pour les patients que pour celle obtenue auprès des locuteurs contrôles. Le même scénario est observable pour le contexte [i-a], où le VOT est de 52,08 ms (12,07 ms) en post-opératoire 2, de 43,92 ms (9,58 ms) en post-opératoire 3 et de 45,40 ms (10,22 ms) en post-opératoire 4. L'intervalle est de 45,75 ms (7,89 ms) pour les locuteurs contrôles.

Notons, encore une fois, que la variabilité, quantifiée à l'aide des écarts-types, est plus marquée pour le contexte [a-i] que pour le contexte [i-a]. Pour [aki], comme attendu, la variabilité est plus importante lors des phases d'enregistrement des patients que pour celle des locuteurs contrôles. Signalons cependant que cette même variabilité se réduit au cours du temps. En revanche, pour [ika], les écarts-types sont faibles pour les deux groupes de locuteurs.

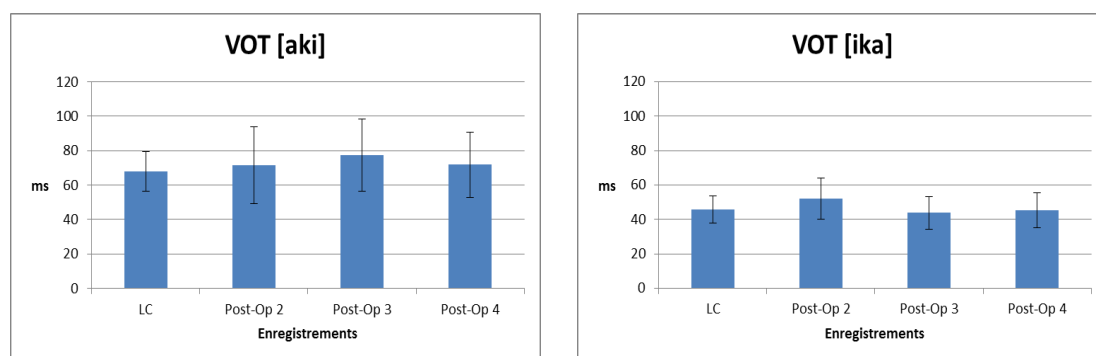


Figure 28 : Valeurs absolues de la durée du VOT (Klatt, 1975) en contexte [k] pour les logatomes [aki] (à gauche) et [ika] (à droite)

VIII.3.1.3. Le VOT comme indice de lieu d'articulation

Rappelons que le paramètre du VOT peut varier en fonction du lieu d'articulation. En effet, il a été observé, entre autres, que plus la closure a lieu à l'arrière de la cavité buccale, plus la durée du VOT est longue (Fischer-Jorgensen, 1954 ; Peterson et Lehiste, 1960). Cela voudrait dire que plus l'aire du contact est étendue, plus le VOT est élevé (Stevens *et al.*, 1986). Ainsi le VOT est généralement plus court pour les consonnes bilabiales et plus long pour le [k], le [t] ayant une durée intermédiaire dans la plupart des cas. Par ailleurs, on sait que le VOT de Klatt (1975) est plus long pour les occlusives non voisées que pour les occlusives voisées (Sack et Benoit, 1986). Traditionnellement en français, la durée du VOT d'un [b] est plus petite que celle du VOT d'un [d] qui est elle-même plus petite que celle d'un [g]. Le VOT est également plus court pour un [p] que pour un [t] ou un [k]. Si les données qui viennent d'être mentionnées valent pour des sujets sans trouble de la parole, la Figure 29 indique que ces constatations se maintiennent chez des locuteurs présentant une paralysie récurrentielle après thyroïdectomie, et cela pour les deux contextes étudiés, ainsi que pour toutes les phases d'enregistrement. Si la durée du VOT est effectivement allongée en post-opératoire 2 pour la production des occlusives voisées des patients, ces dernières n'en restent pas moins quantitativement différenciées.

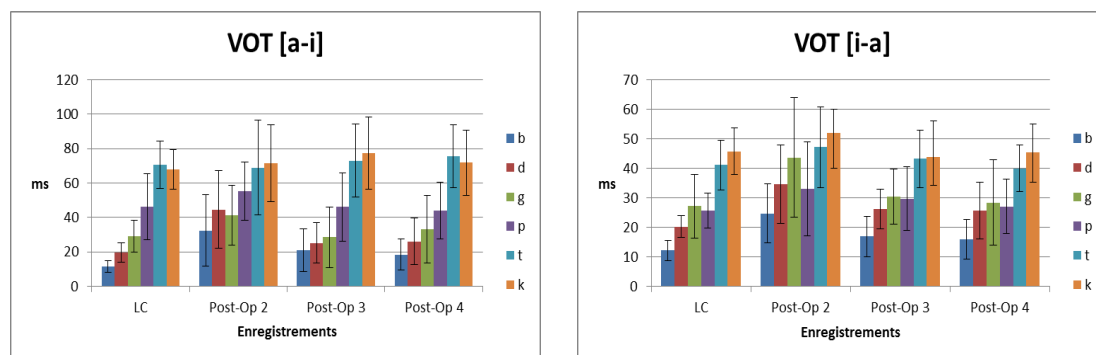


Figure 29 : Durées moyennes du VOT (Klatt, 1975) pour six occlusives du français en contexte [a-i] (à droite) et en contexte [i-a] (à gauche)

Pour résumer :

En post-opératoire 2, la durée du VOT est plus longue pour les trois occlusives voisées [b d g] par rapport à celle des locuteurs contrôles, et cela en contextes [a-i] et [i-a]. Cette phase d'enregistrement est également caractérisée par des écarts-types plus importants, ce qui témoigne d'une certaine variabilité lors de cette phase d'enregistrement. Notons que la différence entre la durée du VOT des patients et des locuteurs contrôles est plus grande dans le contexte [a-i]. Il est donc probable que la transition [C=> i] exige un contrôle plus précis et plus problématique que la transition [C=> a]. En effet, la voyelle [i] requérant une constriction réduite dans la zone palatale, sans qu'il y ait de contact entre le dos de la langue et la voûte palatine, les patients semblent avoir plus de difficulté à réaliser cette transition de la consonne vers la voyelle cible [i]. Cela ne semble pas être le cas pour le [a], dont l'abaissement mandibulaire, accompagné de celui de la masse de la langue avec qui la mandibule est fortement couplée, crée sans difficulté notoire la constriction radico-pharyngale nécessaire pour l'émergence acoustique de cette voyelle.

Les difficultés de voisement rencontrées par les patients suite à l'immobilité laryngée ont donc un impact sur les valeurs de VOT uniquement pour les consonnes voisées. Les difficultés de maintien des vibrations laryngées conduisent les patients à produire des VOT plus longs. En effet, les patients sont contraints d'arrêter l'activité laryngienne périodique pour la consonne voisée plus tôt, permettant ainsi un relâchement consonantique relativement précoce. Cela a pour conséquence d'augmenter la durée du VOT. Avec le temps et la rééducation vocale, une réduction des valeurs de VOT est observée, et cela dès la phase d'enregistrement post-opératoire 3, ce qui laisse

supposer une amélioration du contrôle des mouvements laryngés. Ce phénomène se poursuit jusqu'en post-opératoire 4, où les valeurs des VOT sont alors comparables à celles des locuteurs contrôles, notamment pour les occlusives [b] et [d].

L'étude des productions des patients de façon individuelle a permis de montrer que l'importante variabilité relevée en post-opératoire 2 était le reflet de différences inter et intra-individuelles. De façon générale, cette variabilité diminue lors des phases d'enregistrement plus tardives, ce qui va dans le sens d'un meilleur contrôle des vibrations laryngées : les productions sont alors plus régulières.

La durée du VOT des occlusives non voisées n'a pas permis de mettre au jour des différences significatives entre les productions des patients et des locuteurs contrôles, et ce quelle que soit la phase d'enregistrement étudiée. L'immobilité laryngée ne perturbe donc pas ce paramètre temporel. Rappelons que c'est surtout au niveau spectral que la voyelle est modifiée (*cf.* Chap. VII) et non pas au niveau du démarrage de l'activité glottique.

Signalons enfin que si la durée du VOT est effectivement allongée en post-opératoire 2 pour la production des occlusives voisées des patients, celles-ci n'en restent pas moins quantitativement différenciées les unes par rapport aux autres.

VIII.3.1.1. Durées relatives du VOT

Cette section traitera des durées relatives du VOT. Comme nous l'avons précisé plus haut, si les données temporelles en valeurs absolues permettent, certes, de quantifier l'ampleur d'un paramètre quelconque dans la différenciation de catégories phonétiques et/ou phonologiques, les valeurs relatives, elles, offrent *in fine* la possibilité d'apprécier le maintien et la robustesse ou non de ce paramètre dans des conditions de réalisation différentes. Sur le plan temporel, ce sont essentiellement les différences de vitesses d'élocution intrinsèques aux locuteurs qui sont les sources majeures de l'élasticité du signal de la parole (Gaitenby, 1965 ; Sock, 1998).

La Figure 30 indique que les observations faites pour les valeurs absolues sont confirmées lors de l'étude des valeurs relatives. Les durées relatives du VOT des consonnes voisées restent plus longues pour les patients présentant une immobilité laryngée que celles des locuteurs contrôles. Cela est particulièrement remarquable en

contexte [a-i] (à gauche sur la Figure 30). Les durées les plus importantes sont alors relevées en post-opérateur 2 (en rouge sur la Figure 30), puisque le VOT a été quantifié à 22,93% (10,15%) pour [b], à 31,61% (9,21%) pour [d] et à 26,99% (6,96%) pour [g]. Les durées se réduisent ensuite en post-opérateur 3, les occlusives voisées étant alors mesurées à 17,07% (8,55%), 21,35% (7,68%), 20,49% (8,69%) respectivement. Ces valeurs se maintiennent jusqu'en post-opérateur 4, soit 15,36% (6,67%) pour [b], 21,10% (7,44%) pour [d] et 22,88% (8,28%) pour [g]. Notons que cette dernière phase est également caractérisée par une variabilité réduite.

Les mêmes observations, dans une moindre mesure toutefois ($p=ns$), sont observables en contexte [i-a], à droite sur la Figure 30. Ces constatations étaient notées en valeurs absolues également. Il est donc probable que la transition [$C_{voisée} \Rightarrow i$] exige un contrôle plus précis et plus problématique que la transition [$C_{voisée} \Rightarrow a$]. La durée relative du VOT de [b d g] est plus importante lors de la phase post-opérateur 2 (en rouge sur la Figure 30), étant donné que le VOT de [b] était quantifié à 20,38% (8,65%), celui de [d] à 29,08% (8,14%) et celui de [g] à 32,96% (10,41%) à cette date. Dès la phase d'enregistrement suivante (en vert sur la Figure 30), les durées diminuent pour les trois consonnes : le [b] est à 16,75% (8,33%), le [d] à 26,80% (7,30%) et le [g] à 29,95 (8,51%). Ces valeurs se maintiennent en post-opérateur 4 (en violet sur la Figure 30) : 14,82% (6,89%) pour [b], 25,17% (7,66%) pour [d] et 26,29% (9,81%) pour [g].

Les durées relatives du VOT pour les consonnes non voisées ne permettent pas de distinguer les différentes phases d'enregistrement ($p=ns$), et ce dans les deux contextes [a-i] et [i-a].

Les écarts-types les plus importants sont relevés lors de la phase d'enregistrement post-opérateur 2, la variabilité se réduisant dans les phases d'enregistrement suivantes.

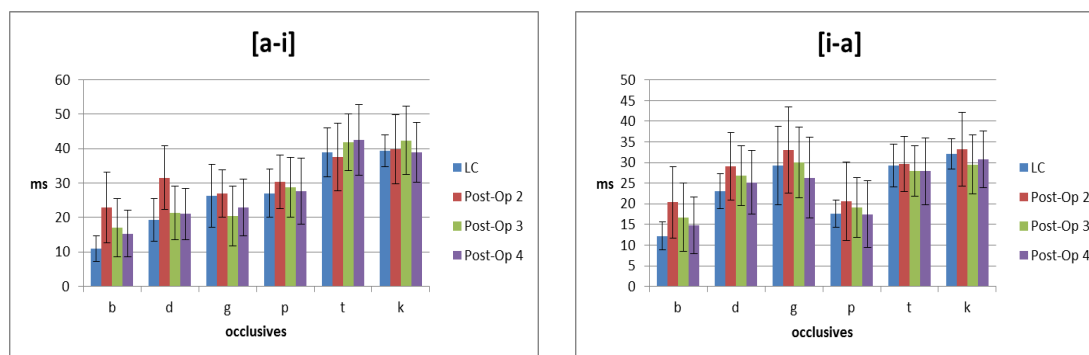


Figure 30 : Durées relatives moyennes du VOT (Klatt, 1975) pour six occlusives du français en contexte [a-i] (à droite) et en contexte [i-a] (à gauche)

Pour résumer :

Les valeurs relatives de VOT ont permis de confirmer les résultats obtenus à partir des valeurs absolues, à savoir que la durée du VOT des consonnes voisées est allongée pour les patients en post-opératoire 2, avant de se réduire à partir de la phase post-opératoire 3, pour s'approcher des valeurs obtenues à partir des productions des locuteurs contrôles. Ce phénomène est plus important dans le contexte [a-i] que dans le contexte [i-a]. Ceci pourrait être expliqué par le contexte lui-même. Il est possible que la transition [$C_{\text{voisée}} \Rightarrow i$] exige un contrôle plus précis et plus problématique que la transition [$C_{\text{voisée}} \Rightarrow a$]. L'étude de cet intervalle pour les consonnes non voisées n'a pas permis de mettre au jour des différences significatives entre les différentes phases d'enregistrement.

Les écarts-types les plus importants sont à remarquer lors de la phase d'enregistrement post-opératoire 2. La variabilité se réduit dans les phases d'enregistrement suivantes.

VIII.3.2. Durées du VTT

Dans cette section, nous proposons d'étudier un autre paramètre intrasegmental, le VTT (Agnello, 1975), considéré comme un indice du timing oro-laryngé. Ce paramètre est lié à l'arrêt des vibrations périodiques, donc, par extrapolation, à l'amortissement des vibrations laryngées (Sock, 1998) après obstruction du conduit vocal. L'étude de cet intervalle dans les séquences prononcées par les locuteurs atteints d'une paralysie récurrentielle, suite à une thyroïdectomie, indique qu'il n'est pas

modifié par l'immobilité laryngée. Les productions des patients, lors des phases post-opératoires, restent en effet proches des valeurs relevées pour les productions des locuteurs contrôles, soit environ 20 ms dans tous les contextes étudiés. C'est pour cette raison que nous proposons de n'étudier le VTT que dans deux contextes, à savoir [ati] et [ita], qui sont représentatifs de l'ensemble de nos données, et qui sont illustrés dans la Figure 31 ; les données et les graphiques pour les autres séquences sont disponibles en annexe.

Dans les séquences [ati] et [ita], l'immobilité laryngée diagnostiquée ne semble pas avoir d'effet sur ce paramètre, puisque le VTT, mesuré à partir des productions des patients, n'est que légèrement (et donc pas significativement) plus important lors des phases d'enregistrement post-opératoires 2, 3 et 4. Le VTT est d'environ 30 ms toutes phases d'enregistrement confondues pour la séquence [ati], prononcée par les locuteurs ayant été opérés, contre 26 ms pour les productions des locuteurs contrôles : ce même paramètre est de 29 ms pour [ita] produits par les patients contre 22,41 ms pour les productions des locuteurs contrôles (voir Figure 31).

Les productions des patients se distinguent toutefois par des écarts-types légèrement plus importants, reflète d'une variabilité plus grande dans l'ensemble des phases d'enregistrement post-opératoires. Pour exemple, les écarts-types, mesurés à partir de la séquence [ati], sont de 9,11 ms en post-opératoire 2, de 9,68 ms en post-opératoire 3 et de 11,77 ms en post-opératoire 4. Cette variabilité perdure dans toutes les phases d'enregistrement, et elle est présente dans les deux contextes. Les productions des locuteurs contrôles sont caractérisées par des écarts-types plus faibles (6,99 ms pour [ati] et 5,15 ms [ita]). Néanmoins, les écarts-types ne permettent pas de dégager des tendances pour ce paramètre (voir Figure 31).

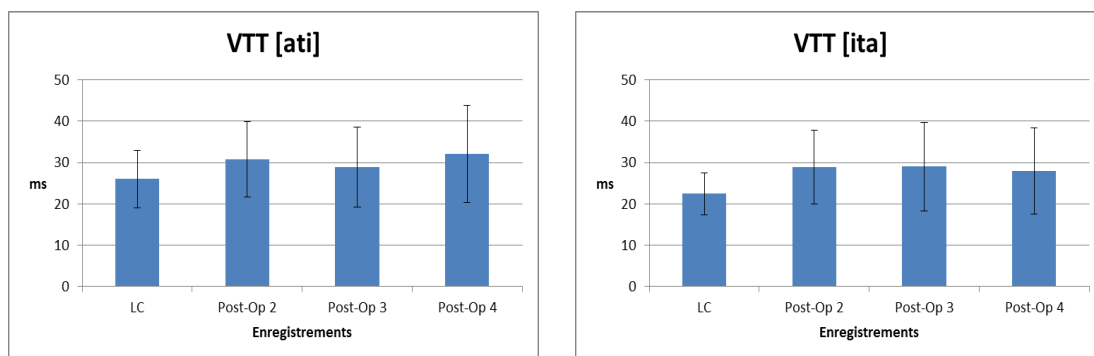


Figure 31 : Valeurs absolues de la durée du VTT (Agnello, 1975) en contexte [t] pour les logatomes [ati] (à gauche) et [ita] (à droite)

Pour résumer :

L'étude du VTT n'a pas permis de montrer de modification significative de ce paramètre suite à l'immobilité laryngée post-thyroïdectomie. Les productions des patients restent proches de celles des locuteurs contrôles, et ce quels que soit le contexte et la phase d'enregistrement considérés. L'amortissement des vibrations laryngées, quantifiable à partir de l'intervalle nécessaire à l'arrêt des vibrations périodiques, soit la mesure du VTT (Sock, 1998), ne varie pas avec l'immobilité laryngée.

Les valeurs des écarts-types sont également comparables pour les deux groupes. La variabilité n'est que légèrement plus marquée dans les productions des patients, par rapport à celle des contrôles.

VIII.3.3. Durées de l'occlusion

Dans cette section, nous souhaitons étudier la durée de l'intervalle allant de la fin de la structure formantique stable de V1, jusqu'au début de l'explosion de la consonne intervocalique. Il est à noter que si Meynadier & Gaydina (2012) considèrent cet intervalle comme la délimitation de la tenue consonantique, nous avons choisi de l'appeler « occlusion consonantique », dans la mesure où il s'étend de la fin de la réalisation vocalique jusqu'au début du relâchement de la consonne. Si l'on accepte que le VOT est une phase de transition CV, après un relâchement supraglottique, il n'est pas incohérent de considérer que la phase purement occlusive de la tenue consonantique s'arrête avant tout relâchement supraglottique. D'après la littérature (voir Vercherand, 2010 ou Meynadier & Gaydina, 2012 par exemple), cet intervalle serait plus long en parole chuchotée qu'en parole modale, et naturellement plus long pour les consonnes non voisées que pour les consonnes voisées. Cette différence entre consonnes voisées et non-voisées, bien qu'affaiblie, serait maintenue en voix chuchotée (voir Malécot, 1955 et Meynadier & Gaydina, 2012).

Des analyses de variances à trois facteurs, *voisement*, *lieu d'articulation* et *phase d'enregistrement* ont été conduites pour la variable *occlusion*. Un effet de *voisement* a pu être mis au jour [(F1,6) = 35,51 p<0, 000999]. Les tests *post-hoc* révèlent que toutes phases d'enregistrement et tous lieux d'articulation confondus, la durée de l'occlusion

des consonnes voisées est significativement différente de celle des consonnes non voisées. La durée de l'occlusion serait également significativement différente [$F(2,12) = 5,51$ $p < 0,020004$] en fonction de la consonne étudiée. Ainsi, la durée de l'occlusion des bilabiales [p b] est significativement différente de celle des autres consonnes [t d] et [k g]. Aucun effet de *phase d'enregistrement* ($p = ns$) n'a pu être mis au jour.

L'interaction des trois facteurs, *voisement*, *lieu d'articulation* et *phase d'enregistrement*, s'est révélée significative [$F(6,36) = 2,64$ $p < 0,031575$] elle repose sur une différence significative entre les productions des locuteurs contrôles et les productions des patients en post-opératoire 2, pour les consonnes voisées. Dès la phase post-opératoire 3, les différences s'estompent. Aucune différence significative n'a pu être observée entre les différentes phases d'enregistrement pour les consonnes non voisées.

L'analyse de nos données révèle que, en contexte [a-i], la durée de l'occlusion n'est pas différente suite à l'immobilité laryngée, pour cinq des six occlusives testées. La durée de l'intervalle est alors d'environ 100 ms pour la consonne [b], et de 90 ms pour la consonne [d]. Le même constat est observable pour les occlusives non voisées, à savoir que les durées de l'occlusion restent proches, que l'on considère les productions des locuteurs contrôles ou celles des patients dans toutes les phases post-opératoires : l'occlusion est d'environ 130 ms pour [p], 110 ms pour [t] et 100 ms pour [k]. Seule l'occlusion de [g] est plus longue en post-opératoire 2, soit 110,16 ms (29,78 ms), que celle mesurée pour le groupe contrôle, soit 76,71 ms (16,85 ms). Par ailleurs, la durée de l'occlusion reste, dans les phases d'enregistrement suivantes, supérieure à celle mesurée pour les locuteurs contrôles, soit 106,12 ms (26,96 ms) et 104,68 ms (20,47 ms) respectivement en post-opératoire 3 et post-opératoire 4 (voir Figure 32).

En contexte [i-a], le scénario est sensiblement le même que celui décrit précédemment. La durée de l'occlusion a néanmoins tendance à être plus importante lors de la phase post-opératoire 2, pour toutes les consonnes testées, par rapport à celles mesurées à partir des productions des locuteurs contrôles. Prenons pour exemple, l'occlusion de la consonne [b] qui est de 88,12 ms (12,72 ms) pour les locuteurs contrôles et de 98,14 ms (19,30 ms) en post-opératoire 2. L'intervalle se réduit ensuite en post-opératoire 3, où il est alors de 87,51 ms (17,00 ms). En post-opératoire 4, l'intervalle est à nouveau plus grand, soit 95,90 ms. L'écart-type est également plus

important lors de cette phase d'enregistrement (25,00 ms), ce qui témoigne d'une certaine variabilité dans la production des patients. En contexte [i-a], les observations faites pour les occlusives voisées sont également confirmées pour les occlusives non voisées. L'occlusion de la consonne [p], par exemple, est mesurée à 123,06 ms (19,96 ms) pour les locuteurs contrôles et à 131 ms (36,59 ms) pour les patients en post-opératoire 2. L'intervalle se réduit ensuite en post-opératoire 3, mesuré alors à 119,22 ms (24,78 ms) avant de ré-augmenter légèrement en post-opératoire 4 pour atteindre 128,58 ms (36,35 ms).

Ce scénario est toutefois à nuancer au regard des écarts-types, puisque leur importance ne permet pas de dégager une différence significative ($p=ns$) entre les productions des locuteurs contrôles et des patients, et ce quelle que soit la consonne étudiée (voir Figure 32 pour une illustration de ce phénomène).

Signalons que les phases d'enregistrement des patients sont caractérisées par des écarts-types plus importants pour les deux contextes, notamment dans la phase d'enregistrement post-opératoire 2. La variabilité reste importante, même lors des enregistrements post-opératoires plus tardifs.

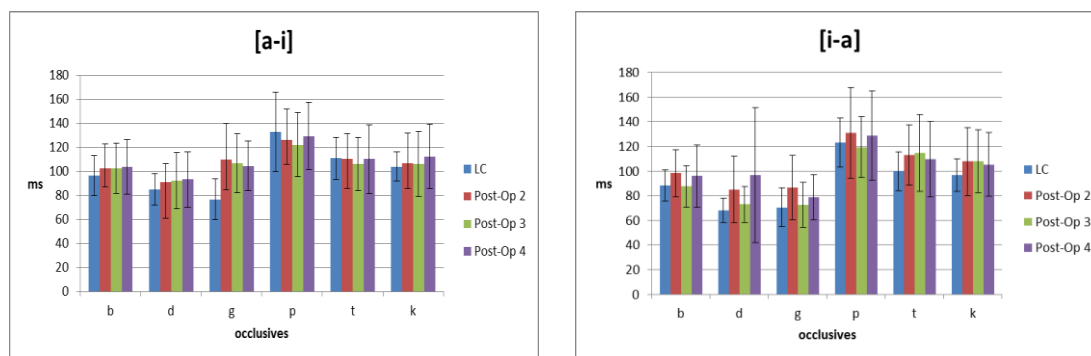


Figure 32 : Durées moyennes de l'occlusion en contexte [a-i] (à gauche) et en contexte [i-a] (à droite)

Pour résumer :

En contexte [a-i], la durée de l'occlusion n'est pas modifiée par l'immobilité laryngée. L'intervalle reste proche des valeurs obtenues auprès des locuteurs contrôles, pour toutes les consonnes et toutes les phases d'enregistrement étudiées.

En contexte [i-a], il y aurait une tendance à l'allongement de l'occlusion pour les productions des patients en post-opératoire 2, par rapport aux locuteurs contrôles. Il est toutefois ici uniquement question de tendance compte tenu de l'importance des écarts-

types. L'intervalle mesuré à partir des productions des patients conserve une durée similaire, quelle que soit la phase d'enregistrement post-opératoire considérée.

Les phases d'enregistrement des patients sont également caractérisées par des écarts-types plus importants. La variabilité reste élevée jusque dans les phases d'enregistrement tardives. Les productions des patients présentant une immobilité laryngée ne correspondent pas aux caractéristiques temporelles attendues en condition de voix chuchotée.

VIII.3.4. Durées de la tenue consonantique

Rappelons que nous considérons la tenue consonantique comme l'intervalle allant de la fin de la structure formantique stable de V1 jusqu'au début de la structure formantique stable de V2. Cette section propose d'étudier les éventuelles modifications que cet intervalle peut subir suite à l'immobilité laryngée.

Les analyses de variance à deux facteurs *phase* et *consonne* ont été conduites pour la variable *tenue consonantique*. Un effet principal de *consonne* [$F(5,30) = 59,31$ $p < 0,000000$], a pu être révélé. En revanche, l'effet de *phase d'enregistrement* ne s'est pas montré significatif ($p = ns$). L'interaction de ces deux facteurs a été significative [$F(15,90) = 2,98$ $p < 0,000677$]. Les tests *post-hoc* révèlent plusieurs différences significatives. Premièrement, les consonnes voisées sont significativement différentes des consonnes non voisées, et cela toutes phases d'enregistrement confondues. Deuxièmement, pour toutes les consonnes voisées, il existe une différence significative entre les productions des patients en post-opératoire 2 et toutes les autres phases d'enregistrement. Il n'y a pas de différence significative entre les productions des locuteurs contrôles et celles des patients en post-opératoire 3 et 4. Troisièmement, il n'existe pas de différence significative entre les productions des consonnes non voisées des locuteurs contrôles et celles des patients, et ce pour toutes les phases d'enregistrement.

En contexte [a-i] ou [i-a], cet intervalle est toujours plus long pour les trois occlusives non voisées [p t k] que pour les trois occlusives voisées [b d g], et ce pour les quatre phases d'enregistrement (voir Figure 33).

Il convient toutefois de remarquer que, lors de la phase d'enregistrement post-opératoire 2 (phase en rouge sur la Figure 33), cet intervalle pour les consonnes voisées

a tendance à être plus long que celui mesuré à partir des productions des locuteurs contrôles. Ainsi, la tenue consonantique est de 108,25 ms (15,90 ms) en contexte [a-i] pour la consonne [b] chez le groupe contrôle, et de 135,13 ms (32,18 ms) pour les patients en post-opératoire 2. La même tendance est observable pour les deux autres consonnes, puisque cet intervalle est mesuré à 105,55 ms (15,42 ms) et à 106,21 ms (16,44 ms) respectivement pour les consonnes [d] et [g] prononcées par les locuteurs contrôles ; ce même intervalle a été quantifié à 135,60 ms (30,67 ms) et à 154,83 ms (50,18 ms) pour les [d] et [g] produits par les patients en post-opératoire 2. Comme on peut le constater, la phase d'enregistrement post-opératoire 2 est caractérisée par des écarts-types plus importants, témoins d'une variabilité accrue. La Figure 33 illustre ces phénomènes.

Dès la phase d'enregistrement post-opératoire 3 (en vert sur la Figure 33), la tenue consonantique se réduit chez les patients pour les trois consonnes voisées étudiées. Il est alors de 123,93 ms (22,88 ms) pour [b], de 118,24 ms (27,59 ms) pour [d] et de 135,43 ms (29,81 ms) pour [g]. Les écarts-types restent importants mais sont également plus petits que ceux mesurés en post-opératoire 2, ce qui laisse penser que la variabilité commence à se réduire.

La tendance amorcée en post-opératoire 3 se poursuit en post-opératoire 4 (en violet sur la Figure 33) pour les consonnes [b] et [d], dans la mesure où l'intervalle étudié et les écarts-types continuent à se réduire : les valeurs moyennes se rapprochent des valeurs des locuteurs contrôles, puisqu'elles ont été mesurées à 122,31 ms (24,02 ms) pour [b] et à 115,27 ms (32,74 ms) pour [d] durant cette phase d'enregistrement. Signalons cependant que les écarts-types restent plus importants. Seul l'intervalle mesuré à partir de l'occlusive [g] indique une durée moyenne proche de celle mesurée en post-opératoire 3, à savoir 137,93 ms (32,74 ms). Ces valeurs restent différentes de celles obtenues à partir des productions des locuteurs contrôles.

En ce qui concerne les trois occlusives non voisées, les valeurs obtenues à partir des productions des patients restent proches des valeurs des locuteurs contrôles, soit environ 180 ms, et ce quelle que soit la consonne étudiée (voir Figure 33).

Les mêmes tendances sont remarquables pour le contexte [i-a] (voir Figure 33, à droite) : la tenue consonantique reste plus longue pour les occlusives non voisées que pour les occlusives voisées. Le contexte sourd ne semble pas modifié par l'immobilité

laryngée, puisque les valeurs obtenues auprès des patients en phase post-opératoire sont proches des valeurs des locuteurs contrôles, soit environ 150 ms, quelles que soient la consonne et la phase étudiée. En outre, le contexte voisé indique le même type de tendances que celles observées en contexte [a-i], à savoir que l'intervalle est plus long en post-opératoire 2 (phase en rouge), avant de diminuer et de se rapprocher des valeurs des locuteurs contrôles à partir de la phase post-opératoire 3 (phase en vert). Notons que cette observation se poursuit jusqu'en post-opératoire 4 (phase en violet). Prenons pour exemple l'intervalle mesuré à partir de la consonne [d] : il est de 88,37 ms (10,01 ms) pour le groupe contrôle, de 119,81 ms (34,64 ms) en post-opératoire 2, de 100,04 ms (18,03 ms) en post-opératoire 3 et 104,21 ms (25,08 ms) en post-opératoire 4, ce qui traduit bien le fait que la durée de la tenue consonantique est plus importante chez les patients, même si celle-ci a tendance à diminuer à mesure que l'on s'éloigne de la date de l'opération. Les écarts-types sont également plus élevés en enregistrements postopératoires.

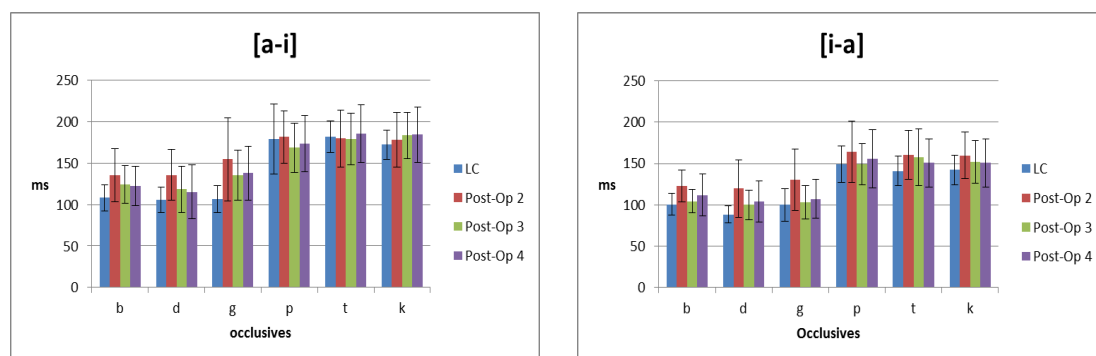


Figure 33 : Durées moyennes de l'intervalle allant de la fin de la structure formantique stable de V1 au début de l'explosion en contexte [a-i] (à gauche) et en contexte [i-a] (à droite) pour six occlusives du français

Pour résumer :

La durée de la tenue consonantique des occlusives voisées a tendance à être plus longue pour les patients en post-opératoire 2 par rapport aux locuteurs contrôles, et ce dans les deux contextes vocaliques étudiés. Ce constat pouvait être attendu au regard des résultats obtenus à partir de la durée du VOT notamment, puisque cet intervalle était allongé pour les occlusives voisées des patients en post-opératoire 2, par rapport à celles produites par les locuteurs contrôles.

En revanche, l'étude de cet intervalle durant la production des occlusives non voisées n'a pas permis de mettre au jour des différences significatives entre les locuteurs opérés et contrôles, ainsi qu'entre les différentes phases d'enregistrement (locuteurs contrôles vs. patients).

Enfin, bien que la tenue consonantique s'allonge dans les productions des patients pour les consonnes voisées, elle reste toujours plus longue pour les occlusives non voisées que pour les occlusives voisées, et ce quelle que soit la phase d'enregistrement étudiée. Il semble donc que cet intervalle, qui participe, dans une certaine mesure, à la distinction entre les occlusives voisées et non voisées, ne soit pas modifié par l'immobilité laryngée comme nous aurions pu l'attendre, compte tenu des irrégularités dans le mouvement des plis vocaux des patients, lors des phases d'enregistrement post-opératoires précoces.

VIII.3.5. Durées vocaliques vs. durées consonantiques

Nous avons jusqu'à présent attaché une attention particulière aux intervalles qui s'étendent de la fin de la structure formantique stable de V1 au début de la structure formantique stable de V2. Cette section s'attachera à décrire la durée de V1 et de V2, délimitée par la structure formantique stable, propre à chacune des deux voyelles. Il s'agira d'observer si la durée de ces voyelles peut également se trouver modifiée suite à l'immobilité laryngée.

L'étude des intervalles vocaliques dans les séquences prononcées par les locuteurs atteints d'une paralysie récurrentielle, suite à une thyroïdectomie, indique qu'ils ne sont pas modifiés ($p=ns$) par l'immobilité laryngée. C'est pour cette raison que nous proposons de n'étudier ces intervalles que dans deux séquences pour chacun des contextes, à savoir [adi] vs. [ati] et [ida] vs. [ita], qui sont représentatives de l'ensemble de nos données. Les résultats, sous forme de représentations graphiques pour les autres séquences, sont disponibles en annexe.

En ce qui concerne la durée de V1 dans la séquence [adi] (à gauche sur la Figure 34), les valeurs des patients restent proches de celles obtenues auprès des locuteurs contrôles, pour toutes les phases d'enregistrement post-opératoires ($p=ns$). La durée de la voyelle est de 96,84 ms (8,99 ms) pour les sujets de contrôle (en bleu sur la Figure

34) ; elle est plus importante en post-opératoire 2 (en rouge sur la Figure 34), 102,31 ms (26,51 ms), diminue en post-opératoire 3 (en vert sur la Figure 34) à 93,24 ms (21,62 ms) et est mesurée à 97,61 ms (20,28 ms) en post-opératoire 4 (en violet sur la Figure 34). Notons que les phases d'enregistrements des patients sont caractérisées par des écarts-types plus importants. La tendance est toutefois à une réduction progressive de la variabilité.

La durée de V2 pour la séquence [adi] (à gauche sur la Figure 34) est légèrement plus importante ($p=ns$) chez les patients en post-opératoire 2, où elle a été mesurée à 87,40 ms (28,25 ms) (en rouge sur la Figure 34), par rapport aux productions des locuteurs contrôles (en bleu sur la Figure 34), pour lesquels V2 a été quantifié à 77,08 ms (11,68 ms). Cet intervalle a tendance à augmenter ($p=ns$) au cours du temps, puisqu'il est mesuré à 93,80 ms (23,27 ms) en post-opératoire 3 (en vert sur la Figure 34) et à 95,90 ms (21,57 ms) en post-opératoire 4 (en violet sur la Figure 34). La variabilité diminue, puisque les écarts-types les plus petits sont observables lors de la dernière phase d'enregistrement post-opératoire.

Les remarques sur la durée de la consonne ont déjà été effectuées en VIII.3.4. Rappelons donc brièvement ici que la consonne est plus longue lors des productions des patients en post-opératoire 2 que lors des productions des locuteurs contrôles. La durée de cet intervalle se réduit ensuite au cours du temps pour se rapprocher des valeurs des contrôles.

En ce qui concerne les durées vocaliques de la séquence [ati] (à droite sur la Figure 34), le scénario pour V1 est identique à celui observé pour la première voyelle de la séquence [adi]. Les valeurs sont proches, aussi bien pour les productions des locuteurs contrôles (83,60 ms, avec un écart-type de 14,76 ms), que pour les productions des patients pour toutes les phases d'enregistrement. En effet, V1 est mesuré à 91,62 ms (26,97 ms) en post-opératoire 2 (en rouge sur la Figure 34), à 84,02 ms (22,76 ms) en post-opératoire 3 (en vert sur la Figure 34) et à 84,35 ms (16,15 ms) en post-opératoire 4 (en violet sur la Figure 34). La variabilité est plus marquée en post-opératoire 2, mais elle commence à se réduire dès la phase d'enregistrement post-opératoire 3, ce phénomène se poursuivant jusqu'en post-opératoire 4, où les écarts-types sont les plus faibles.

La durée de V2 dans la séquence [ati] (à droite sur la Figure 34) est toujours plus importante pour les patients que pour les locuteurs contrôles, pour lesquels cet intervalle a été mesuré à 54,88 ms (12,55 ms). Les phases post-opératoire 2 (en rouge sur la Figure 34) et post-opératoire 4 (en violet sur la Figure 34) indiquent des valeurs relativement similaires, soit 83,55 ms (29,88 ms) et 82,50 ms (29,52 ms) respectivement. Seule la phase d'enregistrement post-opératoire 3 (en vert sur la Figure 34) indique une durée et une variabilité réduites pour V2, à savoir 74,82 ms (16,09 ms).

La durée de l'intervalle consonantique n'est pas modifiée par l'immobilité laryngée. Les valeurs relevées pour les patients sont proches de celles des contrôles, quelle que soit la phase d'enregistrement considérée (voir 4.3. pour plus de détails). La Figure 34 illustre ces remarques.

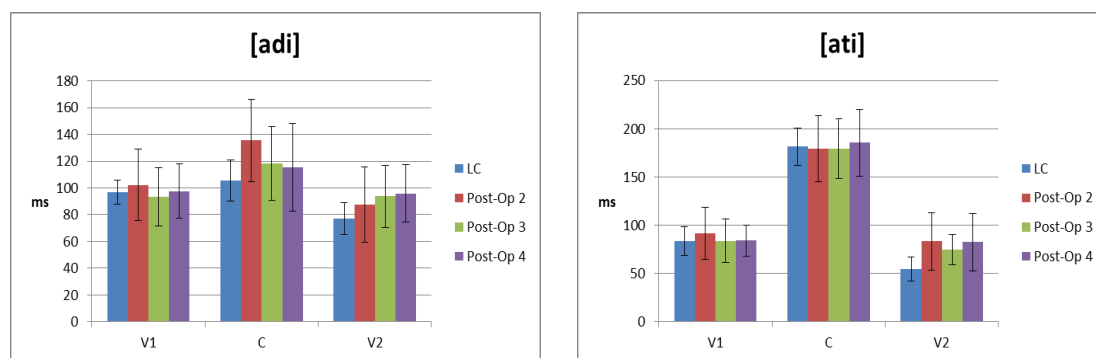


Figure 34 : Durées vocaliques et consonantiques pour la séquence [adi] (à gauche) et pour la séquence [ati] (à droite)

Le contexte [i-a], illustré par la Figure 35, indique des tendances comparables à celles commentées précédemment (voir Figure 34) pour l'occlusive non voisée [t].

En effet, en ce qui concerne la durée de V1, les valeurs restent proches ($p=ns$), que l'on considère les enregistrements des locuteurs contrôles, soit 78,82 ms (12,60 ms), ou les enregistrements des patients, soit 80,75 ms (27,90 ms), 72,24 ms (20,65 ms) et 73,24 ms (17,73 ms) respectivement en phase d'enregistrement post-opératoire 2, 3 et 4. De plus, la variabilité la plus importante dans les productions des patients est relevée lors de la phase d'enregistrement post-opératoire 2, avant de se réduire pour atteindre la plus petite valeur en post-opératoire 4.

La durée de V2 est légèrement plus importante ($p=ns$) en post-opératoire 2 (en rouge sur la Figure 35), soit 104,57 ms (24,15 ms), que celle relevée pour les

productions des locuteurs contrôles (en bleu sur la Figure 35), soit 90,25 ms (18,93 ms). Les durées de V2, pour les productions des patients, restent proches lors des trois phases d'enregistrement post-opératoires, à savoir 101,10 ms (21,34 ms) en post-opératoire 3 (en vert sur la Figure 35) et 110,72 ms (29,67 ms) en post-opératoire 4 (en violet sur la Figure 35).

Etant donné que la durée de la consonne a déjà été commentée précédemment, (voir VIII.3.4. nous rappelons simplement ici que la durée des occlusives non voisées n'est pas modifiée suite à l'immobilité laryngée. Cet intervalle ne permet pas de conclure à des différences de durées entre les productions des locuteurs contrôles et des patients.

Le scénario est en revanche différent si l'on considère les durées vocaliques des séquences VCV, où C est une consonne voisée. En effet, comme le montre la Figure 35, si la durée de V1 diminue pour les productions des patients par rapport à celles des locuteurs contrôles, la durée de V2 tend à augmenter ($p=ns$). En ce qui concerne la durée de V1 pour le groupe de contrôle, elle est de 103,67 ms (17,55 ms). Cet intervalle est plus petit pour les patients, et ce lors des trois phases d'enregistrement post-opératoires. En effet, il est de 92,58 ms (30,58 ms), 84,55 ms (23,13 ms) et 88,27 ms (22,27ms) respectivement en post-opératoires 2, 3 et 4. La variabilité reste importante lors de toutes les phases d'enregistrement, même si elle a tendance à se réduire au cours du temps.

La durée de V2 tend également à être plus importante dans les productions des patients que pour les productions des locuteurs contrôles ($p=ns$). Les écarts-types, qui ne se réduisent pas avec le temps, montrent que la variabilité reste importante dans toutes les phases d'enregistrement post-opératoires. A titre indicatif, V2 mesure 106,71 ms (17,85 ms) en moyenne pour les locuteurs contrôles (en bleu dans la Figure 35) et 113,28 ms (26,04 ms) en post-opératoire 2 (en rouge dans la Figure 35), 112,38 ms (20,06 ms) en post-opératoire 3 (en vert dans la Figure 35) et 120,77 ms (27,67 ms) en post-opératoire 4 (en violet dans la Figure 35).

L'intervalle consonantique, lors des productions des patients en post-opératoire 2, est plus long que pour les locuteurs contrôles. Cet intervalle diminue dès la phase post-opératoire 3, pour s'approcher des valeurs obtenues auprès des locuteurs contrôles.

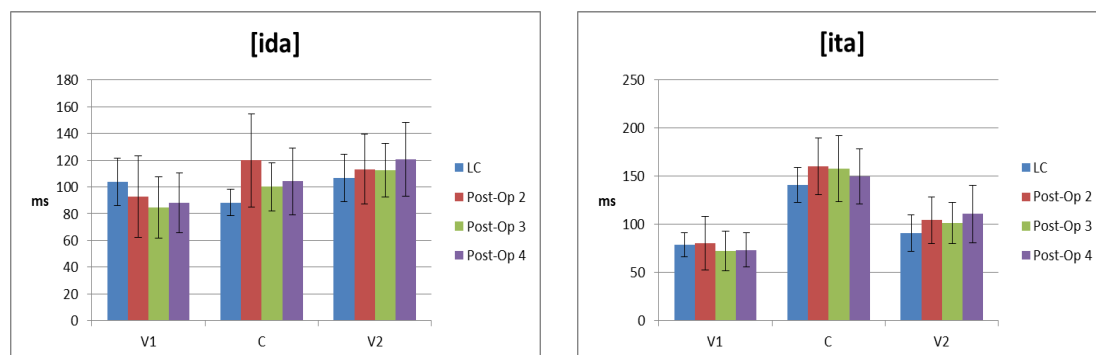


Figure 35 : Durées vocaliques et consonantiques pour la séquence [ida] (à gauche) et pour la séquence [ita] (à droite)

Pour résumer :

Les durées moyennes respectives des voyelles V1 et V2, produites par les patients présentant une paralysie récurrentielle post-thyroidectomie, ne se sont pas révélées différentes de celles des locuteurs contrôles. Il semble donc que l'immobilité laryngée ne perturbe pas la durée des voyelles dans la production des séquences VCV.

Des modifications de durées restent observables dans la tenue des consonnes voisées, notamment en raison de l'allongement du VOT (voir VIII.3.1.1.). Aucune modification de la durée de la consonne n'a pu être observée dans la production des séquences VCV dans lesquelles C est une consonne non voisée.

Le contexte vocalique, [a-i] ou [i-a], n'a pas permis de montrer de différences entre locuteurs contrôles et patients, en ce qui concerne la durée de l'intervalle vocalique.

Enfin, les écarts-types les plus importants sont apparus dans les productions des patients, notamment lors de la phase d'enregistrement post-opératoire 2. La variabilité tend à se réduire au cours du temps.

VIII.4. Discussion

Il convient à présent de vérifier si nos hypothèses initiales ont été confirmées ou infirmées, en ce qui concerne le timing des paramètres retenus.

1. Les difficultés de voisement rencontrées par les patients ont un impact sur la durée du VOT (*Hypothèse 1*). En raison de difficultés de maintien de vibrations

laryngées, le VOT est plus long pour les occlusives voisées produites par les patients que celles produites par les locuteurs contrôles. En effet, les vibrations périodiques s'estompent plus tôt, « laissant ainsi la place », de manière précoce, au relâchement consonantique. Cela est particulièrement vrai lors de la phase d'enregistrement post-opératoire 2. La durée du VOT de Klatt (1975) des occlusives voisées des patients reste quantitativement différente de celle mesurée pour leurs occlusives non voisées ; la distinction entre ces deux catégories devrait rester possible dans une perspective perceptive. Signalons que dans certains cas, les occlusives voisées, produites par les patients dans les phases d'enregistrement post-opératoires précoces, sont plus ou moins désonorisées puisque produites avec une certaine difficulté de maintien de l'activité laryngienne.

2. Le VTT (Agnello, 1975), reflet également du timing oro-laryngé, n'est pas modifié par l'immobilité laryngée. Les productions des patients restent proches de celles des locuteurs contrôles, et ce quels que soit le contexte et la phase d'enregistrement considérés. L'amortissement des vibrations laryngées, quantifiable à partir de l'intervalle nécessaire à l'arrêt des vibrations périodiques, soit la mesure du VTT (Sock, 1998), ne varie pas avec l'immobilité laryngée. Ces résultats infirment *l'hypothèse 2*. Nous supposons qu'en raison des difficultés de maintien du voisement, le VTT des patients serait plus court que celui des contrôles. Toutefois, cet intervalle apparaît parfois dans certaines occlusives voisées produites par les patients, témoin d'une certaine désonorisation de celles-ci. Ce phénomène reste à quantifier.
3. La durée de l'occlusion pourrait également se voir modifiée par l'immobilité laryngée. En parole chuchotée, la durée de l'occlusion serait significativement plus longue que celle observée en parole modale. Contrairement à ce que nous pensions sous notre *hypothèse 3*, les productions des patients présentant une immobilité laryngée ne correspondent pas aux caractéristiques temporelles attendues en condition de voix chuchotée. L'intervalle mesuré à partir des productions des patients conserve une durée similaire, quelle que soit la phase d'enregistrement post-opératoire considérée, et reste ainsi proche des valeurs

obtenues à partir des productions des locuteurs contrôles. Ainsi, les durées de l’occlusion de nos patients ne sont pas plus longues que celles des contrôles.

4. La durée de la tenue consonantique des occlusives voisées a tendance à être plus longue pour les patients en post-opératoire 2, par rapport aux locuteurs contrôles. En revanche, l’étude de cet intervalle durant la production des occlusives non voisées n’a pas permis de mettre au jour des différences significatives. Toutefois, bien que la tenue consonantique s’allonge dans les productions des patients pour les consonnes voisées, elle reste toujours plus longue pour les occlusives non voisées que pour les occlusives voisées, et ce quelle que soit la phase d’enregistrement étudiée. Il semble donc que cet intervalle, qui participe, dans une certaine mesure, à la distinction entre les occlusives voisées et non voisées, ne soit pas modifié par l’immobilité laryngée, comme nous aurions pu nous y attendre (*Hypothèse 4*).
5. Conformément à l’*hypothèse 5*, les différences de durées observées en valeurs absolues dans l’*hypothèse 1* sont maintenues en valeurs relatives. Rappelons que le VOT est plus long en post-opératoire 2 pour les occlusives voisées des patients. Il semble donc que l’allongement du VOT pour les occlusives voisées dans les productions des patients en post-opératoire 2 soit une manœuvre robuste. Cette différence de VOT ne proviendrait pas alors d’une variation de la vitesse d’élocution.
6. Nous avons vu que, quel que soit le paramètre étudié, la variabilité inter et/ou intra-locuteurs est toujours plus marquée pour tous les paramètres chez les patients (*Hypothèse 6*). Cela est vrai même pour les paramètres qui ne sont pas modifiés par l’immobilité laryngée, c’est-à-dire que même si les productions des patients sont apparemment proches de celles des contrôles, leurs productions restent tout de même marquées par une variabilité plus importante. La phase post-opératoire 2, notamment, indique des écarts-types souvent plus importants que ceux mesurés chez les locuteurs contrôles, et que ceux observés dans les phases d’enregistrement post-opératoires suivantes. La variabilité inter et intra individuelle se réduit dès la phase post-opératoire 3. Cette observation, qui se confirme en post-opératoire 4, témoigne d’une régularité retrouvée dans les

productions des patients, et pourrait être un signe de récupération d'un timing oro-laryngé « standard ».

7. Enfin, le temps et la rééducation vocale ont un impact positif sur les paramètres modifiés par l'immobilité laryngée. Les paramètres temporels perturbés, lors de l'enregistrement post-opératoire 2, commencent à ressembler à ceux des locuteurs contrôle, dès la phase d'enregistrement post-opératoire 3 (c'est-à-dire à environ un mois après l'intervention). Ce phénomène se confirme lors de la phase d'enregistrement post-opératoire 4.

CONCLUSION CHAPITRE VIII

Dans ce chapitre, nous avons étudié les modifications spatio-temporelles de la voix de sept patients présentant une immobilité laryngée, suite à une opération de la glande thyroïde. Notre étude se veut longitudinale, puisque les patients ont été enregistrés à plusieurs reprises (en post-opératoire 2, 3 et 4) lors de leur rééducation orthophonique. Afin d'obtenir des valeurs de « référence », les résultats des patients ont été comparés à ceux obtenus à partir de productions de locuteurs contrôles, appariés en sexe et en âge avec la population de patients.

Le corpus était composé de douze logatomes (ou séquence VCV) qui ont permis de quantifier l'étendue de différents intervalles vocalique et consonantique. Une attention particulière a été accordée aux paramètres temporels intrasegmentaux de l'occlusive comme le VOT, le VTT, l'occlusion consonantique et la tenue consonantique.

Il était supposé que l'immobilité laryngée, modifiant le timing oro-laryngé, aurait un impact sur les différents intervalles observés. Le VOT, considéré comme un indice de vibration des plis vocaux et pouvant permettre ainsi la distinction entre consonnes voisées et consonnes non voisées, serait allongé pour les occlusives voisées des patients. Le VTT, lié à l'arrêt des vibrations périodiques, donc par extrapolation à l'amortissement des vibrations des plis vocaux (Sock, 1998), devrait être plus court pour les patients par rapport aux locuteurs contrôles. Compte tenu des difficultés de voisement rencontrées par nos patients, nous pouvions supposer que la tenue consonantique des occlusives voisées pourrait se voir allongée, ce qui impliquerait également que la tenue consonantique des occlusives non voisées le soit également, afin de maintenir le contraste de sonorité nécessaire à la distinction des sons. Par conséquent, la tenue consonantique pourrait également être plus longue dans les productions des patients. Enfin, nous émettions également l'hypothèse que le temps et la rééducation vocale devraient avoir un effet positif sur les paramètres étudiés, ce qui se reflèterait par une amélioration de la qualité vocale des patients. Leurs mesures seraient alors comparables à celles obtenues auprès des locuteurs contrôles.

Les analyses de variance (ANOVA) à trois facteurs indiquent que l'effet principal pour le facteur *phase d'enregistrement* a été significatif ($p < 0,05$) pour la variable VOT.

Seul l'intervalle du VOT pour les occlusives sonores a été modifié. Pour les occlusives voisées, la transition CV implique uniquement un changement d'un état obstrué de conduit vocal à un état suffisamment dégagé de celui-ci pour permettre l'apparition d'une structure formantique claire. En revanche, pour les occlusives non voisées, cette transition CV exige, outre le changement d'état du conduit vocal, une modification de la configuration de la glotte d'une position ouverte pour la consonne non voisée à une position fermée pour la voyelle. En conséquence, en raison de difficultés de maintien de vibrations laryngées, le VOT, qui est un indice de voisement, est alors généralement plus long pour les occlusives voisées que celles produites par les locuteurs contrôles. La durée du VOT de Klatt (1975) des occlusives voisées des patients reste toutefois quantitativement distincte des occlusives non voisées. En conséquence, la durée de la tenue consonantique des patients se voit également allongée dans les phases d'enregistrement précoces ($p < 0,05$).

L'étude des autres paramètres, durées du VTT et des voyelles n'a pas permis de mettre au jour des différences significatives ($p = ns$) entre les productions des locuteurs contrôles et celles des patients.

Notons que c'est dans le contexte [a-i] que se manifestent les plus grandes différences entre les productions des patients et celles des locuteurs contrôles. Il est donc probable que la transition [C => i] exige un contrôle plus précis et plus problématique que la transition [C => a]. En effet, la voyelle [i] requérant une constriction réduite dans la zone palatale, sans qu'il y ait de contact entre le dos de la langue et la voûte palatine, les patients semblent alors avoir plus de difficultés à réaliser cette transition de la consonne vers la voyelle cible [i]. Cela ne semble pas être le cas pour le [a], dont l'abaissement mandibulaire, accompagné de celui de la masse de la langue avec qui la mandibule est fortement couplée, crée sans difficulté notoire la constriction radico-pharyngale nécessaire pour l'émergence acoustique de cette voyelle.

Nos données indiquent que la variabilité, qu'elle soit intra ou interindividuelle, se réduit à partir de la phase post-opératoire 3, révélant ainsi une certaine stabilisation du timing des paramètres temporels retenus.

Avec le temps et la rééducation, l'on observe non seulement une réduction de la variabilité signalée ci-dessus, mais également des valeurs temporelles qui ressemblent de plus en plus à celles des locuteurs contrôles.

SYNTHESE GENERALE

Introduction

L'*objectif général* de ce travail, mené en partenariat avec le Département de Chirurgie Oncologique du Centre Paul Strauss à Strasbourg, était d'évaluer les conséquences d'une chirurgie thyroïdienne sur la voix des patients, afin de déceler les différentes perturbations qu'entraîne cette opération chirurgicale, et de mettre au jour les possibles stratégies de compensation ou de réajustements que le patient peut mettre en place, seul ou à l'aide d'une rééducation orthophonique ; il s'agissait d'une étude longitudinale.

Le *but spécifique* de notre étude était d'analyser les caractéristiques spectrales et temporelles de la voix de patients souffrant de paralysies récurrentielles, mais également la voix de patients pour qui le diagnostic ORL n'a pas détecté de paralysie mais dont la voix est altérée après une opération de la glande thyroïde.

L'évaluation reposait sur l'analyse de signaux acoustiques, à partir desquels une quantité importante d'indices relatifs à la qualité vocale, mais aussi au comportement articulatoire des locuteurs, a pu être extraite. La production de voyelles soutenues a permis de conduire deux études spectrales. L'étude spatio-temporelle a été rendue possible grâce à l'analyse de logatomes, et plus précisément de séquences VCV.

Dans cette discussion, nous reviendrons sur les diverses hypothèses que nous avons formulées. Elle permettra également de rappeler les résultats de nos différentes expériences, afin de faire le point sur les enseignements apportés à nos connaissances sur la production de la parole, grâce à l'étude acoustique des productions de nos patients ayant subi une thyroïdectomie.

Etude de la qualité vocale de patients ayant subi une thyroïdectomie et ne présentant pas d'immobilité laryngée (Expérience 1)

Notre première expérience visait à étudier les perturbations qu'une opération de la glande thyroïde pouvait provoquer sur la voix des patients. En effet, des troubles vocaux peuvent apparaître après ce type d'opération, alors même que les nerfs laryngés sont préservés (Lombardi et *al.* 2006). Les troubles vocaux sont alors généralement transitoires et de courtes durées (Debruyne, et *al.* 1994).

Pour ce faire, nous avons mené une étude acoustique des productions de sept patients ayant été opérés de la glande thyroïde. Ils ont été enregistrés à trois reprises : avant l'opération (phase préopératoire), le lendemain de l'intervention (phase post-opératoire 1) et quinze jours après l'intervention (phase post-opératoire 2). Notre corpus, composé de trois voyelles soutenues [i a u], nous a permis de conduire une étude spectrale sur la voix de ces sept patients. Les paramètres spectraux retenus étaient les suivants : la fréquence fondamentale (F0), le Harmonics-to-Noise Ratio (HNR), le jitter, le shimmer, les formants (F1 et F2) et l'aire de l'espace vocalique, calculé à partir de la formule de Héron.

Nos *hypotheses* étaient les suivantes :

Si après l'opération chirurgicale, l'activité du larynx est irrégulière, alors les valeurs de la fréquence fondamentale et du Harmonics-to-Noise Ratio devraient être modifiées. Ces modifications devraient également se traduire par des modifications de la périodicité de la voix. Les valeurs du jitter et du shimmer pourraient alors se trouver augmentées. Compte tenu des modifications des valeurs de la fréquence fondamentale, il est probable que les valeurs de F1 et de F2 soient également différentes de celles observées en phase préopératoire. La taille et la forme de l'espace vocalique pourraient alors également se voir modifiées. Des écarts-types plus importants devraient être observés dans les phases d'enregistrement post-opératoires et notamment lors de l'enregistrement en post-opératoire 1. Ils seraient alors le reflet d'une variabilité inter et/ou intra individuelle, liée à une coordination oro-laryngée problématique. La récupération de la coordination entre gestes glottiques et supraglottiques devrait aboutir à une réduction de cette variabilité inter et/ou intra individuelle. Le temps devrait

permettre une récupération vocale et une « standardisation » des paramètres précédemment évoqués : les valeurs des paramètres mesurés seraient alors comparables à celles obtenues en phase préopératoire.

Si toutes nos hypothèses ont été confirmées en termes de tendances, seules les variables fréquence fondamentale et aire de l'espace vocalique ont été significativement modifiées par le facteur temps ou phase d'enregistrement ($p < 0,05$). Plus précisément, toutes les mesures sont modifiées en post-opératoire 1. La fréquence fondamentale est significativement modifiée ($p < 0,05$) après l'opération de la glande thyroïde. Ces résultats ne corroborent pas les conclusions des études mentionnées en IV.6.1. En effet, Hong et Kim (1997) ou Stojadinovic et *al.* (2002) indiquaient que la thyroïdectomie ne modifiait pas significativement les mesures de la fréquence fondamentale sur leur cohorte de locuteurs. Nous avons enregistré les patients au lendemain de l'intervention, alors que les résultats des études précédemment citées reposent sur des données de patients enregistrés plus tardivement (un mois après l'intervention pour celle de Hong et Kim, 1997 ; une semaine après l'intervention pour celle de Stojadinovic, *et al.*, 2002). Il est alors probable que la fréquence fondamentale de leurs patients se soit déjà améliorée. Leurs résultats sont alors à rapprocher de ceux que l'on a pu obtenir en post-opératoire 2 (soit deux semaines après l'intervention), où l'on note que la fréquence fondamentale de nos locuteurs est à nouveau proche de celle obtenue en phase préopératoire.

Les mesures de perturbations du signal, tels que le Harmonics-to-Noise Ratio, le jitter et le shimmer, se sont également révélés modifiées en post-opératoire 1. Les mesures du Harmonics-to-Noise Ratio diminuent alors que les mesures du jitter et du shimmer augmentent, ce qui témoigne d'une certaine apériodicité dans la voix des patients. Ces résultats sont à rapprocher de ceux de Debruyne et *al.* (1997).

En ce qui concerne les mesures formantiques, elles ne sont pas significativement modifiées par l'ablation de la glande thyroïde ; en revanche les légères modifications impactent l'aire de l'espace vocalique, qui est alors significativement réduite en post-opératoire 1 ($p < 0,05$).

La phase d'enregistrement post-opératoire 1 est également caractérisée par une variabilité plus importante, qui se traduit par des écarts-types inter et intra individuels plus importants. Cette observation fait consensus dans la littérature. Même lorsque les

mesures ne sont pas significativement modifiées par l'opération chirurgicale, les phases d'enregistrement post-opératoires précoces sont caractérisées par des écarts-types plus importants (voir Stojadinovic, et *al.*, 2002 ou de Netto, et *al.*, 2006 par exemple).

Enfin, le temps permet une normalisation de tous les paramètres précédemment mentionnés. Quinze jours après l'intervention, les caractéristiques spectrales des voyelles soutenues par des patients opérés de la glande thyroïde sont à nouveau comparables à celles observées en phase préopératoire.

Notre étude corrobore donc en partie les résultats attestés dans la littérature. En tout état de cause, une thyroïdectomie peut donc avoir des conséquences sur les caractéristiques acoustiques de la voix, alors même que les nerfs laryngés sont préservés. L'intubation oro-trachéale est souvent invoquée pour expliquer ces dysphonies transitoires, puisqu'il est admis que l'intubation peut provoquer des troubles vocaux (Domino et *al.* 1999). La dysphonie est alors décrite comme un enrouement passager, alors même qu'aucune lésion n'est visible sur les plis vocaux (Yamanaka et *al.* 2009). De plus, ce type de dysphonie régresse rapidement et de façon spontanée (Jones et *al.* 1992). Ces symptômes sont alors comparables à ceux d'une laryngite bénigne : le timbre est rauque et l'intensité vocale est diminuée (Le Huche & Allali, 2010). Ces observations ont été vérifiées par notre étude puisque la modification de la qualité vocale de nos patients ne perdure pas, et la récupération vocale s'effectue de façon spontanée dès quinze jours après l'intervention.

Etude de la qualité vocale de patients ayant subi une thyroïdectomie et présentant une immobilité laryngée (Expérience 2)

Notre deuxième étude avait pour objectif d'étudier la qualité vocale de patients présentant une immobilité laryngée post-thyroïdectomie. Une atteinte des nerfs laryngés ne provoque pas nécessairement de dysphonie (Meek et *al.*, 2008), et lorsque la dysphonie apparaît, les paramètres de la voix peuvent être de légèrement à sévèrement modifiés (Stewart & Allen, 2006). Il s'agissait alors, dans notre étude, d'observer les modifications spectrales et de tenter de les quantifier.

Afin d'évaluer les conséquences de l'immobilité laryngée sur la voix des patients, nous avons conduit une étude acoustique à partir d'enregistrements de sept locuteurs, pour lesquels le diagnostic ORL post-opératoire avait révélé une immobilité

laryngée unilatérale. Notre étude se veut longitudinale, et c'est pour cette raison que les patients ont été enregistrés à trois reprises tout au long de leur rééducation orthophonique : en post-opératoire 2 (quinze jours après l'intervention), en post-opératoire 3 (à environ un mois après l'intervention) et en post-opératoire 4 (à environ deux mois après l'intervention). Afin d'obtenir des mesures de « références », sept locuteurs contrôles, appariés en genre et en âge avec les patients, ont également été enregistrés. Le corpus et les paramètres retenus sont les mêmes que ceux utilisés dans l'expérience 1. Le corpus était composé des trois voyelles soutenues [i a u], à partir desquelles ont été effectuées des mesures de la fréquence fondamentale, du Harmonics-to-Noise Ratio, du jitter, du shimmer, des mesures formantiques (F1 et F2) et l'aire de l'espace vocalique, calculé à partir de la formule de Héron.

Nos hypothèses sont les mêmes que celles formulées pour l'expérience 1, à savoir que nous pensions que tous les paramètres susmentionnés seraient modifiés par l'immobilité laryngée. La fréquence fondamentale serait diminuée, les mesures de perturbations du signal comme le Harmonics-to-Noise Ratio, le jitter et le shimmer seraient modifiés, en raison de l'activité irrégulière de la source glottique. Les valeurs formantiques (soit F1 et F2), également liées à l'activité laryngienne étant donné le couplage entre le larynx et le conduit vocal, seraient modifiées, ce qui entraînerait une diminution de l'aire de l'espace vocalique. Les modifications les plus importantes seraient observables lors de la phase d'enregistrement post-opératoire 2, le temps et la rééducation orthophonique devant permettre une amélioration de la qualité vocale, dès la phase post-opératoire 3. Ce phénomène se poursuivrait jusqu'en post-opératoire 4.

De façon générale, nos hypothèses ont été confirmées.

En post-opératoire 2, la qualité vocale des patients est altérée par rapport à celle des locuteurs contrôles. Ce sont notamment les mesures de perturbations du signal qui le démontrent, puisque ce sont elles qui sont significativement modifiées ($p < 0,05$). Les mesures du Harmonics-to-Noise Ratio sont plus basses, tandis que les valeurs du jitter et du shimmer sont plus importantes. Ces résultats peuvent être rapprochés de l'étude de Hartl et *al.* (2001) qui indiquaient des tendances identiques. En ce qui concerne la fréquence fondamentale des patients en post-opératoire 2, elle est également modifiée par rapport à celle des locuteurs contrôles. Deux tendances se dégagent : 1/ une principale tendance est à la diminution des valeurs comme attendu ; elle concerne six de

nos sept patients. 2/Une locutrice en revanche présente la particularité d'avoir une fréquence fondamentale moyenne plus élevée que celle de sa locutrice contrôle associée. Ce schéma unique dans nos données serait alors le signe d'une atteinte du nerf laryngé supérieur (Remacle, 1999). En ce qui concerne les valeurs formantiques, seules les valeurs de F2 sont significativement modifiées, ce qui affecte l'aire de l'espace vocalique, qui est alors réduit en post-opératoire 2, de manière tendancielle.

La phase d'enregistrement post-opératoire 2 est également marquée par des écarts-types plus importants, reflets d'une variabilité inter et intra individuelle importante. Ces phénomènes sont également présents dans d'autres études comme celles de Hartl et *al.* (2001), Zang et *al.* (2005).

A partir de la phase d'enregistrement post-opératoire 3, c'est-à-dire un mois après l'intervention, et alors que la rééducation orthophonique a commencé de façon intensive, l'altération vocale diminue. Toutes les mesures commencent à ressembler à celles observées pour les locuteurs contrôles. Cette tendance se poursuit jusqu'en post-opératoire 4. Il semble donc que le temps et la rééducation aient effectivement un effet positif sur la qualité vocale de nos patients. A partir de cette dernière phase d'enregistrement, la variabilité peut également être considérée comme négligeable.

De façon générale, la littérature (voir Wagner et Seiler, 1994 par ex.) indique que 60% des patients présentant une immobilité laryngée post-thyroïdectomie récupèrent une mobilité laryngée et donc une qualité vocale satisfaisante. En ce qui concerne notre étude, la qualité vocale de six de nos locuteurs s'améliore significativement après deux mois de rééducation intensive. Seule la locutrice qui présentait une augmentation de sa fréquence fondamentale ne suit pas ce schéma : pour elle, la récupération est beaucoup plus lente et, six mois après l'opération, la rééducation n'étaient pas terminée. Dans la mesure où nos sept patients dysphoniques ne présentaient pas de symptômes associés plus importants, la rééducation orthophonique et le temps ont suffi à améliorer la qualité vocale, comme recommandé dans la littérature (Laccourreye et *al.*, 2003).

Enfin, signalons que la voix de nos patients, certes dégradée en post-opératoire 2, nous a tout de même permis de prendre des mesures. En effet, comme le signale l'étude de Zhang et *al.* (2005), dans certains cas, la voix est rendue tellement aperiodique par l'immobilité laryngée qu'il est alors difficile d'y appliquer les

paramètres d'analyse traditionnelle. Il faut alors recourir à différents types de mesures : acoustiques, aérodynamiques et subjectives afin de pouvoir quantifier l'étendue de la dégradation vocale et sa possible amélioration. Cela n'était pas le cas avec nos données ; la voix de nos patients était certes dégradée mais restait sujette aux mesures traditionnelles de perturbation du signal.

Etude des intervalles temporels de la voix de patients présentant une immobilité laryngée post-thyroïdectomie (Expérience 3).

Notre dernière expérience proposait d'analyser les caractéristiques spatio-temporelles de la voix de patients présentant une immobilité laryngée, suite à une opération de la glande thyroïde. Il s'agit ici d'évaluer les conséquences de l'immobilité laryngée sur le *timing* des articulateurs. Rappelons que les mesures de durées du VTT (Agnello, 1975) et du VOT (Klatt, 1975) permettent de révéler certaines perturbations au niveau de la coordination entre gestes glottiques et supraglottiques.

Pour ce faire, nous avons enregistré sept locuteurs (les mêmes que ceux retenus dans l'expérience 2) présentant une immobilité laryngée post-thyroïdectomie. Le corpus était constitué de logatomes et plus précisément de séquences VCV, où C était l'une des six occlusives [p t k b d g] et V1 [a] et V2 [i], et inversement. Les patients ont été enregistrés à trois reprises (en phases post-opératoires 2, 3 et 4). Afin d'obtenir des valeurs de « référence », sept locuteurs contrôles ont également été enregistrés.

Nous supposons que l'immobilité laryngée aurait un impact sur tous les paramètres suprasegmentaux lors des phases d'enregistrement post-opératoires précoces. Le VOT, la durée de l'occlusion et la tenue consonantique seraient alors plus longs pour les consonnes voisées. En ce qui concerne les consonnes non voisées, le VTT devrait être raccourci. Cette phase d'enregistrement devrait également être caractérisée par des écarts-types plus importants. Les phases d'enregistrement suivantes devraient afficher des valeurs qui se rapprochent des valeurs des locuteurs contrôles, ainsi qu'une variabilité réduite.

L'étude des occlusives non voisées n'a pas permis de dégager des différences entre les productions des locuteurs contrôles et celles des patients, que l'on considère la phase d'enregistrement post-opératoire 2 ou les suivantes. Les difficultés de maintien de voisement n'entraînent pas une réduction du VTT comme nous l'attendions. En

revanche, cet intervalle apparaît dans certaines occlusives voisées ce qui laisse supposer une certaine désonorisation de celles-ci. Ce phénomène reste à quantifier. Les autres intervalles (VOT, occlusion et tenue consonantique) ne sont pas perturbés par l'immobilité laryngée. Les productions des patients restent toujours proches de celles des locuteurs contrôles. Une variabilité plus importante, caractérisée par des écarts-types plus grands, est toutefois observable pour la production des occlusions non voisées des patients, notamment lors de la phase d'enregistrement post-opératoire 2. La variabilité commence à se réduire en post-opératoire 3, jusqu'en post-opératoire 4.

C'est l'étude des occlusives voisées qui a permis de mettre au jour les plus grandes différences entre les productions des locuteurs contrôles et les productions des patients. La durée du VOT, que ce soit en valeurs absolues ou en valeurs relatives, est plus longue pour les productions des patients en post-opératoire 2 que celles relevées lors de toutes les autres phases d'enregistrement. Cela est probablement dû aux difficultés de maintien du voisement, puisque les vibrations périodiques s'estompent plus tôt pour, en quelque sorte, « laisser la place » au relâchement consonantique. La durée de la tenue consonantique des occlusives voisées est alors également plus longue pour les patients en post-opératoire 2 que pour les locuteurs contrôles. Bien que ces intervalles soient plus longs pour les occlusives voisées produites par les patients, ils restent toujours plus courts que ceux mesurés pour les occlusives non voisées. Il est alors probable que, malgré ces allongements, la distinction perceptive entre les consonnes voisées et non voisées soit toujours possible. Il semble alors que les patients adoptent des stratégies compensatoires afin de ne pas « mettre en péril » la distinction catégorielle des consonnes. Cela fait écho à la Théorie de la Viabilité puisqu'on peut alors considérer que les stratégies des patients restent un ensemble de trajectoires viables qui permettent une émergence cohérente d'un système (Aubin & Cellina, 1984), afin d'atteindre, en l'occurrence ici, la « cible » acoustique recherchée (Sock & Vaxelaire, 2001). Si les sons désirés restent reconnaissables au niveau perceptif, c'est alors que les réajustements temporels et spatiaux demeurent pertinents. Il semble alors que des ajustements sur différents paramètres soient possibles pour atteindre une « cible » acoustique désirée, qui sera alors perçue correctement.

Comme lors de nos expériences précédentes, la variabilité inter et intra individuelle la plus importante est relevée lors de la première phase d'enregistrement (ici post-opératoire 2). Cette variabilité se réduit en même temps que les valeurs se rapprochent des valeurs des

locuteurs contrôles. Il est possible de la considérer comme acceptable dès la phase d'enregistrement post-opératoire 4.

De même que la qualité vocale s'améliorait (voir expérience 2) avec le temps et la rééducation orthophonique, les paramètres suprasegmentaux aussi se rapprochent de ceux mesurés à partir des productions des locuteurs contrôles, et ce dès la phase d'enregistrement post-opératoire 3. Ce phénomène se poursuit jusqu'en post-opératoire 4, où les productions des patients deviennent similaires à celles des locuteurs contrôles.

CONCLUSIONS ET PERSPECTIVES

Ce travail, de manière générale, se veut une contribution à l'étude de la parole pathologique et à l'étude de la dysphonie en général. Plus particulièrement, notre recherche tente d'apporter quelques lumières sur les investigations préoccupées par les conséquences acoustiques d'une opération de la glande thyroïde. Les études des conséquences post-opératoires sur la voix des patients ayant été opérés de la glande thyroïde sont peu nombreuses, surtout lorsque les troubles sont transitoires et ne nécessitent pas de réintervention chirurgicale.

Il nous semble important de rappeler ici la difficulté de travailler sur des productions de patients.

La première difficulté réside dans le recrutement des patients ; il s'agit de prendre en considération l'état psychologique du patient qui est souvent inquiet quant aux conséquences post-opératoires et ne mesure pas bien souvent l'importance de sa voix avant qu'un trouble n'apparaisse.

La deuxième difficulté est due au protocole expérimental lui-même : en situation hospitalière, il n'est pas toujours évident de demander aux patients de répéter plusieurs fois le contenu d'un corpus sans le fatiguer outre mesure.

Enfin la troisième difficulté est liée au sujet d'étude lui-même, puisque l'étude de la parole pathologique peut rendre l'analyse et l'interprétation des résultats difficiles. En effet, la variabilité inter et intra individuelle est omniprésente, ce qui rend difficile l'énoncé de généralités et impose en revanche une rigueur d'analyses précises, tout en relativisant la robustesse des résultats statistiques.

Cette investigation a également permis d'observer les stratégies de compensation que les locuteurs sont capables de mettre en place, seul ou à l'aide d'une rééducation orthophonique. Il s'agit notamment d'une prise de conscience proprioceptive des gestes oro-laryngés à déployer pour atteindre une qualité vocale acceptable, cela grâce aussi à une auto-correction par feedback auditif.

Enfin, signalons encore le caractère idiosyncrasique de la production de la parole. Chaque locuteur peut utiliser une multitude de trajectoires articulatoires et de solutions acoustiques pour atteindre une même cible perceptive. C'est l'organisation même de son système qui définit en partie ses propres limites de viabilité. L'existence d'un invariant quelconque, autre que la cible perceptive elle-même, semble alors difficile à déceler. En effet, les locuteurs peuvent faire varier un nombre important de paramètres pour atteindre une même cible acoustico-phonétique. Cela est particulièrement remarquable chez les patients, qui sont capable de palier leur handicap vocal au niveau de la source laryngée en faisant preuve d'une grande habilité de réajustement pour atteindre le plus souvent possible leur cible et garantir ainsi l'émergence d'un message linguistique le plus intelligible possible. Ainsi, les théories orientées « sortie acoustique », qui supposent l'existence d'une cible acoustique nécessaire à l'émergence du message linguistique et qui admettent peu de variabilité sur le plan acoustique (par ex. Stevens, 1972 ou Guenther & Perkell, 1995 a et b), ne semblent pas capturer correctement nos données. Il en va de même des approches théoriques qui prônent l'existence d'invariants gestuels (cf. par ex. Kelso et *al.*, 1986). L'interprétation articulatoire que nous faisons à partir de nos données acoustiques, révèle des perturbations aussi bien des gestes glottiques que des gestes supraglottiques, avec des stratégies de réajustements variables suivant les patients. Le message linguistique est, *in fine*, correctement perçu, malgré l'absence de gestes invariants. En revanche, si l'on accepte que l'unité phonétique ne reste pas stable, puisqu'elle subit les fluctuations en fonction de la situation de communication, de perturbations internes ou externes, on pourrait expliquer une grande partie de la variabilité observée dans nos données par la mise en place, par le patient, de stratégies adaptatives (Lindblom, 1963), lors de la production de la parole. La conception d'une « cible », comme espace de contrôle pour la réalisation de possibles articulatoires et acoustiques perceptivement acceptables, nous semble particulièrement

pertinente, dans la mesure où les locuteurs pathologiques réorganisent leurs productions selon leurs propres contraintes physiologiques et anatomiques provoquées par leur pathologie. Il s'agit ainsi pour eux aussi de respecter un espace dit viable pour l'émergence des phonèmes souhaités.

QUELQUES LIMITES DE CETTE RECHERCHE

Les limites de notre travail, selon nous, sont globalement liées au système d'acquisition de données utilisé. En effet, toutes nos expériences ont été réalisées sur la base unique du signal acoustique. Il est certain qu'un complément de données aérodynamiques aurait été riche d'informations. L'observation des paramètres suivants aurait alors pu être possible : la périodicité de la vibration des plis vocaux, le quotient d'ouverture et de fermeture de la glotte, le timing entre le pic d'ouverture glottique par rapport à des événements supra-glottiques, tel que le relâchement et la closure, etc.

Une investigation en parole pathologique pose généralement le problème des locuteurs et de leur recrutement. Il s'agit de toujours garder à l'esprit que les locuteurs sont des patients dans une situation de stress (liée à leur hospitalisation et/ou à leurs troubles de la parole). Il n'est alors pas toujours aisé d'obtenir une cohorte de locuteurs importante. Pour nos expériences, le nombre de locuteurs restent relativement faible. Nous avons voulu privilégier une analyse fine et précise de nos résultats en ne retenant qu'un nombre relativement restreint de locuteurs afin que les groupes soient les plus homogènes possible, dans la mesure où les productions de différents locuteurs peuvent être considérées comme étant homogènes. Cette méthodologie présente l'inconvénient de ne pouvoir alors affirmer des conclusions fortes qui seraient le résultat de phénomènes robustes corroborés par des statistiques irréfutables. Nos résultats, liés à des données relativement réduites, mais bien contrôlées, indiquent quelques fois plutôt des tendances que des conclusions définitives. La variabilité au niveau inter et intra-individuel est omniprésente dans notre travail. Aussi serait-il intéressant de poursuivre

ce travail avec un nombre plus large de locuteurs afin de pouvoir constituer des sous-groupes et ainsi renforcer et affiner nos premiers résultats.

En raison de contraintes hospitalières, les voix de certains de nos patients (expériences 2 et 3) n'ont pas pu être enregistrées avant l'opération, aussi est-il délicat de parler de détérioration vocale sans avoir pour référence la voix même du patient en phase préopératoire. Le recours à un locuteur contrôle permet de combler ce manque mais peut constituer un léger biais méthodologique. Il serait certainement plus opportun d'obtenir la voix du patient avant l'opération afin de pouvoir évaluer de façon plus fine les modifications vocales qu'une immobilité laryngée post-thyroïdectomie peut provoquer.

Afin de poursuivre nos investigations, des tests de perception pourraient également être réalisés. L'intelligibilité des patients pourrait être évaluée à l'aide de tests d'identification. Il pourrait également être intéressant de mener une étude perceptive plus subjective de la qualité de la voix des patients, notamment à l'aide d'une parole continue. La qualité de vie du patient serait alors au centre de cette investigation. Enfin, nous souhaitons conduire des tests visant à catégoriser le sexe du locuteur, notamment pour les locutrices pour lesquelles la fréquence fondamentale est fortement diminuée. Notre hypothèse est que la dysphonie, évaluée de façon objective sur le signal sonore, perturberait non seulement l'intelligibilité mais également la qualité de vie du patient. Ces tests pourraient être menés par des auditeurs spécialistes et non spécialistes, afin de comparer les résultats d'« experts » à ceux d'auditeurs naïfs.

Signalons, cependant, une étude préliminaire (Fauth, 2012) que nous avons conduite sur l'évaluation de la qualité vocale de patients ayant subi une chirurgie thyroïdienne. Il s'agissait, en outre, de vérifier si des auditeurs étaient en mesure de distinguer une voix d'homme d'une voix de femme, parmi les locuteurs thyroïdectomisés. Il s'agissait également d'évaluer l'intelligibilité des productions de ces patients. Dans cette étude, 14 locuteurs (7 patients et 7 sujets contrôle) avaient été enregistrés en train de produire des séquences [iCa], ainsi que des voyelles soutenues. Les enregistrements des locuteurs opérés avaient eu lieu à différentes phases : à 15 jours après l'intervention chirurgicale (post-op1), à 1 mois (post-op2), à 2 mois (post-op3) et à 3 mois (post-op4). Deux tests perceptifs avaient ensuite été proposés à un jury naïf constitué de 27 auditeurs. Ces derniers avaient (1) à identifier le sexe du locuteur et à

attribuer une note de confiance allant de 1 à 5 à partir de 3 productions vocaliques de la voyelle [a], par locuteur et par phase d'enregistrement (soit 84 stimuli). Les auditeurs devaient également (2) identifier le logatome prononcé et attribuer une note de confiance allant de 1 à 5 à ce choix. Le critère distinctif retenu était celui de la sonorité consonantique. Nous avons abouti ainsi à 168 stimuli, au total. Les résultats relatifs à l'identification du sexe du locuteur ont indiqué que les locuteurs pathologiques étaient correctement catégorisés. Les notes de confiance ne se sont pas révélées significatives. En ce qui concerne l'intelligibilité des locuteurs pathologiques, un effet de phase d'enregistrement a été démontré pour l'identification des logatomes, puisque les auditeurs naïfs perçoivent mieux les logatomes dans les phases d'enregistrement tardives. Les notes de confiance ne se sont pas révélées significatives.

Il pourrait également être intéressant d'étudier un flux de parole plus important, comme une conversation ou un texte lu à différents moments de la journée. Ces données, qui se rapprochent des productions dites spontanées, pourraient permettre d'évaluer les effets de la fatigue vocale, entre autres.

Enfin, d'autres paramètres acoustiques, ainsi que des paramètres articulatoires mériteraient d'être explorés afin de compléter les résultats présentés ici, et d'apporter davantage de poids et de précisions aux interprétations proposées dans ce travail. Cette démarche devrait permettre de consolider nos résultats à l'aide de modèles géométriques du conduit vocal et de modèles biomécaniques (Laprie & Busset, 2011 ; (Perrier, et *al.* 2011) ; cela s'inscrirait alors dans le projet de notre Equipe de Recherche « Percevoir : une histoire sensori-motrice d'événements audibles et visibles », Programme de la MISHA, 2013-2017.

BIBLIOGRAPHIE

- Aasland, W. A., Baum, S. R., & McFarland, D. H. (2006). Electropalatographic, acoustic, and perceptual data on adaptation to a palatal perturbation. *The Journal of the Acoustical Society of America* (Vol. 119, p. 2372-2381).
- Abitbol, J., Abitbol, P., & Abitbol, B. (1999). Sex hormones and the female voice. *Journal of voice: official journal of the Voice Foundation*, 13(3), 424-446.
- Abry, C., Autesserre, D., Barrera, C., Benoit, C., Boe, L. J., Caelen, J., Caelen-Haumont, G., et al. (1985). Propositions pour la segmentation et l'étiquetage des sons du français. *14e JEP du GCP du GALF* (p. 156-163).
- Abry, C., Orliaguet, J.-P., & Sock, R. (1990). Patterns of speech phasing. Their robustness in the production of a timed linguistic task: single vs. double (abutted) consonants in French. *Cahiers de psychologie cognitive*, 10(3), 269-288.
- Ackermann, H., Konczak, J., & Hertrich, I. (1997). The temporal control of repetitive articulatory movements in Parkinson's disease. *Brain and language*, 56(2), 312-319.
- Ackermann, H., & Ziegler, W. (1989). Dysarthrophonia of Parkinson syndrome. *Fortschritte der Neurologie-Psychiatrie*, 57(4), 149-160.

- Ackermann, H., & Ziegler, W. (1991). Articulatory deficits in parkinsonian dysarthria: an acoustic analysis. *Journal of Neurology, Neurosurgery, and Psychiatry*, 54(12), 1093-1098.
- Agnello, J. (1975). Voice onset and voice termination features of stutterers. *Vocal tract dynamics and dysfluency: the proceedings of the first annual Hayes Martin Conference on Vocal Tract Dynamics*. Speech and Hearing Institute.
- Aronson, A., & Bless, D. (2009). *Clinical Voice Disorders, Book + DVD*. Thieme.
- Aubin, J.-P. (1991). *Viability Theory* (1st ed. 2001. 2nd printing.). Berlin: Birkhäuser.
- Auzou, P. (2001). Analyse perceptive de la dysarthrie. *Les dysarthries* (p. 95-99). Marseille: SOLAL.
- Babin, E., Lemarchand, V., Comoz, F., & Edy, E. (2005). Laryngites chroniques de l'adulte. *EMC - Oto-rhino-laryngologie*, 2(4), 420-431.
- Bailly, L. (2010). *Etude articulatoire de la parole produite en environnement bruyant* (Mémoire pour le Master Sciences et Technologies de l'UPMC Mention Sciences de l'Ingénieur Spécialité MIS, Parcours ATIAM). Université Paris 6, Paris.
- Baujat, B., Delbove, H., Wagner, I., Fugain, C., de Corbière, S., & Chabolle, F. (2001). Immobilité laryngée post-thyroïdectomie. *Annales de Chirurgie*, 126(2), 104-110.
- Baum, & McFarland. (1997). The development of speech adaptation to an artificial palate. *The Journal of the Acoustical Society of America*, 102(4), 2353-2359.
- Béchet, M. (2011). *Perturbation de la production des occlusives chez des locuteurs présentant une division palatine ou labio-palatine*. Université de Strasbourg, Strasbourg.

- Benninger, Gillen, & Altaian. (1998). Changing etiology of vocal fold immobility. *The Laryngoscope*, 108(9), 1346-1350.
- Benoit, C., Lallouache, T., Tayeb, M., & Abry, C. (1992). A set of French visemes for visual speech synthesis. *Talking Machines: Theories, Models and Designs* (Elsevier., p. 485–504). Amsterdam.
- Blumstein, S. E., & Stevens, K. N. (1979). Acoustic invariance in speech production: evidence from measurements of the spectral characteristics of stop consonants. *The Journal of the Acoustical Society of America*, 66(4), 1001-1017.
- Boek, W., Wieneke, G. H., & Dejonckere, P. H. (1997). Clinical relevance of the fractal dimension of F0perturbations computed by the box-counting method. *Journal of Voice*, 11(4), 437-442.
- Boersma, P. (1993). Accurate short-term analysis of the fundamental frequency and the harmonics-to-noise ratio of a sampled sound. *IFA Proceedings 17* (p. 97–110). University of Amsterdam.
- Boersma, P. (2001). Praat, a system for doing phonetics by computer. *Glott International*, 9/10(5), 341-345.
- Brasnu, D, Menard, M., Fabre, A., Janot, F., & Laccourreye, H. (1988). Partial supracricoid laryngectomies: techniques, indications and results. *The Journal of Otolaryngology*, 17(4), 173-178.
- Brasnu, Daniel, Ayache, D., Hans, S., Hartl, D., & Papon, J.-F. (2008). *Traité d'ORL*. Médecine Sciences Publications.
- Browman, C., & Goldstein, L. (1989). Articulatory Gestures as Phonological Units. *Haskins Laboratories Status Report on Speech Research, SR-99/100*, Articulatory Gestures as Phonological Units.

- Browman, C., & Goldstein, L. (1990a). Gestural specification using dynamically-defined articulatory structures. *Journal of Phonetics*, (18), 299-320.
- Browman, C., & Goldstein, L. (1990b). Representation and reality: Physical systems and phonological structure. *Journal of Phonetics*, (18), 411-424.
- Browman, C., & Goldstein, L. (1990c). Tiers in articulatory phonology, with some implications for casual speech. *Papers in Laboratory Phonology*, Papers in Laboratory Phonology (Vol. 1, p. 341-376). Cambridge University Press.
- Browman, C., & Goldstein, L. (1992). Articulatory Phonology: An Overview. *Haskins Laboratories Status Report on Speech Research*, SR-111/112, 23-42.
- Brunner, J., Hoole, P., Perrier, P., & Fuchs, S. (2006). Temporal development of compensation strategies for perturbed palate shape in German /S/-production. *7th International Seminar on Speech Production 2006*. Présenté à International Seminar on Speech Production, Ubatuba.
- Clairat, S. (2004). *Compensation articulatoire dans la production des occlusives du français* (Thèse de doctorat). Université Aix Marseille I - Université de Provence, Aix-en-Provence.
- Clements, G. N., & Ridouane, R. (2006). Quantal phonetics and distinctive features: a Review. *Proceedings of the ISCA Tutorial and Research Workshop on Experimental Linguistics, 28-30 August 2006* (p. 17-24). Présenté à ISCA Tutorial and Research Workshop on Experimental Linguistics.
- Crevier-Buchman, L., Tessier, C., Sauvignat, A., Brihaye-Arpin, S., & Monfrais-Pfauwadel, M.-C. (2005). Diagnostic d'une dysphonie non organique de l'adulte. *Revue de laryngologie, d'otologie et de rhinologie*, 5(126), 353-360.

- Crevier-Buchman, Lise. (1999). *La voix et la parole sans cordes vocales après laryngectomie partielle supracricoïdienne: étude de 20 cas avec cricohyoïdoépiglottopexie*. Université Paris III - Sorbonne Nouvelle, Paris.
- Crumley, R. L. (1990). Teflon versus thyroplasty versus nerve transfer: a comparison. *The Annals of otology, rhinology, and laryngology*, 99(10 Pt 1), 759-763.
- Crumley, R. L. (1994). Unilateral recurrent laryngeal nerve paralysis. *Journal of Voice: Official Journal of the Voice Foundation*, 8(1), 79-83.
- Daniloff, R. G., & Hammarberg, B. (1973). On defining coarticulation. *Journal of Phonetics*, 239-248(1).
- Darley, F. L., Aronson, A. E., & Brown, J. R. (1969a). Clusters of Deviant Speech Dimensions in the Dysarthrias. *J Speech Hear Res*, 12(3), 462-496.
- Darley, F. L., Aronson, A. E., & Brown, J. R. (1969b). Differential Diagnostic Patterns of Dysarthria. *J Speech Hear Res*, 12(2), 246-269.
- Darley, F. L., Aronson, A. E., & Brown, J. R. (1975). Hypokinetic dysarthria: disorders of extrapyramidal system. *Motor Speech Disorders* (WB Saunders., p. 171-197). Philadelphia.
- de Pedro Netto, I., Fae, A., Vartanian, J. G., Barros, A. P. B., Correia, L. M., Toledo, R. N., Testa, J. R. G., et al. (2006). Voice and vocal self-assessment after thyroidectomy. *Head & neck*, 28(12), 1106-1114.
- Debruyne, F., Delaere, P., Wouters, J., & Uwents, P. (1994). Acoustic Analysis of Tracheo-Oesophageal Versus Oesophageal Speech. *The Journal of Laryngology & Otology*, 108(04), 325-328.

- Debruyne, F., Ostyn, F., Delaere, P., & Wellens, W. (1997). Acoustic analysis of the speaking voice after thyroidectomy. *Journal of voice: official journal of the Voice Foundation*, 11(4), 479-482.
- Dejonckere, P. H., Bradley, P., Clemente, P., Cornut, G., Crevier-Buchman, L., Friedrich, G., Van De Heyning, P., et al. (2001). A basic protocol for functional assessment of voice pathology, especially for investigating the efficacy of (phonosurgical) treatments and evaluating new assessment techniques. *European Archives of Oto-Rhino-Laryngology*, 258(2), 77-82.
- Dejonckere, P. H., Crevier-Buchman, L., Marie, J. P., Moerman, M., Remacle, M., & Woisard, V. (2003). Implementation of the European Laryngological Society (ELS) basic protocol for assessing voice treatment effect. *Revue de laryngologie - otologie - rhinologie*, 124(5), 279-283.
- Docherty, G. J. (1992). *The Timing of Voicing in British English Obstruents*. Walter de Gruyter.
- Domino, K. B., Posner, K. L., Caplan, R. A., & Cheney, F. W. (1999). Airway injury during anesthesia: a closed claims analysis. *Anesthesiology*, 91(6), 1703-1711.
- Dralle, H., Kruse, E., Hamelmann, W. H., Grond, S., Neumann, H. J., Sekulla, C., Richter, C., et al. (2004). Not all vocal cord failure following thyroid surgery is recurrent paresis due to damage during operation. *Der Chirurg; Zeitschrift für alle Gebiete der operativen Medizen*, 75(8), 810-822.
- Dubreuil, C., & Céruse, P. (2004). *Guide pratique d'ORL*. Masson.
- Duffy, J. R. (2005). *Motor speech disorders: substrates, differential diagnosis, and management*. Elsevier Mosby.

- Ekberg, O., & Nylander, G. (2004). Anatomy and Physiology. In O. Ekberg (Éd.), *Radiology of the Pharynx and the Esophagus* (p. 1-14). Berlin, Heidelberg: Springer Berlin Heidelberg.
- Estienne, F., Piérart, B., & Collectif. (2006). *Les bilans de langage et de voix : Fondements théoriques et pratiques*. Masson.
- Farrag, T. Y., Samlan, R. A., Lin, F. R., & Tufano, R. P. (2009). The Utility of Evaluating True Vocal Fold Motion Before Thyroid Surgery. *The Laryngoscope*, 116(2), 235-238.
- Fauth, C. (2012). Paralysies récurrentielles et perturbation de l'intelligibilité de la parole et de la classification homme/femme. *Colloque International Jeunes Chercheurs (CIJC)* (p. Sous Presse). Présenté à Les Classifications en Linguistique : Problèmes, Méthodologie, Enjeux, Strasbourg.
- Fauth, C., Vaxelaire, B., Rodier, J. F., Volkmar, P. P., Bouarourou, F., Hirsch, F., & Sock, R. (2011). A spatiotemporal prospective study of speech in patients with or without recurrent laryngeal nerve paralysis after thyroid surgery (p. 57-64). Présenté à ISSP'2011 : La production de la parole : du cerveau au comportement, Montréal.
- Fougeron, C. (2005). Phonologie articulatoire. *Phonologie et phonétique* (p. 267-290). Paris: Hermes Science Publications.
- Fougeron, C., Crevier-Buchman, L., Fredouille, C., Ghio, A., Meunier, C., Chevrier-Muller, C., Audibert, N., et al. (2010). Developing an acoustic-phonetic characterization of dysarthric speech in French. In E. L. R. A. (ELRA) (Éd.), *Proceedings of the 7th International Conference on Language Resources, Technologies and Evaluation (LREC)* (Vol. 1, p. 2831-2838). Valletta, Malte.

- Fowler, C. A. (2007). Speech production. *The Oxford Handbook of Psycholinguistics* (In Gaskell, M.G., p. 489-501). Oxford: Oxford University Press.
- Fowler, C. A., & Saltzman, E. (1993). Coordination and coarticulation in speech production. *Language and speech*, 36 (Pt 2-3), 171-195.
- Fowler, C. A., & Turvey, M. (1980). Immediate compensation in bite block speech. *Phonetica*, (37), 306-326.
- Friedrich, T., Hänsch, U., Eichfeld, U., Steinert, M., Staemmler, A., & Schönfelder, M. (2000). Recurrent laryngeal nerve paralysis as intubation injury? *Der Chirurg; Zeitschrift Für Alle Gebiete Der Operativen Medizen*, 71(5), 539-544.
- Gaitenby, J. (1965). The elastic word. *Speech Research*, 2, 1-12.
- Gay, T., Lindblom, B., & Lubker, J. (1981). Production of bite block vowels : acoustic equivalence by selective compensation. *Journal of the Acoustical Society of America*, (69), 802-810.
- Gentil, M., Pollak, P., & Perret, J. (1995). La dysarthrie parkinsonienne. *Revue neurologique*, 151(2), 105-112.
- Giovanni, A., Guelfucci, B., Yu, P., Robert, D., & Zanaret, M. (2002). Acoustic and Aerodynamic Measurements of Voice Production after Near-Total Laryngectomy with Epiglottoplasty. *Folia Phoniatrica et Logopaedica*, 54, 304-311.
- Giovanni, A., Ouaknine, M., & Triglia, J.-M. (1999). Determination of largest lyapunov exponents of vocal signal: Application to unilateral laryngeal paralysis. *Journal of Voice*, 13(3), 341-354.

- Grosjean, & Deschamps. (1975). Analyse contrastive des variables temporelles de l'anglais et du français : vitesse de parole et variables composantes, phénomènes d'hésitation. *Phonética*, (31), 144-184.
- Guenther, F H, Hampson, M., & Johnson, D. (1998). A theoretical investigation of reference frames for the planning of speech movements. *Psychological review*, 105(4), 611-633.
- Guenther, Frank H. (1995b). A modeling framework for speech motor development and kinematic articulator control. *Proceedings of the XIIIth International Congress of Phonetic Sciences*. Présenté à Symposium on Physiological Models for Speech Production, Stockolm.
- Guenther, Frank H. (1994). A neural network model of speech acquisition and motor equivalent speech production. *Biological Cybernetics*, 72, 43–53.
- Guenther, Frank H. (1995a). Speech sound acquisition, coarticulation, and rate effects in a neural network model of speech production. *Psychological Review*, 102, 594–621.
- Guenther, Frank H., & Barreca, D. M. (1997). Neural models for flexible control of redundant systems. *Self-organization, Computational Maps, and Motor Control* (Vol. Volume 119, p. 383-421). North-Holland.
- Guenther, Frank H., & Perkell, J. S. (2004). A Neural Model of Speech Production and Its Application to Studies of the Role of Auditory Feedback in Speech (Oxford University Press., p. 29-49). Oxford: B. Maassen, R. Kent, H. Peters, P. Van Lieshout, and W. Hulstijn.
- Hammarberg, B., Fritzell, B., Gauffin, J., & Sundberg, J. (1986). Acoustic and perceptual analysis of vocal dysfunction. *Journal of Phonetics*, (14), 533-547.

- Hardcastle, W. (2006). Coarticulation. In Editor-in-Chief: Keith Brown (Éd.), *Encyclopedia of Language & Linguistics (Second Edition)* (p. 501-505). Oxford: Elsevier.
- Hardcastle, W. J. (1985). Some phonetic and syntactic constraints on lingual coarticulation during /kl/ sequences. *Speech Communication*, 4(1-3), 247-263.
- Hardcastle, William J., & Hewlett, N. (1999). *Coarticulation: Theory, Data and Techniques*. Cambridge University Press.
- Harmegnies, B., & Poch-Olivé, D. (1992). A study of style-induced vowel variability: Laboratory versus spontaneous speech in Spanish. *Speech Communication*, 11(4-5), 429-437.
- Harmegnies, B., & Poch-Olivé, D. (1994). Formants frequencies variability in French vowels under the effect of various speaking styles. *Le Journal de Physique IV*, 04(C5), C5-509-C5-512.
- Hartl, D. (2004). *Caractérisation de la voix et de la parole dans les paralysies récurrentielles unilatérales*. Université de Paris III-Sorbonne Nouvelle, Paris.
- Hartl, D., & Brasnu, D. (2000). Les paralysies récurrentielles : connaissances actuelles et traitements. *Annales Françaises d'Oto-Rhino-Laryngologie et de Pathologies Cervico-Faciales*, 117(2), 60-84.
- Hartl, D., Hans, S., Vaissière, J., Riquet, M., & Brasnu, D. (2001). Objective Voice Quality Analysis Before and After Onset of Unilateral Vocal Fold Paralysis. *Journal of Voice*, 15(3), 351-361.
- Heaver, L. (1959). Spastic dysphonia : Psychiatric considerations. *Logos*, (2), 15-24.

- Higashikawa, M., Nakai, K., Sakakura, A., & Takahashi, H. (1996). Perceived pitch of whispered vowels-relationship with formant frequencies: A preliminary study. *Journal of Voice*, 10(2), 155-158.
- Hillenbrand, J., Cleveland, R. A., & Erickson, R. L. (1994). Acoustic Correlates of Breathy Vocal Quality. *J Speech Hear Res*, 37(4), 769-778.
- Hirano, M. (1981). Clinical examination of voice (p. 81-84). New: Springer Verlag.
- Hockauf, H., & Sailer, R. (1982). Postoperative recurrent nerve palsy. *Head & Neck Surgery*, 4(5), 380-384.
- Hoffman, H. T., & McCulloch, T. M. (1996). Anatomic considerations in the surgical treatment of unilateral laryngeal paralysis. *Head & Neck*, 18(2), 174-187.
- Hong, K. H., & Kim, Y. K. (1997). Phonatory characteristics of patients undergoing thyroidectomy without laryngeal nerve injury. *Otolaryngology - Head and Neck Surgery*, 117(4), 399-404.
- Hugh-Munier, C., Scherer, K. R., Lehmann, W., & Scherer, U. (1997). Coping strategies, personality, and voice quality in patients with vocal fold nodules and polyps. *Journal of Voice*, 11(4), 452-461.
- Inabnet, W. B., Murry, T., Dhiman, S., Aviv, J., & Lifante, J.-C. (2009). Neuromonitoring of the external branch of the superior laryngeal nerve during minimally invasive thyroid surgery under local anesthesia: a prospective study of 10 patients. *The Laryngoscope*, 119(3), 597-601.
- Inagi, K., Khidr, A. A., Ford, C. N., Bless, D. M., & Heisey, D. M. (1997). Correlation Between Vocal Functions and Glottal Measurements in Patients With Unilateral Vocal Fold Paralysis. *The Laryngoscope*, 107(6), 782-791.

- Jamski, J., Jamska, A., Graca, M., Barczyński, M., & Włodyka, J. (2004). Recurrent laryngeal nerve injury following thyroid surgery. *Przegląd lekarski*, 61(1), 13-16.
- Jatzko, L., & Wette. (1994). Recurrent nerve palsy after thyroid operations - principal nerve identification and a literature review. *Surgery*, 115(2), 139-144.
- Jégoux, F. (2002). *La Papillomatose laryngée*. Université de Nantes, Nantes.
- Jones, M. W., Catling, S., Evans, E., Green, D. H., & Green, J. R. (1992). Hoarseness after tracheal intubation. *Anaesthesia*, 47(3), 213-216.
- Junqua, J.-C. (1992). The Variability of Speech Produced in Noise (p. 43-52). Présenté à SPAC, Cannes-Mandelieu, France.
- Kain, A. B., Hosom, J.-P., Niu, X., Van Santen, J. P. H., Fried-Oken, M., & Staehely, J. (2007). Improving the intelligibility of dysarthric speech. *Speech communication*, 49(9), 743-759.
- Keating, P. A. (1984). Phonetic and Phonological Representation of Stop Consonant Voicing. *Language*, 60(2), 286-319.
- Keilmann, A., & Hülse, M. (1992). Dysphonie nach Strumektomie bei ungestörter respiratorischer Beweglichkeit der Stimmlippen. *Folia Phoniatica et Logopaedica*, 44(6), 261-268.
- Kelchner, L. N., Stemple, J. C., Gerdeman, B., Le Borgne, W., & Adam, S. (1999). Etiology, pathophysiology, treatment choices, and voice results for unilateral adductor vocal fold paralysis: A 3-year retrospective. *Journal of Voice*, 13(4), 592-601.

- Kelso, J. A. ., & Tuller, B. (1983). « Compensatory Articulation » Under Conditions of Reduced Afferent Information: A Dynamic Formulation. *J Speech Hear Res*, 26(2), 217-224.
- Kelso, J. A. ., Tuller, B., Vatikiotis-Bateson, E., & Fowler, C. A. (1984). Fonctionnaly-specific cooperation following jaw perturbation during speech : evidence for coordinative structures. *Journal of Experimental Psychology : Human Perception and Performance*, (10), 812-832.
- Kelso, J. A., Saltzman, E. L., & Tuller, B. (1986). The dynamical perspective on speech production: Data and theory. *Journal of Phonetics*, 14(1), 29-59.
- Kieburzt, K., McDermott, M., Como, P., Growdon, J., Brady, J., Carter, J., Huber, S., et al. (1994). The effect of deprenyl and tocopherol on cognitive performance in early untreated Parkinson's disease. *Neurology*, 44(9), 1756-1756.
- Klatt, D. (1975). Voice Onset Time, Frication, and Aspiration in Word-Initial Consonant Clusters. *J Speech Hear Res*, 18(4), 686-706.
- Klostermann, F., Ehlen, F., Vesper, J., Nubel, K., Gross, M., Marzinzik, F., Curio, G., et al. (2008). Effects of subthalamic deep brain stimulation on dysarthrophonia in Parkinson's disease. *Journal of neurology, neurosurgery, and psychiatry*, 79(5), 522-529.
- Koch, B., Boettcher, M., Huschitt, N., & Hülsewede, R. (1996). Must the recurrent nerve in thyroid gland resection always be exposed? A prospective randomized study. *Der Chirurg; Zeitschrift Für Alle Gebiete Der Operativen Medizen*, 67(9), 927-932; discussion 932.

- Koufman, J. A., Walker, F. O., & Joharji, G. M. (1995). The cricothyroid muscle does not influence vocal fold position in laryngeal paralysis. *The Laryngoscope*, 105(4), 368-372.
- Laboissière, R., Lametti, D., & Ostry, D. (2009). Impedance control and its relation to precision in orofacial movement. *Journal of Neurophysiology*, (102), 523-531.
- Lacau Saint Guily, J., Boisson-Bertrand, D., & Monnier, P. (2003). Lésions liées à l'intubation oro- et nasotrachéale et aux techniques alternatives : lèvres, cavités buccale et nasales, pharynx, larynx, trachée, oesophage. *Annales Françaises d'Anesthésie et de Réanimation*, 22, Supplement 1, 81-96.
- Laccourreye, H. (1993). Partial surgery of epitheliomas at the level of the glottis. An update. *Bulletin De l'Académie Nationale De Médecine*, 177(8), 1333-1340; discussion 1340-1341.
- Laccourreye, O., Weinstein, G., Brasnu, D., Trotoux, J., & Laccourreye, H. (1991). Vertical partial laryngectomy: a critical analysis of local recurrence. *The Annals of otology, rhinology, and laryngology*, 100(1), 68-71.
- Laccourreye, O., Le Clech, G., & Santini, J. (2008). Complications de la chirurgie des affections et tumeurs bénignes de la thyroïde. *Les risques chirurgicaux en oto-rhino-laryngologie : information, prise en charge et prévention* (Société française d'oto-rhino-laryngologie., p. 369-386). Paris: Laccourreye O, Chabolle F.
- Laccourreye, O., Malinvaud, D., Ménard, M., & Bonfils, P. (2009). Immobilités laryngées unilatérales après chirurgie de la glande thyroïde. *Journal de Chirurgie*, 146(6), 553-558.

- Laccourreye, Ollivier, Papon, J., Kania, R., Ménard, M., Brasnu, D., & Hans, S. (2003). Paralysies laryngées unilatérales : données épidémiologiques et évolution thérapeutique. *Presses Medicales*, 32, 781-786.
- Ladefoged. (1983). The linguistic use of different phonation types. *Vocal fold physiology: contemporary research and clinical issues* (p. 351-360). College-Hill Press.
- Ladefoged, P. (1960). The regulation of subglottal pressure. *Folia Phoniatica*, (12), 169-175.
- Laprie, Y., & Busset, J. (2011). Construction and evaluation of an articulatory model of the vocal tract. Présenté à 19th European Signal Processing Conference - EUSIPCO-2011.
- Laukkanen, A., Ilomaki, I., Leppanen, K., & Vilkmán, E. (2008). Acoustic Measures and Self-reports of Vocal Fatigue by Female Teachers. *Journal of Voice*, 22, 283-289. doi:10.1016/j.jvoice.2006.10.001
- Le Huche, F., & Allali, A. (2010a). *La voix : Volume 2, Pathologies vocales d'origine fonctionnelle* (3e éd.). Elsevier Masson.
- Le Huche, F., & Allali, A. (2010b). *La voix : Volume 3, Pathologies vocales d'origine organique* (2e éd.). Elsevier Masson.
- Lecours, A. R., & Lhermitte, F. (1979). *L'aphasie*. Paris: Flammarion Médecine.
- Lee, B. S. (1950). Effects of Delayed Speech Feedback. *The Journal of the Acoustical Society of America*, 22(6), 824.
- Lehiste, I. (1965). Some acoustic characteristics of dysarthric speech. *Bibliotheca Phonetica*, (2).

- Liberman, A M, Delattre, P. C., & Cooper, F. S. (1958). Some cues for the distinction between voiced and voiceless stops in initial position. *Language and Speech*, 1, 153_167.
- Liberman, A M, & Mattingly, I. G. (1985). The motor theory of speech perception revised. *Cognition*, 21(1), 1-36.
- Liberman, Alvin M., Delattre, P. C., Cooper, F. S., & Gerstman, L. J. (1954). The role of consonant-vowel transitions in the perception of the stop and nasal consonants. *Psychological Monographs: General and Applied*, 68(8), 1-13.
- Lieberman, P., Kako, E., Friedman, J., Tajchman, G., Feldman, L. S., & Jiminez, E. B. (1992). Speech production, syntax comprehension, and cognitive deficits in Parkinson's disease. *Brain and Language*, 43(2), 169-189. doi:10.1016/0093-934X(92)90127-Z
- Lindblom, B. (1983). Economy of speech gestures. *The Production of Speech* (p. 217-246). Springer-Verlag.
- Lindblom, B. (1987). Absolute Constancy and Adaptative Variability: two Themes in the Quest for Phonetics Invariance. *Proceedings of the 11th International Congress of Phonetic Sciences*. Tallinn, Estonia.
- Lindblom, B. (1990). Explaining phonetic variation: a sketch of the H&H theory. *Speech production and speech modelling* (p. 403-439). Dordrecht, Netherlands: Kluwer Academic Publishers,: W.J.Hardcastle & A. Marchal.
- Lindblom, B. (1991). The status of phonetic gesture. *Modularity and the Motor Theory of Speech Perception* (Mattingly I. & Studdert-Kennedy M., p. 7-24).
- Lindblom, B. (1996). Role of articulation in speech perception: clues from production. *The Journal of the Acoustical Society of America*, 99(3), 1683-1692.

- Lindblom, B., Lubker, J., & Gay, T. (1979). Formant frequencies of some fixed mandible vowels and a model of speech motor programming by predictive simulation. *Journal of Phonetics*, (7), 147-161.
- Lindblom, B., Lubker, J., & Mc Allister, R. (1977). Compensatory articulation and the modelling of normal speech production behaviour. *Articulatory Modelling and Phonetics* (p. 147-161). Présenté à Symposium at Grenoble, Groupe Aixois Linguistique Française, Grenoble: Carré R Descout R & Wajskop M.
- Lindblom, B., & Sundberg, J. (1971). Acoustical consequences of lip, tongue, jaw and larynx movement. *Journal of the Acoustical Society of America*, (50), 1166-1179.
- Lindeberg, H. (1991). Laryngeal papillomas: histomorphometric evaluation of multiple and solitary lesions. *Clinical Otolaryngology and Allied Sciences*, 16(3), 257-260.
- Lisker, L., & Ambramson, A. (1964). A cross-language study of voicing in initial stops: Acoustical measurements. *Word*, (20), 384-422.
- Liu, Q., Djuricin, G., & Prinz, R. A. (1998). Total thyroidectomy for benign thyroid disease. *Surgery*, 123(1), 2-7.
- Lo, C. Y., Kwok, K. F., & Yuen, P. W. (2000). A prospective evaluation of recurrent laryngeal nerve paralysis during thyroidectomy. *Archives of surgery (Chicago, Ill.: 1960)*, 135(2), 204-207.
- Logemann, J. A., Fisher, H. B., Boshes, B., & Blonsky, E. R. (1978). Frequency and Cooccurrence of Vocal Tract Dysfunctions in the Speech of a Large Sample of Parkinson Patients. *Journal of Speech and Hearing Disorders*, 43(1), 47-57.

- Lombard, E. (1911). Le signe de l'élévation de la voix. *Annales des Maladies de l'oreille et du larynx*, 2(37), 101-119.
- Lombardi, C. P., Raffaelli, M., D'Alatri, L., Marchese, M. R., Rigante, M., Paludetti, G., & Bellantone, R. (2006). Voice and swallowing changes after thyroidectomy in patients without inferior laryngeal nerve injuries. *Surgery*, 140(6), 1026-1034.
- Ludlow, C. L., & Bassich, C. J. (1984). Relationship between perceptual rating and acoustic measures of hypokinetic speech. *The dysarthrias: physiology, acoustic, perception, management* (College Hill Press., p. 163-196). San Diego: McNeil, Mr, Rosenbek, J.-C., Arosen, A.-E.
- Makeieff, M., Venail, F., Garrel, R., de la Breteque, B. A., Giovanni, A., & Guerrier, B. (2004). Voice handicap evaluation after supracricoid partial laryngectomy. *Revue De Laryngologie - Otologie - Rhinologie*, 125(5), 313-317.
- Malécot, A. (1955). An experimental study of force of articulation. *Studia Linguistica*, 9(1-2), 35-44.
- Marchal, A. (1988). Coproduction: Evidence from EPG data. *Speech Communication*, 7(3), 287-295. doi:10.1016/0167-6393(88)90074-X
- Martensson, H., & Terins, J. (1985). Recurrent Laryngeal Nerve Palsy in Thyroid Gland Surgery Related to Operations and Nerves at Risk. *Arch Surg*, 120(4), 475-477.
- Mathieson, L. (2001). *Greene and Mathieson's The Voice and its Disorders*, 6th Ed. (6^e éd.). Wiley.
- McCulloch, T. M., Hoffman, H. T., Andrews, B. T., & Karnell, M. P. (2000). Arytenoid adduction combined with Gore-Tex medialization thyroplasty. *The Laryngoscope*, 110(8), 1306-1311.

- McGurk, H., & MacDonald, J. (1976). Hearing lips and seeing voices. *Nature*, 264(5588), 746-748.
- Meek, P., Carding, P. N., Howard, D. H., & Lennard, T. W. J. (2008). Voice Change Following Thyroid and Parathyroid Surgery. *Journal of Voice*, 22(6), 765-772.
- Meynadier, Y., & Gaydina, Y. (2012). Contraste de voisement en parole chuchotée. *Actes de la conférence conjointe JEP-TALN-RECITAL 2012* (Vol. 1, p. 361-368). Présenté à JEP-TALN-RECITAL 2012, Grenoble.
- Morsomme, D., Jamart, J., Wéry, C., Giovanni, A., & Remacle, M. (2001). Comparison between the GIRBAS Scale and the Acoustic and Aerodynamic Measures Provided by EVA for the Assessment of Dysphonia following Unilateral Vocal Fold Paralysis. *Folia phoniatica et logopaedica: official organ of the International Association of Logopedics and Phoniatrics (IALP)*, 53(6), 317-325.
- Netterville, J. L., Jackson, C. G., & Civantos, F. (1993). Thyroplasty in the functional rehabilitation of neurotologic skull base surgery patients. *The American journal of otology*, 14(5), 460-464.
- Nicollas, R., Giovanni, A., & Triglia, J.-M. (2008). Les dysphonies de l'enfant. *Archives de Pédiatrie*, 15, 1133-1138.
- Niedzielska, G. (2001). Acoustic analysis in the diagnosis of voice disorders in children. *International Journal of Pediatric Otorhinolaryngology*, 57(3), 189-193.
- Nohoma, O., Berthommier, F., & Schwartz, J. L. (2010). Liage et fusion audiovisuelle en perception de la parole : on peut « débrancher » l'effet McGurk par un contexte audiovisuel incohérent. Présenté à XXVIII ème Journées d'Etudes sur la Parole, Mons.

- O'Grady, W., Archibald, J., Aronoff, M., & Rees-Miller, J. (2004). *Contemporary Linguistics: An Introduction* (5th éd.). Bedford/St. Martin's.
- Ohala, J. (1983). The origin of sound patterns in vocal tract constraints. *The production of speech* (New-York: Springer-Verlag., p. 189-216). New York: P. F. MacNeilage.
- Ohman, S. E. (1966). Coarticulation in VCV utterances: spectrographic measurements. *The Journal of the Acoustical Society of America*, 39(1), 151-168.
- Ormezzano, Y. (2000). *Le guide de la voix*. Paris: O. Jacob.
- Ozsancak, C. (2007). Auto-évaluation de la dysarthrie. *Les Dysarthries + CD* (p. 107-110). Marseille: SOLAL.
- Perkell, J S, & Nelson, W. L. (1985). Variability in production of the vowels /i/ and /a/. *The Journal of the Acoustical Society of America*, 77(5), 1889-1895.
- Perkell, J., Guenther, F., Lane, H., Matthies, M., Perrier, P., Vick, J., Tricarico, R., et al. (2000). A theory of speech motor control and supporting data from speakers with normal hearing and with profound hearing loss. *Journal of Phonetics*, 28, 233-272.
- Perkell, Joseph S., Matthies, M. L., Svirsky, M. A., & Jordan, M. I. (1995). Goal-based speech motor control: A theoretical framework and some preliminary data. *Journal of Phonetics*, 23(1-2), 23-35.
- Perrier, P., Payan, Y., Buchaillard, Stéphanie, I., Nazari, M. A., & Chabanas, M. (2011). Biomechanical models to study speech. *Faits de Langues*, 37, 155-171.
- Peterson, G., & Lehiste, I. (1962). Duration of syllable Nuclei in American English. *Journal of the Acoustical Society of America*, (32), 693-703.

- Pinto, S., & Ghio, A. (2008). Troubles du contrôle moteur de la parole : contribution de l'étude des dysarthries et dysphonies à la compréhension de la parole normale. *Revue française de linguistique appliquée*, *XIII*(2).
- Port, R. F. (1980). Linguistic timing factors in combination. *Journal of the Acoustical Society of America*, (69).
- Pribuisiene, R., Uloza, V., Kupcinskas, L., & Jonaitis, L. (2006). Perceptual and Acoustic Characteristics of Voice Changes in Reflux Laryngitis Patients. *Journal of Voice*, *20*(1), 128-136.
- Pröschel, U., & Eysholdt, U. (1993). Short-term changes in the larynx and voice after intubation. *Laryngo- Rhino- Otologie*, *72*(2), 93-97.
- Ramig, L. O., Sapir, S., Fox, C., & Countryman, S. (2001). Changes in vocal loudness following intensive voice treatment (LSVT®) in individuals with Parkinson's disease: A comparison with untreated patients and normal age-matched controls. *Movement Disorders*, *16*(1), 79–83.
- Randolph, G. W., & Kamani, D. (2006). The importance of preoperative laryngoscopy in patients undergoing thyroidectomy: Voice, vocal cord function, and the preoperative detection of invasive thyroid malignancy. *Surgery*, *139*(3), 357-362.
- Recasens. (1999). Coarticulation models in recent speech production theories, in collaboration with E. Farnetani (Chapter 2), Lingual coarticulation (Chapter 4) & Acoustic analysis (Chapter 16). *Coarticulation. Theory, Data and Techniques*, (Cambridge University Press.). Cambridge: Hardcastle, W.J. & Hewlett, N.
- Recasens, D. (2002). An EMA study of VCV coarticulatory direction. *The Journal of the Acoustical Society of America*, *111*(6), 2828-2841.

- Remacle, M. (2000). Troubles de la voix. *Troubles du langage. Bases théoriques, diagnostic et rééducation, nouvelle édition*, Manuel et Trait (p. 437-467). Mardaga.
- Remacle, M., Delos, M., Lawson, G., & Jamart, J. (1999). Correcting vocal fold immobility by autologous collagen injection for voice rehabilitation : A short-term study. *The Annals of otology, rhinology & laryngology*, 108(8), 788-793.
- Remacle, M., Friedrich, G., Dikkers, F. G., & de Jong, F. (2003). Phonosurgery of the vocal folds: a classification proposal. *European archives of oto-rhino-laryngology: official journal of the European Federation of Oto-Rhino-Laryngological Societies (EUFOS): affiliated with the German Society for Oto-Rhino-Laryngology - Head and Neck Surgery*, 260(1), 1-6. doi:10.1007/s00405-002-0507-5
- Robbins, J., Fisher, H. B., Blom, E. C., & Singer, M. I. (1984). A Comparative Acoustic Study of Normal, Esophageal, and Tracheoesophageal Speech Production. *J Speech Hear Disord*, 49(2), 202-210.
- Rousselot, P. J. (1897). *Principes de phonétique expérimentale Tome 1*. H. Welter (Paris).
- Rubin, A. D., Praneetvatakul, V., Gherson, S., Moyer, C. A., & Sataloff, R. T. (2006). Laryngeal Hyperfunction During Whispering: Reality or Myth? *Journal of Voice*, 20(1), 121-127.
- Saltzman, E., & Byrd, D. (2000). Task-dynamics of gestural timing: Phase windows and multifrequency rhythms. *Human Movement Science*, 19(4), 499-526. doi:10.1016/S0167-9457(00)00030-0

- Saltzman, E. L. (1986). Task dynamic coordination of the speech articulators: A preliminary model. *Experimental Brain Research Series 15* (p. 129-144). New York: H. Heuer & C. Fromm.
- Saltzman, E. L., & Munhall, K. (1989). A dynamical approach to gestural patterning in speech production. *Haskins Laboratories Status Report on Speech Research*, (SR-99 /100), 3869.
- Savariaux, C., Perrier, P., & Orliaguet, J. P. (1995). Compensation strategies for the perturbation of rounded vowel [u] using a lip tip : a study of control space in speech production. *Journal of the Acoustical Society of America*, (98), 2428-2442.
- Schoentgen, J. (2001). Stochastic models of jitter. *The Journal of the Acoustical Society of America*, 109(4), 1631-1650.
- Schoentgen, J., & Grenez, F. (2011). Suivi de la qualité de voix d'un locuteur professionnel sous des conditions écologiques. Présenté à Quatrièmes Journées de Phonétique Clinique, Strasbourg.
- Schoentgen, Jean. (2006). Vocal Cues of Disordered Voices: An Overview. *Acta Acustica united with Acustica*, 92(5), 667-680.
- Schulman, R. (1989). Articulatory dynamics of loud and normal speech. *Journal of Acoustical Society of America*, 1(85), 295-312.
- Schwartz, S. R., Cohen, S. M., Dailey, S. H., Rosenfeld, R. M., Deutsch, E. S., Gillespie, M. B., Granieri, E., et al. (2009). Clinical practice guideline: Hoarseness (Dysphonia). *Otolaryngology - Head and Neck Surgery*, 141(3, Supplement 2), S1-S31.

- Schweizer, V., & Dörfl, J. (1997). The anatomy of the inferior laryngeal nerve. *Clinical Otolaryngology and Allied Sciences*, 22(4), 362-369.
- Shiller, D. M., Sato, M., Gracco, V. L., & Baum, S. R. (2009a). Perceptual recalibration of speech sounds following speech motor learning. *The Journal of the Acoustical Society of America*, 125(2), 1103-1113.
- Shiller, D. M., Sato, M., Gracco, V. L., & Baum, S. R. (2009b). Perceptual recalibration of speech sounds following speech motor learning. *The Journal of the Acoustical Society of America*, 125(2), 1103-1113. doi:10.1121/1.3058638
- Sock, R. (1983). *L'organisation temporelle de l'opposition de quantité vocalique en wolof de Gambie. Sa résistivité aux conditions de durée segmentales et suprasegmentales* (Thèse de Doctorat de Troisième Cycle). Institut de Phonétique de Grenoble, Université de Grenoble III, Grenoble.
- Sock, R. (1998). *Organisation temporelle en production de la parole: émergence de catégories sensori-motrices phonétiques*. Presses univ. du Septentrion, Université Stendhal, Grenoble III,.
- Sock, R., & Benoit, C. (1986). VOT et VTT en français. *15 èmes Journées d'Etude sur la Parole du Groupe Communication Parlée* (p. 307-310). Aix-en-Provence.
- Sock, R., & Vaxelaire, B. (2001). Réflexions sur le timing. *Travaux de l'Institut de Phonétique de Strasbourg*, (31), 80-126.
- Sock, R., Vaxelaire, B., Roy, J. P., & Hirsch, F. (2006). Temporal and spatial correlates of quantity contrasts in Wolof. *Towards a better understanding of speech production processes* (Psychology Press., p. 253-274). New York: Harrington, J. & Tabain, M.

- Stathopoulos, E. T., Hoit, J. D., Hixon, T. J., Watson, P. J., & Solomon, N. P. (1991). Respiratory and Laryngeal Function During Whispering. *J Speech Hear Res*, 34(4), 761-767.
- Steinmüller, T., Ulrich, F., Rayes, N., Lang, M., Seehofer, D., Tullius, S. G., Jonas, S., et al. (2001). Surgical procedures and risk factors in therapy of benign multinodular goiter. A statistical comparison of the incidence of complications. *Der Chirurg; Zeitschrift für alle Gebiete der operativen Medizen*, 72(12), 1453-1457.
- Stetson, R. H. (1951). *Motor phonetics: a study of speech movements in action*. North-Holland Pub. Co.
- Stevens, K. (1972). The quantal nature of speech: Evidence from articulatory-acoustic data. *Human communication: A unified view* (p. 51-66). McGraw-Hill.
- Stewart, C., & Allen, E. (2006). Vocal therapy for unilateral fold paralysis. *Vocal fold paralysis* (p. 87-96). Springer-Verlag Berlin and Heidelberg GmbH & Co. K.
- Stojadinovic, A., Shaha, A. R., Orlikoff, R. F., Nissan, A., Kornak, M.-F., Singh, B., Boyle, J. O., et al. (2002). Prospective functional voice assessment in patients undergoing thyroid surgery. *Annals of surgery*, 236(6), 823-832.
- Stuart, A., Kalinowski, J., Rastatter, M. P., & Lynch, K. (2002). Effect of delayed auditory feedback on normal speakers at two speech rates. *The Journal of the Acoustical Society of America*, 111(5), 2237.
- Tartter, V. C. (1991). Identifiability of vowels and speakers from whispered syllables. *Perception & Psychophysics*, 49(4), 365-372.

- Teston, B. (2001). L'évaluation objective des dysfonctionnements de la voix et de la parole; 2e partie: les dysphonies. *Travaux Interdisciplinaires du Laboratoire Parole et Langage d'Aix-en-Provence (TIPA)*, 20, 169-232.
- Thibeault, M., Ménard, L., Baum, S. R., Richard, G., & McFarland, D. H. (2011). Articulatory and acoustic adaptation to palatal perturbation. *The Journal of the Acoustical Society of America*, 129(4), 2112-2120.
- Thomusch, O., Machens, A., Sekulla, C., Ukkat, J., Brauckhoff, M., & Dralle, H. (2003). The impact of surgical technique on postoperative hypoparathyroidism in bilateral thyroid surgery: A multivariate analysis of 5846 consecutive patients. *Surgery*, 133(2), 180-185.
- Timon, C. I., Hirani, S. P., Epstein, R., & Rafferty, M. A. (2010). Investigation of the impact of thyroid surgery on vocal tract steadiness. *Journal of voice: official journal of the Voice Foundation*, 24(5), 610-613.
- Titche, L. L. (1976). Causes of Recurrent Laryngeal Nerve Paralysis. *Arch Otolaryngol*, 102(5), 259-261.
- Trésallet, C., Chigot, J.-P., & Menegaux, F. (2006). How to prevent recurrent nerve palsy during thyroid surgery? *Annales de chirurgie*, 131(2), 149-153.
- Trinidad Ruiz, G., González Palomino, A., Pantoja Hernández, C., Mora Santos, E., Cruz de la Piedad, E., & Blasco Huelva, A. (2007). Influence of non-neuronal factors on post-thyroidectomy dysphonia. *Acta otorrinolaringológica española*, 58(8), 352-357.
- Tsunoda, K., Niimi, S., & Hirose, H. (1994). The Roles of the Posterior Cricoarytenoid and Thyropharyngeus Muscles in Whispered Speech. *Folia Phoniatrica et Logopaedica*, 46(3), 139-151.

- Tsunoda, K., Ohta, Y., Soda, Y., Niimi, S., & Hirose, H. (1997). Laryngeal adjustment in whispering: Magnetic resonance imaging study. *Annals of Otology, Rhinology & Laryngology*, 106(1), 41-43.
- Tucker, H. M. (1980). Vocal cord paralysis: etiology and management. *The Laryngoscope*, 90(4), 585-590.
- Vaissière, J. (2006). *La phonétique*. Presses Universitaires de France - PUF.
- Van den Berg, J. (1988). *The perception of voicing in Dutch two-obstruent sequences*. Catholic University Nijmegen, Nijmegen.
- Van Lierde, K., D'Haeseleer, E., Wuyts, F. L., Baudonck, N., Bernaert, L., & Vermeersch, H. (2010). Impact of thyroidectomy without laryngeal nerve injury on vocal quality characteristics: an objective multiparameter approach. *The Laryngoscope*, 120(2), 338-345.
- Van Lieshout, P. (2011). Keeping it stable: Motor adaptation to changes in lower lip characteristics. *9th International Seminar on Speech Production 2011*. Présenté à International Seminar on Speech Production '11, Montréal.
- Vaxelaire, B. (1993). *Etude comparée des effets des variations de débits - lent, rapide - sur les paramètres articulatoires, à partir de la cinéradiographie (sujets français)*. Université des Sciences Humaines, Strasbourg.
- Vaxelaire, B. (1995). Single vs. double (Abutted) consonants across speech rate. X-Ray and acoustic data for French. *XIIIe Congrès International des Sciences Phonétiques (ICPhS)* (Vol. 1, p. 384-387). Présenté à ICPhS, Stockholm.
- Vaxelaire, B. (2007). *Le geste et la production de la parole. Résultats et implications d'études cinéradiographiques quantitatives*. (Habilitation à Diriger des Recherches). Université Marc Bloch Strasbourg 2, Strasbourg.

- Vaxelaire, B., & Sock, R. (1996). A cineradiographic and acoustic study of velar gestures in French consonant sequences as a function of speech rate. *Proceedings of the 4th Speech Production Seminar* (p. 65-68). Autrans.
- Vercherand, G. (2010). *Production et perception de la parole chuchotée en français : analyse segmentale et prosodique*. Université de Paris 7, Paris.
- Verduyckt, I., Rungassamy, C., Remacle, M., & Dubuisson, T. (2011). Suivi embarqué en temps réel de l'inconfort vocal chez des professionnels de la voix. Présenté à 67ème Congrès de la Société Française de Phoniatrie, Paris.
- Verhulst, J., Traissac, L., Guerrier, B., Bours, A. F., Hamoir, M., Remacle, M., & Van De Heyning. (1998). Les laryngites chroniques. *Revue de laryngologie, d'otologie et de rhinologie*, 119(1), 65-72.
- Viallet, F., & Gentil, M. (2001). Les troubles de production de parole au cours de la maladie de Parkinson : la dysarthrie hypokinétique. *Les dysarthries* (p. 153-160). Elsevier Masson.
- Wagner, H. E., & Seiler, C. (1994). Recurrent laryngeal nerve palsy after thyroid gland surgery. *British Journal of Surgery*, 81(2), 226-228.
- Weismer. (1984). Acoustic description of dysarthric speech: perceptual correlates and physiological inferences. *Seminars in speech and language*. New York: Rosnebeck, JC.
- Wittekindt, C., Streubel, K., Arnold, G., Stennert, E., & Guntinas-Lichius, O. (2007). Recurrent pleomorphic adenoma of the parotid gland: analysis of 108 consecutive patients. *Head & neck*, 29(9), 822-828.
- Wood, S. (1994, 2012). swphonetics. Consulté de swphonetics.com

- Yamanaka, H., Hayashi, Y., Watanabe, Y., Uematu, H., & Mashimo, T. (2009). Prolonged hoarseness and arytenoid cartilage dislocation after tracheal intubation. *British Journal of Anaesthesia*, 103(3), 452 -455.
- Yates, A. (1963). Delayed auditory feedback. *Psychological Bulletin*, 60(3), 213-232.
- Yin, S. S., Qiu, W. W., Stucker, F. J., Hoasjoe, D. K., & Aarstad, R. F. (1997). Evaluation of bilateral vocal fold dysfunction: Paralysis versus fixation, superior versus recurrent, and distal versus proximal to the laryngeal nerves. *American Journal of Otolaryngology*, 18(1), 9-18.
- Yumoto, E., Gould, W. J., & Baer, T. (1982). Harmonics-to-noise ratio as an index of the degree of hoarseness. *The Journal of the Acoustical Society of America*, 71(6), 1544-1549.
- Zellner, B. (1998). *Caractérisation et prédiction du débit de parole en français*. Faculté des Lettres, Lausanne.
- Zhang, Y., Jiang, J. J., Biazzo, L., & Jorgensen, M. (2005). Perturbation and Nonlinear Dynamic Analyses of Voices from Patients with Unilateral Laryngeal Paralysis. *Journal of Voice*, 19(4), 519-528.

TABLES ET INDEXES

Index des auteurs

- Aasland, 56
Abitbol, 76
Abry, 125, 126, 189, 190, 208
Ackermann, 66, 71, 72
Agnello, 126, 189, 193, 195, 216, 217, 229, 240, 315
Aronson, 75
Aubin, 18, 39, 241
Auzou, 67
Babin, 78, 80
Bailly, 60
Baujat, 93
Baum, 54, 56
Béchet, 41
Benninger, 91, 94, 159
Benoit, 60, 199, 212
Blumstein, 29
Boek, 124
Boersma, 120, 123, 124
Brasnu, 80, 81, 93, 95
Browman, 30, 31, 33
Brunner, 56
Clairet, 54
Clements, 30
Crevier-Buchman, 9, 74, 82
Crumley, 96, 97
Daniloff, 38
Darley, 67, 69, 71
de Pedro Netto, 101, 102
Debruyne, 84, 99, 100, 102, 235, 236
Dejonckere, 122, 139, 167
Docherty, 126
Domino, 84, 237
Dralle, 93
Dubreuil, 73
Duffy, 68, 69
Ekberg, 90
Estienne, 76
Farrag, 93
Fauth, 116, 117, 118, 133, 160, 247
Fougeron, 33, 66, 70
Fowler, 24, 37, 38, 39, 53
Friedrich, 93, 159
Gaitenby, 47, 128, 214
Gay, 54
Gentil, 70, 71
Giovanni, 81, 106
Grosjean, 47
Guenther, 24, 31, 32, 244
Hammarberg, 38, 105
Hardcastle, 38
Harmegnies, 9, 49
Hartl, 91, 95, 103, 104, 105, 106, 190, 238, 239

- Heaver, 75
- Higashikawa, 51, 52
- Hillenbrand, 104
- Hirano, 78, 79
- Hoffman, 96
- Hong, 98, 99, 100, 101, 102, 236, 303
- Hugh-Munier, 75
- Inabnet, 101
- Inagi, 103
- Jamski, 93
- Jatzko, 94
- Jégoux, 79
- Jones, 84, 237
- Junqua, 60
- Kain, 70
- Keating, 126
- Keilmann, 99
- Kelchner, 96
- Kelso, 32, 34, 53, 57, 244
- Kieburtz, 71
- Klatt, 126, 189, 192, 194, 199, 200, 203, 204, 207, 209, 210, 211, 212, 215, 228, 232, 240
- Klostermann, 66
- Koch, 94
- Koufman, 92
- Laboissière, 58
- Lacau Saint Guily, 84
- Laccourreye, 80, 81, 92, 94, 96, 103, 190, 239
- Ladefoged, 30, 127
- Laprie, 9, 248
- Laukkanen, 75
- Le Huche, 69, 73, 76, 77, 81, 85, 86, 89, 90, 96, 107, 237
- Lecours, 69
- Lee, 59
- Lehiste, 48, 71, 212
- Liberman, 33, 126, 192
- Lindblom, 24, 26, 28, 29, 33, 35, 36, 37, 50, 53, 54, 244
- Lindeberg, 79
- Lisker, 126
- Liu, 94
- Lo, 94
- Logemann, 71
- Lombard, 59, 60
- Lombardi, 97, 235
- Ludlow, 71
- Makeieff, 83
- Malécot, 218
- Marchal, 38
- Martensson, 95
- Mathieson, 169
- McCulloch, 96, 97
- McGurk, 62, 63
- Meek, 103, 237
- Meynadier, 195, 217
- Morsomme, 106
- Netterville, 97
- Nicollas, 79
- Niedzielska, 79
- Nohoma, 62
- Ohala, 127

- Ohman, 33
Ozsancak, 68
Perkell, 24, 26, 31, 32
Peterson, 48, 212
Pinto, 66, 67, 73, 86, 90
Port, 48
Pribuisiene, 78
Pröschel, 84, 85
Ramig, 71
Randolph, 93
Recasens, 38
Robbins, 84
Rousselot, 45
Rubin, 51
Saltzman, 24, 34, 37, 38, 39
Savariaux, 55
Schoentgen, 74, 124
Schulman, 60
Schwartz, 84
Schweizer, 91
Shiller, 61
Sock, 9, 18, 27, 39, 40, 47, 48, 113, 125, 126, 127, 128, 194, 195, 199, 208, 212, 214, 216, 217, 229, 231, 241
Stathopoulos, 51
Steinmüller, 95, 159
Stetson, 26, 32
Stevens, 24, 29, 30, 212, 244
Stewart, 103, 237
Stojadinovic, 98, 101, 236, 237
Stuart, 61
Tartter, 52
Teston, 86, 124
Thibeault, 56
Thomusch, 95
Timon, 102
Titche, 92
Trinidad Ruiz, 103
Tsunoda, 50
Tucker, 81, 92
Vaissière, 167
Van den Berg,, 126
Van Lierde, 102
Van Lieshout, 57
Vaxelaire, 9, 18, 39, 40, 47, 48, 49, 127, 241
Vercherand, 218
Verduyckt, 75
Verhulst, 78
Viallet, 70
Weismer, 71
Wittekindt, 94
Wood, 123
Yamanaka, 84, 237
Yates, 59, 61
Yin, 89
Yumoto, 122
Zellner, 49
Zhang, 105, 106, 239

Index des figures

Figure 1 : Vues de profil montrant la production des séquences /aka/ et /akka/ (à gauche) et des séquences /ada/ et /adda/ (à droite), en vitesse d'élocution rapide. On observe des étendues de contact plus larges pour les consonnes doubles (pointillés) par rapport aux consonnes simples (lignes continues). Locuteur A.E. (Vaxelaire, 1995)	49
Figure 2 : La glotte en position de phonation à gauche (non voisée et voisée), de chuchotement à droite (O'Grady, et al. 2004)	51
Figure 3 : Terminologie des troubles de la parole, d'après Pinto et Ghio, (2008)	67
Figure 4 : Glande thyroïde, vue antérieure d'après Netter (2011)	91
Figure 5 : Script de Gendrot utilisé pour nos mesures	120
Figure 6 : Valeurs moyennes de la fréquence fondamentale (F0) en Hz pour [i a u] - locutrices du groupe NPP	137
Figure 7 : Valeurs moyennes de la fréquence fondamentale (F0) en Hz pour [i a u] - locuteurs du groupe NPP	139
Figure 8 : Valeurs moyennes de fréquence fondamentale (F0) en fonction du locuteur pour la voyelle [a]	141
Figure 9 : Valeurs moyennes de HNR (dB) pour [i a u] – Groupe NPP	143
Figure 10 : Valeurs moyennes de HNR (dB) en fonction du locuteur pour la voyelle [a]	144
Figure 11 : Espace vocalique – Groupe NPP	153
Figure 12 : Valeurs moyennes de la fréquence fondamentale (F0) en Hz pour [i a u] - locutrices du groupe UPP	165
Figure 13 : Valeurs moyennes de la fréquence fondamentale (F0) en Hz pour [i a u] - locuteurs du groupe UPP	166
Figure 14 : Valeurs moyennes de la fréquence fondamentale (F0) en fonction du locuteur pour la voyelle [a]	168
Figure 15 : Valeurs moyennes de HNR (dB) pour [i a u] – Groupe UPP	171

Figure 16 : Valeurs moyennes de HNR (dB) en fonction du locuteur pour la voyelle [a]	173
Figure 17 : Espace vocalique – Groupe UPP.....	183
Figure 18 : Mesures temporelles pour une séquence VCV dans laquelle C est une consonne non voisée	194
Figure 19 : Mesures temporelles pour une séquence VCV dans laquelle C est une consonne voisée.....	194
Figure 20 : Valeurs absolues de la durée du VOT (Klatt, 1975) en contexte [b] pour les logatomes [abi] (à gauche) et [iba] (à droite)	200
Figure 21 : Valeurs absolues de la durée du VOT (Klatt, 1975) en contexte [b] pour les logatomes [abi] (à gauche) et [iba](à droite) en fonction du patient	201
Figure 22 : Valeurs absolues de la durée du VOT (Klatt, 1975) en contexte [d] pour les logatomes [adi] (à gauche) et [ida] (à droite)	204
Figure 23 : Valeurs absolues de la durée du VOT (Klatt, 1975) en contexte [d] pour les logatomes [adi](à gauche) et [ida](à droite) en fonction du patient	205
Figure 24 : Valeurs absolues de la durée du VOT (Klatt, 1975) en contexte [g] pour les logatomes [agi] (à gauche) et [iga] (à droite)	208
Figure 25 : Valeurs absolues de la durée du VOT (Klatt, 1975) en contexte [g] pour les logatomes [agi] (à gauche) et [iga] (à droite) en fonction du patient	209
Figures 26 : Valeurs absolues de la durée du VOT (Klatt, 1975) en contexte [p] pour les logatomes [api] (à gauche) et [ipa] (à droite)	210
Figure 27 : Valeurs absolues de la durée du VOT (Klatt, 1975) en contexte [t] pour les logatomes [ati] (à gauche) et [ita] (à droite)	211
Figure 28 : Valeurs absolues de la durée du VOT (Klatt, 1975) en contexte [k] pour les logatomes [aki] (à gauche) et [ika] (à droite)	213
Figure 29 : Durées moyennes du VOT (Klatt, 1975) pour six occlusives du français en contexte [a-i] (à droite) et en contexte [i-a] (à gauche).....	214
Figure 30 : Durées relatives moyennes du VOT (Klatt, 1975) pour six occlusives du français en contexte [a-i] (à droite) et en contexte [i-a] (à gauche).....	217
Figure 31 : Valeurs absolues de la durée du VTT (Agnello, 1975) en contexte [t] pour les logatomes [ati] (à gauche) et [ita] (à droite)	219

Figure 32 : Durées moyennes de l'occlusion en contexte [a-i] (à gauche) et en contexte [i-a] (à droite).....	221
Figure 33 : Durées moyennes de l'intervalle allant de la fin de la structure formantique stable de V1 au début de l'explosion en contexte [a-i] (à gauche) et en contexte [i-a] (à droite) pour six occlusives du français.....	224
Figure 34 : Durées vocaliques et consonantiques pour la séquence [adi] (à gauche) et pour la séquence [ati] (à droite)	227
Figure 35 : Durées vocaliques et consonantiques pour la séquence [ida] (à gauche) et pour la séquence [ita] (à droite)	229
Figure 36 : Valeurs absolues de la durée du VTT (Agnello, 1975) en contexte [p] pour les logatomes [api] (à gauche) et [ipa] (à droite)	317
Figure 37 : Valeurs absolues de la durée du VTT (Agnello, 1975) en contexte [t] pour les logatomes [ati] (à gauche) et [ita] (à droite)	317
Figure 38 : Valeurs absolues de la durée du VTT (Agnello, 1975) en contexte [k] pour les logatomes [aki] (à gauche) et [ika] (à droite)	317
Figure 39 : durées vocaliques et consonantiques pour la séquence [abi] (à gauche) et pour la séquence [api] (à droite)	318
Figure 40 : durées vocaliques et consonantiques pour la séquence [adi] (à gauche) et pour la séquence [ati] (à droite)	318
Figure 41 : durées vocaliques et consonantiques pour la séquence [agi] (à gauche) et pour la séquence [aki] (à droite)	318
Figure 42 : durées vocaliques et consonantiques pour la séquence [iba] (à gauche) et pour la séquence [ipa] (à droite)	319
Figure 43 : durées vocaliques et consonantiques pour la séquence [ida] (à gauche) et pour la séquence [ita] (à droite)	319
Figure 44 : durées vocaliques et consonantiques pour la séquence [iga] (à gauche) et pour la séquence [ika] (à droite)	319

Index des tableaux

Tableau 1 : Données numériques pour le jitter et le shimmer d'après Hartl et al. (2001).....	104
Tableau 2 : Données numériques pour le HNR et le CPP d'après Hartl et al. (2001).....	104
Tableau 3 : Données numériques pour les différences H1-H2, H1-L1 et H1-L3 d'après Hartl et al. (2001).....	105
Tableau 4 : Locuteurs présentant une immobilité laryngée retenus pour notre étude	116
Tableau 5 : Locuteurs retenus pour notre étude et qui ne présentaient pas d'immobilité laryngée.....	117
Tableau 6 : Locuteurs contrôles retenus pour notre étude	118
Tableau 7 : Valeurs moyennes de la fréquence fondamentales (F0) en Hz pour [i a u] pour les locutrices du groupe NPP. Les écarts-types sont donnés entre parenthèses	138
Tableau 8 : Valeurs moyennes de la fréquence fondamentales (F0) en Hz pour [i a u] pour les locuteurs du groupe NPP. Les écarts-types sont donnés entre parenthèses	139
Tableau 9 : Valeurs moyennes de fréquence fondamentale (F0) en fonction du locuteur pour la voyelle [a] – Groupe NPP. Les écarts-types sont donnés entre parenthèses.....	141
Tableau 10 : Valeurs moyennes de HNR (dB) pour [i a u] – Groupe NPP. Les écarts-types sont donnés entre parenthèses	143
Tableau 11 : Valeurs moyennes de HNR (dB) en fonction du locuteur pour la voyelle [a] – Groupe NPP. Les écarts-types sont donnés entre parenthèses	145
Tableau 12 : Valeurs moyennes de jitter en % – Groupe NPP. Les écarts-types sont donnés entre parenthèses	146

Tableau 13 : Valeurs moyennes de jitter en % en fonction du locuteur – Groupe NPP. Les écarts-types sont donnés entre parenthèses	147
Tableau 14 : Valeurs moyennes de shimmer en % – Groupe NPP. Les écarts-types sont donnés entre parenthèses	147
Tableau 15 : Valeurs moyennes de shimmer en fonction du locuteur – Groupe NPP. Les écarts-types sont donnés entre parenthèses	148
Tableau 16 : Valeurs formantiques moyennes – Groupe NPP. Les écarts-types sont donnés entre parenthèses	151
Tableau 17 : Espace vocalique (kHz ²) en fonction des locuteurs – Groupe NPP	154
Tableau 18 : Valeurs moyennes de la fréquence fondamentale (F0) en Hz pour [i a u] pour les locutrices du groupe UPP. Les écarts-types sont donnés entre parenthèses	165
Tableau 19 : Valeurs moyennes de la fréquence fondamentale (F0) en Hz pour [i a u] pour les locuteurs du groupe NPP. Les écarts-types sont donnés entre parenthèses	167
Tableau 20 : Valeurs moyennes de la fréquence fondamentale (F0) en fonction du locuteur pour la voyelle [a] – Groupe NPP. Les écarts-types sont donnés entre parenthèses.....	169
Tableau 21 : Valeurs moyennes de HNR (dB) pour [i a u] – Groupe NPP. Les écarts-types sont donnés entre parenthèses	172
Tableau 22 : Valeurs moyennes de HNR (dB) en fonction du locuteur pour la voyelle [a] – Groupe NPP. Les écarts-types sont donnés entre parenthèses	174
Tableau 23 : Valeurs moyennes de jitter en % – Groupe UPP. Les écarts-types sont donnés entre parenthèses	175
Tableau 24 : Valeurs moyennes de jitter en % en fonction du locuteur – Groupe UPP. Les écarts-types sont donnés entre parenthèses	176
Tableau 25 : Valeurs moyennes de shimmer en % – Groupe UPP. Les écarts-types sont donnés entre parenthèses	178
Tableau 26 : Valeurs moyennes de shimmer en fonction du locuteur – Groupe UPP. Les écarts-types sont donnés entre parenthèses	178

Tableau 27 : Valeurs formantiques moyennes – Groupe UPP. Les écarts-types sont donnés entre parenthèses	181
Tableau 28 : Espace vocalique (kHz ²) en fonction des locuteurs – Groupe UPP	184
Tableau 29 : Récapitulatif de la fonction et de l’innervation des différents muscles laryngés intrinsèques.	306
Tableau 30 : Valeurs formantiques (F1 et F2 en Hz) pour le locuteur NPPENS	313
Tableau 31 : Valeurs formantiques (F1 et F2 en Hz) pour le locuteur NPPGER	313
Tableau 32: Valeurs formantiques (F1 et F2 en Hz) pour le locuteurNPPHOE	313
Tableau 33: Valeurs formantiques (F1 et F2 en Hz) pour le locuteurNPPHOF	313
Tableau 34 : Valeurs formantiques (F1 et F2 en Hz) pour le locuteur NPPKAU	314
Tableau 35: Valeurs formantiques (F1 et F2 en Hz) pour le locuteur NPPKRE	314
Tableau 36 : Valeurs formantiques (F1 et F2 en Hz) pour le locuteur NPPLN	314
Tableau 37: Valeurs formantiques (F1 et F2 en Hz) pour le locuteur UPPBAS	315
Tableau 38: Valeurs formantiques (F1 et F2 en Hz) pour le locuteurUPPHEI	315
Tableau 39: Valeurs formantiques (F1 et F2 en Hz) pour le locuteur UPPHUB	315
Tableau 40 : Valeurs formantiques (F1 et F2 en Hz) pour le locuteur UPPLAT	315
Tableau 41 : Valeurs formantiques (F1 et F2 en Hz) pour le locuteur UPPPAI	316
Tableau 42: Valeurs formantiques (F1 et F2 en Hz) pour le locuteur UPPWAL	316
Tableau 43: Valeurs formantiques (F1 et F2 en Hz) pour le locuteur UPPWAN	316

Table des matières

Introduction.....	17
Organisation de la thèse.....	21
Etat de la Question.....	23
I. Théories en production de la parole.....	25
I.1. Planification et Exécution - Variation et Contraintes.....	26
I.2. Théories en production de la parole	27
I.2.1. L'existence d'une cible acoustique ?	28
I.2.1.1. Lindblom : Théorie de la Variabilité Adaptative en parole.....	28
I.2.1.2. Stevens : La Théorie Quantique	29
I.2.1.3. Le modèle Directions Into Velocities of Articulators (DIVA) de la production de la parole	31
I.2.2. L'existence d'une cible articulatoire ?	32
I.2.2.1. Phonologie Articulatoire.....	32
I.2.2.2. Le modèle de la Dynamique des Tâches	34
I.2.3. Invariance et Variabilité	35
I.2.3.1. La Théorie de la Variabilité Adaptative	35
I.2.3.2. Coordination et coarticulation en production de la parole	37
I.2.3.3. La Théorie de la Viabilité.....	39
I.2.3.3.1. Définition générale	39
I.2.3.3.2. Perturbations et réajustements	40
II. Perturbation de la production et de la perception de la parole normale	45
II.1. Perturbation de la production de la parole	46
II.1.1. Perturbations de types écologiques.....	46
II.1.1.1. Variation de la vitesse d'élocution	47
II.1.1.2. Le chuchotement	50
II.1.2. Perturbations avec ajout d'un dispositif.....	53
II.1.2.1. Bite-block.....	53

II.1.2.2. Lip-tube ou tube labial	55
II.1.2.3. Palais artificiel.....	56
II.1.2.4. Perturbation dynamique	57
II.2. Perturbation de la perception de la parole.....	58
II.2.1. Le feedback auditif et son rôle.....	59
II.2.2. Effet Lombard et milieu bruité	60
II.2.3. Feedback différé ou DAF (<i>Delayed Auditory Feedback</i>)	60
II.3. Perturber la perception pour perturber la production.....	61
II.3.1. Effet McGurk	62
III. Troubles de la parole et de la voix	66
III.1. Les dysarthries – troubles du mouvement ?	66
III.1.1. Définition et limites	66
III.1.2. Classification des dysarthries	67
III.1.3. Parole Dysarthrique – Le cas du sujet parkinsonien.....	70
III.2. Dysphonies – Troubles laryngés.....	72
III.2.1. Définition et classification.....	73
III.2.2. Les dysphonies dysfonctionnelles	73
III.2.2.1. Définition de la dysphonie dysfonctionnelle	74
III.2.2.2. Etiologie et développement	74
III.2.3. Les pathologies d’origine organique	76
III.2.3.1. Les laryngites	77
III.2.4. Dysphonie et traumatismes laryngés	80
III.2.4.1. Laryngectomies.....	80
III.2.4.2. Laryngectomies partielles	81
III.2.4.3. Laryngectomies totales	83
III.2.4.4. Intubation et conséquences sur la qualité vocale	84
III.2.4.5. Dysphonie et radiothérapie	86
III.2.5. Les dysphonies d’origine neurologique.....	86
IV. Dysphonie et Immobilités laryngées	89
IV.1. Classification des paralysies laryngées	89
IV.2. Anatomie et physiologie du larynx.....	90
IV.3. Epidémiologie de la paralysie laryngée unilatérale.....	91

IV.4. Thyroïdectomie et paralysie laryngée unilatérale.....	93
IV.4.1. Pronostic de la paralysie laryngée post-thyroïdectomie.....	93
IV.5. Immobilités laryngées post-thyroïdectomie	95
IV.5.1. Données chiffrées.....	95
IV.5.2. Traitement des paralysies laryngées.....	95
IV.5.2.1. La rééducation orthophonique de la paralysie laryngée	95
IV.5.2.2. Les traitements chirurgicaux de la paralysie laryngée.....	96
IV.6. Conséquences acoustiques de la thyroïdectomie.....	97
IV.6.1. Altérations vocales sans paralysies laryngées	97
IV.6.1.1. Fréquence fondamentale.....	98
IV.6.1.2. Temps maximum de phonation	98
IV.6.1.3. Fréquence fondamentale de la voix parlée	99
IV.6.1.4. Amplitude de la fréquence fondamentale	100
IV.6.1.5. Intensité de la voix parlée	100
IV.6.1.6. Jitter et Shimmer.....	100
IV.6.1.7. Evaluation subjective.....	101
IV.6.2. Altérations vocales avec paralysies laryngées.....	103
IV.6.2.1. Mesures acoustiques	103
IV.6.2.2. Mesures aérodynamiques.....	106
IV.6.2.3. Mesures perceptives	106
Protocole expérimental	109
V. Protocole expérimental.....	111
V.1. Conditions d'enregistrement.....	111
V.2. Matériel	112
V.3. Corpus	113
V.3.1. Voyelles	113
V.3.2. Logatomes – Séquences VCV	114
V.4. Locuteurs	115
V.4.1. Locuteurs présentant une immobilité laryngée.....	116
V.4.2. Locuteurs ne présentant pas d'immobilité laryngée post-thyroïdectomie	117
V.4.3. Locuteurs Contrôles.....	117

V.5. Enregistrements	118
V.5.1. Locuteurs présentant une immobilité laryngée post-thyroïdectomie.....	119
V.5.2. Locuteurs ne présentant pas d'immobilité laryngée post-thyroïdectomie	119
V.5.3. Locuteurs Contrôles.....	119
V.6. Mesures – Analyses	120
V.6.1. Voyelles Soutenues.....	121
V.6.1.1. Fréquence Fondamentale	122
V.6.1.2. Harmonics-to-Noise Ratio (HNR)	122
V.6.1.3. Mesures formantiques - F1 et F2	123
V.6.1.4. Jitter.....	124
V.6.1.5. Shimmer	125
V.6.2. Logatomes	125
V.6.2.1. Paramètres.....	125
V.6.2.1.1. Voice Onset Time (VOT).....	126
V.6.2.1.2. Voice Termination Time (VTT).....	126
V.6.2.2. Mesures	126
V.7. Analyses Statistiques – Remarques générales	128
Résultats expérimentaux.....	130
VI. Expérience 1 : Etude acoustique de voyelles soutenues produites par des patients ne présentant pas de paralysie récurrentielle après thyroïdectomie	132
VI.1. Introduction	132
VI.2. Procédure expérimentale	133
VI.2.1. Locuteurs	133
VI.2.2. Corpus	133
VI.2.3. Mesures	133
VI.2.4. Analyses des données.....	133
VI.3. Hypothèses	134
VI.4. Résultats	135
VI.4.1. Remarques Générales	135
VI.4.1. Fréquence fondamentale (F0).....	136
VI.4.2. Harmonics-to-Noise Ratio (HNR)	142

VI.4.3. Jitter et Shimmer	146
VI.4.4. Valeurs formantiques	149
VI.4.5. Aire de l'espace vocalique	152
VI.5. Discussion.....	154
VII. Expérience 2 : Etude acoustique de voyelles soutenues produites par des patients présentant une paralysie récurrentielle après thyroïdectomie.....	160
VII.1. Introduction	160
VII.2. Procédure expérimentale	161
VII.2.1. Locuteurs.....	161
VII.2.2. Corpus	161
VII.2.3. Mesures	162
VII.2.4. Analyse des données	162
VII.3. Hypothèses	162
VII.4. Résultats	164
VII.4.1. Remarques générales	164
VII.4.2. Fréquence fondamentale (F0)	164
VII.4.2.1. Harmonics-to-Noise Ratio (HNR)	171
VII.4.3. Jitter et Shimmer	175
VII.4.4. Valeurs formantiques	180
VII.4.5. Espace Vocalique.....	182
VII.5. Discussion	185
VIII. Expérience 3 : Etude spatio-temporelle de la voix de patients présentant une paralysie récurrentielle après thyroïdectomie	191
VIII.1. Introduction.....	191
VIII.2. Procédure expérimentale.....	192
VIII.2.1. Locuteurs	192
VIII.2.2. Corpus.....	193
VIII.2.3. Mesures.....	193
VIII.2.4. Analyse des données.....	195
VIII.2.5. Hypothèses.....	195
VIII.3. Résultats.....	198
VIII.3.1. Remarques générales	198

VIII.3.1. Durées du VOT.....	198
VIII.3.1.1. Occlusives voisées	200
□ Contexte [b]	200
□ Contexte [d]	204
□ Contexte [g]	206
VIII.3.1.2. Occlusives non voisées	210
□ Contexte [p]	210
□ Contexte [t]	211
□ Contexte [k]	212
VIII.3.1.3. Le VOT comme indice de lieu d'articulation	213
VIII.3.1.1. Durées relatives du VOT	215
VIII.3.2. Durées du VTT	217
VIII.3.3. Durées de l'occlusion	219
VIII.3.4. Durées de la tenue consonantique.....	222
VIII.3.5. Durées vocaliques vs. durées consonantiques	225
VIII.4. Discussion.....	229
Synthèse Générale	236
Conclusions et perspectives	245
Quelques limites de cette recherche	248
Bibliographie	252
Tables et indexes.....	282
Index des auteurs.....	283
Index des figures	286
Index des tableaux.....	289
Table des matières.....	292
Annexes	299
I. Annexes 1 : Anatomie du larynx	300
I.1. Cartilages du larynx.....	300
I.2. Muscles du larynx	301
I.3. Le système membrano-ligamentaire	301
I.4. Les articulations	302
I.4.1. L'articulation crico-thyroïdienne	302

I.4.2. L'articulation crico-aryténoïdienne.....	302
I.5. Les muscles	303
I.5.1. Les muscles intrinsèques du larynx.....	303
I.5.1.1. Le muscle thyro-aryténoïdien ou muscle vocal (TA).....	303
I.5.1.2. Le muscle crico-thyroïdien (CT)	304
I.5.1.3. Le muscle crico-aryténoïdien postérieur (CAP).....	304
I.5.1.4. Le muscle crico-aryténoïdien latéral (CAL).....	304
I.5.1.5. Le muscle inter-aryténoïdien (IA)	305
I.5.1.6. Remarques	305
I.5.2. Les muscles extrinsèques	306
I.5.2.1. Les muscles sus-hyoïdiens	306
I.5.2.2. Les muscles sous-hyoïdiens	307
I.6. La vascularisation.....	307
I.7. L'innervation	307
I.7.1. Le nerf vague.....	307
I.7.2. Le nerf laryngé récurrent.....	308
I.7.3. Le nerf laryngé supérieur : branche externe.....	308
II. Annexes 2 : Fiches de renseignements.....	309
II.1. Fiche signalétique patient.....	309
II.2. Questionnaire Post-Opératoire adapté du Voice Quality Related Life	311
III. Annexes 3 : Résultats.....	313
III.1. Valeurs formantiques Expérience 1	313
III.2. Valeurs formantiques Expérience 2	315
III.3. Durées du VTT – Expérience 3	317
III.4. Durées vocaliques vs. durées consonantiques	318

ANNEXES

I. Annexes 1 : Anatomie du larynx

Le larynx est un assemblage de cartilages articulés reliés entre eux par des muscles et des ligaments. L'ensemble est tapissé de muqueuses. Chez l'adulte, le larynx est situé à la hauteur de la sixième vertèbre cervicale et constitue l'extrémité supérieure de la trachée artère.

I.1. Cartilages du larynx

Quatre cartilages principaux constituent le larynx.

1. Le cricoïde, situé au-dessus de l'anneau trachéal supérieur, a la forme d'une chevalière, dont la partie renforcée et tournée vers l'arrière, est appelée le chaton. Les cartilages aryénoïdes, en forme de pyramide sont posés et articulés sur le chaton coracoïdien.
2. Le thyroïde est le plus grand cartilage du larynx, il a la forme d'un bouclier. Les deux ailes latérales du thyroïde sont prolongées verticalement par des cornes supérieures et inférieures qui s'articulent de chaque côté du cricoïde. L'angle thyroïdien, plus aigu chez l'homme présente des plis vocaux plus longs et plus massifs.
3. L'épiglotte est une lame cartilagineuse fine en forme de fleur ou d'éventail. Sa pointe se situe au niveau de la commissure antérieure des plis vocaux. Il s'agit du cartilage mobile majeur mais n'est toutefois pas indispensable (en cas d'ablation de l'épiglotte, les aryénoïdes ou la base de la langue pourront s'y substituer). Son rôle est toutefois important puisqu'elle s'abaisse comme un clapet lors de la déglutition pour protéger le larynx et éviter ainsi que les aliments ou la salive n'entrent dans les voies aériennes.
4. Les cartilages aryénoïdes sont pairs et symétriques. Leur forme est triangulaire ou pyramidale avec une base articulaire. Ils sont situés au-dessus du cartilage cricoïde et sont reliés entre eux par le ligament vocal.

I.2. Muscles du larynx

Les muscles extrinsèques appartiennent, pour la plupart à des organes voisins du larynx, tout en étant partiellement insérés à celui-ci. Leur fonction est de relier le larynx au squelette, en assurant ainsi son positionnement. Par conséquent, même si ces muscles n'appartiennent pas au larynx à proprement dit, ils participent à la fonction phonatoire, puisqu'ils sont directement reliés aux cartilages. Le sternohyoïdien et le sterno-thyroïdien sont des muscles abaisseurs et sous-hyoïdiens, antagonistes des muscles élévateurs. Pendant la phonation, ils ont pour rôle de stabiliser le larynx. Le génio-hyoïdien, le hyoglosse ou encore le stylo-hyoïdien, ils sont considérés comme les muscles sus-hyoïdiens élévateurs du larynx. Ainsi, pendant la phonation, les muscles sus-hyoïdiens peuvent élever le larynx et, à l'aide des sous-hyoïdiens, ils permettent une stabilisation de sa position. Il est encore à noter que, lorsque la fréquence de vibration des cordes vocales augmente, un accroissement de l'activité des un accroissement de l'activité des muscles élévateurs est observé et le larynx s'élève.

I.3. Le système membrano-ligamentaire

Le système membrano-ligamentaire du larynx comprend un système intrinsèque et un système extrinsèque. Il s'agit de systèmes de suspensions des différentes structures laryngées.

Le système membrano-ligamentaire intrinsèque permet de relier les structures cartilagineuses entre elles. Notons qu'il porte les replis muqueux et à un rôle dans les mouvements fins des aryténoïdes. Le ligament vocal est tendu entre le processus vocal de l'aryténoïde et le cartilage thyroïde au niveau de la commissure antérieure.

Le système membrano-ligamentaire extrinsèque permet de relier les structures du larynx au reste de l'appareil respiratoire.

I.4. Les articulations

Pour permettre la phonation, les muscles doivent déplacer les structures citées précédemment. Cette partie traite donc des mouvements du larynx. Le larynx comprend deux articulations principales : l'articulation crico-thyroïdienne et l'articulation crico-aryténoïdienne. Ces deux articulations sont de type synovial donc librement mobiles.

I.4.1. L'articulation crico-thyroïdienne

L'articulation crico-thyroïdienne est constituée d'une partie fixe, la corne inférieur du thyroïde, et d'une partie mobile, l'extérieur du cricoïde. Le cricoïde peut tourner autour d'un axe de rotation horizontal et transversal. L'articulation permet ainsi la mise en tension des plis vocaux.

Un mouvement vers l'avant de l'articulation crico-thyroïdienne aura pour conséquence d'allonger et de tendre les plis vocaux tandis que le mouvement vers l'arrière permettra leur relâchement et donc leur rétrécissement.

L'articulation crico-thyroïdienne est responsable du changement de registre phonatoire.

I.4.2. L'articulation crico-aryténoïdienne

L'articulation crico-aryténoïdienne relie la base des aryténoïdes à la surface supérieure de la lame postérieure du cricoïde. Cette articulation permet la rotation des aryténoïdes vers l'intérieur ou vers l'extérieur du cricoïde. A ce mouvement de rotation s'ajoute une possibilité de déplacement de type translation.

Le mouvement de rotation de l'articulation crico-aryténoïdienne provoque l'élévation ou l'abaissement des plis vocaux. Lors de du mouvement d'élévation, les plis vocaux s'écartent latéralement, à l'inverse le mouvement d'abaissement provoque leur rapprochement.

Le mouvement de déplacement de l'articulation crico-aryténoïdienne permet une translation médiale ou latérale.

Cette articulation permet l'adduction ou l'abduction des plis vocaux.

I.5. Les muscles

Comme pour le système membrano-ligamentaire, il convient de distinguer deux types de muscles qui affectent les fonctions du larynx. Les muscles sont dits intrinsèques si tous leurs points d'attache sont à l'intérieur de la structure squelettique du larynx. A l'inverse, les muscles dits extrinsèques ont un point d'attachement sur les structures laryngées et un autre sur une structure extérieure au larynx.

I.5.1. Les muscles intrinsèques du larynx

Les muscles intrinsèques du larynx naissent et s'étendent dans les structures laryngées. Ils sont notamment à l'origine des mouvements d'adduction ou d'abduction des plis vocaux. Ils permettent également leur tension ou leur relâchement.

Lors de la phonation, les muscles intrinsèques permettent les ajustements suivants :

1. Ils peuvent faire varier la compression médiale, c'est-à-dire le degré de force avec lequel les plis vocaux se réunissent à la ligne médiale ;
2. Ils sont capables de faire varier la tension longitudinale donc le degré de force d'allongement ou d'étirement des plis vocaux.

Les muscles intrinsèques sont au nombre de cinq.

I.5.1.1. Le muscle thyro-aryténoïdien ou muscle vocal (TA)

Le muscle thyro-aryténoïdien est un muscle pair qui forme l'essentiel de la masse des plis vocaux. Il participe donc à la formation de la « valve » laryngée, notamment responsable de la protection des voies respiratoires.

Ce muscle trouve son origine antérieure sur la surface interne du thyroïde et se prolonge postérieurement pour s'insérer sur les aryténoïdes, entre l'apophyse vocale et l'apophyse musculaire.

Son action permet de pousser l'apophyse vocale vers l'intérieur et crée donc une constriction. Elle permet également l'abaissement de l'épiglotte.

Le muscle thyro-aryténoïdien est innervé par la division antérieure (ou nerf laryngé inférieur) du nerf récurrent laryngé.

I.5.1.2. Le muscle crico-thyroïdien (CT)

Le muscle crico-thyroïdien est situé sur la face antérieure du larynx. Il s'agit d'un muscle pair tendu entre la face antérieure du chaton du cricoïde et le bord inférieur du cartilage thyroïdien au niveau de sa surface quadrilatère.

Le muscle crico-thyroïdien est le muscle tenseur des plis vocaux. Il permet de faire basculer le thyroïde sur le cricoïde, son action consistant en une augmentation de la longueur et de la tension des plis vocaux.

Ce muscle peut être divisé en deux portions : la *pars* oblique (portion inférieure oblique) effectue une translation antérieure du cartilage thyroïde par rapport au cricoïde. La *pars* recta serait responsable du mouvement de rotation antérieure du thyroïde sur le cricoïde.

Lorsque ces deux parties agissent en synergie et se contractent simultanément, elles produisent un maximum d'effet sur la tension des plis vocaux (Hong, 1998).

Le muscle crico-thyroïdien est le seul muscle innervé par la branche externe du nerf laryngé supérieur.

I.5.1.3. Le muscle crico-aryténoïdien postérieur (CAP)

Ce muscle est situé entre la face postérieure du processus musculaire de l'aryténoïde et la face postérieure du chaton du cricoïde.

Le muscle crico-aryténoïdien postérieur est un muscle en forme d'éventail. Il s'agit d'un muscle pair considéré comme le seul muscle abducteur des plis vocaux. Il a pour but de tirer les apophyses musculaires vers l'arrière, ce qui provoque un écartement des plis vocaux et donc une ouverture de la glotte.

Ce muscle a une fonction respiratoire quasi permanente.

Le muscle crico-aryténoïdien postérieur est innervé par la branche postérieure du nerf récurrent laryngé.

I.5.1.4. Le muscle crico-aryténoïdien latéral (CAL)

Le muscle crico-aryténoïdien est un muscle pair latéral qui est attaché à la surface supérieure de la partie latérale de l'arc cricoïde. Il se prolonge postérieurement et vers le haut pour s'insérer sur la portion antérieure de l'apophyse musculaire des aryténoïdes.

Certaines de ses fibres sont continues avec celles du muscle thyro-aryténoïdien, qui partage avec ce dernier le fait d'être un muscle à contraction rapide (Goding, 1991).

Le muscle crico-aryténoïdien latéral est l'abducteur principal. Sa contraction tire les apophyses musculaires et fait pivoter les apophyses vocales vers la ligne médiane, ce qui provoque l'abduction des plis vocaux.

Ce muscle intervient lors du reniflement.

Il est innervé par la division antérieure du nerf récurrent laryngé.

I.5.1.5. Le muscle inter-aryténoïdien (IA)

Le muscle inter-aryténoïdien est le seul muscle laryngé intrinsèque impair. Il est situé entre les faces postérieures des corps aryténoïdiens.

Sa tâche principale est de rétrécir la glotte lors de la phonation. Pour ce faire, il tire les deux cartilages aryténoïdiens ensemble sur la ligne médiane ce qui provoque l'adduction des plis vocaux.

La partie transverse serait responsable du rapprochement de la base des aryténoïdes alors que la partie oblique permettrait de rapprocher les apex.

Les expériences de Choi (1995) sur un modèle canin ont permis de montrer que la phonation dépend de l'effet de l'inter-aryténoïdien sur la glotte postérieure et de l'effet des autres adducteurs sur la glotte membraneuse. De fait, aucune phonation n'est possible avec une stimulation isolée de l'inter-aryténoïdien à cause de la fermeture incomplète de la glotte membraneuse. A l'inverse, aucune phonation n'est possible en l'absence de stimulation de l'inter-aryténoïdien du fait de la fermeture incomplète de la glotte postérieure.

Ce muscle est également sollicité dans la régulation de la pression médiale entre les plis vocaux.

Le muscle inter-aryténoïdien est innervé par la division antérieure du nerf récurrent laryngé.

I.5.1.6. Remarques

Notons que le muscle constricteur inférieur et le muscle crico-pharyngien sont également des muscles intrinsèques du larynx. Ils participent au rôle de support cartilagineux du larynx pour la partie inférieure du pharynx, de l'hypopharynx et du

sphincter supérieur de l'œsophage mais n'ont pas d'action directe dans la phonation c'est pourquoi nous ne les avons pas décrits en détail.

Tableau 29 : Récapitulatif de la fonction et de l'innervation des différents muscles laryngés intrinsèques.

Muscle	Action principale	Innervation
Thyro-aryténoïdien	Adducteur	Nerf laryngé récurrent
Crico-thyroïdien	Tenseur du pli vocal	Branche externe du nerf laryngé supérieur
Crico-aryténoïdien postérieur	Abducteur	Nerf laryngé récurrent
Crico-aryténoïdien latéral	Adducteur	Nerf laryngé récurrent
Inter-aryténoïdien	Adducteur	Nerf laryngé récurrent

I.5.2. Les muscles extrinsèques

Les muscles extrinsèques au larynx ont un point d'attachement sur un organe voisin du larynx tout étant partiellement insérés sur celui-ci. Leur structure influence la position et le mouvement du larynx. Ils participent à la suspension, au support et à la mobilité du larynx. Lors de la phonation, ils ont pour effet d'abaisser la fréquence fondamentale.

Les muscles extrinsèques peuvent être différenciés suivant qu'ils sont situés au dessus ou au dessous de l'os hyoïde. Les muscles localisés au dessus sont appelés sus-hyoïdiens et participent à l'élévation du larynx tandis que les muscles situés au dessous de l'os hyoïde sont appelés sous-hyoïdiens et permettent l'abaissement du larynx.

I.5.2.1. Les muscles sus-hyoïdiens

Les muscles sus-hyoïdiens permettent l'ouverture de la cavité buccale en tirant la mandibule vers le bas, en élevant l'os hyoïde et en déplaçant le larynx vers le haut, l'avant et l'arrière. Ils ont tous un point d'attache sur le crâne ou la mandibule et sont reliés à l'os hyoïde.

Ces muscles sont au nombre de cinq : le digastrique, le mylo-hyoïdien, le génio-hyoïdien, le stylo hyoïdien et le hyoglosse. Ils sont innervés par la branche mandibulaire du nerf trijumeau (nerf crânien V) ou par le premier nerf spinal (nerf crânien XII).

I.5.2.2. Les muscles sous-hyoïdiens

Les muscles sous-hyoïdiens permettent le mouvement vers le bas, l'avant et l'arrière du larynx. Ils sont au nombre de quatre : le muscle thyro-hyoïdien (peut également être classifié comme un muscle élévateur), le sterno-hyoïdien, l'omo-hyoïdien et le sterno-thyroïdien.

Ce dernier pourrait raccourcir la longueur des plis vocaux permettant ainsi de diminuer leur tension et leur fréquence de vibration.

I.6. La vascularisation

La vascularisation artérielle est assurée par les artères thyroïdiennes supérieure et inférieure. En conséquence, la partie supérieure du larynx est irriguée par la branche laryngée supérieure de l'artère thyroïdienne supérieure. La vascularisation du cartilage crico-thyroïdien est assurée par la branche crico-thyroïdienne de l'artère laryngée supérieure. Enfin la branche laryngée inférieure de l'artère thyroïdienne est responsable de l'irrigation de la partie inférieure du larynx.

I.7. L'innervation

I.7.1. Le nerf vague

Le nerf vague est le X^{ème} nerf crânien, il assure de chaque côté du larynx son innervation motrice.

A droite, le nerf vague sort de la base du crâne et suit le trajet de la veine jugulaire interne. Le nerf laryngé supérieur dont la branche externe est motrice pour le muscle crico-thyroïdien naît à cet endroit, au niveau d'un/du ganglion plexiforme.

A gauche, le nerf vague sort de la base du crâne pour suivre le trajet de l'artère carotide primitive jusqu'à la crosse de l'aorte. C'est à cet endroit que naît le nerf laryngé inférieur ou récurrent. Il sous-croise ensuite l'aorte pour rejoindre l'angle trachéo-œsophagien.

I.7.2. Le nerf laryngé récurrent

Dans la mesure où le nerf récurrent naît dans le thorax, sous la crosse de l'aorte, des lésions, des traumatismes cardio-thoraciques ou médiastinaux peuvent avoir des conséquences d'atteinte récurrentielle au même titre que les lésions directement liées au larynx.

Chaque nerf récurrent est responsable de l'innervation de l'hémi-larynx dans lequel il pénètre.

A gauche, le nerf laryngé récurrent pénètre dans le larynx entre l'aile thyroïdienne et le chaton cricoïdien.

A droite, le nerf récurrent naît sous l'artère sous-clavière. Son trajet, ascendant, est plus oblique que celui du récurrent gauche. Conséquemment, il passe derrière le lobe thyroïdien avec un angle trachéo-bronchique plus latéral que celui du récurrent gauche.

La pénétration intra-laryngée, la bifurcation et les branches des nerfs récurrents sont identiques à droite et à gauche.

Rappelons que le nerf récurrent permet l'innervation motrice pour quatre des cinq muscles intrinsèques du larynx (à l'exception du crico-thyroïdien). Il a également une fonction sensitive pour les membranes muqueuses situées en dessous des plis vocaux.

I.7.3. Le nerf laryngé supérieur : branche externe

La branche externe du nerf laryngé supérieur innerve principalement le muscle crico-thyroïdien. Elle permet également, l'innervation sensitive de muqueuse sous glottique antérieure (Woodson, 1996) après avoir pénétré/sa pénétration dans le larynx (en passant) à travers/en traversant la membrane crico-thyroïdienne

Toutefois si le nerf laryngé supérieur est supposé n'innerver que le muscle crico-thyroïdien, certaines positions du pli vocal paralysé laissent à penser qu'il existerait une voie nerveuse motrice commandée par ce nerf-ci.

II. Annexes 2 : Fiches de renseignements

II.1. Fiche signalétique patient

Nº : _____ Date : _____

Nom et Prénom :

Sexe :

Date de naissance :

Age au moment de l'intervention :

Langue maternelle :

Autre(s) langue(s) pratiquées dont dialecte(s) :

Résidence principale pendant la durée du traitement :

Téléphone :

Milieu : socio-professionnel :

Pathologie :

Chirurgie envisagée :

Chirurgien :

Traitement envisagé :

II.2. Questionnaire Post-Opératoire adapté du Voice Quality Related Life

LOCUTEUR :

DATE :
Séance Ortho n° :

A cause de ma voix :

J'ai des difficultés à parler fort ou à être entendu.	Jamais	Parfois
Toujours		
Je ne sais pas ce qui va se passer quand je vais prendre la parole.	Jamais	Parfois
Toujours		
Je suis anxieux ou frustré à cause de ma propre voix.	Jamais	Parfois
Toujours		
Je me sens déprimé à cause de mon problème de voix.	Jamais	Parfois
Toujours		
J'ai des problèmes pour parler au téléphone.	Jamais	Parfois
Toujours		
J'ai des difficultés dans mon travail à cause de ma voix.	Jamais	Parfois
Toujours		
Je dois répéter pour me faire entendre.	Jamais	Parfois
Toujours		
Je sors moins de chez moi à cause de ma voix	Jamais	Parfois
Toujours		
Je suis à court de souffle et je dois reprendre fréquemment ma respiration quand je parle.	Jamais	Parfois
Toujours		
Douleurs similaires à une angine :	Jamais	Parfois
Toujours		
Problèmes de déglutition :	Jamais	Parfois
Toujours		
Problèmes de mobilité de la tête :	Jamais	Parfois
Toujours		
Gorge sèche :	Jamais	Parfois
Toujours		
Glaires :	Jamais	Parfois
Toujours		

Annexes 2

Fatigue générale :	Jamais	Parfois
Toujours		
Moral :	Jamais	Parfois
Toujours		
Voix enrouée :	Jamais	Parfois
Toujours		
Fatigue vocale :	Jamais	Parfois
Toujours		
Changements perçus par rapport à sa propre voix :		
Lesquels :	+ grave : ☐	+ aigue : ☐
☐		idem :
Pathologie :		
Remarques :		

III. Annexes 3 : Résultats

III.1. Valeurs formantiques Expérience 1

NPPENS

Tableau 30 : Valeurs formantiques (F1 et F2 en Hz) pour le locuteur NPPENS

	Préop		Post-Op 1		Post-Op 2	
	F1	F2	F1	F2	F1	F2
i	298	2209	309	2282	285	2267
a	718	1305	838	1406	927	1406
u	330	730	460	859	594	1296

Tableau 31 : Valeurs formantiques (F1 et F2 en Hz) pour le locuteur NPPGER

	Préop		Post-Op 1		Post-Op 2	
	F1	F2	F1	F2	F1	F2
i	298	2475	312	2452	294	2486
a	852	1334	727	1381	759	1200
u	313	789	410	1425	333	908

Tableau 32: Valeurs formantiques (F1 et F2 en Hz) pour le locuteur NPPHOE

	Préop		Post-Op 1		Post-Op 2	
	F1	F2	F1	F2	F1	F2
i	317	2296	309	2388	312	2364
a	655	1230	397	1129	528	1141
u	353	831	364	842	323	752

Tableau 33: Valeurs formantiques (F1 et F2 en Hz) pour le locuteur NPPHOF

	Préop		Post-Op 1		Post-Op 2	
	F1	F2	F1	F2	F1	F2

i	341	2587	370	2582	318	2626
a	812	1433	787	1457	788	1464
u	381	694	407	824	387	804

Tableau 34 : Valeurs formantiques (F1 et F2 en Hz) pour le locuteur NPPKAU

	Préop		Post-Op 1		Post-Op 2	
	F1	F2	F1	F2	F1	F2
i	258	2125	230	2143	249	2158
a	600	1164	443	1130	550	1099
u	311	777	310	1208	310	811

Tableau 35: Valeurs formantiques (F1 et F2 en Hz) pour le locuteur NPPKRE

	Préop		Post-Op 1		Post-Op 2	
	F1	F2	F1	F2	F1	F2
i	368	2854	331	2455	374	2805
a	740	1343	605	1311	815	1386
u	399	909	323	876	395	1046

Tableau 36 : Valeurs formantiques (F1 et F2 en Hz) pour le locuteur NPPLEN

	Préop		Post-Op 1		Post-Op 2	
	F1	F2	F1	F2	F1	F2
i	366	2406	354	2403	313	2449
a	781	1292	870	1372	711	1334
u	303	715	365	807	385	578

III.2. Valeurs formantiques Expérience 2

Tableau 37: Valeurs formantiques (F1 et F2 en Hz) pour le locuteur UPPBAS

	LC		Post-Op 2		Post op 3	
	F1	F2	F1	F2	F1	F2
i	339	2595	365	2501	408	2474
a	810	1437	870	1286	857	1286
u	373	701	361	914	380	866

Tableau 38: Valeurs formantiques (F1 et F2 en Hz) pour le locuteurUPPHEI

	LC		Post-Op 2		Post op 3	
	F1	F2	F1	F2	F1	F2
i	298	2652	298	2686	327	2651
a	676	1350	475	1223	705	1297
u	318	665	356	915	346	798

Tableau 39: Valeurs formantiques (F1 et F2 en Hz) pour le locuteur UPPHUB

	LC		Post-Op 2		Post op 3	
	F1	F2	F1	F2	F1	F2
i	341	2587	430	2685	277	2154
a	812	1433	1143	1974	787	1452
u	381	694	449	1137	298	728

Tableau 40 : Valeurs formantiques (F1 et F2 en Hz) pour le locuteur UPPLAT

	LC		Post-Op 2		Post op 3	
	F1	F2	F1	F2	F1	F2
i	275	2075	280	2177	306	2142
a	614	1089	737	1174	735	1250
u	296	951	299	828	319	702

Tableau 41 : Valeurs formantiques (F1 et F2 en Hz) pour le locuteur UPPPAI

	LC		Post-Op 2		Post op 3	
	F1	F2	F1	F2	F1	F2
i	275	2077	250	1944	265	1973
a	614	1137	898	1341	767	1085
u	296	781	360	1128	440	918

Tableau 42: Valeurs formantiques (F1 et F2 en Hz) pour le locuteur UPPWAL

	LC		Post-Op 2		Post op 3	
	F1	F2	F1	F2	F1	F2
i	341	2587	386	2574	358	2688
a	812	1433	796	1589	808	1482
u	381	694	435	951	417	772

Tableau 43: Valeurs formantiques (F1 et F2 en Hz) pour le locuteur UPPWAN

	LC		Post-Op 2		Post op 3	
	F1	F2	F1	F2	F1	F2
i	337	2581	332	2555	294	2798
a	805	1418	939	1594	764	1410
u	371	709	487	1667	327	812

III.3. Durées du VTT – Expérience 3

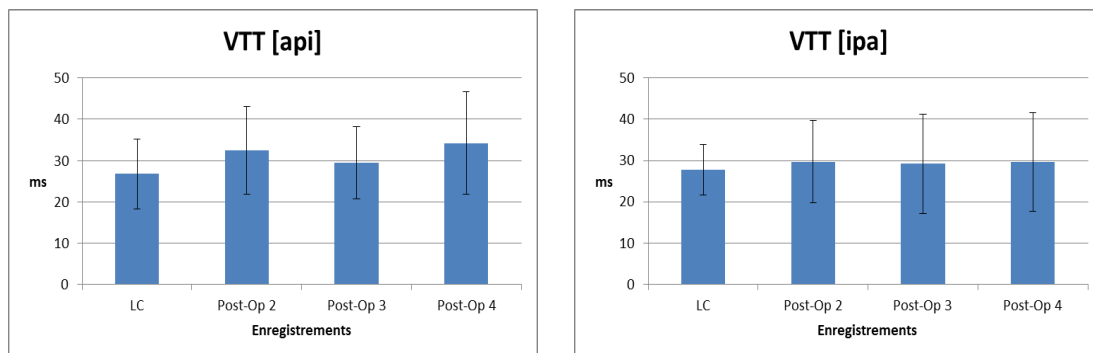


Figure 36 : Valeurs absolues de la durée du VTT (Agnello, 1975) en contexte [p] pour les logatomes [api] (à gauche) et [ipa] (à droite)

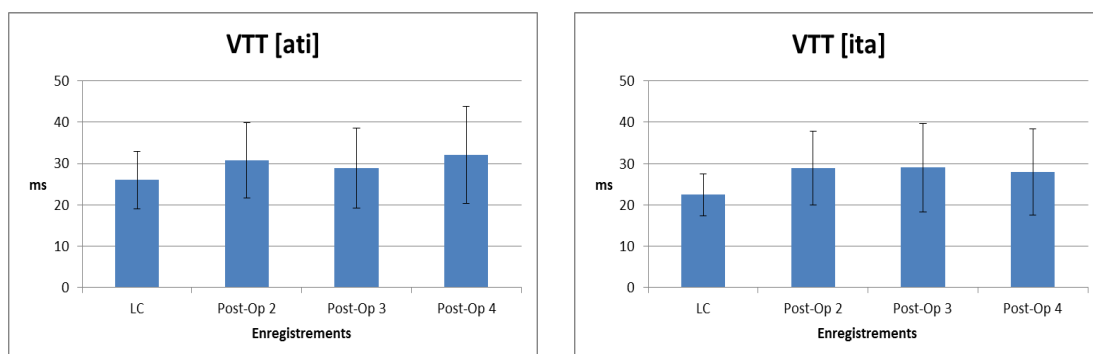


Figure 37 : Valeurs absolues de la durée du VTT (Agnello, 1975) en contexte [t] pour les logatomes [ati] (à gauche) et [ita] (à droite)

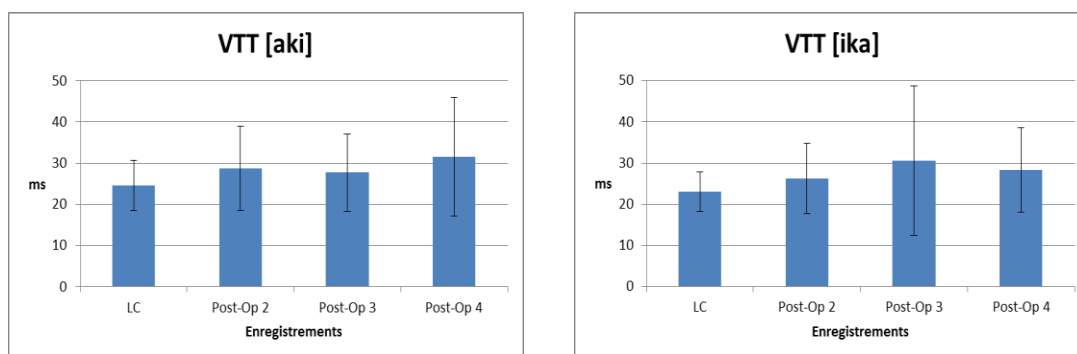


Figure 38 : Valeurs absolues de la durée du VTT (Agnello, 1975) en contexte [k] pour les logatomes [aki] (à gauche) et [ika] (à droite)

III.4. Durées vocaliques vs. durées consonantiques

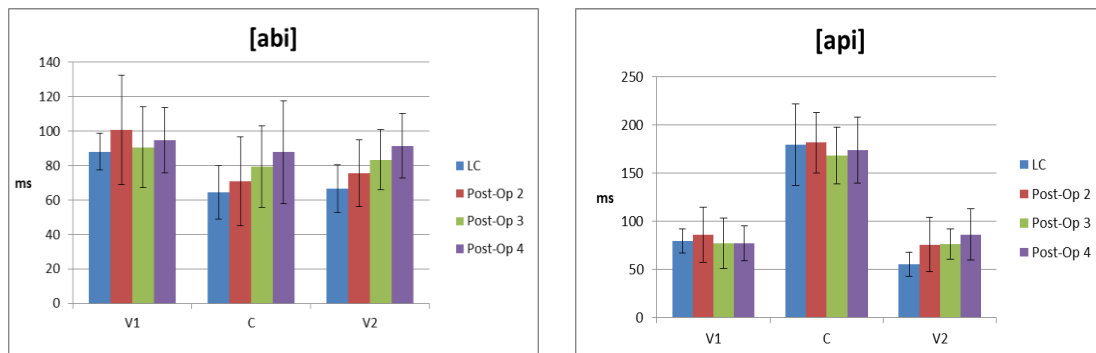


Figure 39 : durées vocaliques et consonantiques pour la séquence [abi] (à gauche) et pour la séquence [api] (à droite)

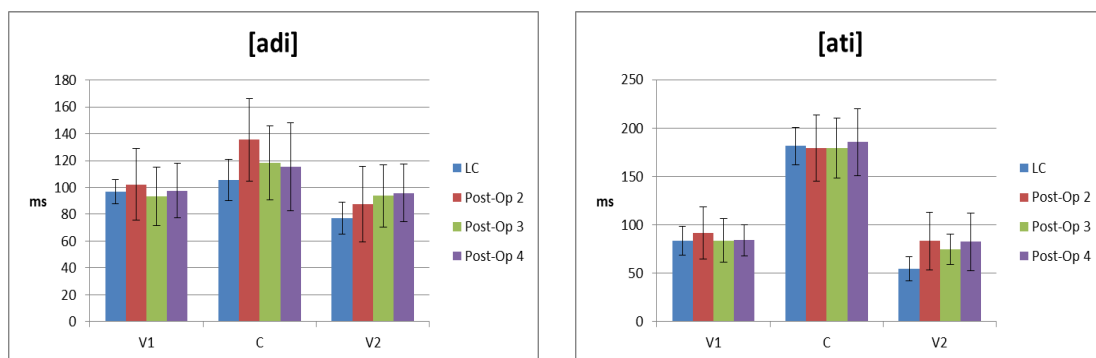


Figure 40 : durées vocaliques et consonantiques pour la séquence [adi] (à gauche) et pour la séquence [ati] (à droite)

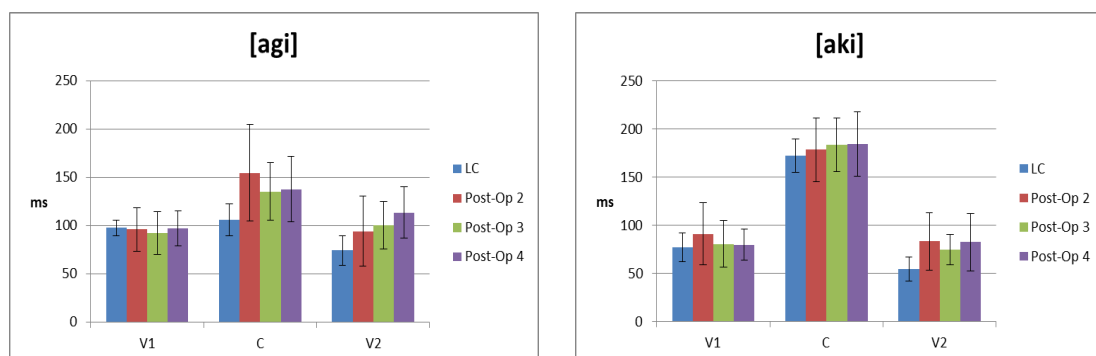


Figure 41 : durées vocaliques et consonantiques pour la séquence [agi] (à gauche) et pour la séquence [aki] (à droite)

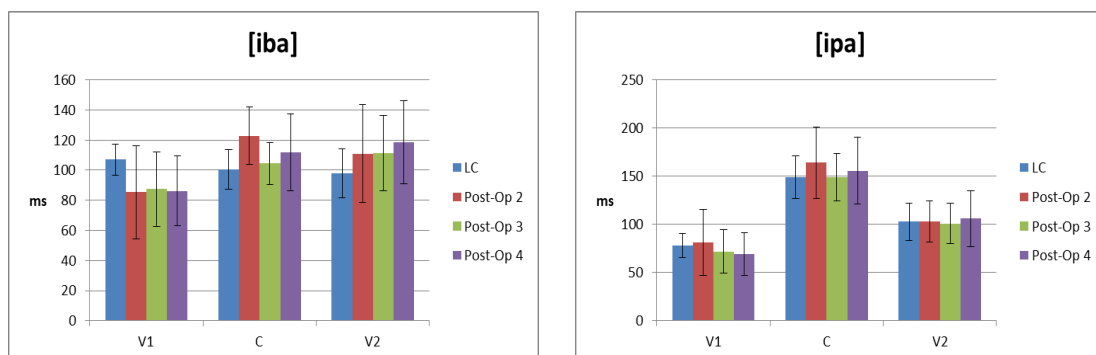


Figure 42 : durées vocaliques et consonantiques pour la séquence [iba] (à gauche) et pour la séquence [ipa] (à droite)

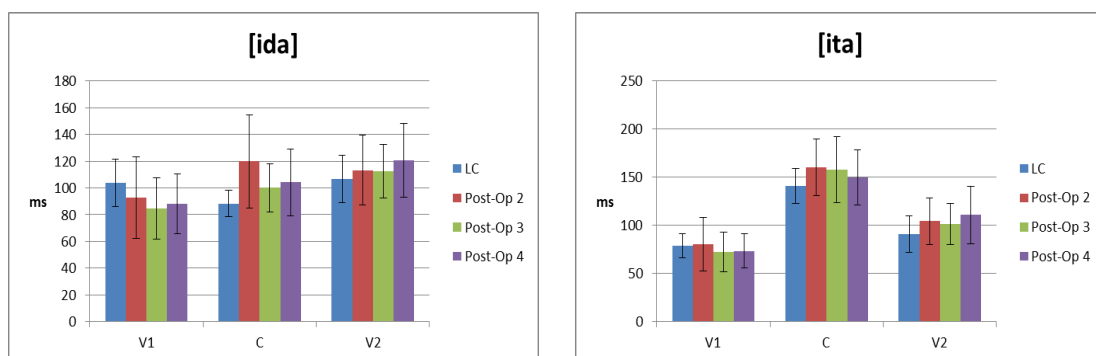


Figure 43 : durées vocaliques et consonantiques pour la séquence [ida] (à gauche) et pour la séquence [ita] (à droite)

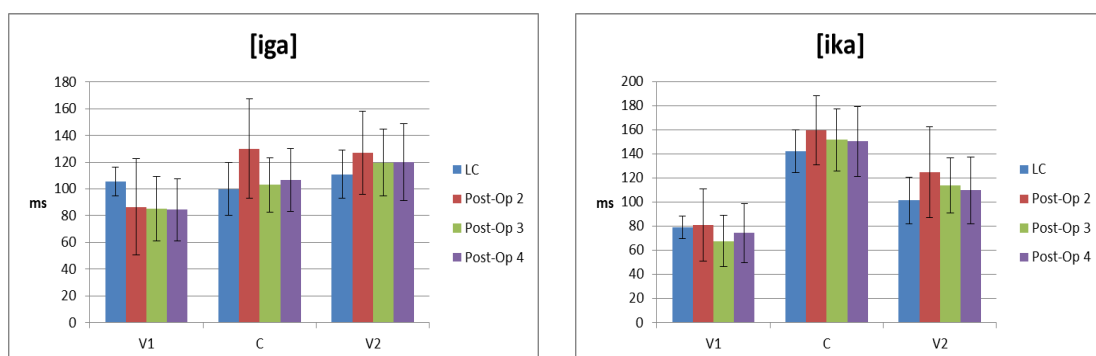


Figure 44 : durées vocaliques et consonantiques pour la séquence [iga] (à gauche) et pour la séquence [ika] (à droite)

PERTURBATION DE LA PRODUCTION DE LA PAROLE SUITE A UNE OPERATION DE LA GLANDE THYROÏDE

RESUME

L'objectif de ce travail, mené en partenariat avec le Département de Chirurgie Oncologique du Centre Paul Strauss à Strasbourg, est d'évaluer les conséquences d'une chirurgie thyroïdienne sur la voix des patients, afin de déceler les différentes perturbations qu'entraîne cette opération chirurgicale, et de mettre au jour les possibles stratégies de *compensation* ou de *réajustements* que le patient peut mettre en place, seul ou à l'aide d'une rééducation orthophonique ; il s'agit d'une étude longitudinale. L'évaluation reposait sur l'analyse de signaux acoustiques, à partir desquels une quantité importante d'indices relatifs à la qualité vocale, mais aussi aux comportements articulatoires des locuteurs, a pu être extraite. La production de voyelles soutenues a permis de conduire deux études spectrales. L'étude spatio-temporelle a été rendue possible grâce à l'analyse de logatomes, et plus précisément de séquences VCV. Nous avons mené trois expériences sur le plan acoustique. La première expérience traite des caractéristiques spectrales de la voix de patients ayant subi une thyroïdectomie et ne présentant pas d'immobilité laryngée. La deuxième étude est menée à partir des caractéristiques spectrales de la voix de locuteurs présentant une immobilité laryngée post-thyroïdectomie. Enfin, notre dernière expérience a pour objet, l'étude des caractéristiques spatio-temporelles de la voix de locuteurs qui ont subi une thyroïdectomie et présentant une immobilité laryngée. L'interprétation articulatoire, effectuée à partir des données acoustiques, révèle des perturbations aussi bien des gestes glottiques que des gestes supraglottiques, avec des stratégies de réajustements variables suivant les patients. La conception d'une « cible », comme espace de contrôle pour la réalisation de possibles articulatoires et acoustiques perceptivement acceptables, semble particulièrement pertinente ici, dans la mesure où les locuteurs pathologiques réorganisent leurs productions selon leurs propres contraintes physiologiques et anatomiques, provoquées par leur pathologie.

MOTS CLES : production de la parole, parole pathologique, phonétique clinique, glande thyroïde, perturbations, réajustements, viabilité, mesures spectrales, durées, VTT, VOT.

PERTURBATION OF SPEECH PRODUCTION AS A RESULT OF THYROID GLAND SURGERY

ABSTRACT

The aim of this study, conducted in partnership with the Department of Surgical Oncology, Centre Paul Strauss in Strasbourg, is to assess the consequences of thyroid surgery on the voice of patients, in order to identify the various perturbations caused by this surgery, and also to reveal possible compensatory strategies or readjustments that the patient may develop, alone or with speech therapy; this is a longitudinal study. The assessment was based on the analysis of acoustic signals, from which a large amount of cues related to voice quality, but also to articulatory behaviour of speakers was extracted. Production of sustained vowels allowed carrying out two spectral studies. The spatiotemporal study was made possible through the analysis of nonsense words, and more specifically through VCV sequences. We conducted three experiments on the acoustic level. The first experiment deals with the spectral characteristics of the voice of patients who underwent thyroidectomy, with no laryngeal paralysis. The second study was conducted based on spectral characteristics of the voice of speakers with laryngeal post-thyroidectomy paralysis. Our final experiment was to study spatio-temporal characteristics of the voice of speakers who underwent thyroidectomy, with laryngeal paralysis. Articulatory interpretations, made from the acoustic data, reveal perturbations of gestures on both the glottal and supraglottal levels, with readjustment strategies varying according to patients. The concept of a "target" as a control space for execution of possible articulatory and acoustic entities, which are perceptually acceptable, seems particularly relevant in this study, since disordered speakers reorganize their productions according to their own physiological and anatomical constraints, caused by the disorder.

KEYWORDS: speech production, speech disorder, clinical phonetics, thyroid gland, perturbations, readjustments, viability, spectral measurements, durations, VTT, VOT.