



HAL
open science

Apprécier sa voix en rééducation vocale : apports du spectrogramme. Étude de cas

Florence Parmentier

► To cite this version:

Florence Parmentier. Apprécier sa voix en rééducation vocale : apports du spectrogramme. Étude de cas. Sciences cognitives. 2013. dumas-00873952

HAL Id: dumas-00873952

<https://dumas.ccsd.cnrs.fr/dumas-00873952v1>

Submitted on 17 Sep 2019

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

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

ACADEMIE DE PARIS
UNIVERSITE PARIS VI PIERRE et MARIE CURIE
MEMOIRE POUR LE CERTIFICAT DE CAPACITE D'ORTHOPHONISTE

Apprécier sa voix en rééducation vocale :
apports du spectrogramme
Étude de cas

Directrice de mémoire : Isabelle MARIE-BAILLY
Co-directrice de mémoire : Claire PILLOT-LOISEAU

ANNEE UNIVERSITAIRE : 2012/2013

PARMENTIER Florence

Née le 28 août 1978

Remerciements

Je remercie Isabelle Marié-Bailly, directrice de mon mémoire, pour sa disponibilité, son écoute, le partage de sa vision du soin, et pour avoir donné sens à mon étude.

Un grand merci également à Claire Pillot-Loiseau, co-directrice de mon mémoire, qui a su orienter mes recherches et m'aider à finaliser mon travail.

Je remercie également Sylvain Lamesch, pour m'avoir encouragée au début de ma recherche et pour avoir accepté d'être mon rapporteur.

Je remercie les orthophonistes et les patients qui ont eu la gentillesse d'accepter de participer à mon étude, et qui ont, chacun, apporté leur pierre à mon travail.

Merci à Mathilde Descamps-Juhasz, qui m'a accompagnée dans ma recherche de patients, qui m'a donné l'occasion de me jeter à l'eau, et qui me donne encore l'occasion de beaucoup apprendre.

Merci à Françoise Rivière du service de prêt entre Bibliothèques, et merci aux bibliothécaires de la Bibliothèque Inter-Universitaire de Médecine.

Merci à Brigitte Lutrand Pezant pour son aide pour la rédaction de l'abstract.

Je tiens à citer particulièrement les étudiants osteovox : chaque session de formation a été l'occasion de la maturation de mon projet de mémoire et de mon projet professionnel.

Merci aux maîtres de stages qui ont su partager leurs savoir-faire et leurs savoir-être, et qui ont pu également me donner confiance, sur le long et parfois tortueux chemin de ma reconversion en orthophonie. Je tiens à citer notamment Michèle Plana, Valérie Gardeux-Zanotti et Nicole Oudin.

Merci à Emmanuelle Trinquesse avec qui je découvre les multiples facettes de la voix chantée, avec la simplicité, la joie et l'extrême compétence qui la caractérisent.

Merci à Isabelle Retailleau et à tous les chanteurs du Chœur de Chambre Agapanthe (d'hier et d'aujourd'hui) : vous m'êtes précieux et irremplaçables.

Merci en particulier à Marine Legrand, parce qu'un jour elle m'a parlé de sa rééducation vocale avec une orthophoniste formidable...

Merci à la fondatrice et aux membres du « Club des Mamans en Or(thophonie) », qu'elles soient mamans ou non, solidaires dans les difficultés.

Merci à Blandine Choret-Duru pour l'intelligence et la sensibilité de son accompagnement.

Un grand merci à mes relecteurs pugnaces et attentifs.

*Ce travail est dédié mes deux enfants, et à celui qu'on attend ; à leur père
Sans leur amour, rien n'a de sens.*

Sommaire

Introduction	1
1 Assises théoriques	2
1.1 Dysphonie dysfonctionnelle et rééducation vocale	2
1.2 Le spectrogramme	10
1.3 Spectrogramme et rééducation vocale : avantages et limites	18
2 Problématique	31
3 Hypothèses	32
4 Méthodologie	33
4.1 Matériel	33
4.2 Méthode.....	35
5 Analyse des résultats	51
5.1 Auto-analyse : échelles visuelles analogiques	51
5.2 Évolution des patients du groupe « protocole »	66
5.3 Commentaires autour du spectrogramme	71
6 Discussion	87
6.1 Validation des hypothèses de travail	87
6.2 Limites.....	94
6.3 Perspectives	98
7 Conclusion	100
8 Table des matières	101
9 Bibliographie	109
10 Annexes	117
10.1 Classifications des images spectrographiques caractérisant la qualité vocale	117
10.2 Évolution objective des patients du groupe protocole	118
10.3 Le spectrogramme en pratique	119
10.4 Documents du protocole.....	124
10.5 Questionnaire en ligne sur le spectrogramme	129
10.6 Exemples d'images spectrographiques	130

Index des tableaux

Tableau I : lecture des phonèmes sur le spectrogramme.....	11
Tableau II : grille d'évaluation des qualités vocales à partir de l'image spectrographique (Espanol, 2011) [26].....	13
Tableau III : calendrier des séances pour les patients du groupe protocole.....	36
Tableau IV : définition des différents items du GRBAS (Revis <i>et al.</i> 2012, p.59) [91].....	37
Tableau V : activités proposées avec les patientes du groupe protocole. Temps passé en minutes et temps total.....	39
Tableau VI : plan du site internet pour l'utilisation du spectrogramme en autonomie.....	41
Tableau VII : recueil des données : nombre d'évaluations par groupe.....	49
Tableau VIII : corrélations entre les échelles de clarté, richesse, attaque et intensité.	53
Tableau IX : corrélations entre les des différentes échelles avec l'échelle geste	53
Tableau X : fiabilité intra-juge selon les différents feedbacks et effet d'apprentissage	55
Tableau XI : fiabilité inter-juges avant / après harmonisation et selon le type de feedback.....	56
Tableau XII : conclusion sur les fiabilités intra- et interjuge.....	57
Tableau XIII : corrélations entre les critères subjectifs de clarté et de richesse et les mesures objectives selon les groupes et le type de feedback	57
Tableau XIV : significativité de l'indépendance des évaluations avant et après feedback selon les groupes.	58
Tableau XV : aide, gratification et utilisation du feedback additionnel pour l'échelle richesse	61
Tableau XVI : aide, gratification et utilisation du feedback additionnel pour l'échelle clarté.....	62
Tableau XVII : aide, gratification et utilisation du feedback additionnel pour l'échelle qualité d'attaque	62
Tableau XVIII : aide, gratification et utilisation du feedback additionnel pour la différence d'intensité	62
Tableau XIX : aide, gratification et utilisation du feedback additionnel pour l'échelle geste.....	63
Tableau XX : aide, gratification et utilisation du feedback additionnel pour la différence de hauteur	63
Tableau XXI: impact du feedback additionnel : résumé.....	65
Tableau XXII : évolution de l'évaluation objective des patients du groupe protocole	66
Tableau XXIII : accord inter-juges pour le GRBASI	67
Tableau XXIV : accord inter-juges pour l'évaluation subjective des progrès.....	67
Tableau XXV : évolution des questionnaires d'auto-évaluation	68

Tableau XXVI : évolution des caractéristiques citées dans la liste d'adjectifs.	68
Tableau XXVII : changements positifs et négatifs dans les caractéristiques citées, classées selon leur nature.....	69
Tableau XXVIII : nombre d'adjectifs cités par les patientes avant et après protocole, et nombre de changements positifs et négatifs.....	70
Tableau XXIX : statistiques de visites du site internet pour les patients en autonomie	84
Tableau XXX : classification de Yanagihara et ses adaptations ultérieures (Yanagihara, 1967, pp.533-534, (Remacle & Morsomme, 2009, p.129, Wolfe & Steinfatt, 1987, p.233, Cooper, 1974, p.291) [116] [86] [114] [16].....	117
Tableau XXXI : classification de Remacle & Morsomme (2009, p.130) [86] à gauche, et son pendant à droite par Delacroix et al. (2009, p.427) [24], des tracés spectrographiques caractérisant la raucité vocale	118
Tableau XXXII : déroulement de l'exercice d'auto-analyse.....	125

Index des figures

Figure I : groupes (groupe protocole, groupe hors protocole, groupe en autonomie et jury d'écoute) et déroulement général du recueil des données.....	35
Figure II : reproduction de la feuille de passation et distribution des réponses pour chacune des échelles tous sons confondus et tous feedbacks confondus, et pour chaque groupe.....	51
Figure III : Utilisation des échelles visuelles analogiques selon les groupes.....	52
Figure IV: erreur avant feedback pour les patients	52
Figure V : synthèse des corrélations tous groupes et tous feedbacks confondus.....	54
Figure VI : Utilisation du feedback additionnel.....	58
Figure VII : effet gratifiant ou dévalorisant du feedback additionnel.....	59
Figure VIII : effet aidant ou gênant du feedback additionnel selon les groupes.....	61
Figure IX : effet du feedback additionnel sur les corrélations en fonction du groupe.....	64
Figure X : Evolution des répartitions des adjectifs choisis	71
Figure XI : changements positifs et négatifs dans les adjectifs choisis	71
Figure XII : spectrogrammes jugés comme type I, III et IV par les deux juges (Yumoto <i>et al.</i> , 1984, p. 4) [118].....	117
Figure XIII : représentation schématique des 4 degrés de raucité sur le spectrogramme (Yanagihara, 1967, p.535) [116].....	117
Figure XIV : évolution de la dynamique d'intensité	118
Figure XV : évolution de l'étendue vocale	118

Figure XVI : évolution du TMP sur [a].....	118
Figure XVII : évolution du rapport z /s.....	118
Figure XVIII : Evolution de la mesure du jitter	119
Figure XIX : évolution de la mesure du shimmer	119
Figure XX : évolution de la mesure du singing formant.....	119
Figure XXI : évolution de la mesure du rapport signal / bruit	119
Figure XXII : évolution de la hauteur tonale	119
Figure XXIII : feuille de passation de l'auto-analyse de 2 [a].....	124
Figure XXIV : formulaire de consentement libre et éclairé.....	125
Figure XXV : évaluation perceptive : comparaison extraits A et extraits B.....	126
Figure XXVI : Evaluation Perceptive : GRBAS(I).....	126
Figure XXVII : entretien semi-dirigé pré-protocole	126
Figure XXVIII: entretien semi-dirigé post-protocole (première partie)	127
Figure XXIX : entretien semi-dirigé post-protocole (deuxième partie : évaluation du spectrogramme)	127
Figure XXX : VHI-10	128
Figure XXXI : questionnaire relatif à sa relation à sa voix.....	128
Figure XXXII : liste d'adjectifs sur la voix	128
Figure XXXIII : sirène sur [a] (patiente D) ; échelle 8kHz	130
Figure XXXIV : f non efficace puis f efficace (patiente C) ; échelle 8kHz.....	130
Figure XXXV : "mammographie" sur l'exercice de coloration du [s] (patiente B) ; échelle 8kHz....	130
Figure XXXVI : essai infructueux de fry puis fry réussi après une sirènes descendante (patiente C), echelle 8kHz.....	130
Figure XXXVII : "ti té tè ta" toniques (patiente B), échelle 8kHz	130
Figure XXXVIII : sirène en battements de lèvres, présentant des ruptures (patiente C) ; échelle 4 kHz	130
Figure XXXIX : visualisation du spasme sur une sirène sur [a] (patiente B) ; échelle 4kHz.....	130
Figure XL : fry sur [i] et [a] alternés (patiente D) ; échelle 4kHz.....	130
Figure XLI : extrait chanté (patiente A) ; échelle 4 kHz.....	130

Introduction

La rééducation vocale des dysphonies a pour objectif de redonner au patient un geste vocal efficace et confortable permettant une production satisfaisante.

Le plus souvent, le thérapeute doit attirer l'attention du patient sur les différents paramètres vocaux de ses productions (intensité, hauteur, timbre, début de son ...) afin que le patient prenne conscience de ses capacités et qu'il sache adapter son geste à la fois grâce à son ressenti intérieur et à ce qu'il perçoit de sa voix.

Dans un contexte où les outils informatiques sont de plus en plus utilisés en orthophonie, nous nous proposons d'observer l'impact de l'utilisation du spectrogramme en temps réel en séance d'orthophonie, notamment en ce qui concerne la prise de conscience et l'auto-contrôle des paramètres vocaux par le patient.

Effectivement, le spectrogramme en temps réel matérialise par un feedback visuel à la fois l'intensité, la hauteur, le timbre (au travers des harmoniques), et permet d'objectiver l'instabilité, le coup de glotte, le forçage vocal, la présence de souffle, la prosodie, le vibrato naturel, etc. Cette analyse peut être partagée avec le patient en séance, et constitue un bon complément aux évaluations subjectives. Quelle que soit l'approche rééducative choisie, l'outil permet de comparer des productions, constater des progrès, soutenir l'intérêt et la motivation du patient pour la rééducation.

Au travers d'une étude de 4 cas, complétée par l'analyse statistique d'exercices d'auto-analyse dans une population plus nombreuse et par un large recueil de témoignages, nous étudierons l'impact d'un feedback auditif et visuel sur la prise de conscience des paramètres vocaux (intensité, hauteur, qualité de l'attaque, clarté et richesse du timbre, geste vocal). Nous nous intéresserons également aux impacts positifs sur l'image qu'a le patient de sa voix, permis par la visualisation de sa propre voix, et sur la relation thérapeutique.

1 Assises théoriques

Après nous être intéressés à la dysphonie dysfonctionnelle, aux aspects multiples de la rééducation vocale et aux différents rétro-contrôles (feedbacks) existant pour la phonation et en rééducation vocale, nous présenterons l'outil constitué par le spectrogramme.

1.1 Dysphonie dysfonctionnelle et rééducation vocale

Voyons d'abord en quoi la dysphonie dysfonctionnelle est liée au psychisme et souvent à une mauvaise image de soi du patient. Dans ce cadre, nous décrirons, d'une part, les objectifs multiples d'une rééducation vocale ; nous nous intéresserons d'autre part au rétrocontrôle de façon théorique, puis au travers des différents feedbacks utilisés lors de la phonation, et particulièrement en rééducation vocale.

1.1.1 Dysphonie dysfonctionnelle et image de soi

1.1.1.1 Définition

« La dysphonie est une anomalie de la voix portant sur un ou plusieurs paramètres vocaux, la hauteur, l'intensité ou le timbre » Dans les dysphonies dysfonctionnelles, il « existe une perturbation du geste vocal à l'origine du trouble, ou le pérennisant. Elles peuvent être accompagnées ou pas d'une laryngopathie ». (Amy de la Bretèque, 2012, p. 130) [3]. « C'est à Le Huche que l'on doit le terme de dysfonctionnel qui doit être préféré à fonctionnel, car il traduit mieux l'utilisation défectueuse des moyens vocaux » (Remacle, 2009, p. 243) [85].

Les symptômes peuvent concerner tous les paramètres vocaux. (Fugain, 1990, p. 101) [37] :

- le timbre peut être voilé, enroué, éraillé, rauque, sourd, étouffé...
- la hauteur de la voix peut être aggravée ou au contraire trop aiguë, ou encore monotone,
- l'intensité de la voix peut être trop faible.

1.1.1.2 Forçage vocal et dérèglement fonctionnel

Une dysphonie dysfonctionnelle provient d'une attitude vocale défectueuse. Cette attitude est apparue en réaction à des modifications qui peuvent être d'origines diverses (laryngite aiguë, intubation prolongée, hypoacousie, hypotonie générale...). « Ces modifications directes ou indirectes peuvent avoir disparu et laisser comme séquelle l'attitude vocale défectueuse adoptée en réaction » (Remacle, 2009, p. 245) [85]. Une compensation inadéquate du patient peut le faire entrer dans ce que Le Huche a appelé « le cercle vicieux du forçage vocal ». Le geste vocal se modifie et devient de plus en plus inefficace. Le modèle Hypo et Hyper de Gray et Roy décrit un comportement hyperfonctionnel chez le sujet présentant une dysphonie dysfonctionnelle. Le cercle vicieux du forçage vocal est décrit par Giovanni et Aumelas (2004, p. 39) [40], à la suite de Le Huche.

Françoise Estienne précise : « dans certains cas [...], le dysfonctionnement est issu d'une attitude, d'un comportement du sujet. Cette attitude et ce comportement résultent d'un système de valeurs, d'une représentation de soi-même. Cette façon d'être génère le geste vocal qui, à son tour, conforte l'attitude du sujet et son système de valeurs. » (Estienne, 1998, p. 8) [27].

1.1.1.3 *Dysphonie et mauvaise image de soi*

La voix reflète l'état émotionnel d'une personne, elle peut également révéler des aspects de sa personnalité¹. Dans la littérature scientifique, la dysphonie est souvent reliée à des facteurs psychologiques² (Deary *et al.*, 2003, p. 375) [20]. Les patients dysphoniques peuvent être anxieux, dépressifs, ils ont souvent une moins bonne qualité de vie. Lorsque la cause de la dysphonie est organique, comme lors d'une paralysie récurrentielle, les patients peuvent également voir leur image de soi se détériorer suite à l'atteinte de leur fonction vocale (Siupsinskiene *et al.*, 2011, p. 281) [101].

1.1.2 Rééducation vocale

1.1.2.1 *Objectifs*

Selon l'enquête Kalliope, les orthophonistes pratiquant une rééducation vocale définissent une « belle voix » selon trois dimensions : une belle voix « *selon des paramètres mesurés objectivement et relevant de l'anatomie et de la physiologie* », une voix « *agréable à entendre en société* » une voix « *en harmonie avec l'être de la personne* ». Ces trois dimensions permettent de dégager trois modèles de pratique : « *soin thérapeutique* », « *conformité sociale* » et « *cohérence de la personne* » (Fortin *et al.* 2007, p. 237) [34].

L'objectif d'une rééducation peut effectivement être mis en parallèle avec ce qu'on peut mesurer lors d'un bilan vocal. Dans le cadre de l'amélioration du geste vocal, la rééducation vocale a classiquement pour objectifs :

- d'améliorer les paramètres aéro-acoustiques (temps maximum de phonation, facteurs d'instabilité jitter et shimmer, fréquence fondamentale, dynamique d'intensité, tessiture...),
- d'améliorer le jugement perceptif, correspondant à l'impression subjective faite par une voix sur un auditeur (l'échelle GRBAS est la plus utilisée),
- de diminuer le handicap ressenti par le sujet (l'échelle VHI est la plus utilisée).

Selon Françoise Estienne, « *l'objectif de la rééducation est d'enrichir la voix, de lui restituer le maximum de ses potentialités en fonction du désir du sujet et de ses limites organiques [...]. Ce n'est en aucun cas l'embrigader dans un modèle préfabriqué. [...] L'objectif thérapeutique va consister à entrer dans le vécu vocal de la personne avec sa conception de la voix, son système de valeurs pour lui proposer de l'élargir, de l'assouplir, de le revoir ou de le compléter en présentant la voix sous une forme concrète, accessible, un geste sur lequel on a prise et que l'on peut modeler* ». (Estienne, 1998, p. 9) [27]. Il s'agit ici plus précisément de donner au patient la possibilité de jouer avec sa voix et d'avoir par conséquent avec elle une relation apaisée.

Plus profondément encore, constatant qu'« *une voix qui va mieux est une porte d'entrée à un remaniement de toute la personne* », et que, d'autre part, « *un état émotionnel négatif peut être causé et/ou entretenu par un problème vocal* », Françoise Estienne et Dominique Morsomme écrivent que « *l'objectif d'une rééducation peut être de travailler le geste vocal ce qui va entraîner d'autres sensations et émotions et recadrer [la] façon de [se] percevoir et de percevoir [sa] voix [du patient]* » (Estienne & Morsomme, 2009, p. 267) [29].

¹ Green, M, Mathieson L (2001) *The voice and its disorders*. Whurr, London

² Wilson J.A., Deary I.J, Scott S. et al. (1995) *Functional dysphonia*. Br. Med. J 311, 1039-1040

Peut-être peut-on considérer, en orthophonie comme en musicothérapie, qu'« *une aide à un mieux-être, à une meilleure gestion du handicap* » est également un soin. (Blayac *et al.*, 2012, p. 149) [12].

1.1.2.2 Complémentarité des conceptions scientifiques et traditionnelles

Comme le rapporte Caroline Leclerc, les professions d'orthophoniste et de phoniatre utilisent des modèles étiologico-thérapeutiques à la fois ontologiques (investigation objective de l'anatomie et de son fonctionnement physiologique) et fonctionnalistes (exploration du déséquilibre inhérent au sujet) (Leclerc, 2007, pp. 107–108) [60]. « *La pratique orthophonique en rééducation vocale se fonde sur les savoirs récents de la biomédecine, bénéficiant de surcroît de la légitimité du paradigme scientifique dans notre société* » (Bergeras & Tain, 2007, p. 177) [10]. Cependant les orthophonistes désireux de redonner une voix qui soit en cohérence avec la personne recherchent des alternatives à la biomédecine (Fortin *et al.*, 2007, p. 238) [34]. Les savoirs traditionnels, non médicaux et non scientifiques, restent incontournables : « *la formation auprès de chanteurs, professeurs de chant, musicothérapeutes, spécialistes Feldenkrais, professeurs de yoga, etc. représente environ un quart des stages de formation effectués par les orthophonistes* » (Bergeras & Tain, 2007, p. 177) [10].

L'orthophoniste partenaire du médecin et du chirurgien

« *La dysphonie, et plus particulièrement la dysphonie dysfonctionnelle, est plurifactorielle. Pour la traiter, il faut donc lutter contre tous les facteurs pouvant la provoquer ou l'aggraver* ». (Amy de la Bretèque, 2012, p. 132) [3]. Benoît Amy de la Bretèque rappelle que le rééducateur doit être le relais vers le médecin pour parvenir à limiter reflux gastro-oesophagiens, allergies respiratoires, troubles hormonaux, habitudes toxiques, etc. La phonochirurgie peut aussi s'avérer nécessaire. On peut prévoir également d'autres thérapies : kinésithérapie, yoga, sophrologie, psychothérapie (Amy de la Bretèque, 2012, p. 130) [3].

La dimension psychologique

La relation thérapeutique revêt, dans le cadre de la rééducation vocale, une dimension psychologique dont il est important de tenir compte. La voix émise par le sujet joue un « *rôle éminent dans la constitution de la personnalité et parfois peut-être aussi dans sa désorganisation. Le patient dysphonique ou aphonique, que des causes organiques aient été ou non trouvées à son symptôme, va expérimenter une des situations les plus traumatisantes qui soit, il a perdu un lien essentiel avec lui-même et avec le monde qui l'entoure. C'est dire l'intérêt d'une prise en charge globale de toute perturbation de l'émission vocale* » (Chevrie-Muller, 1990, p. 64) [15]. Lorsqu'il propose neuf éléments ou principes de base afin de mesurer et comparer les traitements orthophoniques, Philippe Dejonckere intitule le 4^e élément «*prise en charge psychologique*» (Dejonckere, 2001, p. 353) [21]. L'auteur note : « *c'est l'individu dans sa totalité, et en particulier avec ses dimensions psychologique et émotionnelle, qui devra être assumé. La composante psychogène peut bien entendu être primaire, mais elle peut aussi résulter de la perturbation sociale, professionnelle ou émotionnelle qu'occasionne par répercussion une affection primairement organique* ».

Comme le précisent certains auteurs, il est nécessaire cependant que « *l'orthophoniste ne soit pas pris pour un psychothérapeute, dont il n'a d'ailleurs pas la formation, bien que ses qualités d'écoute et d'empathie soient à l'évidence fondamentales pour assurer le succès de l'entreprise.* » (Amy de la

Bretèque, 2012, p. 133) [3]. Parfois c'est le travail sur la voix qui permet à la personne d' « *expérimenter qu'elle peut quand même agir sur sa voix et que celle-ci est prête à changer* » (Estienne & Morsomme, 2009, p. 267) [29].

Technique de rééducation orthophonique

Selon Benoît Amy de la Bretèque notamment, « *la rééducation soigne la voix plus que les cordes* », et l'objectif est de trouver en premier lieu un bon accord pneumo-phonatoire, c'est-à-dire un équilibre aérodynamique qui permette une bonne vibration laryngée, puis un accord phono-résonantiel, qui enrichisse le timbre de la voix.

Dans leur précis d'audiologie et de déglutition, Françoise Estienne et Dominique Morsomme proposent un tour d'horizon de la littérature francophone et anglo-saxonne sur le traitement orthophonique des pathologies vocales. (Estienne & Morsomme, 2009, pp. 258–261) [29]. Elles concluent : « *les ouvrages que nous avons relevés décrivent minutieusement une série d'exercices très précis centrés sur la posture, la respiration, la relaxation, l'entraînement vocal proprement dit, axé sur les divers paramètres de la voix : tenue, hauteur, étendue, timbre, mélodie... Les entraînements sont effectués à partir de certaines consonnes et voyelles choisies en fonction de leur effet sur la mécanique vocale, exercices progressifs de comptage, lecture indirecte et directe, mise en situation, projection vocale, etc... Il y a donc une étonnante convergence entre les auteurs de langue française et anglo-saxonne pour présenter une rééducation vocale comme un parcours minutieux, décortiqué, applicable dans les divers types de pathologies vocales* ».

1.1.2.3 Conclusion

La rééducation vocale comprend à la fois une partie technique et une partie plus psychologique s'intéressant notamment à l'image de soi du patient. Notre travail va donc s'attacher à étudier les conséquences objectives et subjectives de l'utilisation d'un outil de rétro-contrôle visuel : le spectrogramme. Voyons d'abord en quoi consiste le rétro-contrôle en rééducation vocale.

1.1.3 Rétro-contrôle (feedback) en rééducation vocale

1.1.3.1 Définition de l'apprentissage par feedback

Graham F. Welch (1985, p. 3) [109] rappelle que c'est Wiener en 1948³ qui appliqua au corps humain le modèle physique de rétro-contrôle appliqué habituellement aux machines.

« *Dans la littérature traditionnelle sur l'acquisition d'habileté, le feedback correspond à l'information sensorielle (auditive, visuelle, kinesthésique, etc.) que le sujet qui apprend traite, relativement à sa propre performance* »⁴ (Verdolini & Krebs, 1999, p. 234) [107].

Le terme de feedback désigne l'ensemble du dispositif de régulation d'un système par une rétroaction. Ce système de régulation met en œuvre :

- une **rétro-information** (ou information de retour, feedback) : qui donne à un événement une valeur informationnelle. Pour cela cet événement doit être capté transmis et analysé. Il est soit

³ Wiener, N. (1948) *Cybernetics or Control and Communication in the Animal and the Machine*. New York: Wiley

⁴ "in the traditional skill-acquisition literature, feedback refers to sensory information –auditory, visual, kinesthetic, and so forth- that learners process relative to their own performance"

d'origine interne (intéroceptive ou proprioceptive) soit d'origine externe (extéroceptive = KR = « Knowledge of Result » = feedback externe),

- un **rétro-contrôle** : qui compare ce qui a été réalisé à ce qui était prévu,
- une **rétro-commande** (ou centre de décision) qui assure une correction en cas d'écart.

« Pour que l'acte réalisé soit conforme à son projet, il est nécessaire que des informations en retour (rétro-information) renseignent sur la qualité de son exécution (rétro-contrôle) et permettent un ajustement (régulation) c'est-à-dire une correction (rétro-commande) aboutissant à la réalisation de l'acte décidé (rétro-action) » (André *et al.*, 1986, p. 290) [4].

Le « *knowledge of results* » (connaissance des résultats ou feedback externe) « correspond à l'information fournie à celui qui apprend par une source externe, un professeur par exemple, à propos des résultats de la production, relativement à un certain but »⁵ (Steinhauer & Grayhack, 2000, p. 137) [102].

Deux modalités de connaissance des résultats doivent être associées pour permettre un apprentissage : le **feedback d'action** permettant de modifier l'action en cours, et le **feedback terminal** ou d'apprentissage indiquant les modifications devant intervenir lors des actions futures. (André *et al.*, 1986, p. 298) [4]

1.1.3.2 Trois phases de l'apprentissage moteur

Selon Fitts et Postner 1967⁶ (in Volin, 1999, pp. 242–243) [108], l'apprentissage moteur connaît trois phases qui se chevauchent : la phase cognitive, la phase associative et la phase d'autonomisation.

- La **phase cognitive** est relativement brève et permet des progrès importants dans la réalisation d'une tâche motrice. Dans cette phase, le sujet utilise des feedbacks pour analyser et comprendre les différentes composantes de l'activité à apprendre.
- La **phase associative** demande plus de temps. Durant la phase associative, les sujets font des ajustements fins et les erreurs diminuent en fréquence et en importance. Le sujet acquiert des mécanismes internes de détection et de correction d'erreur qui remplacent progressivement les éléments de feedback externe. Afin d'assurer un apprentissage à long terme qui ne soit pas dépendant du feedback externe, il est important de réduire progressivement de feedback externe (Volin, 1999, p. 243) [108].
- La **phase d'autonomisation** est atteinte après de longues heures de pratique. L'attention consciente n'est plus nécessaire pour contrôler l'action et la corriger. Le sujet peut se concentrer sur des tâches de plus haut niveau.

1.1.3.3 Particularité du rétro-contrôle pour la production vocale

Rétro-contrôles habituels pour la voix

Sensoriel, proprioceptif, postural

Des mécanorécepteurs laryngés donnent aux centres corticaux des informations sur l'état de tension et la position des différents muscles et des différentes articulations. « *En cours de phonation, ces*

⁵ "Knowledge of results (KR) refers to information provided to the learner by an external source, such as a teacher, about performance results relative to some goal"

⁶ Fitts, Postner (1967) *Human performance*. Belmon, CA ; Brooks/Cole.

informations permettent les ajustements instantanés nécessités par le maintien d'une configuration glottique donnée. D'autres arcs réflexes d'origine abdominale, thoracique, cervicale, linguale, etc. contribuent probablement aux ajustements permanents agissant par feedback sur le larynx au cours de la phonation » (Giovanni *et al.* 2012, p. 16) [42]. Le feedback proprioceptif concerne ainsi la position des principales structures du larynx, notamment la position des plis vocaux et des cartilages aryténoïdes, la position de la langue, du palais mou et des lèvres (Welch, 1985, p. 8) [109].

Larson et ses collègues ont étudié le comportement du larynx lorsqu'il était anesthésié, et ont conclu que les feedbacks auditif et somatosensoriel sont tous les deux utilisés pour le contrôle de la hauteur tonale (Larson *et al.* 2008, p. 619) [57].

Auditif: boucle audio-phonatoire

« *La voix est sous le contrôle des centres auditifs comme le montre la voix perturbée et non modulée des sourds profonds. Il existe probablement des commandes induites volontairement par les voies cortico-bulbaires en réponse aux informations acoustiques parvenues jusqu'au cortex auditif de même qu'un ensemble de réflexes acoustico-laryngés* » (Giovanni *et al.*, 2012, p. 16) [42].

Citant Johan Sundberg, Garyth Nair (1999, p. 14) [78] explique la **différence de perception auditive interne et externe** par le fait que la conduction osseuse, par laquelle notre propre voix arrive à notre oreille, transmet moins bien les hautes fréquences que la conduction aérienne. Il est rejoint en cela par Welch (1985, p. 9) [109] citant Békésy⁷, et par Verdolini et Krebs (1999, p. 234) [107]. On note qu'on peut augmenter la perception externe d'un locuteur ou d'un chanteur à moindres frais en utilisant ses mains comme porte-voix entre la bouche et l'oreille. Garyth Nair pense que cette technique tronque et modifie le son arrivant à l'oreille, il cite pour sa part le système ARTF (Aural Real-Time Feedback) qui renvoie grâce à un micro et des haut-parleurs la voix du sujet à ses oreilles (Nair, 1999, pp. 66–67) [78].

Le feedback auditif, sous la forme **d'enregistrements écoutés** a posteriori est facilement accessible (Le Huche & Allali, 1989, p. 151) [59], mais, n'étant **pas en temps réel**, il ne relève pas directement de la boucle audio-phonatoire. Garyth Nair recommande, pour les chanteurs, d'utiliser d'abord le « *post-performance aural feedback* » avec un enregistrement, ce qui permet d'affiner l'acuité de l'oreille de l'élève, et d'utiliser ensuite le « *aural real-time feedback* », c'est-à-dire le retour instantané de la production vocale à l'oreille de l'élève, afin de réduire l'écart entre ce que l'élève perçoit et ce qui est émis à l'extérieur (Nair, 1999, p. 67) [78].

Feedback verbal

Nous relevons ici des remarques provenant de recherches conduites sur les leçons de chant, mais qui pour une grande part s'appliquent également au feedback verbal fourni aux patients dysphoniques et dysodiques par les orthophonistes.

Selon Garyth Nair, la pédagogie vocale passe souvent par la modélisation par le professeur du geste et l'imitation par l'élève, et ce processus comprend la plupart du temps des commentaires de la part du professeur qui « *incluent des indications sur l'amélioration de la production autant que sur les*

⁷ von Békésy, G. (1971); *The Ear. In The biological Bases of Behaviour*. London: Open University Press, 93-105

problèmes qui nécessitent une correction [...] ces commentaires [...] constituent un feedback dans le meilleur sens du terme »⁸ (Nair, 1999, p. 66) [78].

Donald G. Miller souligne pour sa part que les termes utilisés lors des leçons de chant manquent d'objectivité, et que l'expression des impressions subjectives données par les perceptions du chant (proprioceptives et auditives) constitue un défi pour le langage (Miller, 2008, p. 7) [75]. Selon Graham F. Welch et son équipe, citant une étude de Callaghan⁹ en 1997, les termes des professeurs de chant ne correspondent pas à une réalité physiologique ou acoustique, et cela peut conduire à de mauvaises informations durant une leçon de chant (in Welch *et al.* 2005, p. 227) [110].

Imitation

Le thérapeute ou le professeur de chant peut également renvoyer à la personne qu'il accompagne une image globale de son comportement vocal et du son qui en résulte en imitant ce que la personne vient de faire, éventuellement en caricaturant pour que les traits dont il veut faire prendre conscience soient facilement identifiables (Le Huche & Allali, 1989, p. 151) [59].

Suppléer les feedbacks auditifs, proprioceptifs et verbaux

On note avec Robert W. Bastian et Matthew J. Nagorsky qu' « *il est très difficile de modifier la production vocale quand les outils traditionnels utilisés pour produire des changements dans le comportement vocal, les feedbacks auditif et somesthésique, ne sont pas bien développés chez un patient donné* »¹⁰ (Bastian & Nagorsky, 1987, p. 1346) [6]. Effectivement : « *Parfois, il est impossible d'utiliser les systèmes sensoriels disponibles, il est alors nécessaire de faire appel à un dispositif de rétro-information externe capable de saisir le phénomène à analyser* » (in Zonder, 2002, p. 20) [119]. Garyth Nair théorise la pédagogie du chant en soulignant les limites de la méthode traditionnelle de modèle et d'imitation. Pour lui, « *les chanteurs rencontrent deux entraves majeures [...] : pour la plus grande partie, il n'y a pas d'entrée visuelle pour aider le processus d'apprentissage. La plupart des comportements cibles pour la musculature du chant sont internes [...] la perception interne des chanteurs de leur propre son est manifestement différente de la réalité perçue par les auditeurs* »¹¹ (Nair, 1999, pp. 13–14) [78].

Dans le modèle pédagogique standard des cours de chant, « *le professeur s'efforce de traduire la performance de l'élève, par exemple en transformant le geste musical de l'élève en langage, et l'élève s'efforce alors de traduire le feedback verbal et visuel de son professeur dans une performance*

⁸ “*such comments include indications about improved production as well as problems in need of correction. [...] these statements give students an external indication of the status of their progress, and, in the best sense of the word, constitute feedback*”

⁹ Callaghan, J. (1997) *The relationship between scientific understandings of voice and current practice in the teaching of singing in Australia*. Unpublished PhD thesis, University of Western Sydney, Nepean.

¹⁰ “*modifying vocal production can be very difficult when the traditional tools used to make changes in vocal behavior, auditory and somatosensory feedbacks, are not well developed in a given patient*”

¹¹ “*singers encounter two major impediments [...] : for the most part, there is no visual input to help the learning process. Most of the target behaviors for the singing musculature are internal [...] singers (internal perception of their own sound is markedly different from the external reality experienced by listeners)*”

musicale plus adaptée. Cela conduit à la possibilité d'une double mauvaise interprétation de l'information »¹² (Howard *et al.*, 2007, p. 21 ; Welch *et al.*, 2005, p. 227) [50] [110].

Selon Verdolini et Krebs, l'analyse spectrale permet de régler le problème du feedback auditif, qui est mauvais lorsque l'élève s'écoute lui-même directement, et qui n'est pas suffisamment stable lorsque c'est un professeur qui le donne. L'analyse spectrale, elle, est fiable, reproductible, et extrêmement rapide (Verdolini & Krebs, 1999, pp. 235–236) [107].

Selon Isabelle Marié-Bailly, « *l'écoute des harmoniques du timbre de la voix permet l'optimisation de la perception sensorielle et de l'émission du geste vocal, par une activation de la boucle audiophonatoire. Celle-ci est renforcée par la visualisation du spectre harmonique vocal sur le sonagramme proposé comme outil pédagogique et de rééducation de la voix parlée et chantée* » (Marié-Bailly, 2004) [66].

Particularité de l'apprentissage vocal en tant que comportement moteur

Bien que la production vocale soit une forme de comportement moteur, Kimberly Steinhauer et Judith P. Grayhack posent la question de la particularité du feedback (entraînement avec Knowledge of Result) pour la production vocale (Steinhauer & Grayhack, 2000, p. 138) [102]. Citant l'étude de Ferrand, les auteurs expliquent que celui-ci n'a pas obtenu d'effet d'apprentissage sur le shimmer et le jitter d'un [a] tenu contrairement à ce qui est attendu pour des tâches motrices classiques (Ferrand, 1995, p. 421) [31].

Dans leur étude sur l'adhésion au traitement en rééducation vocale, Eva van Leer et ses collègues notent que dans l'apprentissage de la technique vocale, le feedback est un point souvent mis en avant par les patients. Selon certains participants, la production vocale est plus difficile à juger que d'autres tâches motrices comme apprendre à faire du vélo ou faire du jogging par exemple. Cela demande plus de concentration. De plus, les tâches vocales demandées semblent représenter un défi sensori-moteur unique, et les objectifs demandés dans les exercices vocaux sont difficiles à évaluer par rapport à d'autres programmes de santé (comme la réduction du nombre de cigarettes fumées, le poids perdu ou gagné, etc.). Les participants indiquent une perte de motivation dans le cas où ils ne savent apprécier leur prestation. Les patients ont besoin d'un retour pour faire leurs exercices vocaux (Van Leer & Connor, 2010, pp. 463, 467) [106].

1.1.3.4 Feedback acoustique en rééducation vocale

Bastian et Nagorsky citent diverses études montrant que des outils électroniques de feedback sont d'un grand intérêt : un affichage de l'intensité vocale pour les patients parkinsoniens¹³, pour le traitement des nodules vocaux, polypes, ulcères¹⁴ (in Bastian & Nagorsky, 1987, p. 1346) [6].

On trouve également de nombreux dispositifs de visualisation de la fréquence fondamentale, notamment le logiciel Sing & See développé par Callaghan et son équipe¹⁵ (Wilson *et al.*,

¹² “The teacher is engaged in a psychological translation of the student's performance, for example, by turning musical gestures into language, and the student is engaged in a further translation of the teacher's verbal and visual feedback into adapted singing performance. A dual possibility thereby exists for the misinterpretation of information”

¹³ Rubow, R., Swift, A. : A microcomputer-based waearable biofeedback device to improve transfer of treatment in parkinsonian dysarthrie. *Journal of Speech and Hearing Disorders*, 50 :178-185, 1985

¹⁴ Holbrook, A., Rolnick, M.I., Bailey, C.Q. : Treatment of vocal abuse disorders using a vocal intensity controller. *Journal of Speech and Hearing Disorders*, 39 :298-303, 1974

2007) [111]. Les différents modules du SpeechViewer offrent, dans les années 1990 une façon de jouer en temps réel avec les différents paramètres de sa voix : l'amplitude, le voisement, la fréquence fondamentale, les formants vocaliques (Peloux, 1993, pp. 15–18) [81].

On peut citer d'autres logiciels de rééducation comme Dr Speech, Diana, PCLX, la suite LingWaves (Allemagne), le logiciel PreLingua (Rodríguez *et al.*, 2012) [95], Vocalab, qui proposent différentes sortes de visualisations et d'actions à partir des caractéristiques acoustiques de la production vocale. Plusieurs de ces logiciels sont destinés à des enfants présentant une surdité.

1.1.3.5 Conclusion

Toute production vocale est contrôlée par un système d'autorégulation nécessitant des feedbacks internes (proprioceptifs et audition par voie osseuse) comme externes (audition par voie aérienne). Ces feedbacks ne renvoient pas au locuteur une image exacte du son perçu par les auditeurs, et parfois ne sont pas fiables (par exemple en cas de douleur laryngée ou de surdité). Les feedbacks habituels en rééducation vocale (feedback verbal, imitation) peuvent être utilement complétés par un outil rationalisant le son externe perçu. Contrairement aux autres outils présentés ci-dessus, le spectrogramme fournit une image plus complète de ces données acoustiques.

1.2 Le spectrogramme

Après avoir expliqué ce qu'affiche un spectrogramme et donné des clés de lecture, nous nous intéresserons aux utilisations du spectrogramme en tant que feedback visuel, telles qu'elles sont décrites dans la littérature.

1.2.1 Présentation du spectrogramme

Une brève définition de l'image spectrographique en termes d'acoustique sera suivie de clés de lectures du spectrogramme pour la parole, le timbre et l'attaque.

1.2.1.1 Notions d'acoustique

Un son pur est décrit mathématiquement par une fonction sinusoïdale. C'est une fonction périodique, c'est-à-dire qui se répète dans le temps, caractérisée par son amplitude et sa fréquence. Selon l'analyse harmonique de Fourier, connue depuis 1822, tout son peut être décomposé en une somme de sons purs. « *Un phénomène périodique quelconque [se décompose] en une somme de sinusoides élémentaires (harmoniques), dont les fréquences respectives sont des multiples entiers de la composante la plus grave appelée fondamentale. [...] Les composantes qui ne font pas partie de la série harmonique s'appellent des partiels* » (Leipp, 2010, p. 25) [62].

Donald G. Miller rappelle que la série harmonique correspond pour une fréquence fondamentale à des intervalles musicaux bien définis (au moins pour les premières) : la première harmonique correspond à l'octave de la fréquence fondamentale, la deuxième à la quinte au-dessus, etc. (Miller, 2008, p. 7) [75].

Les spectrogrammes à bande étroite (45 Hz par exemple) permettent de bien visualiser les harmoniques, mais masquent certains phénomènes très courts dans le temps, tandis que les spectrogrammes à bande large (300 Hz par exemple) reflètent mieux les phénomènes de la parole

¹⁵ Callaghan, J., Thorpe, W., van Doorn, J. (2004). The science of singing and seeing. In *Conference on Interdisciplinary Musicology (CIM04)*, Graz, Austria.

(Howard, 2002, p. 498; Martin, 2008, p. 69) [47] [68]. Dans notre étude, nous nous intéresserons principalement au spectrogramme en bande étroite, qui est mieux adapté pour l'étude de la voix, tandis que le spectrogramme en bande large est mieux adapté pour l'étude de la parole.

1.2.1.2 Définition

« Un spectrogramme réalise une représentation en trois dimensions : le temps sur l'axe horizontal, la fréquence des composantes harmoniques [...] sur l'axe vertical et l'intensité des différentes composantes sur un axe perpendiculaire à la feuille, qui est encodée par le niveau de noir [ou des codages couleur] » (Martin, 2008, p. 88) [68]. Le spectre est en deux dimensions (fréquence / intensité). Le spectrogramme, en ajoutant la variable temps, est un graphique en 3 dimensions (Nair, 1999, p. 44) [78].

1.2.1.3 Lecture de spectrogramme : la parole

Les caractéristiques de chacun des modes de production de la parole sont visibles sur le spectrogramme et permettent de repérer les phonèmes de la parole sur une image spectrographique, et ceci plus facilement si le spectrogramme est réalisé en bande large (Leipp, 2010, p. 280; Martin, 2008, pp. 53–62) [62] [68]. (Voir Tableau I)

Phonèmes	Phénomènes physiques	Non voisé	Voisé : vibration des cordes vocales
Voyelles , liquides « l m n r » et nasales « ng et ng »	Vibration des cordes vocales et diverses configurations du tractus vocal		spectre de raies horizontales correspondant aux harmoniques
Fricatives	Constriction dans le conduit vocal	Nuage correspondant au bruit de friction	
Explosives = occlusives	Fermeture temporaire du conduit vocal	Trait vertical correspondant au bruit d'explosion	

Tableau I : lecture des phonèmes sur le spectrogramme

D'autres modes de production sont possibles, comme les **clics**, mais non utilisés en Français. Les **modes mixtes** concernent les consonnes voisées où les cordes vocales vibrent alors que la langue produit une constriction dans le conduit vocal (fricatives voisées) ou qu'elle produit une explosion (occlusives voisées) : dans ce cas on observe sur le spectrogramme à la fois les composantes caractéristiques de la consonne non voisée correspondante et un spectre de raies (spectrogramme à bande étroite) ou une barre de voisement (spectrogramme à bande large) résultant du son émis par la vibration des cordes vocales.

Les différentes consonnes peuvent être lues (dans une certaine mesure) sur le spectrogramme car le point d'articulation d'une consonne correspond à une distribution de l'amplitude des composantes du bruit (Martin, 2008, p. 59) [68]. De même, les formants vocaliques dépendent de la position des articulateurs et permettent d'identifier une voyelle (voir timbre vocalique).

1.2.1.4 Lecture de spectrogramme : le timbre

Définition

« Le timbre est la qualité permettant de distinguer un son parmi d'autres qui ont la même intensité et la même hauteur » (Leipp, 2010, p. 149) [62]. Il est constitué du timbre vocalique qui détermine

(lorsque c'est pertinent) l'identité de la voyelle entendue, et le timbre extra-vocalique, correspondant à la couleur de la voix.

Timbre vocalique

Lorsqu'un locuteur émet une voyelle, il met ses plis vocaux en vibration. Le son source (débit d'air pulsé au niveau glottique) est modifié par le filtre constitué par le tractus vocal, qui atténue certaines fréquences et en privilégie d'autres, selon la théorie de la résonance, de façon à renforcer les formants constitutifs de la voyelle (Leipp, 2010; Miller, 2008, pp. 23–33; Vaissière, 2011, p. 61) [62] [75] [105]. « *La position des articulateurs impacte fortement les fréquences et les largeurs de bande des résonances acoustiques du conduit vocal* » (Henrich, 2012, pp. 18–19). « *L'élégante correspondance entre la représentation des voyelles par leurs deux premiers formants et le triangle vocalique est cependant trompeuse : elle néglige l'effet décisif des lèvres, et des formants supérieurs à F2¹⁶, sur le timbre des voyelles antérieures* » (Vaissière, 2011, p. 75) [105].

Les formants sont les barres horizontales plus intenses visibles sur le spectrogramme. Elles correspondent aux bandes de fréquences mises en valeur par le conduit vocal, changeant selon la disposition des articulateurs et définissant une voyelle précise.

Donald G. Miller (2008, p. 30) [75] propose la visualisation des deux premiers formants avec le spectrogramme. Scott McCoy consacre également un chapitre de son livre (destiné aux chanteurs) aux règles de bases de phonétique et de formants vocaliques (McCoy, 2004, pp. 38–52; Welch et al., 2005, p. 232) [70] [110].

Les harmoniques

Lorsque le son émis est périodique, c'est-à-dire qu'il est émis sur une certaine hauteur tonale, ces différents partiels sont des multiples de sa fréquence fondamentale. Dans ce cas « *la sensation de timbre dépend [...] du nombre de composantes, de leur situation dans l'air audible, de leur « consonance » avec le fondamental et de leurs intensités relatives* » (Leipp, 2010, pp. 150–154) [62]. Sur le spectrogramme à bandes étroites, les harmoniques d'un son tenu sont représentées par des traits horizontaux équidistants.

Le bruit entre les harmoniques

Un son peut avoir une composante harmonique (périodique) à laquelle s'ajoute du bruit, apériodique. Le spectrogramme représente ce bruit suivant les fréquences auxquelles il apparaît, comme un nuage entre les harmoniques.

Dans leur étude, Wolfe et Steinfatt (1987, p. 238) [114] concluent que le bruit spectrographique (évalué visuellement sur un spectrogramme) est le meilleur paramètre dans le cadre de leur étude pour rendre compte de la sévérité d'une dysphonie.

La Société Européenne de Laryngologie note que la fiabilité des mesures acoustiques chiffrées (comme le calcul de la fréquence fondamentale par exemple) baisse lorsque le signal vocal est apériodique, contient des irrégularités, des sous-harmoniques (« *subharmonics* ») ou des modulations.

¹⁶ F2 = le deuxième formant

Dans ce cas, Dejonckere et ses collègues conseillent de contrôler visuellement le signal sur un spectrogramme (Dejonckere *et al.*, 2001, p. 80) [23].

La forme du spectrogramme peut révéler diverses impressions subjectives de la voix. Par exemple, Jan W. Martens et son équipe précisent que le grade G du GRBAS correspond à des distorsions dans le spectrogramme. Les sous-harmoniques sont visibles également sur un spectrogramme, et sont caractéristiques d'une certaine raucité (R). Enfin, le souffle se présente sur le spectrogramme comme du bruit qui peut apparaître dans les hautes fréquences, entre les harmoniques ou à la place des harmoniques. Ces particularités sont visibles sur un spectrogramme à bande étroite (Dejonckere, 2006, p. 350; Martens *et al.*, 2007, p. 179) [22] [67].

Atsushi Yasui présente en 2004 un article donnant des exemples d'analyse de troubles vocaux par le spectrogramme. (voix asthénique, voix rauque, serrée...) (Yasui, 2004, pp. 40–45) [117].

Classifications

Critère	Objectifs	Cotation	
Richesse du spectre	Déterminer la qualité du timbre en évaluant la richesse en harmoniques : on distinguera les harmoniques situés dans les basses fréquences (< 1500 Hz) et ceux > 1500 Hz dits extra-vocaliques, qui viennent donner toute sa richesse au timbre.	0 : bon	Présence d'harmoniques répartis sur l'ensemble du spectre.
		1 : moyen	Présence d'harmoniques dans les basses fréquences uniquement
		2 : mauvais	Présence rare d'harmoniques même dans les basses fréquences.
Stabilité du spectre	Sur la partie stable du spectre (on ne tient pas compte de l'attaque) on recherche des variations de fréquence (instabilité dans le tracé des harmoniques) ou d'intensité (instabilité de couleur).	0 : bon	Pas ou peu de variations.
		1 : moyen	Variations légères de fréquence ou d'intensité.
		2 mauvais	Variations importantes de fréquence ou d'intensité.
Energie du spectre	Appréhender l'intensité de la F0 et des premiers harmoniques au travers de l'énergie du spectre (représentée par la couleur).	0 : bon	F0 et 1ers harmoniques de bonne intensité
		1 : moyen	F0 de bonne intensité mais 1 ^{ers} harmoniques relativement faible
		2 : mauvais	F0 de faible intensité et 1 ^{ers} harmoniques faible ou absent.
Rapidité de dégradation	Rechercher des signes de dégradation rapide du spectre (perte d'énergie rapide qui serait liée à l'incompétence glottique).	0 : bon	Le spectre se maintient sur toute la durée de l'émission.
		1 : moyen	Le spectre se dégrade légèrement de façon progressive.
		2 : mauvais	Toute l'énergie est concentrée sur l'attaque.
Propreté du spectre	Rechercher la présence de traces erratiques (traces qui se superposent aux harmoniques).	0 : bon	Absence de traces erratiques.
		1 : moyen	Présence de traces erratiques occasionnelles.
		2 : mauvais	Présence de traces erratiques importantes.

Tableau II : grille d'évaluation des qualités vocales à partir de l'image spectrographique (Espanol, 2011) [26]

Yangihara, en 1967, utilisant la spectrographie en bande étroite propose une classification des tracés en 4 types, correspondant au degré de raucité (Yanagihara, 1967, p. 531) [116]. Voir Figure XIII et Tableau XXX en annexe.

En 1974, Cooper et son équipe utilisent la classification de Yanagihara en ajoutant un Type 0 indiquant une voix claire, et un type 0+ indiquant une trace d'érailement, mais moins importante que dans le type I. L'étude du degré d'érailement est faite sur des voyelles tenues uniquement ([a], [i] et [u]) (Cooper, 1974, p. 291) [16]. Voir Tableau XXX en annexe.

En 1984, Yumoto et ses collègues (1984, p. 3) [118] ont utilisé sensiblement la même échelle pour juger de l'érailement à partir de spectrogrammes à bande étroite. Ils y ont également ajouté un type 0 pour une phonation claire. En 1987, Wolfe et Steinfatt (1987, p. 233) [114] réutilisent la classification de Yanagihara, en ajoutant une première catégorie sans composante de bruit. De la même manière, Remacle et ses collègues en 1989 proposent une classification de 0 à 4 qui reprend l'idée de Yanagihara (Remacle & Morsomme, 2009, p. 130) [86].

Un mémoire d'orthophonie (Espanol, 2011) [26] propose des critères de cotation à partir d'une grille de lecture du spectre (voir Tableau II).

Exemples en voix chantée

Singing formant

Le renforcement des harmoniques dans une zone autour des 3000 Hz est caractéristique de la voix du chanteur entraîné (Pillot-Loiseau & Vaissière, 2009, p. 246; Sundberg, 2001, pp. 185–186) [83] [104]. La présence et la richesse du singing formant sont facilement identifiables sur un spectrogramme. Pour mettre en évidence le singing formant en pratique, Schutte et Miller préconisent d'observer un spectre moyenné alors qu'on émet une voyelle sur une quinte montée chromatiquement (Miller, 2008, p. 37; Miller & Schutte, 1999a, p. 202) [75] [72].

Voix soufflée

Garyth Nair (1999, p. 105) [78] donne cet exemple pour la voix chantée. A partir d'un son d'une bonne qualité, l'ajout d'un souffle réduit le nombre d'harmoniques.

Voix serrée

Garyth Nair (1999, p. 105) [78] donne un exemple en voix chantée. Partant d'un son non serré, il « écrase » le son, et cela est particulièrement bien visible sur le spectrogramme. Robert A. Volin utilise également la visualisation spectrographique comme reflet de la clarté du timbre, (Volin, 1999, p. 254) [108]. Selon Fisher et ses collègues, en voix chantée, une voix serrée se traduit par des harmoniques plus fortes en-dessous de 500 Hz, sans augmentation de la partie supérieure (Fisher *et al.*, 2003, p. 10) [32].

1.2.1.5 Lecture de spectrogramme : l'attaque

L'attaque dure se traduit sur le spectrogramme par « une zone verticale sombre pendant les premières millisecondes de la phonation. Cette bande sombre représente les vibrations chaotiques des plis

vocaux qui ont lieu jusqu'à ce que le larynx puisse prendre le mode vibratoire désiré »¹⁷ (Nair, 1999, p. 111) [78]. L'attaque dure est plus forte et plus longue (plus de 160 ms.). L'explosion résultant de la fermeture vocale génère une instabilité de la voyelle (Fisher et al., 2003, p. 8) [32].

L'attaque soufflée ("*aspirate onset*") se traduit sur le spectrogramme par une forme de bruit, pendant tout le temps que la glotte met pour se réduire afin de pouvoir commencer le son (Nair, 1999, p. 112) [78]. Pour ce qui concerne l'attaque "*floue*", elle se traduit par une accumulation lente d'harmoniques sur le spectrogramme, avec un vibrato qui ne s'installe pas tout de suite.

Pour Garyth Nair (1999, p. 109) [78], le début de son équilibré est appelé "*ASAI : articulators-set/Abdominally Initiated*", et il est caractérisé sur le spectrogramme par un départ simultané de toutes les harmoniques du son. Le vibrato peut être présent dès le début, montrant que les muscles laryngés et pharyngés ne sont pas tendus pendant l'attaque.

1.2.1.6 Conclusion

Le spectrogramme est un outil intéressant pour voir la voix. Il en représente de nombreux paramètres, et notamment le timbre et l'attaque.

1.2.2 Utilisation du spectrogramme en temps réel

Il existe de nombreux domaines d'application du spectrogramme car il permet de comprendre quels sont les paramètres acoustiques qui sont importants dans la production et la perception humaine (Fisher et al., 2003, p. 2; Howard, 2005, p. 329) [32] [49].

1.2.2.1 Formation vocale des chanteurs

Dans la droite ligne de Richard Miller et Sataloff, Garyth Nair se fixe comme objectif de démythifier le langage de la pédagogie vocale et de promouvoir l'utilisation d'outils informatisés pour servir de feedback permettant d'accélérer le développement vocal des chanteurs (Nair, 1999, p. 12) [78]. Pour lui, le spectrogramme en temps réel constitue à la fois « *un moyen d'accélérer le processus d'apprentissage du chant et un moyen pour aider les professeurs et les élèves à mieux comprendre la science de la voix.* »¹⁸. Scott McCoy (2004) [70] et Donald G. Miller (2008) [75] publient également des livres donnant aux professeurs de chant les compétences pour utiliser le spectrogramme. Gillyanne Kayes et Jeremy Fisher¹⁹ présentent par exemple un guide de lecture des spectrogrammes pour les professionnels de la voix dans leur présentation au PEVoC (Pan European Voice Conference) en 2003. Cette visualisation, précisant les différentes qualités de la voix, est également proposée dans la méthode Jo Estill (Estill Voice Training), présentée par Kimberly M. Steinhauer et Cari M. Tellis²⁰.

L'expérience de David M. Howard et ses collègues à Londres et York est particulièrement intéressante. Elle a utilisé une méthodologie d'« *action research* », mettant en position de chercheurs deux professeurs de chant et quatre élèves utilisant le spectrogramme en temps réel. Le projet VOXed, réalisé durant l'année scolaire 2003-2004 (Welch et al., 2005, pp. 232–233) [110], démontre qu'un

¹⁷ “[a] dark vertical area seen in the first milliseconds of phonation. This dark band is representative of chaotic vocal fold vibration that continues until the larynx can settle into the desired vibrational mode”

¹⁸ “a means of speeding up the process of voice training and [...] a means of helping both teachers and students achieve a greater understanding of voice science”

¹⁹ http://www.vocalprocess.co.uk/training/training_computervoice.htm consulté le 25 janvier 2013

²⁰ http://www.estillvoice.com/files/file/courses/Estill%20Voice%20Training%20Flyer_2012_MU.pdf consulté le 25 janvier 2013

outil de feedback en temps réel facile d'accès, valide, robuste et utile en cours de chant est disponible. L'étude s'intéresse également à l'impact de cet outil sur la pédagogie vocale (Howard et al., 2007, p. 22) [50].

En France, Michel Hart écrit pour le Journal de l'AFPC deux articles décrivant des utilisations possibles du spectrogramme en cours de chant (Hart, 2002a, 2002b) [43] [44]. Il cite pour l'ARIAM (Association Régionale d'Information et d'Actions Musicales) en 2011²¹ de nombreux professeurs de chant utilisant les nouvelles technologies. On y trouve notamment, en France, Allan Wright (the Vocalist, The Modern Vocalist), professeur utilisant régulièrement le spectrogramme.

En pratique, Nair et ses collègues donnent des indications sur la manière d'introduire le spectrogramme en leçon de chant (Nair, 1999, pp. 83–91) [78]. Le spectrogramme permet de travailler le timbre (Nair, 1999, p. 181) [78], l'attaque (Nair, 1999, pp. 183–186) [78], la respiration (Nair, 1999, p. 181) [78], la justesse (Nair, 1999, pp. 129–132) [78], la précision des phonèmes (Nair, 1999, pp. 127–128, 155–156) [78], la zone de passage (Miller & Schutte, 1999a, pp. 195–201; Nair, 1999, p. 170) [72] [78], le vibrato (Nair, 1999, p. 177) [78], le singing formant.

Une étude d'Anna-Maria Laukkanen et son équipe se propose d'étudier l'efficacité du biofeedback spectral (la visualisation du spectre en temps réel) par rapport à un entraînement conventionnel. Elle conclut que le feedback visuel semble améliorer l'efficacité de l'entraînement vocal (Laukkanen *et al.*, 2004, p. 66) [58].

1.2.2.2 Langue étrangères et accent, prosodie

Anna-Maria Laukkanen et ses collègues notent que « *le spectrographe et les analyseurs de spectres [...] ont été principalement utilisés dans l'entraînement de l'articulation* » (Laukkanen *et al.*, 2004, p. 67) [58]. En France, la possibilité de travail de la prononciation par un système de spectrogramme en temps réel est également notée par Cazade. (2000, p. 20) [13].

A titre d'exemple, l'article de Pamela Pearson et ses collègues décrit l'utilisation du spectrogramme dans un protocole d'apprentissage par des Vietnamiens des débuts et fin de syllabes en anglais. Ils observent une amélioration de la prononciation ainsi que des réactions positives à l'utilisation de cette technique, autant de la part des professeurs que de la part des élèves. Lors de 6 séances de 30 minutes, les élèves étaient entraînés à reconnaître les composantes d'un spectrogramme d'un mot monosyllabique en Anglais, puis devaient reproduire au mieux ce mot à l'aide d'un spectrogramme en temps réel (Pearson *et al.*, 2011, p. 172) [80].

Une utilisation des modules graphiques du logiciel SpeechViewer II pour la rééducation de la prosodie est suggérée dans le mémoire d'Anne Bouquay et Pascale Beauchene : « *les modules graphiques, enveloppe et fréquence d'une part, signal et spectrogramme d'autre part, permettent [...] d'apprécier les variations de fréquence et d'amplitude* » (Beauchene & Bouquay, 1994, p. 114) [7]. Garyth Nair explicite également comment on peut se servir du spectrogramme pour produire les voyelles non accentuées (Nair, 1999, pp. 119–122) [78].

²¹ <http://www.ariam-idf.com/pdf/MichelHart-Ariam2011.pdf> consulté le 25 janvier 2013

1.2.2.3 Parole dans le cas de surdités

Pour Garyth Nair, la visualisation des voyelles sur le spectrogramme est une aide précieuse pour l'apprentissage de la production des voyelles, parce que leurs formants sont clairement visibles sur le spectre ou sur le spectrogramme. Cependant, dans le cas de surdités, c'est le spectre en temps réel (visualisation des intensités des différentes fréquences) qui est le plus efficace, et non le spectrogramme (visualisation des intensités des différentes fréquences, en fonction du temps). Garyth Nair reprend l'idée de Donald G. Miller en recommandant la production des formants à partir du registre fry afin de visualiser plus nettement les formants vocaliques. (Miller & Schutte, 1999a, pp. 194, 208; Nair, 1999, pp. 98–99) [72] [78]. La même chose est possible pour les consonnes, notamment pour les consonnes fricatives, par la visualisation d'un modèle et la reproduction des hauteurs des composantes spectrales du bruit (Nair, 1999, pp. 164–165) [78].

Pour M. Maulet, « *la visualisation immédiate de son articulation offre au patient des points de repères très efficaces qui vont lui permettre d'améliorer rapidement ses réalisations* » (Maulet, 1997, p. 84) [69]. En essayant de reproduire un spectre donné dans le SpeechViewer, les sujets travaillent la précision articulatoire (Peloux, 1993, p. 18) [81].

1.2.2.4 Dysarthrie

Robert A. Volin fait un parallèle entre l'apprentissage du chant et la rééducation de la dysarthrie. Selon lui, le fait que les fonctions de la parole automatisées interfèrent avec l'acquisition de nouveaux modèles et le fait que les patients ne perçoivent pas leur propre voix comme les autres l'entendent à l'extérieur donnent à la rééducation les mêmes défis que l'apprentissage du chant (Volin, 1999, p. 242) [108].

Volin décrit l'utilisation d'un logiciel d'analyse spectrale en temps réel pour améliorer la précision de l'articulation (Volin, 1999, pp. 246–250) [108] : le spectrogramme peut également être utilisé pour le contrôle de la longueur d'une phrase (Volin, 1999, p. 250) [108]. Ce même auteur suggère également qu'on peut utiliser cet outil pour demander un contrôle de l'intensité vocale, un contrôle prosodique de la fréquence fondamentale (Volin, 1999, pp. 250–253) [108].

Comme le décrivent Anne Bouquay et Pascale Beauchene dans leur mémoire d'orthophonie en 1994, les modules graphiques du logiciel SpeechViewer II peuvent être utilisés dans le cadre de rééducation de dysarthries post-traumatiques. Dans leur travail, elles ont utilisé ces modules avec pour objectif de réduire le temps de production d'une phrase (coordination pneumophonique et débit). Le module signal et spectrogramme « *permet [...] de montrer au patient à quel endroit de la phrase il doit reprendre son souffle ou au contraire accélérer son débit* » (Beauchene & Bouquay, 1994, p. 90) [7] « *Nous pouvons donc montrer au sujet de manière objective ce qu'il doit s'efforcer de modifier* », notamment en terme de prolongement de phonème ou de pause inappropriée.

1.2.2.5 Rééducation vocale

Verdolini et Krebs affirment que « *l'habitude de l'efficacité de l'utilisation des outils de biofeedback dans les rééducations physiques fournit un précédent pour l'utilisation d'outils similaires, comme l'analyse spectrale, dans la pédagogie vocale* »²² (Verdolini & Krebs, 1999, p. 239) [107].

En 1993, Valérie Peloux décrit une utilisation des modules du SpeechViewer et notamment de son module graphique « spectre » dans le cadre de l'évaluation et de la rééducation vocale (Peloux, 1993, pp. 45–47) [81]. L'utilisation du module spectre reste cependant anecdotique. Caroline Coudière reprend les conclusions de C. Routier et A.S. Vanleys²³ : « *suite à un questionnaire soumis à 175 orthophonistes en libéral de la région Nord-Pas de Calais, C. Routier concluait que l'analyse spectrale était une aide appréciable lors du bilan vocal, pour distinguer les problèmes de hauteur et de placement vocal des troubles du timbre* » (Coudière, 2003, p. 125) [19]. Cependant elle ajoute que « *la visualisation spectrale serait [...] plus importante en rééducation, où elle peut servir de support et de moyen de contrôle de la production vocale, que l'analyse précise des différents paramètres vocaux, utile en bilan* » (Coudière, 2003, p. 156) [19] :

Le mémoire de Caroline Coudière liste de nombreux domaines où la visualisation spectrale peut permettre un feedback intéressant : (Coudière, 2003, p. 82) [19]

- timbre (visualisation de l'enrichissement harmonique, investissement des cavités de résonance..),
- respiration (notion de durée, tenue du soutien, sensibilisation aux pauses inspiratoires) et attaque (coups de glotte),
- hauteur (éducation de l'oreille, sensibilisation à la prosodie, maîtrise et stabilité de la hauteur de la voix),
- intensité (éducation de l'oreille, levée des inhibitions, maîtrise de la dynamique vocale et de la stabilité en intensité).

De même, Volin utilise l'image spectrographique pour travailler sur la clarté du timbre (Volin, 1999, p. 254) [108]. Il propose également divers exercices avec le spectrogramme pour réduire l'hyperfonction vocale et l'attaque en coup de glotte (Volin, 1999, pp. 254–256) [108].

1.2.2.6 Conclusion

Les auteurs décrivent de nombreuses applications du spectrogramme. Voyons en détail quels avantages et quelles limites ces auteurs identifient pour une utilisation en rééducation vocale.

1.3 Spectrogramme et rééducation vocale : avantages et limites

Nous reprenons dans cette partie les avantages et limites cités pour la rééducation vocale, mais également pour d'autres situations où le spectrogramme est utilisé, lorsque les remarques faites nous semblent facilement transposables au contexte d'une rééducation vocale.

²² "The routine successful use of biofeedback tools in physical therapy provides precedent for the use of similar tools such as the spectral analyzer in voice pedagogy"

²³ Routier, C., Vanleys, A.S. Analyse objective des troubles de la voix chez l'adulte, apport au niveau de la rééducation orthophonique, Mémoire d'orthophonie de l'université de Lille, 1993

1.3.1 Avantages

Dans l'étude menée par D. M. Howard et ses collègues, les professeurs de chant comme les élèves ont été très enthousiastes ; ils ont même été surpris de découvrir tout le potentiel de l'outil technologique (Howard *et al.*, 2004, p. 142) [48]. Plusieurs études soulignent effectivement les aspects positifs de cet outil, dans les divers domaines où il est utilisé.

Caroline Coudière conclut son mémoire en disant que « *le spectre est un outil particulièrement intéressant dans la prise en charge de la voix. En effet, il apporte des données objectives et une visualisation de l'évolution de la production vocale qui, malgré des difficultés d'interprétation, servent l'orthophoniste et son patient* » (Coudière, 2003, p. 155) [19].

1.3.1.1 Prise de conscience

Importance de la prise de conscience en rééducation vocale

L'auto-contrôle de la voix passe par une prise de conscience : selon David M. Howard et son équipe, « *la thérapie a notamment pour but d'augmenter la conscience du locuteur de sa voix à la fois de façon proprioceptive et auditive, afin de permettre l'auto-contrôle* »²⁴ (Howard *et al.*, 2012, p. 619) [51].

Jocelyne Sarfati indique qu'une ou plusieurs séances peuvent être consacrées à l'analyse d'un enregistrement que le patient réalise de préférence en situation professionnelle mettant en évidence ses difficultés vocales. Ce travail est décisif dans le sens où il permet une analyse du comportement vocal et une modification de l' « *image sonore* » qu'a le patient de lui-même (Sarfati, 1998, pp. 53–54) [99]. De même, les grilles détaillées d'auto-analyse de la voix proposées au début d'une rééducation ont notamment pour finalité de rendre la personne consciente de son geste (Estienne, 1998, p. 92) [27]. « *Analyser son geste vocal* » en détail (Amy de la Bretèque, 2012, p. 132) [3] et « *prendre contact* » avec ses sensations proprioceptives en les « *traduisant avec finesse* » (Roch, 2001, p. 61) [94], sont considérés par certains auteurs comme les premiers objectifs de la rééducation.

Apport du spectrogramme

L'analyse des réponses aux questionnaires de Caroline Coudière (2003, p. 146) [19] permet de dire que « *pour 69% des orthophonistes [interrogées], [la visualisation spectrographique] permet [pour le patient] une meilleure prise de conscience de ses progrès. Trois orthophonistes ajoutent que cette meilleure prise de conscience intéresse aussi les paramètres vocaux, leur mécanisme et leur maîtrise, notamment en ce qui concerne la hauteur et ses modulations* »

Garyth Nair cite le spectrogramme comme un outil idéal pour la prise de conscience du chanteur de certains problèmes de justesse pour les attaques sur des consonnes voisées ou de manque de soutien en fin de phrase (Nair, 1999, pp. 159, 161) [78].

Cependant, lorsque Caroline Coudière cite les limites de l'utilisation des outils informatiques en orthophonie, elle précise qu'« *il faut tenir compte des réactions du patient. Certains se sentent véritablement mal à l'aise avec cet outil, et peuvent le rejeter, ou craindre de se retrouver trop brutalement confronté à l'objectivation de leurs difficultés* » (Coudière, 2003, p. 62) [19].

²⁴ « *a major goal in therapy is to increase the speaker's awareness of his/her voice both proprioceptively and auditorily to enable self-monitoring* »

Cazade considère la visualisation par le spectrogramme comme une mise à distance nécessaire : « *Il faut reconnaître que l'affichage de courbes sonores permet de présenter de façon visible et non subjective (avec la supériorité d'impact que le visuel a sur le sonore pour beaucoup) une transcription du problème que l'apprenant rencontre dans son travail d'approche de la production orale, processus mal maîtrisé dans notre société [...] Cet affichage lui offre la possibilité, d'une certaine manière, de transformer la production orale en un objet-symbole permettant à sa sensibilité autant sensorielle qu'intellectuelle de commencer à avoir prise sur lui : le premier pas pour pouvoir envisager de progresser* » (Cazade, 2000, p. 14) [13].

1.3.1.2 Motivation

Dans le cadre d'une amélioration de la prononciation de l'anglais, Pamela Pearson et son équipe notent que le feedback visuel a un effet sur la motivation des sujets : le spectrogramme permet aux élèves de constater l'écart entre leur production et ce qui est attendu. Les professeurs apprécient de ne pas avoir à donner de feedback verbal négatif mais une cible visuelle positive (Pearson *et al.*, 2011, p. 174) [80].

Selon Caroline Coudière, « *le feedback visuel est un média non négligeable, malgré ses limites, dans le travail sur la voix et ses paramètres. Il permet une prise de conscience qui ne peut se faire sur une simple écoute, une objectivation de l'évolution, une visualisation du but à atteindre, et ainsi un soutien à la motivation.* » (Coudière, 2003, p. 113) [19]. « *65% des orthophonistes notent un regain de motivation chez certains de leurs patients, lorsque le logiciel est utilisé en milieu ou fin de rééducation. Certains observent, non pas un regain, mais une motivation qui transparait au cours de l'utilisation du logiciel* » (Coudière, 2003, p. 144) [19]. « *La moitié des orthophonistes interrogés pensent que cet outil agit[positivement] sur la motivation du patient* » (Coudière, 2003, p. 147) [19].

Dans le cadre de la rééducation de la dysarthrie post-traumatique, « *le fait d'avoir un objectif visuel (augmenter son débit pour que tout l'échantillon de parole s'inscrive dans un cadre) est motivant pour le patient.* » (Beauchene *et al.*, 1994, p. 111) [7].

Dans le cadre de la rééducation d'enfants sourds, M. Maulet précise que « *toutes les possibilités d'analyse décrites précédemment, loin d'être fastidieuses pour le patient, lui procurent un plaisir nouveau, une motivation réelle dans son appropriation lente et difficile de la parole* » (Maulet, 1997, p. 86) [69].

L'étude de Laukkanen a pour objectif d'améliorer le singing formant chez des acteurs en utilisant un feedback en temps réel constitué de l'analyse spectrale. Les auteurs notent que les sujets ont trouvé que « *s'exercer avec le feedback visuel était intéressant et que sa nature concrète était motivante* »²⁵. A l'inverse, un sujet du groupe contrôle, dont la qualité vocale a empiré pendant l'expérience, « *s'est plaint qu'il n'avait aucune idée claire du but des exercices* »²⁶, sa sélectivité auditive semblant insuffisante (Laukkanen *et al.*, 2004, p. 69) [58].

²⁵ "According to the subjects' comments, exercising with visual feedback was interesting and its concrete nature was reported to be motivating"

²⁶ "on the other hand, the subject in the control group whose voice was evaluated to be clearly worse after training, complained that he had not had any clear idea of the goal of the exercises"

1.3.1.3 Autonomisation

Selon Françoise Estienne, l'efficacité d'un traitement sur la voix suppose l'autonomie de la personne traitée. Ses critères permettant d'évaluer l'efficacité d'une thérapie comprennent notamment : « *la personne est capable d'évaluer ces moyens et de les ajuster* », « *la personne vérifie sa propre influence sur les résultats, est consciente de ce qu'elle fait, des stratégies qu'elle emploie* » (Estienne, 2001, p. 39) [28]. L'auteur pense que l'invitation à l'autonomie des patients permet l'intégration du geste vocal plus rapidement lors d'une thérapie vocale. Dans cette optique, le thérapeute doit « *résister au charme du pouvoir, de la toute puissance et de la dépendance ; [...] accepter la critique, le refus comme l'approbation* » (Estienne, 1998, p. 185) [27].

Dans la cadre du biofeedback (électromyographie), les auteurs soulignent que le sujet « *participe activement à sa propre rééducation et passe sur le plan relationnel d'une situation de dépendance et de passivité à une autonomie, à un contrôle, qui constituent en fait le terme souhaité de toute rééducation* » (in Zonder, 2002, p. 22) [119].

Dans ce domaine, Donald G. Miller pense que le spectrogramme donne une certaine autonomie à l'élève : « *il peut arriver que l'élève soit dans une meilleure position que le professeur pour évaluer les progrès techniques, et même pour juger de la désirabilité des interventions du professeur sur la technique vocale* »²⁷ (Miller, 2008, p. 107) [75]. David M. Howard rapporte « *qu'en outre, les étudiants peuvent souvent utiliser la technologie pendant les exercices entre les leçons* »²⁸ (Howard, 2005, p. 329) [49].

1.3.1.4 Enrichissement de compétences, diversification du travail orthophonique

On note que certaines études ont montré que la pédagogie du chant était peu documentée de façon scientifique (in Welch *et al.*, 2005, p. 226) [110]. La vulgarisation des notions d'acoustique pour les professeurs de chant par McCoy et Garyth Nair est saluée dans l'article de Barnes-Burroughs et ses collègues (2008, p. 590) [5]. La visualisation spectrographique donne au professeur de chant l'opportunité d'introduire des notions d'anatomie, de physiologie et de concepts physiques de base (Nair, 1999, pp. 7, 69) [78].

La situation en orthophonie est différente : l'anatomie, la physiologie, les spectrogrammes et la phonétique sont enseignés en école d'orthophonie. L'utilisation d'un spectrogramme en temps réel en rééducation donne aux orthophonistes l'occasion d'enrichir leur pratique.

Caroline Coudière note que « *le spectre [...] va aussi objectiver les progrès réalisés par le patient et pour l'orthophoniste, lui permettre de diversifier son travail et la conforter dans sa compétence* » (Coudière, 2003, p. 83) [19].

1.3.1.5 Conclusion

L'utilisation du spectrogramme en rééducation vocale est décrite comme pouvant favoriser la prise de conscience des paramètres de la voix, la motivation et l'autonomisation du patient. Quelles sont cependant les limites de cette utilisation ?

²⁷ "it can happen that the pupil will be in a better position than the teacher to assess technical progress, and even to judge the desirability of the teacher's interventions in vocal technique"

²⁸ "Systems such as these provide quantitative real-time visual feedback during lessons, and in addition, students are often able to make use of the technology during their intervening practice time"

1.3.2 Limites de l'analyse spectrographique

1.3.2.1 Fiabilité de l'image

Yumoto et ses collègues pensent que le rapport signal sur bruit est plus adapté à la pratique clinique que l'analyse de l'image spectrographique. Parmi les divers inconvénients de la classification spectrographique de Yanagihara, ils notent qu' « *en ajustant certains paramètres du spectrogramme, l'aspect nuageux de l'ombre entre les harmoniques peut être modifié. Au contraire, pour mesurer le rapport signal sur bruit, l'ordinateur traite automatiquement l'onde acoustique de l'extrait phonatoire* »²⁹ (Yumoto *et al.*, 1984, p. 5) [118].

De même, certains auteurs notent que « *beaucoup d'analyseurs de spectre nécessitent des réglages qui vont déterminer la façon dont l'échantillon va être analysé. La forme, la présentation du spectrogramme, l'amplification de certains phénomènes aux dépens d'autres, varient en fonction de ces réglages et donc des choix de l'examineur* » (in Remacle & Morsomme, 2009, p. 129) [86].

Atsushi Yasui indique que le spectrogramme ne permet pas de quantification exacte, et ne peut pas remplacer l'observation directe du larynx (Yasui, 2004, pp. 40–45) [117].

Enfin la fiabilité de l'analyse dépend également d'autres paramètres, comme par exemple la distance du microphone à la bouche et son angle (Verdolini & Krebs, 1999, pp. 236–238) [107].

1.3.2.2 Complémentarité des approches classiques et du feedback visuel

« *Le biofeedback n'est pas une panacée. [...] Ce qui est réellement une panacée, c'est le pouvoir inscrit dans l'être humain de s'auto-réguler, de s'auto-guérir, de se rééquilibrer constamment. Le biofeedback par lui-même ne fait rien à l'individu, il n'est qu'un instrument qui permet de mettre à profit ce potentiel* » (Rémond *et al.*, 1994, p. 32) [87]. « *Le biofeedback n'a de valeur que s'il est choisi comme traitement en connaissance de cause, et associé sans exclusive à tout autre traitement jugé nécessaire* » (in Zonder 2002, p.21) [119].

Le thérapeute de la voix garde toute sa place dans la rééducation vocale, même si les sensations subjectives de l'oreille humaine ne sont pas faciles à partager. Selon Garyth Nair, l'outil informatique permettant d'analyser finement les composantes de chaque son nécessite de toute façon une **interprétation humaine** pour corrélérer ce qui est affiché à ce qui est entendu et pour décider ce que doit faire l'élève. Il n'est donc pas question de remplacer l'humain par une machine. De même, Donald G. Miller précise que les outils informatiques ne sont pas là pour remplacer le professeur, mais pour lui donner des informations supplémentaires (Miller, 2008, p. 35) [75].

Les résultats objectifs délivrés par un appareil de feedback doivent impérativement être présentés de façon positive par le thérapeute. La motivation dépend du comportement du thérapeute à cet égard, de la qualité de la rétro-information et des résultats obtenus (André *et al.*, 1986, p. 299,302; Rémond *et al.*, 1994, p. 13) [4] [87].

La prudence s'impose dans l'analyse de l'image spectrographique car elle **risque d'être mal interprétée**. D.M. Howard a consacré un article en 2002 sur ce sujet, pointant le fait que le spectrogramme ne peut être précis à la fois dans le domaine temporel et dans le domaine fréquentiel,

²⁹ “*Cloudiness of the shadow between harmonics could be modified to some extent by adjusting controls of the sound spectrograph (e.g., "mark level" and "dynamic range"). In contrast, in order to measure the H/N ratio a computer automatically processes the acoustic wave of the phonatory sample*”

et qu'en conséquence, il ne peut rendre compte de toutes les subtilités du son perçues par une oreille humaine (Howard, 2002, p. 500) [47]. Les professeurs de chant mettent souvent en garde contre l'utilisation du spectrogramme sans supervision, les élèves pouvant ne pas comprendre entièrement ce qu'ils y voient. (Hoppe *et al.*, 2006, p. 312) [46].

« *Si le logiciel permet un travail autonome de l'apprenant, le rôle de l'enseignant n'en est pas [pour autant] exclu. Ce dernier va pouvoir s'appuyer sur les difficultés objectivées par le logiciel pour amener l'apprenant à porter un regard auto-évaluatif sur son discours oral. Le professeur a en charge de préciser et de compléter le feedback rendu par les courbes, qui va lui permettre d'individualiser concrètement son discours* » (Coudière, 2003, p. 110) [19].

La qualité du professeur de chant réside également dans **la relation** entretenue avec l'élève, et la façon d'entretenir judicieusement la confiance en soi de celui-ci. (Miller, 2008, p. 106) [75]. Le professeur de chant apporte également des compétences qui n'ont pas à proprement parler de rapport avec la production acoustique (Howard *et al.*, 2004, p. 143) [48]. On fera facilement le parallèle avec l'orthophoniste dont le rôle est également de conforter le patient dans ses diverses situations de communication, travailler la pragmatique, redonner confiance en soi, donner des conseils d'hygiène vocale, sur la relaxation, sur la posture, sur la respiration, etc.

Dans le cadre d'un travail avec ce feedback visuel, il est très important que l'orthophoniste présente les images spectrographiques obtenues de façon positive, voire ludique, pour maintenir l'intérêt et la motivation du patient à sa rééducation. Dans le questionnaire aux orthophonistes de Caroline Coudière, « *plusieurs personnes ont [...] pondéré leur utilisation des données du logiciel en relevant le fait de devoir toujours reprendre avec le patient les différents résultats et de les relativiser* » (Coudière, 2003, p. 134) [19]. Le thérapeute de la voix joue également un rôle actif avec le patient pour le transfert de ce feedback visuel sur les sensations acoustiques et proprioceptives.

1.3.2.3 Pertinence de l'information donnée par le spectrogramme

Dans la rééducation par biofeedback, les auteurs notent généralement que le rétro-contrôle doit être retransmis de manière précise, claire et attrayante. La présentation de ces informations doit revêtir une valeur de sanction et comporter une charge émotionnelle à des fins de motivation (in Zonder, 2002, p. 20) [119]. David M. Howard rappelle que les applications technologiques « *représentent un grand potentiel de bénéfices s'ils sont faciles à utiliser par les non spécialistes et si l'information qu'ils fournissent est complètement comprise, qu'elle a du sens, qu'elle est valide et utile* »³⁰ (Howard, 2005, p. 329) [49].

Donald G. Miller pense que « *le spectrogramme est l'image la plus susceptible de paraître familière au lecteur moyen* »³¹, parmi les images en temps réel proposées par Vocevista (spectrogramme, spectre, électroglottographe, signal auditif, forme d'onde) (Miller, 2008, p. 15) [75].

L'utilisation du spectrogramme ne permet pas une telle simplicité. Certes l'affichage offert par le spectrogramme « *a l'avantage de fournir un feedback qui ne se contente pas d'être binaire (la voix est soit bonne soit non) mais qui a une forme. Le locuteur peut voir où se situe la difficulté dans son*

³⁰ "our experience suggests that technological applications are of greatest potential benefit if they are easy to use by non-specialists and provide information that is readily understood, meaningful, valid and useful."

³¹ "Of the displays used in Vocevista, the sound spectrogram is likely to be the most familiar to the average reader"

énoncé »³² (Howard et al., 2012, p. 619) [51] mais cela se fait au détriment d'une utilisation simple de l'outil.

Pour Callaghan et ses collègues, repris dans la revue de littérature de Hoppe et ses collègues, les professeurs de chant ont trouvé que le feedback visuel aurait dû être plus pertinent musicalement (Hoppe *et al.*, 2006, p. 312) [46]. David M. Howard souligne que le spectrogramme Gamma Tone³³ évite des confusions qui peuvent apparaître avec les spectrogrammes traditionnels, qui ne tiennent pas compte de la perception humaine des sons (Howard et al., 2012, p. 612) [51].

1.3.2.4 Nécessité d'une formation spécifique

L'étude de Pat H. Wilson et ses collègues tend à montrer que l'efficacité de la forme du feedback visuel (en l'occurrence fréquence fondamentale en fonction du temps versus visualisation sur un clavier de piano) dépend des capacités des sujets à l'interpréter facilement (Wilson *et al.*, 2007, p. 7; Wilson *et al.*, 2005, p. 9) [111] [112]. Selon Garyth Nair, pour pouvoir utiliser le feedback visuel en temps réel (VRTF : Visual Real-time Feedback), les professeurs de chant doivent, notamment, savoir reconnaître auditivement et visuellement une bonne et une mauvaise production vocale, et savoir donner des exercices pour renforcer le comportement voulu et le rendre habituel (Nair, 1999, p. 68) [78].

Ainsi « *l'analyse de spectre est encore jugée très complexe pour des thérapeutes peu formés en acoustique. [...] raucité, fatigabilité, souffle[...] ne sont pas aisés à identifier sur un spectre sans une formation spécifique* » (Coudière, 2003, p. 152) [19]. De même, Sylvain Lamesch pointe qu'il est difficile au professeur de chant d'établir des connections entre la représentation du signal acoustique et ses propres représentations internes (Lamesch, 2012, p. 64; Maulet, 1997, p. 86) [56] [69].

Harm K. Schutte et Donald G. Miller confirment encore que la quantité d'information produite par le spectre en temps réel est importante. Selon eux, cela explique la difficulté pour les professeurs de chant d'interpréter l'information. « *Ceux qui pourraient profiter le mieux de son utilisation n'ont pas su [jusqu'ici] quoi rechercher dans cette surabondance d'information* »³⁴. Selon eux, cet état de fait provient du manque de vocabulaire commun entre les professionnels de la voix et les acousticiens. (Miller & Schutte, 1999a, p. 207) [72]. Pour Garyth Nair, si l'on introduit doucement la nouvelle technologie en commençant par les concepts de base, les professionnels de la voix pourront passer outre la peur que la technologie révèle l'incomplétude de leurs connaissances (Nair, 1999, pp. 6–9) [78].

Lors de l'expérience VOXed, les professeurs de chant ont utilisé préférentiellement le spectrogramme en temps réel, qui est le plus complexe des affichages proposés dans le logiciel. La plupart des interprétations des professeurs étaient donc basées sur une approche holistique plutôt que sur une analyse des détails. Ainsi la « surabondance des informations » des images affichées n'a pas gêné les professeurs de chant dans le cadre de cette étude, ceux-ci s'appuyant sur leurs connaissances pour interpréter les images (Howard *et al.*, 2007, p. 31) [50].

³² "This sort of display has the advantage of providing feedback – knowledge of results – that is not simply binary (the voice is either good or not) but feedback that is patterned. The speaker can see where in the utterance the difficulty occurs"

³³ Le Gamma Tone est un spectrogramme construit en respectant la perception humaine des sons. L'échelle des fréquences n'est pas linéaire et les intensités affichées respectent la sensibilité de l'oreille selon les fréquences.

³⁴ "those who could be profiting most from its use have not known what to look for in the glut of information"

L'utilisation de l'outil, en elle-même, peut constituer un début de formation : « *pour l'orthophoniste, la meilleure compréhension des paramètres acoustiques et leur maîtrise seraient dues au fait de voir l'évolution du spectre simultanément aux modifications du geste ou de la production vocale. [...] En effet, on peut imaginer que le manque de formation est pallié par les expériences empiriques* » (Coudière, 2003, p. 155) [19].

1.3.2.5 Conclusion

L'utilisation du spectrogramme peut être ressentie comme manquant de fiabilité du fait des réglages informatiques, d'aspects techniques comme le placement du micro, et du fait de son apport qualitatif plutôt que quantitatif. Le spectrogramme reste un outil qui ne remplace pas les approches classiques, et qui fournit un nombre d'informations importantes qu'il est important de savoir décoder. Ces limites peuvent être contournées en prenant les précautions nécessaires.

1.3.3 Aspects ambivalents

1.3.3.1 Concentration et charge cognitive

Pat H. Wilson et ses collègues mettent en avant une revue de littérature menée par Wulf et Prinz³⁵, concluant que « *l'efficacité avec laquelle les habilités motrices sont acquises peut être diminuée si le sujet fait trop attention à sa performance [...] ; demander au sujet de maintenir son attention à l'extérieur plutôt que de penser à sa propre performance et sa prise de conscience de ses mouvements permet une acquisition d'habilités optimale* »³⁶. Selon les auteurs, cela sous-entend que « *l'apprentissage est plus facile quand il y a une plus grande distance physique entre le cours et l'effet produit par la tâche* »³⁷ (Wilson et al., 2005) [112].

Le cas de la production vocale reste ambigu ; cependant on peut considérer que le son produit par le geste vocal est un phénomène extérieur, notamment lorsqu'il est perçu par l'interlocuteur. De ce point de vue, David Hoppe et ses collègues déduisent que « *le feedback visuel des performances chantées dirigé vers les mouvements du sujet (c'est-à-dire le conduit vocal) peut être moins efficace que le feedback visuel sur la sortie acoustique (c'est-à-dire une information spectrale en temps réel)* »³⁸ (Hoppe et al., 2006, p. 316) [46]. De ce point de vue, le feedback visuel serait alors préférable au feedback proprioceptif en rééducation vocale.

« *L'utilisation du logiciel favorise pour [...] 70% [des orthophonistes interrogés] la concentration des patients au cours des exercices sur ordinateur par rapport aux mêmes effectués sans le support de la visualisation* » (Coudière, 2003, p. 144) [19]. Welch et ses collègues rapportent également que les images fournies par le spectrogramme permettent « *de cibler l'attention de leur élèves sur la façon*

³⁵ Wulf, G., Prinz, W. (2001). Directing attention to movement effects enhances learning: A review. *Psychonomic Bulletin & Review*, 8, (4), 648 - 660.

³⁶ "the effectiveness with which motor skills are acquired may be adversely affected if the learner pays too much attention to her/his performance [...] instructing the learner to maintain an external focus, rather than think about their own performance and maintain self-awareness of their movements, yielded optimal skills acquisition"

³⁷ "Wulf and Prinz add that a consequence of this observation is that motor skills learning is further enhanced when there is a greater physical distance between the body and the effect produced by its task."

³⁸ "VFB on singing performance that is directed to one's own movements (e.g. the vocal tract) may be less effective than VFB on the acoustical output (e.g. real-time spectral information)."

dont ces figures pouvaient être manipulées systématiquement »³⁹ (Welch *et al.*, 2005, p. 242) [110]. Cette référence commune à un objet visuel, que l'on peut cadrer et délimiter, favorise la concentration. Pour Isabelle Marié-Bailly (2004, p. 51) [66], la visualisation spectrographique a un impact positif sur la concentration et la présence à soi.

Cependant, dans l'étude de Howard et ses collègues, les étudiants ont parfois trouvé que le feedback visuel les empêchait de se concentrer sur quelque chose de plus utile (la partition, le professeur, l'accompagnateur) (Howard *et al.*, 2004, p. 143) [48]. Il convient donc de trouver le bon équilibre entre les différents outils, en fonction des objectifs du travail en rééducation.

1.3.3.2 Nécessité du temps réel

Le feedback externe en temps réel doit faire le lien entre ce qui est entendu et ce qui est ressenti, qui sont les sensations sur lesquelles doit se faire un apprentissage durable et autonome. Dans un apprentissage avec rétro-contrôle, pour que le conditionnement soit opérant, il est nécessaire que le renforcement soit délivré aussi rapidement que possible. On voit ici tout l'intérêt du temps réel. Selon M Dutel⁴⁰, cité par Caroline Coudière, l'efficacité d'un feedback visuel repose notamment sur « *une quasi-simultanéité de l'affichage visuel correspondant au signal sonore pour un véritable « feedback visuel de suppléance* » » (Coudière, 2003, pp. 103–104) [19].

Selon Graham F. Welch, le feedback auditif seul n'est pas suffisant dans le chant, notamment pour la justesse : « *il n'est pas possible de « voir » si la réponse chantée est adéquate, puisque le stimulus et la réponse sont séparés dans le temps* »⁴¹ (Welch, 1985, p. 11) [109]. Si l'on demande à un sujet de reproduire un modèle, il est indispensable qu'il se rappelle du modèle après avoir produit l'imitation afin de pouvoir comparer ce qu'il a fait et ce qui était demandé. David M. Howard et ses collègues (Howard *et al.*, 2007, p. 21–22) [50] citent Welch : dans le cas habituel, l'élève produit quelque chose, attend le feedback externe du professeur, puis recommence. Un intérêt majeur du feedback visuel en temps réel est de fournir le feedback externe pendant la production de l'élève.

Sylvain Lamesch note également que « *l'élève n'arrive pas nécessairement à identifier le même indice que son professeur* », et que, de plus, « *la plupart du temps, [le professeur] donne son commentaire une fois le geste vocal effectué* », ce qui induit un « *retard entre le commentaire du professeur et le ressenti de l'élève* » (Lamesch, 2012, p. 63–34) [56].

Graham F. Welch explicite l'intérêt du temps réel dans l'apprentissage de la justesse dans le chant. Le groupe d'enfants bénéficiant de feedback visuel en temps réel a amélioré la justesse de son chant par rapport à l'autre groupe bénéficiant d'un feedback verbal habituel (in Welch *et al.*, 2005, p. 228) [110].

Pour Harm K. Schutte et D. Miller, la possibilité de voir la composition de la voix en temps réel permet de « *faire le pont entre la théorie et la pratique* »⁴². L'intérêt est de pouvoir connecter les trois éléments : affichage acoustique, son perçu, production du son (Miller & Schutte, 1999a, p. 207) [72].

³⁹ “Both teachers tended to treat the on-screen information as visual patterns associated with particular vocal behaviours in order to focus their students attention on how these patterns could be manipulated systematically”

⁴⁰ Dutel, M.M., Labaeye, P., Lamotte, M. SAVANE : système d'aide visuelle aux non-entendants, première expérimentation d'une approche vocale, avril 1985 (source : Mme Dutel).

⁴¹ “it is not possible to “see” if the sung reponse is accurate, as the stimulus and response happen and are separated over time”

⁴² “real-time spectrum analysis provides a bridge between theory and practice, allowing both aspects to gain”

Ces auteurs mettent en évidence l'intérêt de la visualisation lors de la recherche de la zone de passage par exemple : « *le fait d'explorer visuellement [ses modifications lorsqu'on change de registre] peut contribuer à la prise de conscience de la « sensation » de ses modifications* »⁴³.

Cependant, le feedback fourni par le spectrogramme n'est pas immédiat : l'affichage de l'information nécessite entre 40 et 120 ms, et le système sensitif humain demande encore 200 ms pour pouvoir en tenir compte. Ainsi on observe un décalage d'un quart de seconde à une demi-seconde. Dans certains cas, cela est bien supérieur à la durée des phénomènes à observer, c'est pourquoi Verdolini et Krebs considèrent dans ces cas précis que le spectrogramme ne constitue pas un « biofeedback » (ce qui suppose des corrections on-line), mais un « knowledge of results » (feedback externe différé) (Verdolini & Krebs, 1999, pp. 236–238) [107].

1.3.3.3 Analyse fine de sa production vocale

Selon Verdolini et Krebs, l'apprentissage du chant nécessite des processus lents, pour conscientiser les mécanismes du chant, afin d'empêcher les automatismes de la voix parlée de prendre le dessus dans la production de l'élève. Dans ce cadre, les auteurs pensent que l'observation de la voix à partir d'une analyse spectrale permet l'attitude cognitive adéquate pour l'apprentissage du chant (Verdolini & Krebs, 1999, p. 235) [107].

De même, lors de la rééducation, Françoise Estienne préconise parfois de demander au patient ce qu'il pense de ce qu'il vient d'émettre. Par exemple : « *on invite la personne à dire un [a] sur le fondamental [...] On interroge ensuite la personne sur ce qu'elle pense de ce son-là* ». (Estienne, 1998, p. 119) [27]. Il s'agit ici également de décortiquer au plus fin le geste vocal et son résultat. Cependant, pour Françoise Estienne, l'étape finale d'une rééducation reste de faire oublier le fonctionnement de la voix au patient : « *une voix qui se fait ignorer est le meilleur garant de son bon fonctionnement et de sa bonne santé* » (Estienne, 1998, p. 117) [27].

1.3.3.4 Lien avec les feedbacks auditifs et proprioceptifs

La fascination de pouvoir faire le lien entre ce qu'on entend et l'information donnée par l'ordinateur fait partie des réactions des professeurs de chant lors de l'introduction du spectrogramme (Nair, 1999, p. 17) [78]. Cependant selon Donald G. Miller, « *l'obstacle le plus important pour introduire le feedback d'analyse spectrale dans la pratique des cours de chant a été la difficulté de le relier avec la perception ordinaire* »⁴⁴ (Miller, 2008, p. 5) [75].

Lorsque ce lien n'est pas établi, « *une utilisation trop systématique et inappropriée des contrôles visuels [...] peut induire « des forçages ou une tonicité excessive* » (Manteau, 2004, p. 66; Romand, 1999) [65] [96], comme le note également Caroline Coudière dans les résultats de son enquête auprès d'orthophonistes (Coudière, 2003, p. 134) [19]: « *le renforcement des perceptions kinesthésiques au moyen du feedback visuel est efficace pour 46% des orthophonistes [interrogés], mais il est loin d'être évident pour les autres* » (Coudière, 2003, p. 147) [19]. De même, M. Maulet constate « *que les risques de forçage vocal ou d'interprétation erronée des courbes demeurent importants* » (Maulet, 1997, p. 86) [69].

⁴³ “the process of exploring these visually can also contribute to one's awareness of the “feel” of the changes”

⁴⁴ “a more important obstacle to incorporating feedback from spectrum analysis into practical voice instruction has been the difficulty in relating it to ordinary perception”

Jean-Blaise Roch insiste particulièrement sur le retour kinesthésique du geste vocal, qui est pour lui la seule manière pour le patient d'être « *aux manettes de commandement* ». Selon cet auteur, le retour auditif seul ne permet pas de changer le son émis, et seul le retour kinesthésique permet de guider le son (Roch, 2001, pp. 62–63) [94].

La mémoire implicite, qui joue un rôle important dans l'automatisation des apprentissages, est très dépendante de la modalité d'acquisition des connaissances. « *Les comportements entraînés dans une modalité (visuelle) ne se transfèrent pas nécessairement complètement dans une modalité différente (auditive et kinesthésique par exemple) [...] ; cependant nous reconnaissons que les habiletés motrices sont souvent enseignées avec succès dans le domaine de la réhabilitation physique en utilisant des biofeedback visuels et auditifs* »⁴⁵. Il devient donc très important, après un temps d'entraînement avec le feedback visuel, de renforcer les acquisitions sans le feedback visuel (“*withdrawal trials*”). Les deux types de feedback doivent s'alterner (Verdolini & Krebs, 1999, pp. 236–238) [107].

1.3.3.5 Nécessité de retirer progressivement le feedback

Selon Robert A. Volin, le spectrogramme est utile dans la phase cognitive d'apprentissage d'un nouveau geste. Il permet de rendre visible des comportements et permet donc une action sur ces comportements⁴⁶ (Volin, 1999, p. 243) [108]. Afin de ne pas utiliser constamment le feedback, il est facile de tourner l'écran pour retirer le feedback visuel. Après un certain nombre d'occurrences consécutives correctes sans le feedback, on montre au sujet un enregistrement de ses performances sur l'écran ; cela constitue le “*post-trial knowledge of result*” qui renforce la représentation interne du sujet et permet de l'apprendre à long terme. Le feedback est retiré progressivement par le praticien en fonction du degré de réussite du sujet (Volin, 1999, p. 244) [108]. Cela permet de prendre des repères proprioceptifs et auditifs et de rendre le geste automatique (Miller, 2008, p. 105; Nair, 1999, p. 69) [75] [78].

1.3.3.6 Objectivité / subjectivité

Evaluation et confirmation de l'intérêt de la rééducation

Garyth Nair pense que l'outil informatique objective l'efficacité de l'enseignement d'un professeur (Nair, 1999, p. 20) [78]. En cela il est rejoint par Donald G. Miller et Harm K. Schutte qui pensent que le feedback visuel permet d'évaluer objectivement les affirmations des professeurs de chant à propos de la technique vocale, qui, en temps normal, ne peuvent pas être prouvées (Miller & Schutte, 1999a, p. 210) [72].

« *Le caractère objectif et indubitable des résultats affichés sur l'écran permet aux orthophonistes d'avoir un moyen de convaincre de l'efficacité de tel geste, d'affirmer un progrès sans que le patient ne puisse remettre en cause son jugement. Cela permet aussi de conforter l'assurance professionnelle du thérapeute qui y trouve un justificatif à sa démarche et une preuve de son efficacité* » (Coudière, 2003, p. 155) [19].

⁴⁵ “*behaviors trained in one modality (visual, for example) do not necessarily transfer completely to a different modality (auditory-kinesthetic, for example). [...] We do recognize, however, that motor skills are often successfully taught in physical rehabilitation using visual and auditory biofeedback*”

⁴⁶ Basmajian 1982 : Clinical use of biofeedback in rehabilitation. *Psychosomatics*, 23(1), 67-69

Feedback verbal facilité

Le spectrogramme peut être utilisé en temps réel pendant la réalisation, mais également être observé après la réalisation, afin de discuter sur ses qualités de façon rétrospective. Ces différentes utilisations ont été observées lors de l'expérience VOXed (Howard *et al.*, 2007, p. 30) [50]. « *L'utilisation de la technologie multimedia permet de capturer chaque moment du comportement de l'élève et de le figer dans le temps. Cela facilite la possibilité de faire correspondre la compréhension de la cible d'un professeur [...] et la compréhension personnelle d'un élève [...] Dans notre étude [...] le feedback technologique a fourni un centre d'attention externe commun et stable sur lequel baser les interactions* »⁴⁷ (Welch *et al.*, 2005, pp. 242–243) [110].

Selon Garyth Nair, la possibilité d'objectiver des phénomènes acoustiques, auparavant perçus subjectivement, permet aux professeurs de chant de communiquer plus facilement avec les collègues à ce sujet. Cela est selon lui effectivement ressenti par les professeurs de chant lors de l'introduction du spectrogramme (Nair, 1999, p. 17) [78]. « *En signalant les preuves sur l'écran qui représentent l'amélioration de la technique vocale, le professeur gagne un puissant complément dans le processus de renforcement positif des améliorations continues* »⁴⁸ (Nair, 1999, p. 19) [78]. Pour cet auteur, les élèves ne perçoivent pas toujours ce que les professeurs veulent leur faire entendre pour pouvoir le corriger. Le feedback objectif fourni par l'ordinateur permet de rendre visible les subtilités du son, dont on ne peut plus penser qu'elles sont simplement imaginées par l'auditeur. Le fait de pouvoir les identifier est un premier pas vers une explication de leur production par le geste vocal (Nair, 1999, pp. 159, 161) [78].

Le fait de donner la possibilité d'échanger sur un support visuel aide l'étudiant à mieux prendre conscience du feedback verbal de son professeur (Welch *et al.*, 2005, p. 243) [110]. Effectivement, l'étude de Callaghan et ses collègues rapporte que les élèves « *ont ressenti que le feedback visuel a amélioré leur compréhension de la production désirée du modèle cible qu'ils étaient amenés à imiter, puisque le feedback était immédiat et non équivoque* »⁴⁹ (Hoppe *et al.*, 2006, p. 312) [46].

L'interprétation reste subjective

Cependant la notion de qualité de voix est une notion complexe, et en ce sens la validité de l'image spectrographique est à remettre en question. L'interprétation de l'image est très importante. Il n'existe pas aujourd'hui de technique permettant de faire une corrélation directe entre l'image spectrographique et la position de la langue ou la vibration des plis vocaux (Verdolini & Krebs, 1999, pp. 236–238) [107].

⁴⁷ “*The application of multimedia technology allows any particular moment of the student's singing behavior to be captured and 'frozen' in time. This facilitates the possibility of a better relative match between a teacher's target understanding [...] and the personal understanding of an individual student. [...] In the present study [...] the feedback technology provided a stable external and shared focus on which to base [the] interactions of the participants*”

⁴⁸ “*By pointing out evidence on the monitor that represent improvements in technique, the teacher gains a power ful adjunct in the process of positively reinforcing continued growth*”

⁴⁹ “*The VFB improved their understanding of the desired outcome of the target model that they were required to imitate, as the feedback was immediate and unambiguous*”.

1.3.3.7 Conclusion

Les auteurs décrivent des aspects ambivalents de l'utilisation de l'outil spectrographique en rééducation vocale. Il convient pour le thérapeute de trouver un juste équilibre afin de permettre aux patients de se concentrer sans arriver à une surcharge cognitive, de connaître à la fois les apports et les limites du temps réel, de tirer profit de l'analyse fine de la production vocale sans mettre en danger la spontanéité de la parole, de faire les liens entre les feedbacks auditifs et proprioceptifs, de gérer la nécessité de se détacher de l'outil et les possibilités d'objectivation à partir d'une interprétation subjective.

2 Problématique

Toute production vocale est contrôlée par un système d'autorégulation utilisant des feedbacks internes comme externes (audition par voie aérienne), qu'il est parfois nécessaire de renforcer par des feedbacks externes renvoyant une image de sa propre voix au patient.

Le spectrogramme est un outil intéressant pour voir la voix. Il en représente de nombreux paramètres, et notamment, le timbre et l'attaque. L'utilisation du spectrogramme en rééducation vocale est décrite comme pouvant favoriser la prise de conscience des paramètres de la voix, la motivation et l'autonomisation du patient. Le spectrogramme reste cependant un outil qui ne remplace pas les approches classiques, et qui fournit un grand nombre d'informations qu'il est important de savoir décoder. Il convient pour le thérapeute de trouver un juste équilibre dans l'utilisation de cet outil afin de permettre aux patients et à la relation thérapeutique d'en tirer le meilleur profit.

Les patients dysphoniques peuvent avoir une mauvaise image de leur voix et d'eux-mêmes. Au-delà de son côté technique, la rééducation vocale s'intéresse également à restaurer l'image de soi du patient.

Notre travail va s'attacher à étudier les conséquences objectives et subjectives de l'utilisation du spectrogramme en rééducation vocale. Dans quelle mesure le spectrogramme permet-il aux patients de modifier l'image de leur propre voix ?

Le spectrogramme, utilisé en parallèle avec un feedback auditif (réécoute de sa propre voix) permet-il au patient une meilleure prise de conscience de ses paramètres vocaux ? Dans ce cas, est-il suffisant pour améliorer également l'auto-contrôle de ces paramètres ? Comment le spectrogramme modifie-t-il la relation du patient à sa propre voix ? L'utilisation du spectrogramme a-t-il un impact bénéfique sur la relation thérapeutique ?

3 Hypothèses

L'utilisation du spectrogramme en rééducation vocale revêt deux aspects : l'utilisation en temps réel d'une part, et la réécoute d'extraits vocaux en visualisant le spectrogramme. En lien avec un feedback auditif, le feedback visuel constitué par le spectrogramme permet :

- une meilleure *prise de conscience* des paramètres de la voix de la part du patient (intensité, hauteur, richesse en harmoniques...) (hypothèse 1)
- un meilleur *auto-contrôle* de ces paramètres (hypothèse 2)
- la construction par le patient d'une *image positive* de sa propre voix (hypothèse 3)
- un impact positif sur la *relation thérapeutique* (hypothèse 4)

Pour chacune de ces hypothèses, dans le cas du feedback différé qui comprend à la fois un feedback auditif (réécoute de sa voix) et visuel (spectrogramme) on cherche à distinguer la part du feedback visuel de celle du feedback auditif.

4 Méthodologie

4.1 Matériel

4.1.1 Présentation générale

Dans cette recherche, nous avons étudié l'utilisation du spectrogramme en rééducation vocale dans diverses situations :

- en consultation de phoniatrie avec Isabelle Marié-Bailly au CHR d'Orléans,
- en autonomie pour les patients d'Isabelle Marié-Bailly qui avaient téléchargé le logiciel et pour lesquels était mis à disposition un site internet didacticiel,
- au sein d'un protocole mené pour quatre patientes où les séances étaient conduites par nous-même (groupe « patients protocole »).

En outre, un exercice d'auto-analyse, défini au paragraphe 4.2.3, a été conçu afin d'évaluer plus précisément les effets de la visualisation du spectrogramme sur la prise de conscience par les patients de différents paramètres de leur voix. Cet exercice a été proposé aux patients du groupe protocole et aux patients d'Isabelle Marié-Bailly (« patients hors protocole ») lors de leurs séances, en notre présence (voir Figure I). On note qu'une patiente du groupe « patients hors protocole » ne faisait pas partie des patients d'Isabelle Marié-Bailly et ses exercices ont été réalisés à son domicile.

Les épreuves du bilan vocal pour les patients protocole ont été analysées perceptivement par un jury d'écoute. En outre, les extraits vocaux de tous les exercices d'auto-analyses ont été fournis au jury. Les membres du jury d'écoute ont été invités à faire le même exercice d'analyse que les patients, sur les sons des patients, afin d'en constituer une évaluation perceptive.

4.1.2 Description des populations

4.1.2.1 *Patients protocole*

Généralités

Dans un premier temps, ils ont été recherchés parmi des patients d'orthophonistes en libéral qui auraient accepté d'intégrer le protocole dans leur rééducation. Malgré des contacts avec 7 orthophonistes qui font régulièrement des rééducations vocales, il n'a pas été possible de recruter des patients par ce biais, ni par l'intermédiaire d'un message mis en évidence sur le forum Orthomalin.

Nous avons donc modifié le protocole pour recruter des personnes en difficulté par rapport à leur voix, mais qui n'étaient pas en rééducation orthophonique, et intéressées par un travail avec nous-même autour du spectrogramme. Ce recrutement induit un biais méthodologique important, et a conduit à une étude de cas multiples. Toutes les patientes sont chanteuses. Elles répondent toutes « toujours » aux questions « la voix est un sujet qui m'intéresse » et « j'accorde de l'importance à ma propre voix ».

La patiente A

A 55 ans, la patiente A est professeur à temps partiel. Elle a diminué son activité pour des raisons personnelles et notamment pour protéger sa voix. Sa voix devient parfois un problème, 5% du temps sur une année. Elle n'est pas dans une période pathologique en ce moment, elle est mieux armée et

plus vigilante. Elle travaille en parallèle une méthode respiratoire avec une orthophoniste. Elle a bénéficié précédemment d'une rééducation orthophonique suite à des nodules sur ses plis vocaux.

La patiente B

A 30 ans, cette patiente est professeur de chant, chanteuse de jazz et chanteuse lyrique. Elle est très intéressée par la voix et les mécanismes de la phonation. Elle souffre d'une dysphonie spasmodique depuis 5 ans, suite à un accident subi il y a 7 ans, dont les traumatismes ont conduit à une opération du voile du palais. La dysphonie spasmodique s'entend beaucoup plus en voix parlée qu'en voix chantée, qui reste cependant fragile au niveau du passage de registre. Elle se soigne elle-même pour réussir à contrôler sa voix. Elle est soutenue par un coach vocal et un ostéopathe certifié Osteovox.

La patiente C

A 31 ans, la patiente C est journaliste de presse. Suite à un kyste intracordal et à des nodules surajoutés à un sulcus, elle a subi une opération des plis vocaux en juin 2007 suivis de 2 ans de rééducation. Elle chante dans un chœur de chambre mais s'interdit les cours de chant du fait de son problème de voix.

La patiente D

A 32 ans, la patiente D a entamé en septembre une reconversion vers les métiers du chant à l'*American School*. Elle est chanteuse de gospel. Très peu de temps après la rentrée, elle a souffert d'un coup de fouet laryngé. Elle n'a pas entrepris de rééducation orthophonique mais prend soin de sa voix auprès d'un coach vocal.

4.1.2.2 Patients hors protocole

Ce groupe comprend initialement 6 hommes et 5 femmes. Les patients hors protocole sont ceux qui consultent en phoniatrie avec Isabelle Marié-Bailly au CHR d'Orléans. Ils ont des plaintes vocales ou auditives. Les pathologies sont diverses : paralysie récurrentielle, dysphonie dysfonctionnelle, dysarthrie, hypophonie, hyperacousie, acouphènes, mue faussée, paralysie cérébrale. Une patiente supplémentaire, souffrant d'un sulcus, a été incluse dans ce groupe : elle a eu deux séances d'une heure avec nous-même. Le groupe comprend donc finalement 6 hommes et 6 femmes.

4.1.2.3 Personnes répondant au questionnaire en ligne.

10 personnes ont répondu au questionnaire en ligne. Toutes les patientes du protocole ont répondu au questionnaire en ligne, et 4 patients hors protocole. Tous les patients ont été invités à répondre à ce questionnaire, par deux fois par mail.

Le site internet ayant été présenté également à des ateliers vocaux organisés par Isabelle Marié-Bailly, leurs participants ont également été invités à répondre au questionnaire. 2 personnes l'ont fait, dont une étudiante en deuxième année d'orthophonie à Tours (voir Figure I).

4.1.2.4 Jury d'écoute

Le jury d'écoute est constitué de 4 femmes, dont 2 juges expérimentés (une phoniatre, une orthophoniste) et 2 orthophonistes plus débutantes. Ces 4 auditeurs étaient préalablement formés à la lecture de spectrogramme, et ne présentaient aucun trouble auditif.

4.2 Méthode

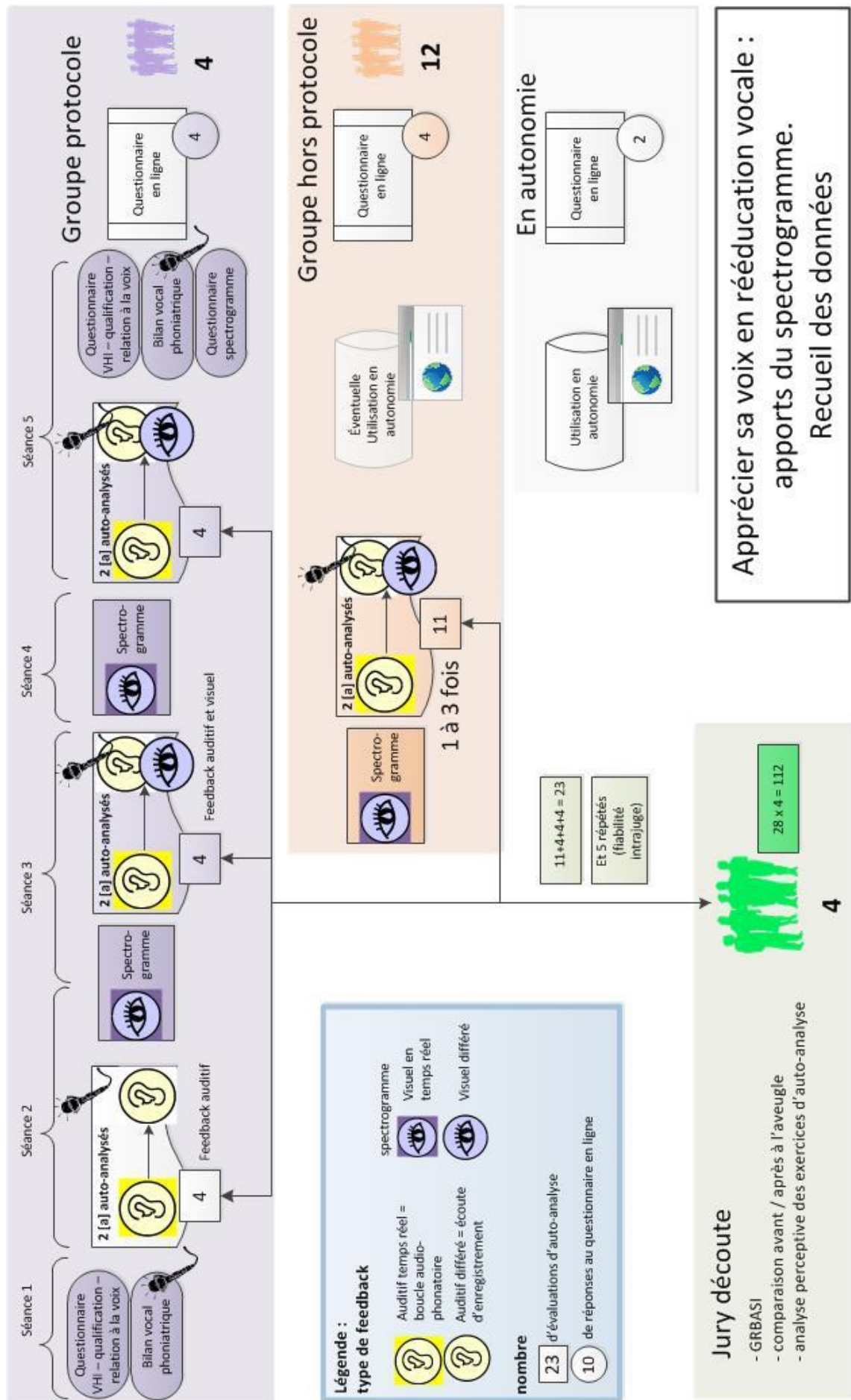


Figure I : groupes (groupe protocole, groupe hors protocole, groupe en autonomie et jury d'écoute) et déroulement général du recueil des données

4.2.1 Protocole

4.2.1.1 Déroulement général

Voir le diagramme général du protocole (Figure I).

	Séance 1	Séance 2	Séance 3	Séance 4	Séance 5	
Patiente A	20 nov	27 nov	8 jan	22 jan	5 fev	
Patiente B	3 dec	10 dec	17 dec	14 jan	4 fev	
Patiente C	27 dec	2 jan	3 jan	18 jan	21 jan	11 fev
Patiente D	7 déc	4 jan	8 jan	22 jan	1 ^{er} fev	

Tableau III : calendrier des séances pour les patients du groupe protocole

Les rencontres ont eu lieu au domicile des patients selon leurs disponibilités (voir Tableau III). Pour la patiente A cependant, les rencontres ont eu lieu le 20 et le 27 novembre avec l'orthophoniste qui prenait en charge cette patiente, par la suite nous avons rencontré la patiente à son domicile.

Les séances duraient environ une heure. Le découpage a été sensiblement différent pour la patiente C, qui a fait 6 séances dont 2 plus courtes, pour des raisons de disponibilité.

4.2.1.2 Evaluation pré- et post- protocole

L'évaluation pré- et post- protocole consiste en une évaluation subjective du handicap ressenti, et une évaluation objective et perceptive au travers d'un bilan vocal orthophonique. Il s'intéresse également, de façon moins classique, à la relation du patient à sa voix. A priori, l'évaluation concerne la voix parlée. Le groupe protocole n'étant constitué que de chanteuses, nous avons adapté a posteriori les questionnaires à la voix chantée. Celle-ci n'est donc pas évaluée avec des outils spécifiques.

Auto-évaluation du handicap ressenti

Les patientes ont répondu au VHI-10 (voire en annexe Figure XXX). Cette échelle a été élaborée par Rosen et ses collègues (2004) [97] à partir du VHI-30. Les auteurs ont demandé à différents spécialistes en voix (phoniatries, orthophonistes, spécialistes en voix chantée) de relever les dix items qui leur paraissaient les plus pertinents. Leur étude statistique ne montre pas de différence significative entre VHI-30 et VHI-10. Le VHI-10 serait donc une bonne représentation du VHI-30 (Morsomme & Estienne, 2006, p. 228) [77].

Après avoir testé des passations du VHI-30 en consultation de phoniatrie d'Isabelle Marié-Bailly, nous avons préféré le VHI-10 au VHI complet car la passation en est plus rapide (Giovanni & Fenollar, 2004, p. 60) [41] et de ce fait moins pénible pour la personne lorsqu'elle touche du doigt son handicap.

Bilan vocal

Après un entretien anamnestique (Estienne, 1998, p. 15) [27], le bilan vocal comprenait divers enregistrements qui ont été l'occasion d'observer le comportement vocal du patient (conduite respiratoire, posture, zones de tension) (Amy de la Bretèque, 2012, p. 137; Estienne & Piérart, 2006, p. 238; Estienne, 1998, p. 90; Woisard, 2004, p. 205) [3] [30] [27] [113].

Les extraits vocaux demandés sont les suivants :

- un extrait d'un texte lu. Nous avons choisi la phrase « c'est une affaire intéressante, qu'en pensez-vous, il faut la faire sans aucun délai », selon les recommandations de la Société

Européenne de Laryngologie demandant l'enregistrement d'une phrase simple (Dejonckere et al., 2001, p. 78) [23]

- la voix d'appel
- une sirène ascendante
- la réalisation de 3 attaques différentes : soufflée / équilibrée / dure
- une dynamique d'intensité (on demande un son le moins fort possible puis le plus fort possible)
- le Temps Maximum de Phonation (TMP) sur « a » et le rapport des TMP z/s

L'analyse perceptive reste la référence dans un bilan vocal, notamment car la parole est produite pour être perçue et que cette approche est accessible à tout clinicien à un faible coût (Ghio, 2012, p. 70) [38]. Les extraits « voix parlée », « voix d'appel » et « sirènes » ont permis la cotation du GRBASI par le jury d'écoute.

L'échelle de classe en 4 points d'Hirano (voir Tableau IV) reste le « meilleur compromis méthodologique pour la communication des informations concernant un patient » (Revis et al., 2012, p. 62) [91]. Nous avons voulu étendre cette échelle en 7 niveaux (0 / 0,5 / 1 / 1,5 / 2 / 2,5 / 3) notamment pour augmenter la précision de l'évaluation des dysphonies légères (voir la feuille de cotation en annexe Figure XXVI)

Les extraits « voix parlée », « voix d'appel », « sirènes » et « 3 attaques » ont fait l'objet d'une comparaison avant / après à l'aveugle dont l'objectif était de déterminer si la voix avait été modifiée pendant les séances du protocole (voir la feuille de cotation en annexe, Figure XXV). Les auditeurs du jury d'écoute ont eu accès aux enregistrements en .wav sur un site internet dédié. Les écoutes sont faites avec un casque fermé Sennheiser HD200 et une carte son externe PureAudio USB d'Andrea.

L'évaluation objective est réalisée grâce aux logiciels Praat (shimmer, jitter, NHR) sur l'extrait « attaque équilibrée » des trois attaques demandées, Audacity (mesure des Temps Maximum de Phonation) et VoceVista (mesure de l'étendue vocale, car Praat s'est avéré non fiable pour cette mesure).

	Terme	Signification	Définition
G	Grade	Grade global de dysphonie	Impression globale du degré d'anormalité de la voix
R	Rough	Raucité	Impression d'irrégularité des vibrations de la corde vocale qui correspond aux fluctuations irrégulières de la F ₀ et/ou à l'amplitude du son glottique
B	Breathy	Caractère soufflé	Impression d'une fuite d'air assez importante à travers les cordes vocales, relative à des turbulences
A	Asthenic	Asthénie	Manque de puissance de la voix relatif à une intensité faible du son et/ou un manque des harmoniques élevés
S	Strained	Forçage	Impression d'un état hyperfonctionnel de phonation relatif à une F ₀ anormalement haute (bruit dans les hautes fréquences et/ou richesse en harmoniques dans les hautes fréquences)

Tableau IV : définition des différents items du GRBAS (Revis et al. 2012, p.59) [91]

Court questionnaire sur la relation à sa voix

Les sept questions (voir en annexe Figure XXXI) sont inspirées de questionnaires de la bibliographie (Borel, 2002; Gillie, 2007) [12] [39] et ont été quelque peu modifiées suite aux pré-tests effectués en

consultation de phoniatrie avec Isabelle Marié-Bailly afin d'éviter des malentendus. (Par exemple « ma voix correspond à ce que je suis » est devenu « ma voix correspond à ce que je suis aujourd'hui »).

Entretien semi-dirigé

Pour Françoise Estienne (1998, p. 14) [27], l'examen orthophonique comprend un versant s'intéressant à « *la voix ressentie du dedans* ». Il a notamment pour objectif de « *relever le système de valeurs de la personne en matière de voix en lui proposant des moyens de préciser ses critères de référence, sa représentation de la voix, en évaluant ses ressources et ses lacunes ; vérifier ses connaissances au sujet de la voix [...] préciser ce que la personne veut obtenir et ce qu'elle croit devoir faire pour résoudre son problème vocal* ».

Les questions de l'entretien semi-dirigé ont été inspirées par différents questionnaires s'intéressant au ressenti du patient par rapport à sa voix et à la prise de conscience (Aimadeddine & Kerlan, 2003; Cotelle, 1993; Coudière, 2003; Estienne, 1998; Rousseaux, 1989; van Leer & Connor, 2010) [1] [18] [19] [27] [98] [106]. Elles ont été modifiées suites aux pré-tests en consultation de phoniatrie avec Isabelle Marié-Bailly afin de clarifier les questions et les adapter à un entretien oral. Enfin une nouvelle adaptation a été réalisée lorsque le protocole s'est recentré sur des patients sans suivi orthophonique ou phoniatrique (voir en annexe Figure XXVII)

Qualification de sa propre voix telle qu'elle est et telle qu'on voudrait qu'elle soit

La liste d'adjectifs sur la voix est citée dans la thèse de Claire Gillie (2007) [39] comme donnée dans le Guide de la Voix d'Ormezzano (2000) [79] (voir en annexe Figure XXXII).

4.2.1.3 Utilisation du spectrogramme

La découverte du spectrogramme se fait à partir de sons tenus (ceux de l'exercice d'auto-analyse qui se déroule au début de la séance n°2). Ces sons simples permettent d'introduire les différents paramètres affichés : l'intensité, les harmoniques, les attaques.

Par la suite, le spectrogramme est toujours affiché et retient l'attention du thérapeute et du patient. Nous suivons un programme de découverte du spectrogramme plutôt qu'un programme thérapeutique. Les questions posées autour du spectrogramme sont proches de celle que rapporte Garyth Nair pour les professeurs de chant : « *est-ce que je peux refaire cette image avec ma propre voix ? Qu'est-ce que je peux dire à mon élève pour faire que l'image et donc le son soient meilleurs ?* »⁵⁰ (Nair, 1999, p. 19) [78] On s'efforce donc d'être à la fois attentif au spectrogramme et au comportement vocal du patient.

4.2.1.4 Exercices proposés

Ces exercices ont pour objectif la réduction du forçage et l'amélioration de la voix, mais également la prise de conscience des différents paramètres de la voix, et leur visualisation sur le spectrogramme. Nous avons choisi de privilégier les exercices ayant pour matériel des phonèmes tenus (recto tono). En effet, ces exercices sont souvent utilisés en rééducation, et de plus ils permettent plus facilement qu'un matériel plus complexe (parole continue par exemple) d'apprendre à observer sur le spectrogramme

⁵⁰ "Can I replicate that image and sound with my own voice? What can I tell the student to change in the voice to make the image, and therefore the sound, better?"

les paramètres étudiés : attaque et timbre. De plus, « l’affichage spectrographique est plus intéressant avec des sons continus, comme les voyelles, les fricatives, et les semi-voyelles. Ces sons s’étendent sur une durée considérable, ils sont donc facilement représentés sur un écran et accessibles pour un monitoring et un ajustement en temps réel »⁵¹ (Volin, 1999, p. 246) [108].

Temps passé en minutes					Patiente A	Patiente B	Patiente C	Patiente D
Séance 2	Séances 2 et 3	Séance 3	Séances 3 et 4	Séance 4				
Introduction et lecture du spectrogramme					1	4	2	
Pole attaque et clarté du timbre	Consonnes voisées	[ʃ ʒ], [f v], [s z]	9	13	5	6		
		[ʃ ʒ u], [f v o/y], [s z i]	12	15	5	3		
		[m n ŋ a]			2			
	Consonnes non voisées	[f s ʃ]	2		4	7		
		[ʃ ʃ ʃ]				1		
		[fi fa fe fy], [pa pe], [ti te ta]...		4	3	2		
		[s] dans la forme du [i], [s] dans la forme du [y] : « s_i, s-u »			5	3		
	Attaques	Dure/équilibré/soufflé	15		3	5		
		attaques piquées				3		
	Fry	fry et voisement alterné	7	4	15	9		
puff d'air		3	4	2	2			
fry [a o i], [a o u], [i e a]...		4	4	10	18			
a chuchoté / voisé					2			
Pole richesse harmoni- que	Sirène	Sur une voyelle, [ʒ], battement de lèvres ou paille		4	9	3		
	Voyelles	Formants	13	8	12	5		
		Transitions vocaliques	17	1	5	11		
		Transitions nasales / orales		14	6	4		
		Formant du chanteur	1	10		3		
	Voix chantée	Vibrato, serrage...				7		
		extrait chanté	9	9	11	12		
Harmoniques / formants		2	1	4	15			
Autre (hors spectrogramme)					2	25	4	10
total					1h37	1h31	1h47	2h11

Tableau V : activités proposées avec les patientes du groupe protocole. Temps passé en minutes et temps total

Certains de ces exercices sont cités dans la littérature pour l’intérêt spécifique que l’outil spectrographique leur apporte : alternance [s]/[ʃ], observation du spectre pour l’identité de la voyelle (Nair, 1999, pp. 90, 100) [78], ajustement du conduit vocal pour explorer les modifications des formants (Miller & Schutte, 1999b, p. 209) [73], distinction entre consonnes voisées et non voisées, distinction entre les consonnes fricatives selon leur point d’articulation (par exemple [f] [s] [ʃ] grâce aux différentes bandes d’énergie mises en jeu dans le bruit entendu et affiché sur l’écran (Volin, 1999, pp. 246–250) [108]), lecture des formants à partir du fry (Miller & Schutte, 1990, p. 330 ; Miller &

⁵¹ “spectrographic displays are most helpful with continuant sounds, such as vowels, fricatives, and glides. Because these sounds extend over a considerable duration, they are readily represented on a screen and are accessible for on-line monitoring and adjustment based on RTF”

Schutte, 1999a, p. 194; Miller, 2008, p. 23) [74] [72] [75]. On propose également des exercices plus spécifiques à la voix chantée, où le spectrogramme est indispensable, comme l'adaptation phono-résonnante (Heinrich, 2012, p. 27; Miller, 2008, pp. 96–105) [45] [75] et la visualisation de l'impact du geste vocal sur le singing formant et la richesse harmonique (Howard *et al.*, 2004, p. 140) [48].

D'autres prennent leur source dans des exercices classiques de rééducation orthophonique : coordination pneumophonique par la « *sonorisation continue de consonnes fricatives* », ou l'expiration fluide sonorisée dans une paille (Amy de la Bretèque, 2004; Revis & Cayreyre, 2004, p. 101; Sarfati, 1998, pp. 63–64) [2] [90] [99], les exercices mettant en jeu le fry : « *qui consiste à utiliser une pression sous-glottique minimale pour permettre un relâchement des cordes vocales* » (Estienne & Morsomme, 2009, p. 261) [29], les différentes résonances de la voix (Estienne, 1998, p. 119; Revis & Cayreyre, 2004, p. 102) [27] [90], l'alternance de voyelles orales et nasales (Estienne, 1998, p. 122) [27], les transitions vocaliques s'inspirant du triangle vocalique (Estienne, 1998, p. 154; Fournier & Dupessey, 2001, p. 240; Klein-Dallant, 2001, p. 123) [27] [36] [53].

La plupart des exercices se basent sur la voix chantée. En effet, « l'entraînement vocal incluant systématiquement un travail en voix chantée est intéressant, à plusieurs titres, en rééducation comme en éducation » (Fournier & Dupessey, 2001, p. 244) [36].

Comme Volin le précise : « *au fur et à mesure que le thérapeute se familiarise avec le programme en l'utilisant en séance, la créativité spontanée qui est une part du processus thérapeutique fera inmanquablement émerger d'autres méthodes et d'autres applications* »⁵² (Volin, 1999, p. 256) [108]. « *On évite ainsi que le patient se considère comme fondamentalement maladroit ou incompetent pour rentrer dans le rôle d'un sujet actif, voire créatif, vivant des situations de découverte* » (Amy de la Bretèque, 2012, p. 147) [3].

La découverte de l'outil spectrogramme est adaptée à chacun selon sa curiosité et son niveau de connaissance et ne suit pas un protocole préétabli (Zonder, 2002, p. 38) [119]. Les exercices proposés sont différenciés suivant les patients et proposés au fil des séances selon les besoins et les découvertes du patient. La plupart des exercices se basent sur la voix chantée. En effet, « l'entraînement vocal incluant systématiquement un travail en voix chantée est intéressant, à plusieurs titres, en rééducation comme en éducation » (Fournier & Dupessey, 2001, p. 244) [36].

reporte tous les types d'exercices proposés. Le paragraphe 10.3 en annexe détaille l'utilisation du spectrogramme en pratique sur ces exercices.

4.2.1.5 Ressenti face au spectrogramme

Les commentaires autour du spectrogramme sont recueillis :

- lors du jeu de découverte du spectrogramme et des exercices d'auto-analyse avec les patients protocole et les patients hors protocole,
- par des commentaires écrits spontanés de deux patients hors protocole (patients-pilote),

⁵² “*As the clinician becomes familiar with the program by using it in therapy, other methods and other applications will undoubtedly spring from the natural creativity that is part of the therapeutic process*”

- par l'entretien post-traitement pour les patients protocole, inspiré de l'étude de Callaghan et ses collègues⁵³ (in Hoppe *et al.*, 2006, p. 312) [46] (voir Figure XXIX en annexe),
- par le questionnaire en ligne qui a été proposé à tous (patients protocole / hors protocole et personnes utilisant le spectrogramme en autonomie) (voire paragraphe 10.5 en annexe).

Ils sont ensuite dépouillés méthodiquement de façon à faire ressortir des thématiques communes aux différents commentaires.

4.2.2 Patients en autonomie

Téléchargement et mode d'emploi du logiciel Vocevista.			
Lecture de spectrogramme	Attaque (début de son)	Exemple de 3 attaques + explications physiologiques sur le début du son et son lien avec les pathologies vocales	
	Timbre	Timbre de voix chantée, clair avec vibrato	
		Timbre soufflé	
		Timbre éraillé, irrégulier	
		Son pur	
		Construction d'une voix chantée harmonique par harmonique	
		Différentes voyelles synthétisées (formant vocalique)	
		Comparaison son pur (fréquence fondamentale) / son pur accompagné de ses harmoniques	
	Comparaison partiels multiples ou non de la fondamentale		
	Curiosités	Une cloche	
Synthèses de sons musicaux			
Idées d'exercices (les exercices sont décrits et des conseils sont donnés succinctement)	Intensité, hauteur	Lien entre harmoniques et intensité (volume)	
		Les sirènes = modification de la hauteur	Son dans la paille « j » « a »
			« s / z »
	Attaque (début de son)	Alternance fricative sourde / sonore	Son dans la paille « ch / j »
			« ch j ou »
		Constrictive guide	« s z i »
		Reproduire 3 attaques	
		Puff d'air, Fry sur « a » puis « a » voisé	
		Fry alterné avec son voisé	Homme Femme
	Timbre (richesse harmonique)	Château des voyelles « o u i u o »	Lent et grave
			Rapide, un peu plus aigu
		Attention aux composantes aiguës dans les consonnes	« s / ch »
			« ch-i » « ch-ou »
		Enrichissement des harmoniques	
		Alternance voyelles orales / nasales	« a /an »
			« o /on »
	« è / in »		
Mise en évidence des formants en fry	« a é i é a » (femme)		
	« a/o »	Homme / femme	
Banque de son	Femme / homme	Sévérité de la dysphonie	Timbre clair / altéré / très altéré
		Richesse en harmoniques	Timbre riche / pauvre / très pauvre

Tableau VI : plan du site internet pour l'utilisation du spectrogramme en autonomie

⁵³ Callaghan J., Thorpe W., Van Doorn J. (2004) The science of singing and seeing. In *Proceedings of the Conference on Interdisciplinary Musicology (CIM04)*, Graz, Australia, April 15–18.

4.2.2.1 Site internet

Le Tableau VI présente la cartographique du site internet, disponible au lien suivant :

<http://florence.parmentier.free.fr/acoustique/spectrogramme/>

Les exemples ont été construits à partir d'extraits vocaux personnels, des extraits vocaux (Foisneau *et al.*, 2003) [33], des extraits vocaux de Garyth Nair, des synthèses de sons par John Pierce (1984) [82]. Chaque exemple montre l'image du spectrogramme, le plus souvent de 0 à 4 kHz (jusqu'à 8kHz pour certains exercices le nécessitant). Lorsqu'on clique sur cette image, le son correspondant (en mp3⁵⁴) est entendu et un curseur suit l'affichage du son sur le spectrogramme. Dans certains cas, des commentaires s'affichent sur l'image spectrographique à la fin de l'animation.

La cartographie du site est présentée sur le Tableau VI (le matériel phonétique est présenté ici comme il l'est sur le site internet, où nous n'avons pas utilisé les symboles phonétiques de l'API)

Le site Internet a été proposé aux patientes du protocole à la fin de la séance 4, et à tous les patients hors protocole. En outre il est accessible depuis les moteurs de recherche.

4.2.2.2 Questionnaire en ligne

Le questionnaire (en annexe paragraphe 10.5) a été construit selon les recommandations de construction de questionnaire (Py, 2010, p. 36; de Singly, 2008, pp. 64–84) [84] [100], afin d'éviter les doubles questions, être attentif aux connotations subjectives des qualificatifs employés, établir un équilibre entre les modalités positives et les modalités négatives, ne pas utiliser de négation, etc.

Il reprend certaines idées du questionnaire sur les outils informatiques du mémoire de Caroline Coudière (2003) [19] ; il a été guidé par les entretiens de fin de protocole. Les questions sont parfois redondantes afin de juger de la cohérence des réponses. Il favorise l'expression personnelle notamment par la possibilité de s'exprimer librement après chaque groupe de questions. Il propose des échelles d'accord où il faut répondre si l'on est « tout à fait d'accord », « plutôt d'accord », « ne sais pas », « plutôt pas d'accord » ou « pas d'accord du tout » à certaines affirmations.

Le questionnaire comprend une première partie permettant de situer la personne qui répond (expérience avec le spectrogramme en terme de connaissance et d'utilisation, utilisation de la voix chantée, problèmes de voix), puis des questions précises sur :

- l'image spectrographique et les sensations auditives et proprioceptives,
- les paramètres vocaux que le spectrogramme permet d'observer (attaque, richesse du timbre, parole),
- les gênes qui peuvent être induites par l'usage du spectrogramme dans le comportement vocal (posture, respiration, geste vocal...) et l'intérêt ou non du temps réel,
- les utilisations du spectrogramme : analyse de la voix, preuve objective, amusement. Envie d'utilisation, avec un orthophoniste ou un professeur de chant,
- l'utilisation du spectrogramme en autonomie et l'utilisation du site internet.

⁵⁴ Cette compression en mp3 concerne uniquement les exemples didactiques, en aucun cas les extraits sonores des bilans vocaux ni des exercices d'auto-analyse utilisés pour l'analyse perceptive.

4.2.3 Exercice d'auto-analyse

4.2.3.1 Description de l'exercice d'auto-analyse

Principe de l'exercice

Les patients sont invités à porter attention à diverses caractéristiques de leur voix (intensité, hauteur, attaque soufflée / dure, clarté du timbre, richesse du timbre, comportement vocal (geste)) lorsqu'ils émettent deux [a] d'intensités différentes. Le premier objectif est de savoir retranscrire après-coup sur les échelles visuelles analogiques de la feuille d'auto-évaluation ce qu'on a ressenti et entendu. La consigne comprend également l'objectif secondaire de faire deux sons d'intensités différentes qui soient sur la même hauteur, avec une attaque la plus équilibrée possible, un son le plus clair et riche possible, et un comportement vocal adapté.

La voyelle [a] a été choisie notamment parce que d'après Yanagihara (1967, pp. 532–533) [116], les voyelles [a], [ɛ] et [i] permettent de mieux faire apparaître les caractéristiques de la dysphonie sur le spectrogramme.

Les patients doivent coter leurs deux sons consécutifs sur chacune des échelles visuelles analogiques, une première fois grâce au feedback interne (proprioception) et des feedback externes limités (écoute en temps réel), une deuxième fois avec un feedback externe, auditif et visuel. Le protocole de passation de l'exercice est disponible en annexe (voir Tableau XXXII).

Pour les patients hors protocole, ce feedback externe est toujours constitué du spectrogramme (feedback visuel) en plus du feedback auditif (réécoute de sa voix). Pour les patients du groupe protocole, en revanche, la première passation de cet exercice a lieu juste après le début du protocole, donc avant qu'ils ne connaissent le spectrogramme. Le feedback additionnel consiste uniquement en la réécoute de sa voix.

L'exercice d'auto-analyse effectué ici par les patients leur demande justement de porter attention à leur geste vocal émis et enregistré dans des conditions pouvant générer du stress : un premier [a] tenu, suivi d'un deuxième [a] tenu, qu'ils doivent faire plus fort, sans pouvoir s'aider d'un changement de hauteur de la voix.

Intérêt de l'exercice

D'après Françoise Estienne « *une bonne voix est capable d'émettre des sons sans fatigue et avec qualité à des intensités très faibles (pianissimo) et très fortes (fortissimo) allant jusqu'au cri. Elle est capable de passer d'une intensité à l'autre (contraste) en un laps de temps très court en émettant des sons de qualité* » (Estienne, 1998, p. 57) [27]. C'est ce que cet exercice tend à travailler. En outre l'exercice est difficile et fait dans des conditions très contraignantes (notamment de temps : il faut que les 2 [a] soient visibles sur une fenêtre de spectrogramme de 6 secondes). Le patient est mis en difficulté, l'intérêt de l'exercice résidant alors non plus dans la bonne réalisation de l'exercice, mais dans la qualité de son attention à sa propre voix.

Plusieurs études montrent que les patients ayant une dysphonie ancienne ont une mauvaise connaissance de leur voix (Berhman *et al.* 2004, in Lescarmontier, 2012, p. 22) [8] [63], et une qualité de perception musicale amoindrie, en terme de « *hauteur tonale, de rythme, d'intensité et de timbre* » (Julien, 2009, p. 30) [52].

Philippe Dejonckere (2001, p. 354) [21] indique que le travail de discrimination auditive constitue un pré-requis nécessaire à la bonne efficacité des exercices. Nous pensons que ce travail n'est pas seulement un préalable mais qu'il est nécessaire d'affiner la perception du patient tout au long du travail orthophonique.

On note qu'une auto-évaluation, adaptée des échelles de douleur, de l'effort fourni pour produire un certain résultat vocal (une certaine hauteur, une certaine qualité de voix et un type d'attaque) a été utilisée par Kimberly M. Steinhauer et ses collègues : le Borg Category-Ratio Scale (CR10)⁵⁵ (Steinhauer *et al.*, 2004, p. 435) [103]. On rapproche cette méthode avec notre dernière échelle d'auto-évaluation (comportement vocal adapté / inadapté).

Dans leur proposition d'un protocole basique d'évaluation des voix pathologiques, La Société Européenne de Laryngologie propose de faire émettre plusieurs [a] tenus à une hauteur et une intensité confortables, puis un [a] légèrement plus fort afin d'évaluer les éventuels changements dans la qualité de la voix. Le matériel que nous utilisons s'approche également de cette proposition (Dejonckere *et al.*, 2001, p. 78) [23].

Échelles visuelles analogiques

Dans les années 1990 sont apparues les échelles visuelles analogiques. Ce sont des échelles où l'on appréhende la variable à mesurer de façon visuelle, spatiale ; elles ont été adaptées des méthodes de mesure utilisées dans les études réalisées en psychologie. Selon Bele, les échelles visuelles analogiques (échelles continues) sont devenues les échelles les plus utilisées pour l'évaluation perceptuelle de la qualité vocale (Bele, 2005, p. 556) [9].

« *L'auditeur porte son évaluation sur une droite de 10 cm par une marque. La note attribuée à la voix correspond à la mesure chiffrée en cm entre le 0 et la croix. Cette échelle combine les avantages de l'échelle ouverte (finesse d'évaluation) et d'une échelle de classe (les valeurs sont standardisées puisque l'échelle est de dimension prédéfinie, ainsi que l'unité de mesure. Il s'agit donc d'une échelle semi-ouverte* » (Revis *et al.*, 2012, p. 61; Revis, 2004, p. 74) [91] [89].

Les échelles utilisées pour le protocole mesurent toutes 10 cm et les extrêmes sont nommés (voir Figure XXIII en annexe). Nous nous sommes inspirés notamment de l'étude de Bele (2005, p. 572) [9].

Des qualifications libres (impression générale, attaque, timbre et comportement vocal) sont également encouragées.

La mesure des attaques évaluées comme soufflées est faite à partir du 0 (équilibré), en négatif.

Cette façon pour le patient de juger de la clarté de son timbre sur une échelle visuelle analogique est utilisée notamment dans le protocole basique de la Société Européenne de Laryngologie, où le patient note son impression de la qualité de sa voix sur une échelle visuelle analogique (Dejonckere *et al.*, 2001, p. 81; Remacle & Morsomme, 2009, p. 130) [23] [86]. Laurence Delacroix et ses collègues, pour l'autoévaluation dans le cas d'une paralysie récurrentielle, citent une échelle visuelle analogique où le patient doit juger de sa qualité de voix : « *le patient doit se situer sur une ligne de dix*

⁵⁵ Borg G. (1998) *Borg's Perceived Exertion and Pain Scales*. Champaign, IL: Human Kinetics.

centimètres « comment considère-t-il sa voix ? » 0 étant une voix tout à fait mauvaise jusqu'à 10 où le patient trouve sa voix tout à fait bonne » (Delacroix *et al.*, 2009, p. 426) [24].

Pour Kempster⁵⁶ (in Revis, 2004, p. 83) [89] « il apparaît que la plupart des auditeurs appliquent une stratégie d'écoute semblable en utilisant trois dimensions principales : intensité, fréquence fondamentale et perturbations ». Dans notre cas, les perturbations sont évaluées sur les échelles « attaque » et « clarté du timbre ». Nous avons ajouté les paramètres « richesse en harmoniques » et « geste vocal ». Nous avons différencié le timbre en deux échelles « clarté du timbre » et « richesse en harmoniques » car, la plupart des patients inclus dans l'étude ayant une dysphonie légère voire pas de dysphonie, l'échelle clarté seule n'était pas la plus pertinente. Françoise Estienne propose au patient, lors de l'évaluation subjective de la voix, d'évaluer, notamment, la qualité de sa voix, selon une échelle discrète à 5 niveaux (Estienne, 1998, p. 51) [27], qui regroupe à la fois des qualités de hauteur tonale (trop aiguë, trop grave), des qualités de timbre (souffle, raucité, grésillement, nasonnement, voix mal posée) et des qualités de geste vocal (fatigue, sensations désagréables) (Estienne, 1998, p. 51) [27].

Banque de sons

Dans le but de contrôler les références des patients et du jury lors de leurs évaluations, nous proposons des références vocales (Kreiman *et al.*, 1993) [54]. « Un protocole d'ancrage auditif permet une amélioration potentiellement significative de la fiabilité des évaluations perceptives de la voix dans les pratiques cliniques et expérimentales » (Revis, 2004, pp. 86–87) [89]. En cela, nous nous sommes inspirés notamment de l'étude de Wolfe et Steinfatt, dans laquelle 5 juges devaient évaluer 25 spectrogrammes. Les 5 juges ont d'abord appris la classification par la présentation de spectrogrammes modèles, qui continuaient d'être utilisés comme référence durant la tâche de classification, d'une durée d'une heure environ (Wolfe & Steinfatt, 1987, p. 233) [114].

Le protocole nécessite de faire entendre (et voir pour tous sauf les patients du groupe protocole la première fois) des sons qui constituent des ancrages. Il y a 3 exemples pour chacune des deux échelles « clarté » et « richesse », déclinés en une voix d'homme et une voix de femme, ce qui représente 12 sons dans la banque de sons, disponibles sur le site internet⁵⁷. Les exemples pour la richesse du timbre ont été générés en filtrant les harmoniques aiguës sur une voix chantée claire⁵⁸. Chan et Yiu concluent que les ancrages par des extraits de voix synthétisés peuvent être plus efficaces, pour les caractères souffle et raucité, que les extraits de voix naturelles (Chan & Yiu, 2002) [14]. Cependant il s'avère

⁵⁶ Kempster G, Kistler D. et al. (1991) Multidimensional scaling analysis of dysphonia in two speaker groups. *Journal of Speech and Hearing Research*, 34, 534-543

⁵⁷ Les exemples de timbres altérés sont tirés des extraits vocaux de «La voix dans tous ses états » (Foisneau *et al.*, 2003) [33]. Nous avons extraits les [a] tenus en voix medium, éliminé les enregistrements trop faibles, saturés, avec trop d'instabilité ou une trop grande composante dysarthrique. Lorsque c'était nécessaire, nous avons coupé les enregistrements à 5s avec un fondu de fermeture avec le logiciel Audacity. Tous les extraits ont été classés selon leur grade de dysphonie de 0 à 3, pour les hommes et pour les femmes, en écoutant le son et en regardant le spectrogramme. Une fois les enregistrements voulus choisis (note fondamentale à peu près similaire et image spectrographique parlante), les enregistrements sont tous recalibrés à 74dB avec Praat et une animation est créée à partir de l'image spectrographique fournie par Vocevista.

⁵⁸ Extraits de voix normalisé avec Audacity, puis 4 fois le filtre passe bas : Pente de -6dB de 1000 à 2000 Hz pour le timbre pauvre en harmoniques (femme), et 2 fois le filtre passe bas pente -6dB de 400 à 1000 Hz pour le timbre très pauvre en harmoniques (femme). Filtre passe bas pente de -6dB dB de 300 à 9000Hz pour les hommes (appliqué 1 fois pour la voix pauvre en harmoniques, et trois fois pour la voix très pauvre en harmoniques). Normalisation à 74 dB avec Praat et création de l'animation à partir de l'image spectrographique fournie par Vocevista.

que nos extraits synthétisés pour les ancrages de richesse vocale ne sont pas complètement satisfaisants au niveau perceptif (Revis, 2004, pp. 87–88) [89]. En effet, il a été difficile de baisser significativement le nombre d’harmoniques aiguës tout en gardant un caractère vocal au son. Ainsi, pour la voix de femme, l’exemple « très pauvre en harmoniques » est peu différent perceptivement de l’exemple « pauvre en harmoniques ».

Conditions d’enregistrement et d’écoute

Le même matériel (micro serre-tête ME-3 Sennheiser, haut-parleur WM-Audio, carte son externe PureAudio USB d’Andrea) a été utilisé tout au long de l’étude (sauf une fois pour le patient D où on a utilisé les haut-parleurs de sa chaîne Haute-Fidélité pour écouter les extraits servant d’ancrages suite à un problème technique).

Le micro serre-tête (comme proposé par Volin, 1999, p. 244) [108] permet d’atténuer les effets du « filtrage par la salle », qui, d’après Emile Leipp, intervient dans la perception du timbre (Leipp, 2010, p. 156) [62]. Le micro utilisé doit être de bonne qualité (Ghio, 2012, p. 77) [38]. Nous avons utilisé un micro recommandé aux Etats-Unis pour sa précision lorsqu’on l’utilise pour la reconnaissance vocale. On a également coupé toute correction d’intensité induite par l’ordinateur ou la carte son (Martin, 2008, pp. 143–144; Volin, 1999, p. 244) [68] [108].

Conditions de passation

La fiche des consignes est disponible en annexe. Des pré-tests ont été effectués en septembre et octobre 2012 en consultation de phoniatrie avec Isabelle Marié-Bailly avant d’adapter les consignes et la feuille de passation.

En moyenne⁵⁹, l’exercice a duré 22 minutes, avec environ 8 minutes d’explications, d’écoute des ancrages et d’essai avant la production des deux sons, puis 4 minutes pour l’évaluation sans feedback et 10 minutes avec ajout des feedbacks auditifs et visuels.

Condition de l’évaluation perceptive

Les auditeurs du jury d’écoute ont eu accès aux sons en format .wav sur un site Internet dédié. Les écoutes sont faites avec un casque fermé Sennheiser HD200 et la carte son externe PureAudio USB d’Andrea. Pour les auditeurs du jury d’écoute, la première condition de feedback correspond à l’écoute des deux sons, la deuxième à l’ajout du spectrogramme à cette écoute.

Les différents exercices sont présentés dans un ordre aléatoire. Les auditeurs ont le loisir d’écouter les extraits autant qu’ils le désirent. Ils font apparaître le spectrogramme en cliquant sur une partie de la page web pour passer à l’évaluation avec le feedback additionnel. Ils peuvent alors réécouter l’extrait, ou non. Les auditeurs du jury d’écoute ont également l’accès libre à la banque de sons pour l’ancrage des échelles richesse et clarté du timbre.

⁵⁹ Calculé sur 19 passations mesurées sur les 23 (les 4 non mesurées concernent toutes les patients hors protocole)

4.2.3.2 *Méthode d'analyse des résultats de l'exercice d'auto-analyse*

Mesures / vérifications / recherche des valeurs aberrantes.

Toutes les séances sont enregistrées sur un dictaphone numérique. Cela a permis de préciser les mesures lorsque les marques sur les exercices d'auto-analyse étaient ambiguës.

Toutes les mesures ont été rentrées sous Excel. Après avoir calculé les différences entre deux sons et les différences induites par le feedback pour les deux sons, les données ont été vérifiées en mesurant ces écarts sur les feuilles de passation. De nombreuses erreurs ont ainsi été repérées (échange entre deux sons, entre les deux feedbacks, erreurs de mesure, erreurs de frappe). Seuls les occurrences où les deux sons sont situés au même endroit et où le feedback ne modifie rien n'ont pas pu être vérifiées de cette manière⁶⁰.

Calcul du paramètre « qualité d'attaque »

La qualité de l'attaque correspond à la valeur absolue de la mesure de l'attaque sur l'échelle bidirectionnelle soufflée / équilibrée / dure. Le 0 correspond à une attaque équilibrée, et la mesure en valeur absolue correspond à l'écart entre l'attaque observée et une attaque équilibrée, l'attaque observée pouvant être indifféremment dure ou soufflée.

Mesures objectives

Tous les extraits vocaux des exercices d'auto-analyse ont été découpés en 2 sons distincts avec Audacity. Afin de mesurer différentes caractéristiques acoustiques sur les 60 extraits de voix, nous avons écrit un script Praat donnant automatiquement dans un tableau (et réduisant les risques d'erreurs humaines) : dB moyen, fréquence fondamentale moyenne⁶¹, jitter (instabilité à court terme de la f_0), shimmer (instabilité à court terme de l'amplitude)⁶², HNR (Harmonic to Noise Ratio = rapport entre l'énergie du bruit et l'énergie de l'onde moyennée), et différence d'énergie dans la bande de fréquences 60 Hz - 2000 Hz par rapport à la bande de fréquences 2000 Hz – 5000 Hz. Devant la complexité de l'objectivation de la qualité de l'attaque, et malgré la possibilité d'une méthode de mesure (Revis *et al.*, 2000, p. 293) [88], nous avons préféré renoncer à mesurer objectivement la qualité de l'attaque.

Homogénéisation des évaluations du jury

Nous avons centré les réponses des auditeurs du jury de façon à ce que pour chaque paramètre et chaque auditeur, la moyenne et l'écart-type observés soient ceux observés pour l'ensemble des auditeurs du jury d'écoute. Cette modification des données harmonise les différentes évaluations subjectives.

⁶⁰ Cela représente seulement 1% des cas pour l'échelle intensité, 2% pour l'échelle richesse du timbre, 8% pour l'attaque, 10% pour la clarté. Cela devient cependant significatif pour la hauteur (56%) et le geste vocal (49%)

⁶¹ L'intensité n'étant pas calibrée, cette mesure seule n'est pas valide. Cependant nous avons utilisé un micro serre-tête de façon à ce que la distance à la bouche ne varie pas pendant l'exercice d'auto-analyse, ce qui nous permet de mesurer la différence d'intensité entre les deux sons émis

⁶² jitter et shimmer étant calculé sur la totalité de l'extrait, la durée du son pris en compte n'est pas harmonisée et cela remet en cause la validité de ses mesures

L'échelle « attaque » étant bidirectionnelle, nous avons uniquement modifié les variables pour que chaque groupe de données, par auditeur, ait le même écart-type, sans modifier la moyenne, de façon à ce que la valeur « 0 » garde le sens « attaque équilibrée ».

Choix des paramètres étudiés

Pour les échelles intensité et hauteur, la différence entre les deux valeurs est retenue comme la valeur la plus significative (plutôt que la valeur de chacun des deux sons) : effectivement, la consigne est de faire la plus grande différence d'intensité et la moins grande différence de hauteur. Pour les échelles qualité d'attaque, clarté, richesse et geste, la valeur brute est retenue comme la plus significative. Effectivement la consigne est de produire une attaque la plus équilibrée possible, un son riche et clair et un geste adapté.

Variables gratification / aide / erreur / corrélation

Gratification

Le meilleur résultat attendu pour les patients (le plus gratifiant) est d'obtenir une différence d'intensité plus grande, une différence de hauteur plus petite, une attaque plus proche de l'attaque équilibrée, une plus grande clarté, une plus grande richesse, et un meilleur comportement vocal.

Lorsque l'ajout du feedback change l'évaluation d'un paramètre, nous nous intéressons au fait qu'il la modifie de façon gratifiante pour le patient, ou non. Nous avons construit avec Excel pour chaque échelle les variables « gratifiantes » et « dévalorisantes » en reprenant la mesure de la différence mesurée entre les évaluations sans feedback additionnel et celles avec feedback additionnel.

Chaque fois qu'une mesure de départ manque (il est arrivé qu'un patient ou un auditeur ne note rien sur l'échelle visuelle analogique), il est important de neutraliser cette mesure, pour qu'elle ne s'apparente pas à un « 0 », ce qui n'aurait pas de sens, notamment lorsque la mesure sert pour calculer des écarts.

Les écarts en mm proposés dans les tableaux ci-après ont été calculés en faisant la moyenne des différences entre les variables gratifiantes et dévalorisantes. Nous proposons également le pourcentage d'occurrences pour lesquelles le feedback a modifié, dans un sens ou dans l'autre, le paramètre mesuré.

Erreur / aide

Le même cheminement nous conduit à calculer le nombre de fois où le feedback aide à s'approcher de la médiane des évaluations harmonisées des auditeurs du jury d'écoute (évaluation avec feedback visuel).

Les calculs ont été effectués avec les valeurs non harmonisées pour tous les patients (il n'était pas possible de les harmoniser car il y a trop peu de passations par patient), et avec les valeurs harmonisées pour le jury.

Résumé utilisation / gratification / aide / erreur

On évalue l'effet du feedback additionnel sur les plans

- de l'utilisation : nombre de fois en % où le feedback induit un changement sur au moins un des sons de l'échelle considérée,

- de la gratification (en positif) ou la dévalorisation (en négatif) constatée entre la condition sans et avec feedback additionnel
 - o en mm (moyenne gratification - pénalisation),
 - o en % (nombre de fois où le changement induit par le feedback additionnel va dans le sens d'une gratification du patient sur le nombre total de fois où le feedback induit un changement),
- de l'erreur : la moyenne des valeurs absolues des écarts entre l'évaluation du patient et la médiane des évaluations des auditeurs du jury d'écoute ,
- de l'aide (en positif) ou la gêne (en négatif) constatée entre la condition sans et avec feedback additionnel
 - o en mm (moyenne aide - gêne),
 - o en % (nombre de fois où le changement induit par le feedback additionnel va dans le sens de l'aide sur le nombre total de fois où le feedback induit un changement).

Corrélations

On étudie également les modifications des corrélations entre les différents paramètres lorsqu'on ajoute le feedback. Les modifications sont supposées significatives lorsque $p < 0,05$ (noté par *) et très significatives lorsque $p < 0,01$ (noté par **) dans les corrélations par paires (Pearson).

Nature du feedback additionnel et nombre de données

Le feedback additionnel n'est pas toujours le même : pour les patients en auto-évaluation, on ajoute les feedback auditifs et visuels, sauf pour la première passation des patients du groupe protocole pour lesquels on ajoute uniquement le feedback auditif. Pour les auditeurs du jury d'écoute, la différence observée correspond à ce que les auditeurs ont pu modifier lorsqu'ils observent le spectrogramme en plus d'écouter le son. Dans tous les cas, la différence est due à un feedback externe additionnel : pour les patients, son et spectrogramme, et pour les auditeurs du jury d'écoute, spectrogramme.

		Auditeurs du jury d'écoute	Patients			
			Hors protocole		protocole	
				Sauf 1 ^{ère} passation	1 ^{ère} passation	
Condition feedback 1	(variables)	Son enregistré	Son qui vient d'être produit (feedback interne)			
Condition feedback 2		Son enregistré et spectrogramme			son enregistré	
Nombre de sons	270	224	38	22	16	8
Nombre de différences de 2 sons	135	112	19	11	8	4

Tableau VII : recueil des données : nombre d'évaluations par groupe

Remarque : les moyennes ont été faites sans corriger le poids respectif de chaque groupe (voir Tableau VII) : les données pour « tous » sont donc très dépendantes des données provenant du jury d'écoute (224 occurrences contre 38). De même, le poids des patients hors protocole est sensiblement plus fort que celui des patients du groupe protocole pour le total « patients » (22 occurrences contre 16). Enfin, le nombre d'occurrences où l'on ajoute seulement le son est très faible (4 différences de deux sons), ce qui nous interdit de généraliser nos résultats.

Fiabilité des évaluations du jury d'écoute pour l'exercice d'auto-analyse

Les fiabilités intra-juge et inter-juges sont mesurées par des statistiques non paramétriques (rang signés de Wilcoxon et test de Kruskal-Wallis) où l'on recherche une non significativité ($p > 80\%$ par exemple) de l'indépendance des auditeurs. Pour avoir une idée de la fiabilité intra-juge, nous avons introduit dans la procédure d'écoute 5 enregistrements qui sont écoutés deux fois. Nous comparons les évaluations des deux écoutes.

L'effet d'apprentissage est évalué en mesurant la corrélation (Pearson) entre les paramètres mesurés et le numéro de l'exercice.

5 Analyse des résultats

5.1 Auto-analyse : échelles visuelles analogiques

5.1.1 Utilisation des échelles

5.1.1.1 Distribution des marques sur les échelles visuelles analogiques

Nous nous intéressons ici aux évaluations toutes conditions de feedback confondues, avant l'harmonisation des mesures des évaluations du jury d'écoute.

Remarque préliminaire : un des juges a noté l'échelle de la clarté non pas comme une échelle entre timbre clair et timbre altéré, mais comme une échelle entre timbre clair et timbre sourd, ce qui conduit à une très grande corrélation entre cette échelle et celle de la richesse (corrélation = 0,88 au lieu de 0,38 ; 0,05 et 0,09 pour chacun des autres auditeurs) ; par la suite nous avons neutralisé ses réponses afin de garder une échelle de clarté qui soit bien distincte de celle de la richesse, puisque nous avons l'assurance que tous les patients ont bien fait la distinction, les passations se faisant en notre présence.

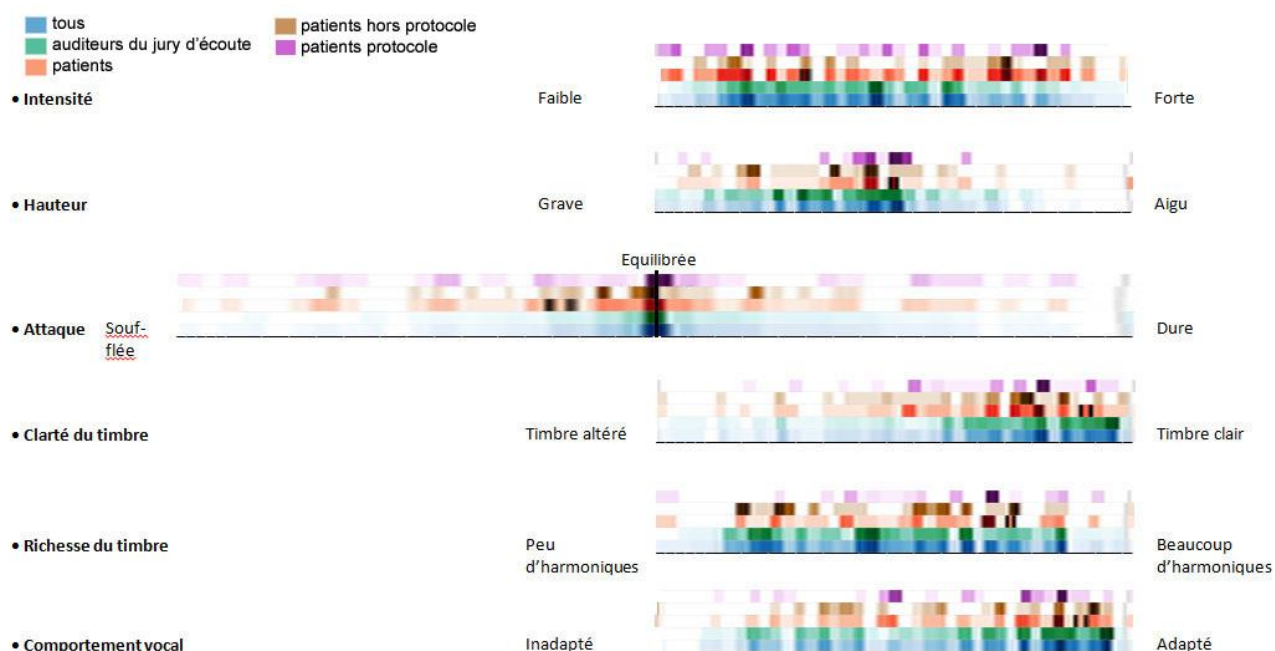
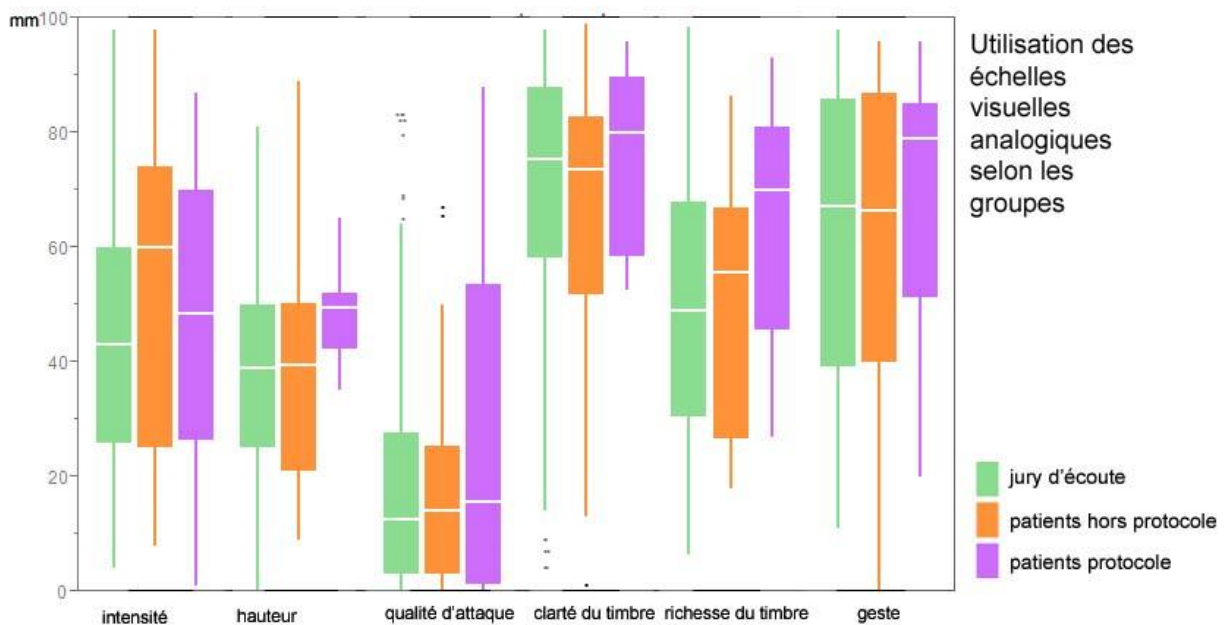


Figure II : reproduction de la feuille de passation et distribution des réponses pour chacune des échelles tous sons confondus et tous feedbacks confondus, et pour chaque groupe.



**Figure III : Utilisation des échelles visuelles analogiques selon les groupes
(pour la qualité d'attaque : 0 = attaque équilibrée)**

Ce sont les patients du groupe protocole qui s'éloignent le plus de la façon d'évaluer du jury (voir Figure III et Figure IV). Ils se distinguent particulièrement en jugeant toujours leur hauteur de façon très compacte vers le milieu de l'échelle. En outre ils utilisent l'échelle d'attaque plus souvent avec de plus mauvaises qualité d'attaque, et ils pensent que leur timbre est plus riche que les patients hors protocole et que le jury d'écoute.

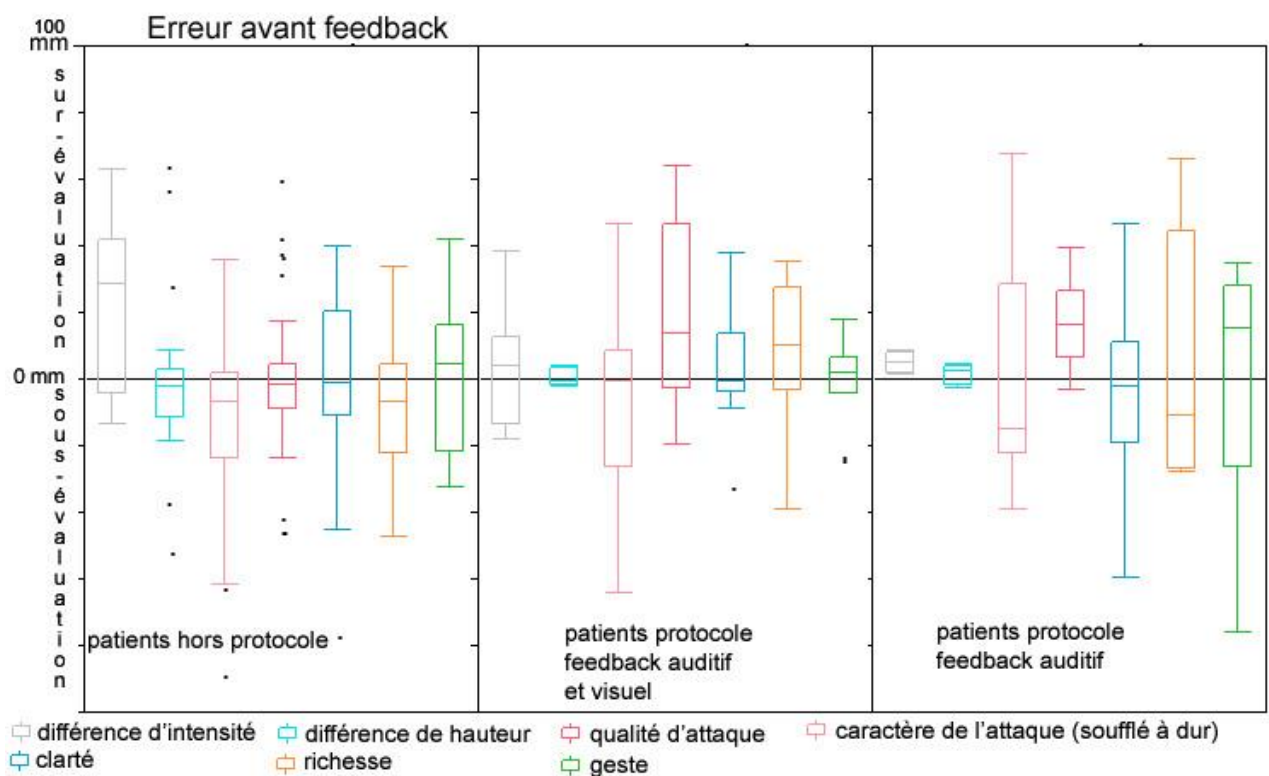


Figure IV: erreur avant feedback pour les patients. Pour la qualité d'attaque, l'échelle est inversée car on représente la mauvaise qualité d'attaque (0 = attaque équilibrée)

On note que les patients hors protocole surévaluent leur différence d'intensité et leur clarté de timbre et sous-estiment la richesse de leur timbre, tandis que les patients du groupe protocole évaluent bien leur différence de hauteur, sous-estiment leur qualité d'attaque et surestiment leur richesse de timbre, ceci par rapport aux évaluations des membres du jury d'écoute.

5.1.1.2 Corrélations entre les différentes échelles (tous groupes confondus et tous feedbacks confondus)

Le **Erreur ! Source du renvoi introuvable.** donne les corrélations menant à définir les deux pôles : clarté / geste / qualité d'attaque » et « richesse intensité / caractère de l'attaque ».

Pôle clarté / geste / qualité d'attaque

La clarté est fortement corrélée au geste (**0,77** ; $p < 0,0001^{**}$), car un geste vocal sain est souvent synonyme d'une bonne qualité vocale. La clarté n'est pas corrélée au caractère de l'attaque (soufflée à dure), mais faiblement corrélée à la mauvaise qualité de l'attaque (**-0,31** ; $p < 0,0001^{**}$) : effectivement la qualité de l'attaque (qu'elle soit soufflée ou dure) a un retentissement sur la qualité du timbre de la voyelle tenue. La clarté est faiblement corrélée à la richesse (0,27 ; $p < 0,0001^{**}$).




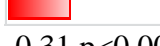




Corrélation	clarté	richesse	intensité
clarté		0,27 $p < 0,0001^{**}$	0,07 $p = 0,13$
geste	 0,77 $p < 0,0001^{**}$	 0,30 $p < 0,0001^{**}$	0,06 $p = 0,16$
hauteur	 0,30 $p < 0,0001^{**}$	0,24 $p < 0,0001^{**}$	0,14 $p = 0,0002^{**}$
mauvaise qualité d'attaque	 -0,31 $p < 0,0001^{**}$	0,15 $p = 0,0005^{**}$	0,25 $p < 0,0001^{**}$
richesse	0,27 $p < 0,0001^{**}$		 0,71 $p < 0,0001^{**}$
intensité	0,11 $p = 0,02^*$	 0,71 $p < 0,0001^{**}$	
Caractère de l'attaque (soufflée à dure)	0,07 $p = 0,13$	 0,39 $p < 0,0001^{**}$	 0,51 $p < 0,0001^{**}$

Tableau VIII : corrélations entre les échelles de clarté, richesse, attaque et intensité.

Pôle richesse /intensité / caractère de l'attaque (soufflée à dure)

La richesse est fortement corrélée à l'intensité (**0,71** ; $p < 0,0001^{**}$) et au caractère de l'attaque (soufflée à dure) (**0,39** ; $p < 0,0001^{**}$) beaucoup plus qu'à la qualité de l'attaque. Tout se passe comme si une intensité plus forte était synonyme d'une attaque dure et une bonne richesse de timbre.

L'intensité est corrélée au caractère de l'attaque (soufflée à dure) (**0,51** ; $p < 0,0001^{**}$) et dans une moindre mesure à la qualité de l'attaque (0,25 ; $p < 0,0001^{**}$). Cela s'explique par le fait qu'une intensité faible et une attaque soufflée sont parfois liées et qu'une intensité forte se traduit souvent par une attaque dure (Steinhauer et al., 2004, p. 433) [103].

La richesse est également faiblement corrélée au geste (**0,30** ; $p < 0,0001^{**}$), une bonne richesse de timbre pouvant être liée à un geste vocal sain.



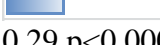

Corrélation	geste
clarté	 0,77 $p < 0,0001^{**}$
richesse	 0,30 $p < 0,0001^{**}$
hauteur	 0,29 $p < 0,0001^{**}$
mauvaise qualité d'attaque	 -0,28 $p < 0,0001^{**}$
intensité	0,06 $p = 0,16$
attaque	-0,02 $p = 0,67$

Tableau IX : corrélations entre les des différentes échelles avec l'échelle geste

Intensité et hauteur sont très faiblement corrélées au global (0,14 ; $p=0,0017^{**}$) au contraire de ce qu'on aurait pu attendre.

Geste

En plus de sa relation avec la clarté, dont nous avons déjà parlé, le geste est corrélé négativement avec la mauvaise qualité d'attaque (voir Tableau IX) : effectivement une attaque équilibrée est liée à un geste vocal sain.

Hauteur

Du fait de la grande subjectivité de l'échelle de hauteur, il est difficile d'interpréter ces résultats. Elle est corrélée à la clarté (0,30 ; $p<0,0001^{**}$) à la richesse et au geste (0,30 et 0,29 ; $p<0,0001^{**}$). Cela tendrait éventuellement à montrer que les notes (relativement) plus hautes sont en moyenne plus faciles à produire, plus riches et plus claires.

5.1.1.3 Synthèse des corrélations

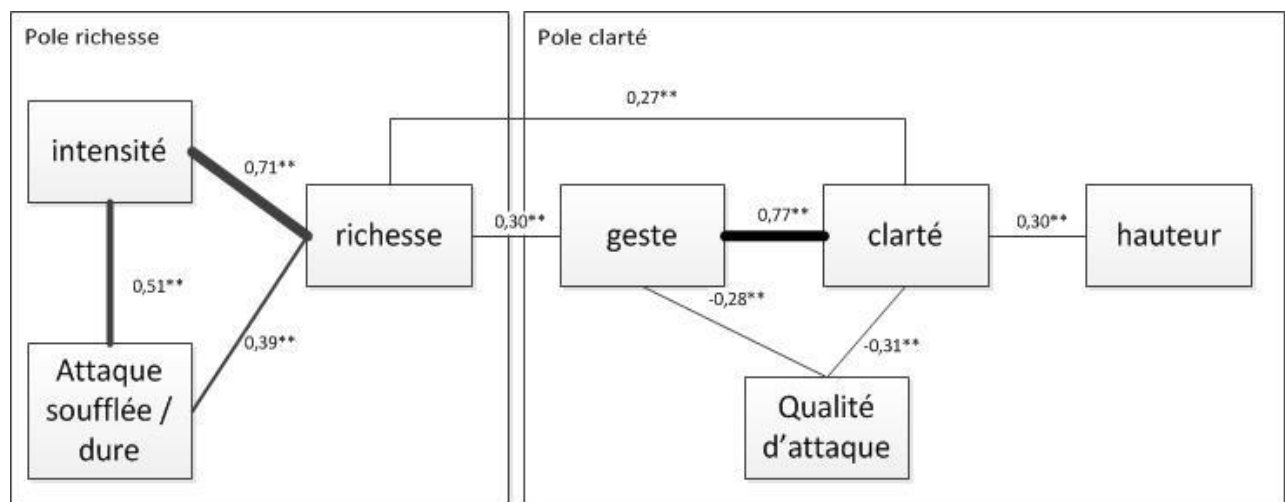


Figure V : synthèse des corrélations tous groupes et tous feedbacks confondus

5.1.2 Fiabilité de l'évaluation subjective

La fiabilité des échelles visuelles analogiques est plus difficile à obtenir (Wuyts *et al.*, 1999) [115].

5.1.2.1 Fiabilité intra-juge

On cherche ici à obtenir des évaluations a et b dépendantes (première et deuxième évaluation d'une même paire de sons), c'est-à-dire avec un p grand (voir Tableau X).

Le fait de centrer les variables n'apporte pas plus de fiabilité intra-juge, ce qui était attendu puisque l'harmonisation se fait juge par juge.

L'ajout du feedback visuel au feedback auditif apporte une plus grande fiabilité intra-juge. Cet effet est particulièrement sensible pour les échelles **d'intensité, de caractère d'attaque** (et qualité de l'attaque), et de **clarté**.

Conclusion sur la fiabilité intra-juge

On observe pour que l'ajout du feedback visuel au feedback auditif fiabilise les notes pour **l'intensité, le caractère de l'attaque et la clarté**.

Effet d'apprentissage supposé : test/retest

Lorsque la fiabilité intra-juge est moins bonne, nous pouvons supposer l'existence d'un effet d'apprentissage lors du passage de l'évaluation par les auditeurs, puisque les sons qui ont été jugés deux fois ont tous été présentés une première fois au début du test et une deuxième fois à la fin du test. Cela tendrait à montrer que les auditeurs se basent beaucoup sur l'image pour corriger leur impression auditive et qu'ils apprennent pendant le test à juger différemment l'intensité, le caractère de l'attaque et la clarté à l'oreille. Ce serait cette éducation auditive qui expliquerait la différence de jugement entre la première et la deuxième écoute du même son.

La fiabilité intra-juge est largement améliorée par l'ajout du feedback visuel. Cependant elle reste faible pour la qualité de l'attaque, la richesse et le geste vocal, ce qui s'explique éventuellement par un effet d'apprentissage.

5 comparaisons a / b	Valeurs brutes		Valeurs harmonisées				Corrélation avec le numéro de passation (apprentissage)	
	Fb auditif		Fb auditif et visuel					
Wilcoxon	p	Z	p	z	p	Z	Corr.	p
différence d'Intensité	0,3803	-0,877	0,4009	-0,840	0,9108	-0,112	-0,21	0,0017**
différence de hauteur	0,8860	-0,140	0,8860	-0,140	0,7794	-0,280		
type d'attaque	0,5391	-0,614	0,4561	-0,745	0,9687	-0,039		
mauvaise qualité d'attaque	<0,0001	-4,193	<0,0001	-4,226	0,0262	-2,223	-0,19	<0,0001**
clarté	0,0242	-2,254	0,0345	-2,114	0,6688	-0,428	0,16	0,0006**
richesse	0,0261	-2,225	0,0190	-2,346	0,0554	-1,915		
Geste vocal	0,0011	-3,254	0,0005	-3,482	0,0008	-3,361		

Tableau X : fiabilité intra-juge selon les différents feedbacks et effet d'apprentissage

Effet d'apprentissage

Pour rechercher un effet d'apprentissage, nous mesurons les corrélations entre les variables et le numéro de l'évaluation (elles sont passées dans l'ordre).

- **Qualité de l'attaque** (corrélation -0,19 ; $p < 0,0001^{**}$) : les auditeurs du jury d'écoute ont jugé les attaques légèrement meilleures au fil de la passation. Or nous verrons (paragraphe 5.1.4) que l'attaque est justement un des paramètres dont l'appréciation évolue favorablement grâce au feedback visuel additionnel. Il est donc probable ici que les auditeurs ont été influencés par le feedback visuel et ont appris pendant la passation à adapter leurs sensations auditives à ce qu'ils pensaient voir par la suite sur le spectrogramme, c'est-à-dire une attaque plus équilibrée que ce qu'ils auraient jugé en premier lieu.
- **Geste vocal** (corrélation 0,16 ; $p = 0,0006^{**}$) : le geste vocal a légèrement tendance à être de mieux en mieux noté. **Concernant la différence de geste vocal entre les deux sons**

(corrélation -0,22 ; p=0,0169*), ces différences ont tendance à être jugées de moins en moins grandes.

- **Intensité** (corrélation -0,21 ; p=0,0017**) : les auditeurs ont tendance à modifier de moins en moins leur jugement d'intensité grâce au feedback visuel additionnel.

Il n'y a pas d'autres corrélations significatives entre le numéro de l'évaluation et les différents autres critères.

5.1.2.2 Fiabilité inter-juges

28 évaluations par 4 auditeurs (3 pour la clarté)	Valeurs brutes		Valeurs centrées			
	Feedback auditif		Feedback auditif et visuel			
Kruskal-Wallis	p	H	p	H	p	H
Différence d'Intensité	0,2294	4,315	0,5471	2,124	0,620	1,750
Différence de Hauteur	0,2703	3,920	0,3148	3,547	0,3163	3,534
Attaque	<0,0001	23,8	<0,0001	22,3	0,0108	11,180
Valeur absolue attaque	0,0020	14,8	0,0117	11	0,0049	12,869
Clarté	<0,0001	56,629	0,6730	0,792	0,8684	0,282
richesse	<0,0001	30,8	0,8376	0,850	0,9316	0,441
Geste vocal	<0,0001	66,0	0,8269	0,9	0,9616	0,291

La fiabilité des évaluations avec le feedback visuel ajouté au feedback auditif est toujours meilleure que celle des évaluations avec feedback auditif seul. (voir Tableau XI)

Le fait de centrer les variables améliore grandement la fiabilité inter-juges pour les variables « clarté », « richesse » et « geste vocal », et dans une moindre mesure « différence d'intensité ».

Tableau XI : fiabilité inter-juges avant / après harmonisation et selon le type de feedback

Le fait d'harmoniser les variables augmente peu la fiabilité pour la différence de hauteur. Cela est certainement dû au faible nombre de valeurs différentes de 0 pour cette variable. La fiabilité n'est pas non plus augmentée pour l'attaque. Cela peut être dû au caractère bidirectionnel de cette échelle, qui nous a interdit de centrer les mesures (nous avons seulement harmonisé les écarts-type).

Conclusion sur la fiabilité inter-juges de la mesure subjective par des auditeurs du jury d'écoute

La fiabilité inter-juges est bonne lorsqu'on harmonise les variables, ce qui permet de s'affranchir quelque peu des ancrages internes des auditeurs. La fiabilité inter-juges est largement améliorée par l'ajout du feedback visuel, ce qui est en accord avec l'étude de Dejonckere montrant que la fiabilité de l'évaluation subjective (GRBAS) est améliorée par l'ajout du spectrogramme au son entendu (Dejonckere, 2006) [22]. Cependant il faut rester prudent : qu'il soit plus facile de s'accorder sur des normes à partir d'une image qu'à partir d'un son ne garantit aucunement que l'image soit plus représentative de la voix que le son. Le Tableau XII synthétise les fiabilités pour chacune des échelles.

	Fiabilité intra-juge	Fiabilité inter-juges	remarques
Différence d'Intensité	très bonne	62 %	
Différence de Hauteur	bonne	32 %	La différence de hauteur est très souvent minime voire nulle
Caractère de l'attaque	très bonne	1 %	La moyenne de cette variable bidirectionnelle n'a pas été harmonisée, car nous avons voulu conserver le sens « équilibré » à la valeur « 0 »
Qualité d'attaque	mauvaise	0,5%	
Clarté	moyenne	87 %	L'ajout du feedback visuel n'améliore pas la fiabilité pour la différence entre deux sons.
Richesse	mauvaise	93 %	
Geste vocal	mauvaise	96 %	L'ajout du feedback visuel n'améliore pas la fiabilité pour la différence entre deux sons.

Tableau XII : conclusion sur les fiabilités intra- et interjuge

5.1.3 Corrélation avec les critères acoustiques

Le script Praat donne des résultats parfois aberrants pour le jitter pour les extraits à faible intensité, car il interprète des parties du silence avant et après le son tenu comme ayant une certaine périodicité. A défaut de retraiter chacun des sons manuellement, nous nous contentons de ces résultats. Cela explique en partie les difficultés à lier le jitter avec un des paramètres subjectifs. Shimmer et jitter ne sont pas corrélé significativement avec les évaluations de richesse et de clarté du timbre.

On observe (Tableau XIII) une corrélation (Pearson) significative entre NHR (signal / bruit) et clarté (0,2907 ; $p < 0,0001^{**}$) et entre la mesure du singing formant et la richesse du timbre (0,6003 ; $p < 0,0001^{**}$).

		NHR / clarté		Singing formant / richesse	
		Feedback 1	Feedback 2	Feedback 1	Feedback 2
Jury		0,2657 $p=0,0005^{**}$	0,3063 $p < 0,0001^*$	0,5773 $p < 0,0001^{**}$	0,6826 $p < 0,0001^{**}$
Patients		Pas de corrélation	0,4599 $p=0,0037^*$	0,5071 $p=0,0008^*$	0,5165 $P=0,0009^*$
hors protocole			0,3399 $p=0,1217$	0,4896 $p=0,0543$	0,307 $P=0,1646$
Proto-cole	Feedback auditif et visuel	0,4603 $p=0,0236^*$	0,7607 $p=0,0006^{**}$	0,5151 $p=0,01^{**}$	0,7479 $P=0,0009^{**}$
	Feedback auditif		Pas de corrélation		Pas de corrélation

Tableau XIII : corrélations entre les critères subjectifs de clarté et de richesse et les mesures objectives selon les groupes et le type de feedback

Lorsque le feedback additionnel est uniquement auditif (mesuré sur 8 évaluations seulement), les corrélations entre NHR et clarté et entre singing formant et richesse ne sont pas significatives. En revanche, lorsque le feedback additionnel comprend le spectrogramme, les corrélations deviennent plus fortes, sauf pour les patients hors protocole, pour lesquels ses corrélations ne sont pas significatives.

5.1.4 Influence de l'image sur la perception de la qualité de la réalisation

5.1.4.1 Utilisation du feedback additionnel

Généralités

On mesure ici le nombre d'auto-évaluations où les patients dans leur ensemble (sauf la première passation des patients du groupe protocole, pour laquelle le spectrogramme n'est pas utilisé) ont modifié au moins un des deux sons sur les différentes échelles visuelles (Figure VI). Le feedback externe constitué du spectrogramme et du son amène à modifier plus de 7 fois sur 10 au moins un des deux sons pour l'échelle de clarté, et presque 7 fois sur 10 pour l'échelle de richesse. On note que ce pourcentage est bien plus élevé que pour la première passation des patients du groupe protocole, qui, eux, modifient systématiquement l'attaque et le geste en réentendant leur voix, mais la clarté et la richesse seulement dans 50% des cas. On peut donc penser que c'est le spectrogramme qui donne des informations pertinentes aux patients sur la clarté et la richesse de leur timbre.

Un test rang de Wilcoxon (Tableau XIV) entre les données avant / après feedback nous permet également de quantifier la significativité de l'indépendance des évaluations avant / après feedback. On recherche ici une indépendance entre les évaluations avant et après feedback, c'est-à-dire un p petit ($p < 0,01$).

p		intensité	hauteur	Caractère de l'attaque	Qualité d'attaque	clarté	richesse	Geste
jury		0,003	0,17	0,0004	0,02	0,37	0,004	0,80
hors protocole		1,00	0,37	0,12	0,01	0,01	0,14	0,24
Proto- cole	feedback auditif et visuel	0,29	-	0,06	0,50	0,31	0,72	0,83
	feedback auditif	0,18	-	0,35	0,50	0,65	0,11	0,18

Tableau XIV : significativité de l'indépendance des évaluations avant et après feedback selon les groupes. La différence pour la hauteur n'est pas significative pour les patients du groupe protocole, sans qu'on puisse calculer p, car les patients de ce groupe n'ont jamais utilisé le feedback additionnel pour modifier leurs évaluations

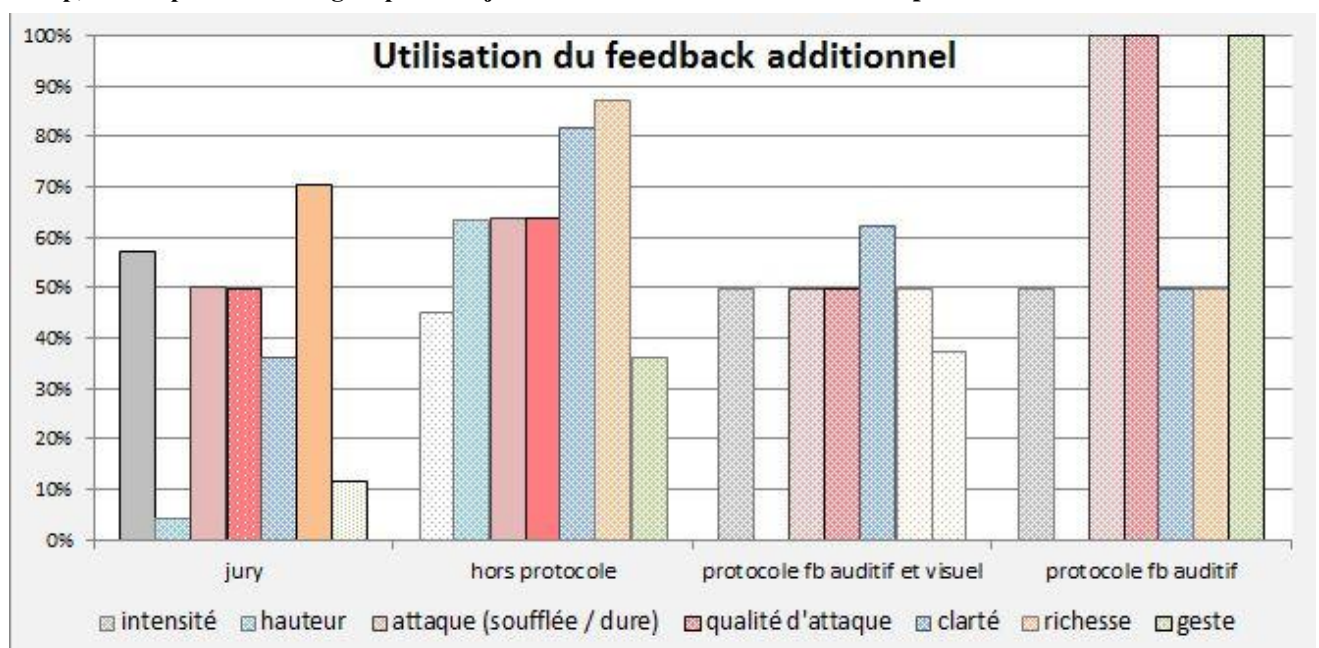


Figure VI : Utilisation du feedback additionnel. Le remplissage est plein lorsque la différence entre les évaluations avant et après feedback est significative ($p < 0,01$)

Dans 3 cas sur les 11 (27%), les patients hors protocole n'ont pas pu évaluer la richesse de leur timbre avant le feedback auditif et visuel. Pour les 8 cas restants, le feedback est utilisé à 88% c'est-à-dire 7 fois sur les 8. Au total, le feedback auditif et visuel apporte un gain 10 fois sur 11 (91%). Dans un cas sur les 8, une patiente n'a pas évalué la clarté avant le feedback auditif et visuel, ce qui augmente le nombre de cas où le feedback auditif et visuel modifie l'évaluation.

Pour le feedback auditif et visuel, effet de l'appartenance à un groupe (jury d'écoute, patients hors protocole, patients protocole)

Les auditeurs du jury d'écoute se distinguent par une beaucoup plus faible utilisation du spectrogramme pour modifier leurs évaluations sur l'échelle de la clarté, du geste et de la hauteur. En revanche ils l'utilisent sensiblement plus pour l'échelle de l'intensité.

Les patients hors protocole ont globalement plus modifié leurs évaluations. L'exercice était d'ailleurs en moyenne plus long pour ces patients.

5.1.4.2 Effet gratifiant du feedback additionnel

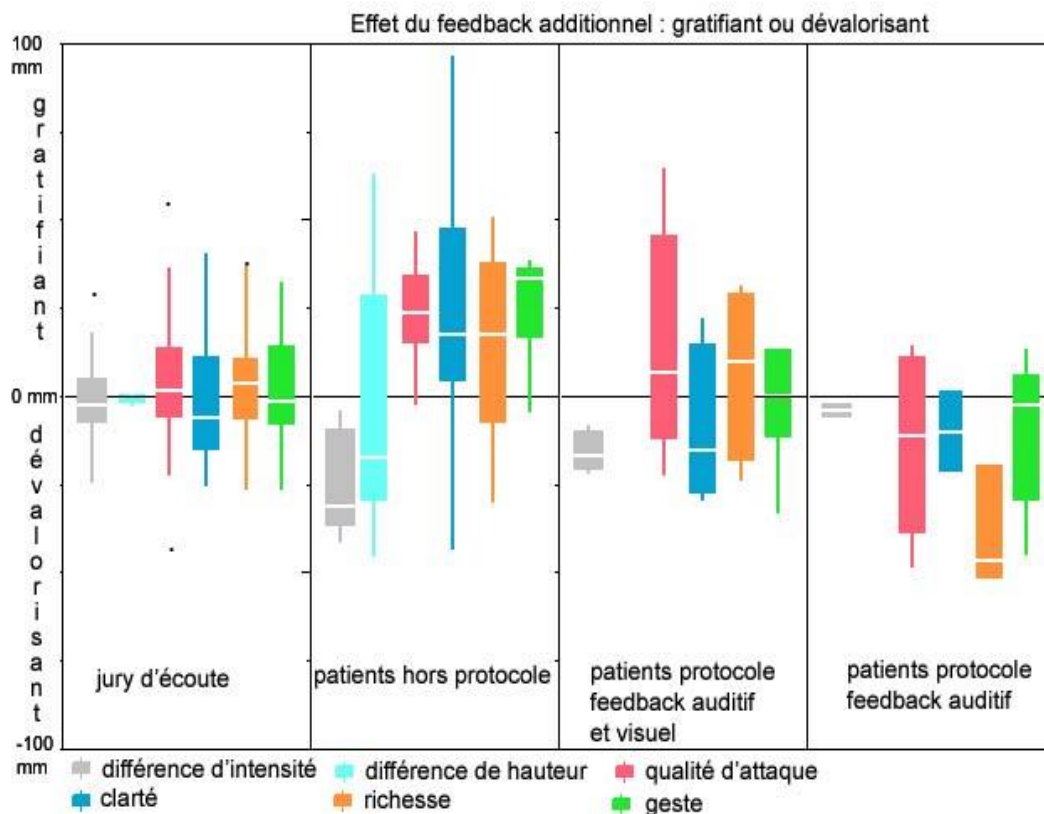


Figure VII : effet gratifiant ou dévalorisant du feedback additionnel

Sur la Figure VII, les effets gratifiants sont en positifs et les effets dévalorisants en négatif

Prise de conscience d'un défaut

Pour les patients, les modifications apportées à l'échelle « **intensité** » s'éloignent toujours de l'objectif de l'exercice : après avoir pensé avoir fait des sons bien différenciés en terme d'intensité, les patients révisent leur jugement lorsqu'ils bénéficient du feedback (auditif et visuel). La hauteur également est modifiée, pour les patients hors protocole, qui réalisent également, en moyenne, que leurs deux sons ne sont pas toujours à la même hauteur, contrairement à l'objectif de l'exercice et ce qu'ils avaient perçu sans feedback additionnel. Cet effet n'existe pas chez les patients du groupe protocole, qui ont un meilleur contrôle de leur voix.

Prise de conscience d'une qualité, confiance en soi :

Dans tous les groupes, le feedback avec spectrogramme permet de juger de façon plus bienveillante sa qualité d'attaque et dans une moindre mesure la richesse de son timbre. Pour les patients hors protocole, la clarté et le geste sont impactés positivement également.

Pour le feedback auditif et visuel, effet de l'appartenance à un groupe (jury d'écoute, patients hors protocole, patients protocole)

Les auditeurs du jury d'écoute sont moins impactés par cet effet de gratification. On l'observe uniquement, dans une faible mesure, pour la richesse du timbre. On peut penser qu'ici le facteur personnel joue beaucoup dans le changement d'attitude face à un son (Lescarmontier, 2012) [63].

Les comportements des patients du groupe protocole sont plus proches en moyenne des comportements des auditeurs du jury d'écoute que ceux des patients hors protocole.

Différence pour les patients du groupe protocole, entre la première passation sans spectrogramme, et celles avec spectrogramme

L'effet du feedback uniquement auditif est toujours en moyenne dévalorisant (sauf pour la hauteur qui n'est jamais impactée par le feedback). Il est même systématiquement dévalorisant pour les échelles « intensité » et « richesse ».

La variable « richesse » est significativement modifiée lorsque le feedback additionnel est auditif et visuel plutôt que seulement auditif. Les patients jugent la richesse de leur timbre meilleure alors que ce n'était pas le cas sans le spectrogramme. La qualité de l'attaque bénéficie également de ce mouvement vers une évaluation gratifiante.

5.1.4.3 Aide du feedback additionnel

La Figure VIII présente l'aide (en positif) ou la gêne (en négatif) représentées par le feedback additionnel dans le cadre de la différence de deux sons pour l'intensité et la hauteur, et pour les autres paramètres pour les sons eux-mêmes. Par définition, le feedback additionnel aide lorsque la modification apportée permet à la différence des deux sons de se rapprocher de la différence observée par les auditeurs du jury d'écoute, pour l'intensité et la hauteur, ou lorsqu'il permet aux évaluations des sons eux-mêmes d'être plus proches de la médiane de celles des auditeurs du jury d'écoute.

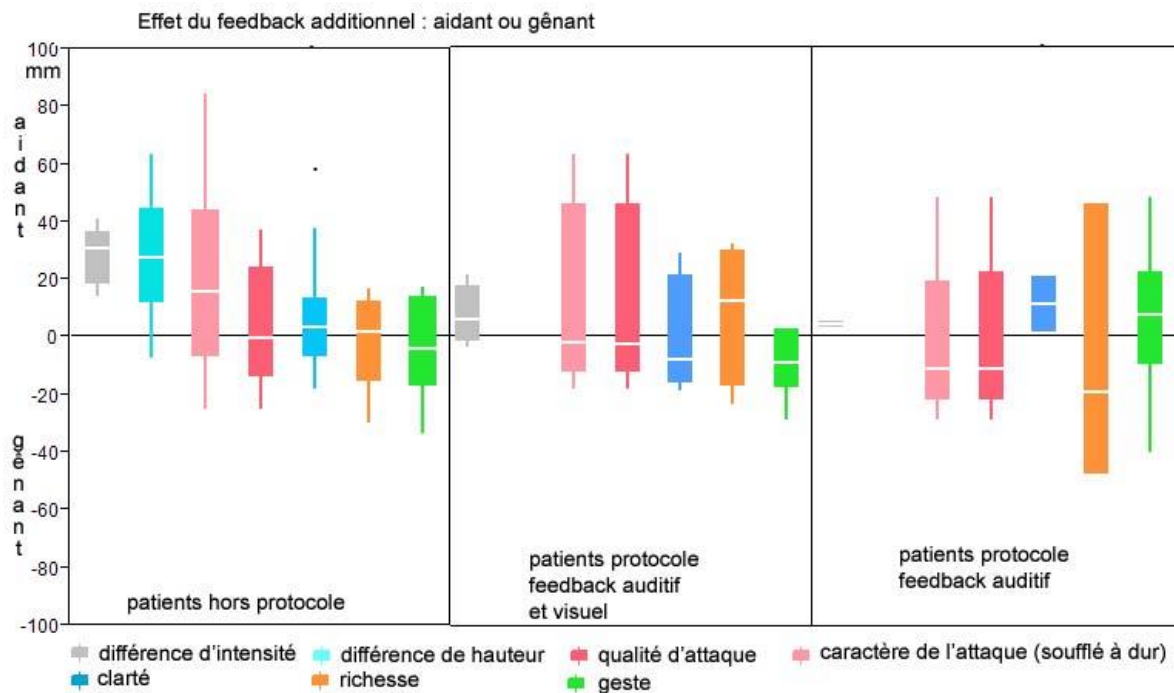


Figure VIII : effet aidant ou gênant du feedback additionnel selon les groupes

En moyenne, lorsque l'ajout du feedback auditif et visuel modifie les évaluations, il permet de s'approcher de l'évaluation médiane (avec feedback visuel) des auditeurs du jury d'écoute. Cela est particulièrement vrai pour la différence d'intensité, et, pour les patients hors protocole, pour la différence de hauteur, le caractère de l'attaque (soufflée à dure) et dans une moindre mesure la clarté du timbre.

La comparaison des passations avec et sans spectrogramme pour les patients du groupe protocole tend à montrer que c'est le spectrogramme qui permet d'évaluer plus justement la richesse du timbre. En revanche il est plutôt gênant pour la clarté du timbre et pour le geste.

5.1.4.4 Mise en perspective erreurs / utilisation / aide / gratification

Richesse, clarté et qualité d'attaque

Richesse du timbre	Aidant ou gênant (notes brutes)		Gratifiant ou dévalorisant		Utilisation (modification de la différence entre deux sons en %)	
	mm	%	mm	%		
Jury	3,5	73%	2,1	64%	71%	
patients	1,1	64%	4,5	57%	69% => 84 %	
Hors protocole	-1,1	60%	8,2	60%	88% => 91%	
Pro- cole	Feedback auditif et visuel	3,3	75%	0,7	50%	50%
	Feedback auditif	-2,5	33%	-14,6	50%	50%

Tableau XV : aide, gratification et utilisation du feedback additionnel pour l'échelle richesse

Les patients hors protocole ont en moyenne davantage sous-évalué la richesse de leur timbre que les patients du groupe protocole (Figure IV). C'est pourquoi, en moyenne, les feedback additionnels sont plus gratifiants que dans le cas des patients du groupe protocole.

<i>Clarté du timbre</i>		Aidant ou gênant (notes brutes)		Gratifiant ou dévalorisant		Utilisation (modification de la différence entre deux sons en %)
		Mm	%	mm	%	
Jury		3,4	79%	-0,2	40%	36%
patients		3,7	57%	8,5	65%	74% => 89 %
Hors protocole		6,1	63%	17,6	81%	82% => 91 %
Proto- cole	Fb auditif et visuel	0,4	43%	-3,9	50%	63% => 75 %
	Fback auditif	5,1	100%	-0,1	0%	50%

Tableau XVI : aide, gratification et utilisation du feedback additionnel pour l'échelle clarté

L'effet est encore plus net pour la clarté du timbre (Tableau XVI). Les patients hors protocole utilisent beaucoup le feedback pour corriger l'échelle « clarté », et dans 63 % des cas, la modification est gratifiante pour eux, avec une moyenne de 17,6 mm vers un timbre peu altéré, qui est la plus grande modification de cette échelle. La modification permet également à leurs évaluations de s'approcher de celles du jury d'écoute. Dans une moindre mesure, le feedback aide les patients du groupe protocole à revenir au niveau du jury d'écoute. La différence entre ces deux groupes provient surtout du fait que les patients du groupe protocole surestiment en moyenne beaucoup plus la clarté de leur timbre que les patients hors protocole (Figure IV).

<i>Qualité d'attaque</i>		Aidant ou gênant (notes brutes)		Gratifiant ou dévalorisant		Utilisation (modification de la différence entre deux sons en %)
		mm	%	mm	%	
Jury		2,3	68%	1,8	56%	50%
patients		3,3	47%	9,5	74%	58%
Hors protocole		1,5	50 %	12,4	92%	64%
Proto- cole	Fb auditif et visuel	5,8	43%	5,6	43%	50%
	Fback auditif	-1,2	20%	-1,9	20%	100%

Tableau XVII : aide, gratification et utilisation du feedback additionnel pour l'échelle qualité d'attaque

Pour l'attaque (Tableau XVII), le même effet est de nouveau observé : les patients hors protocole surestiment la qualité de leur attaque lorsqu'ils utilisent le feedback auditif et visuel par rapport aux évaluations des auditeurs du jury d'écoute. Cet effet n'est pas retrouvé pour les patients du groupe protocole.

<i>Différence d'intensité</i>		Aidant ou gênant (différence)		Gratifiant ou dévalorisant		Utilisation (modification de la différence entre deux sons en %)
		mm	%	mm	%	
Jury		0,5	50 %	-0,5	42%	57%
patients		3,5	71 %	-11,1	0%	47%
Hors protocole		1,5	75 %	-13,5	0%	45%
Proto- cole	Fb auditif et visuel	6,2	67 %	-7,7	0%	50%
	Fback auditif	-1,0	0%	-4,0	0%	50%

Tableau XVIII : aide, gratification et utilisation du feedback additionnel pour la différence d'intensité

On observe systématiquement que lorsqu'il est utilisé pour l'échelle intensité (environ une fois sur deux), le feedback additionnel entraîne les patients à estimer que les deux sons qu'ils ont émis sont

plus proches en intensité qu'ils ne l'avaient d'abord jugé (Tableau XVIII). Le feedback auditif et visuel tend à compenser l'erreur faite au départ, lorsque, juste après avoir réalisé l'exercice, les patients pensent avoir réalisé deux sons bien distincts en termes d'intensité, ce qui était demandé. L'effet dévalorisant du feedback est le plus fort constaté parmi toutes les autres variables et tous groupes confondus.

On note particulièrement une différence dans l'erreur commise par les patients (16,3 mm en moyenne pour la différence entre les sons) avant le feedback par rapport à l'erreur du jury (7,6 mm en moyenne). Cela montre que le fait d'entendre le son, pour les auditeurs du jury d'écoute, permet de bien qualifier l'intensité des sons entendus.

Geste		Aidant ou gênant (notes brutes)		Gratifiant ou dévalorisant		Utilisation (modification de la différence entre deux sons en %)
		mm	%	mm	%	
Jury		0,8	71%	0,1	44%	12%
patients		-2,3	38%	3,9	69%	37%
Hors protocole		-1,5	43%	7,1	86%	36%
Proto -	Fb auditif et visuel	-3,4	33%	-0,5	50%	38%
cole	Fback auditif	8,2	80%	-12,4	20%	100%

Tableau XIX : aide, gratification et utilisation du feedback additionnel pour l'échelle geste

Les auditeurs du jury d'écoute ont globalement peu utilisé cette échelle (Tableau XIX). Les patients l'ont utilisée d'avantage, mais la plupart du temps, la modification de l'évaluation les éloignait de l'évaluation des auditeurs du jury d'écoute.

L'effet positif pour le jury d'écoute est un effet de construction de la mesure (les auditeurs du jury d'écoute s'approchent forcément, en moyenne et pour les notes brutes, de la médiane des auditeurs du jury d'écoute). En revanche, l'effet positif constaté pour les patients du groupe protocole lorsqu'ils n'utilisent pas le spectrogramme (première passation) tend à montrer que les patients sont gênés par l'image pour juger au mieux de leur geste vocal. Cependant on peut nuancer cette hypothèse car pour la première passation l'erreur faite par les patients (28,7 mm, c'est la plus grande erreur constatée) est beaucoup plus importante que lors des passations suivantes.

Différence de hauteur		Aidant ou gênant (différence)		Gratifiant ou dévalorisant		Utilisation (modification de la différence entre deux sons en %)
		mm	%	mm	%	
Jury		0,1	33%	0,0	60%	4%
patients		3,1	73%	-1,0	43%	37%
Hors protocole		5,3	73%	-1,7	43%	64%
Protocole		Non utilisé				0%

Tableau XX : aide, gratification et utilisation du feedback additionnel pour la différence de hauteur

Les patients protocole n'ont jamais modifié leur évaluation de la hauteur (Tableau XX). Les patients du groupe protocole ont presque toujours émis des sons à la même hauteur, les données ne sont donc pas pertinentes ici. De même les auditeurs du jury d'écoute ont fait des erreurs minimales à l'écoute, ce qui entraîne des données difficilement exploitables en termes d'apport du spectrogramme.

En revanche, les patients hors protocole ont fait des erreurs d'appréciation importantes (22 mm en moyenne), qui ont nécessité l'aide du feedback, permettant de corriger en moyenne 5,3 mm de cette erreur.

5.1.4.5 Corrélations entre les évaluations des différents paramètres

La Figure IX présente les corrélations entre les différents paramètres, pour les auditeurs du jury d'écoute (en violet) et pour les patients (en vert). Les doubles flèches indiquent l'évolution de la corrélation entre la première et la deuxième condition de feedback.

Effet du feedback visuel (jury d'écoute) / du feedback auditif et visuel (patients)

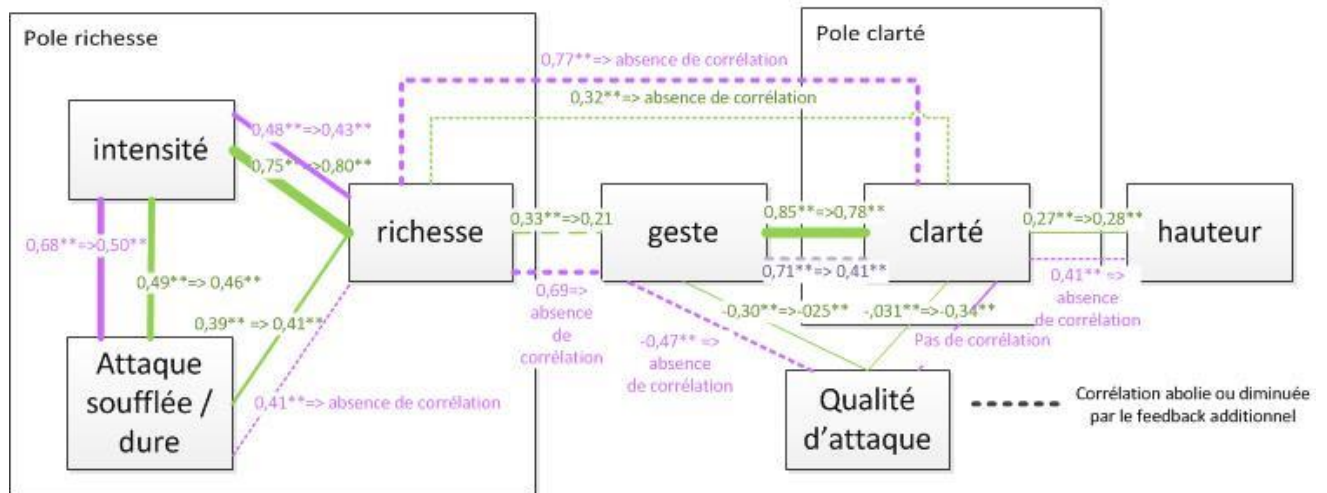


Figure IX : effet du feedback additionnel sur les corrélations en fonction du groupe

Richesse et intensité

Le feedback externe abolit presque toutes les corrélations à la richesse du timbre, sauf celle de l'intensité. Le feedback semble permettre aux patients de distinguer ce qui relève de l'attaque de ce qui relève du son tenu, ce qui fait chuter la corrélation (**0,41**** à **0,28****) entre richesse et attaque soufflée / dure. Le même effet s'observe dans la corrélation entre richesse et clarté (**0,77**** sans feedback, aucun avec le feedback auditif et visuel). On suppose que le fait d'objectiver par l'image, d'une part la richesse et d'autre part la clarté, permet de dé-corréler ces deux paramètres.

La corrélation entre l'intensité et la richesse du son s'explique par le fait qu'on visualise davantage d'harmoniques lorsque le son est fort, mais également parce qu'un son fort implique un quotient de fermeture glottique augmenté, ce qui d'après Leipp augmente le nombre d'harmoniques « *plus le rapport entre durée d'obturation et d'ouverture de la glotte est grand, plus le timbre comportera d'harmoniques* » (Leipp, 2010, p. 283) [62].

Intensité et attaque

Les patients ont plus tendance à lier intensité et attaque dure (**0,68****), au risque de les confondre. Le feedback auditif et visuel remet le niveau de la corrélation (**0,50****) au niveau de celle des auditeurs du jury d'écoute (**0,46****). En même temps, le feedback auditif et visuel rend la corrélation significative et forte pour les patients entre une intensité forte et une mauvaise qualité d'attaque (**0,38****), ce qui, de nouveau, met la corrélation au niveau de celle des auditeurs du jury d'écoute (**0,35****).

Intensité et clarté

Les données de la littérature laissaient attendre que le son fort serait plus difficile à produire que le son faible. Les auteurs décrivent un quotient de fermeture glottique augmenté et un effort vocal plus important pour les sons d'intensité forte (Dejonckere *et al.*, 2001, p. 79; Giovanni & Aumelas, 2004, pp. 34, 41) [23] [40]. Le contraire a été observé dans notre étude : l'exercice d'auto-analyse a été l'occasion par deux fois d'une prise de conscience qu'il était plus facile d'émettre une voix claire à intensité plus forte. De plus, on ne trouve pas de corrélation importante entre les paramètres clarté et intensité (corrélations : 0,11 ; $p=0,02^*$).

Clarté et hauteur

Les patients lient beaucoup la clarté avec la hauteur (**0,41****) avant de bénéficier d'un feedback auditif et visuel (0,28). Celui-ci modifie beaucoup leur jugement et a tendance à rapprocher leur comportement de celui des auditeurs du jury d'écoute. On peut penser que c'est le feedback auditif ici qui joue le plus, en permettant aux patients de prendre conscience de la hauteur du son qu'ils ont produit.

Isolation de l'évaluation du geste

Pour les patients, le feedback externe fait beaucoup baisser la corrélation (de **0,71**** à 0,41**) entre le geste et la clarté. L'effet existe aussi chez les auditeurs du jury d'écoute, mais celui-ci garde tout de même une corrélation de **0,78****. Ici le feedback externe conduit les patients à s'éloigner du comportement des auditeurs du jury d'écoute.

Pour tous les groupes, l'ajout du feedback fait perdre toute significativité de la corrélation entre le geste et la richesse du timbre. L'effet le plus frappant est observé chez les patients pour lesquels l'évaluation du geste avant feedback découle directement de celle de la richesse (corrélations **0,69**** sans feedback, sans corrélation significative avec). Le fait de disposer des feedbacks entraîne également une moins bonne corrélation entre le geste et la clarté du timbre, pour les patients (**0,71**** à 0,41**), et dans une moindre mesure pour les auditeurs du jury d'écoute (**0,85**** à **0,78****).

5.1.4.6 Conclusion

	Aidant ou gênant		Gratifiant ou dévalorisant	
	Hors protocole	Protocole	Hors protocole	protocole
Différence d'intensité	++	+	--	--
Différence de hauteur	++	0	-	0
Caractère (soufflé / dure) de l'attaque	+	(+)		
Qualité de l'attaque	(+)	(+)	++	+
Clarté	0	0	++	-
Richesse	0	0	+	(+)
Geste	-	--	++	0

Tableau XXI: impact du feedback additionnel : résumé

Les variables les plus impactées par un feedback auditif et visuel sont la richesse, la clarté et la qualité d'attaque. Ces différentes variables sont jugées plus finement indépendamment les unes des autres. Les feedbacks auditifs et visuels rapprochent les corrélations observées dans les variables évaluées par les patients de celles observées dans les variables évaluées par les auditeurs du jury d'écoute.

L'ajout des feedbacks externes permet de se valoriser (pour la richesse du timbre, la qualité d'attaque pour tous les patients, pour la clarté et le geste pour les patients hors protocole). Quand ce n'est pas le cas (différence d'intensité et de hauteur), cela s'explique par une meilleure évaluation de ces paramètres, en particulier pour les patients hors protocole.

5.2 Évolution des patients du groupe « protocole »

5.2.1 Bilan vocal

5.2.1.1 Évaluation objective

Les graphiques présentant les résultats sont disponibles en annexe. (Figure XIV à Figure XXII). Le Tableau XXII synthétise les résultats.

	Patiente A	Patiente B	Patiente C	Patiente D
Dynamique d'intensité	-	--	--	-
Étendue vocale	++	+	+	++
Jitter	+	+	++	+
Shimmer	0	0	-	-
Rapport signal sur bruit	0	0	+	+
Évaluation du singing formant	-	+	++	+
Rapport des TMP z / s	-	0	-	-
TMP sur /a/	-	-	+	+

Tableau XXII : évolution de l'évaluation objective des patients du groupe protocole

+ représente une amélioration du paramètre : par exemple, un jitter moindre, ou un TMP sur [a] plus grand, 0 une absence d'évolution – une involution.

L'évaluation post-protocole montre pour toutes les patientes une meilleure étendue vocale et un jitter moins important. Le rapport signal sur bruit a une tendance à l'amélioration. Cependant, ces données sont tempérées par d'autres paramètres qui se détériorent : dynamique d'intensité, rapport z/s, shimmer. On note une augmentation de la mesure du singing formant pour toutes les patientes sauf la patiente A. Pour l'étendue vocale, on fait l'hypothèse que les conditions de passation (sirène ascendante) influencent les résultats dans le cadre de patientes chanteuses qui n'ont pas de restriction de leur étendue vocale. Il nous a semblé que le fait d'être plus en confiance la deuxième fois permettait d'atteindre des notes plus aiguës, que les patientes semblaient ne pas se permettre la première fois.

En conclusion, on observe un effet bénéfique du travail effectué lors du protocole, assez bien distribué selon les patientes, avec cependant des résultats contrastés suivant les paramètres retenus, et un bénéfice plus limité pour la patiente A.

5.2.1.2 Évaluation perceptive

Les évaluateurs ont été gênés par les conditions d'enregistrement qui étaient différentes d'un enregistrement à l'autre. On ne note pas d'évolution significative (Tableau XXIV). Les voix parlées ont été évaluées légèrement moins bien en post-protocole et les 3 attaques plutôt mieux sauf pour la patiente B. Cependant, la voix d'appel est plutôt améliorée en fin de protocole, et la patiente D est systématiquement mieux notée en fin de protocole. On ne peut pas affirmer que ces observations résultent directement et exclusivement de notre travail auprès des patients.

	Accord interjuge de Kruskal-Wallis		Patiente A		Patiente B		Patiente C		Patiente D	
			Moyenne	Ecart-type	Moyenne	Ecart-type	Moyenne	Ecart-type	Moyenne	Ecart-type
	p	H								
G	0,6861	1,483	0,4	0,3	1,7	0,7	1	0,4	0,1	0,3
R	0,5847	1,941	0,1	0,3	0	0	0,3	0,3	0	0
B	0,6459	1,660	0,1	0,3	1,6	0,8	1	0,7	0,4	0,3
A	0,7288	1,301	0	0	1,1	0,3	0,6	0,8	0	0
S	0,5395	2,162	0,4	0,3	0,6	0,9	0,4	0,5	0	0
I	0,3426	3,336	0,6	0,6	1,1	0,8	0,8	0,5	0,1	0,3

Tableau XXIII : accord inter-juges pour le GRBASI

Evaluation subjective des progrès	Accord interjuge (Kruskal-Wallis)		Patiente A		Patiente B		Patiente C		Patiente D	
			Moyenne	Ecart-type	Moyenne	Ecart-type	Moyenne	Ecart-type	Moyenne	Ecart-type
	p	H								
Voix parlée	0,6263	1,748	-0,1	0,3	-0,3	0,5	-0,5	0,6	0	0,8
Voix d'appel	0,5562	2,079	1,5	1	-0,3	0,5	1,1	0,9	1,3	1
3 attaques	0,2714	3,910	0,5	1,3	-1,3	0,5	0,5	1	0,5	0,6

Tableau XXIV : accord inter-juges pour l'évaluation subjective des progrès

5.2.2 Évolution de l'exercice d'auto-analyse

Les graphiques présentant les résultats sont disponibles en annexe (Figure XVIII à Figure XXII). Les mesures objectives (jitter / shimmer / rapport signal sur bruit / mesure du singing formant), observées sur le son le plus fort de chaque exercice d'auto-évaluation, montrent une évolution positive pour le paramètre « rapport signal sur bruit », qui augmente 6 fois sur 8, et reste quasiment constant une fois. On ne trouve pas d'effet d'apprentissage vis-à-vis du spectrogramme. On ne trouve pas d'évolution systématique vers un exercice mieux réussi, ou des sons plus adaptés. On ne trouve pas d'évolution systématique de l'erreur, ni de l'aide ni de la gratification.

5.2.3 Modification de la façon de voir sa voix

5.2.3.1 Évolution des questionnaires d'auto-évaluation

À la première évaluation, c'est surtout la voix parlée qui a été évaluée, du fait notamment que certains auteurs comme Rosen et Murry (2000) notent que le VHI n'est pas sensible pour les sujets dysodiques (pour lesquels la plainte concerne la voix chantée). Ils relèvent 5 items pertinents sur les 30.

Après les séances du protocole et du fait que la population du protocole était au moins autant dysodique que dysphonique, il est apparu nécessaire de faire une distinction entre les deux usages. Les scores sont plutôt meilleurs pour le VHI-10, mais souvent moins bons pour le questionnaire de relation à sa voix (Tableau XXV). En particulier les réponses « j'ai peur que ma voix me trahisse » sont systématiquement moins bonnes à l'issue du protocole.

En revanche, 3 patientes sur 4 sont passées, pour l'affirmation « je dois forcer pour produire la voix » de « parfois » à « jamais », et 3 patientes sur 4 sont moins tracassées par leur problème de voix. La quatrième (patiente D) avait répondu avant le protocole qu'elle ne se tracassait jamais pour son problème de voix, elle a modifié son jugement après le traitement concernant la voix chantée. Cela

s'explique par le fait qu'elle explore sa voix en parallèle de façon très précise du fait de son changement d'orientation professionnelle.

		Avant		Après	
		voix parlée	voix chantée	voix parlée	voix chantée
Patiente A	VHI-10	11	?	6	8
	Relation à sa voix	3		4	
	% de satisfaction	80 %			
Patiente B	VHI-10	21	?	15	9
	Relation à sa voix	11	10	8	5
	% de satisfaction	40 %	50 %	50 %	60 %
Patiente C	VHI-10	25	?	24	?
	Relation à sa voix	6		8	9
	% de satisfaction	60 % (10 - 100)		65 %	
Patiente D	VHI-10	2		0	5
	Relation à sa voix	3		2	3
	% de satisfaction	80 %	65 %	90 %	78 %

La patiente B est celle qui est la plus positive à l'issue du protocole. La patiente C a des résultats au questionnaire « relation à sa voix » bien moins bons à l'issue du protocole, qui peuvent être dus à un contexte personnel indépendant de notre étude au moment de l'évaluation de fin de protocole.

Tableau XXV : évolution des questionnaires d'auto-évaluation

Les scores au VHI-10 et au questionnaire « relation à sa voix » sont cohérents : plus les patientes se sentent handicapées par leur voix, plus leur relation à leur voix est difficile (sauf pour les patientes B et C, avant le protocole). Le pourcentage de satisfaction est corrélé négativement aux scores, sauf pour la patiente B qui se donne un pourcentage de satisfaction relativement bas par rapport à ses scores aux questionnaires.

5.2.3.2 Liste d'adjectifs

	Caractéristique gagnée (donnée à l'évaluation de fin alors qu'elle n'était pas donnée à l'évaluation de début de protocole)	Caractéristique perdue (donnée à l'évaluation de début, mais pas à l'évaluation de fin de protocole)	Caractéristique gagnée par une personne et perdue par une autre	Caractéristique inchangée
Positive	Energique, ronde, nuancée claire, douée, agréable, ample, riche, émouvante, assurée sonore, sucrée, chantée, régulière, ferme, forte, spontanée, animée,	Gaie , moelleuse, chaude, caressante, attirante, belle, apaisante, douce, éclatante	puissante, musicale, calme	brillante, chaleureuse, chantante, timbrée, résonnante, sexy, décidée, naturelle, poignante, mélodieuse
Négative	agitée, effrayante, saccadée, voilée, âpre	Atone, blanche , insipide, angoissée, chuchotée, discordante, étouffée, irrégulière, mielleuse, râpeuse, enrouée, mal posée, serrée, désagréable	tendue	épillée, étranglée, faible, monotone, nasale, plate, soufflée, spasmée, terne
Neutre	métallique,	grave, haute, moyenne, mordante, pointue, humble		de soprano, aigüe, de colorature

Tableau XXVI : évolution des caractéristiques citées dans la liste d'adjectifs. Caractéristique citée par deux, par trois /

Pour les caractéristiques gagnées : **souhaitées par la personne qui l'a gagné, par au moins une autre personne**

Nous avons classé les adjectifs dans des caractéristiques « positives » systématiquement si au moins un patient l'avait cité comme caractéristique souhaitée (« la voix que j'aimerais avoir »). Nous y avons ajouté les adjectifs suivants, qui n'avaient pas été cités comme caractéristiques souhaitées, que nous avons considérés comme plutôt positives : « animée », « nuancée », « énergique », « ferme », « brillante ». Les autres adjectifs (qui n'ont jamais été cités comme souhaitables pour la voix) ont été considérés comme négatifs, à l'exclusion des termes suivants qui ont été considérés comme neutres : « grave », « moyenne », « haute », « aigue », « de colorature », « de soprano », qui correspondent à une description de tessiture, et de « humble », « métallique », « pointue », « mordante », qui peuvent être considérées alternativement comme positives ou négatives.

On constate au global 33 caractéristiques (Tableau XXVI) qui, lorsqu'elles font l'objet d'un changement, sont toujours modifiées de façon positive (une caractéristique positive gagnée ou une caractéristique négative perdue), pour 14 caractéristiques qui, lorsqu'elles font l'objet d'un changement, sont toujours modifiées de façon négative. Les patients ont donc tendance à évaluer leur voix, après les séances, avec globalement plus de qualificatifs positifs et moins de qualificatifs négatifs.

Catégorie de l'adjectif	Changement positif		Nombre		Changement négatif		Nombre	
	Adjectif positif gagné	Adjectif négatif perdu	brut	Importance relative	Adjectif positif perdu	Adjectif négatif gagné	brut	Importance relative
Intensité	Forte, sonore	Chuchotée, étouffée	4	12 %	Eclatante		1	7 %
Clarté et régularité du timbre	Claire, régulière	Enrouée, serrée, irrégulière	5	15 %		Saccadée, voilée	2	14 %
Richesse du timbre	Ronde, ample, riche,	Insipide, atone, blanche	6	18 %			0	0 %
Geste	Assurée, douée, ferme,	Mal posée	4	12 %		Agitée	1	7 %
Emotions	Agréable, émouvante, spontanée	Angoissée, discordante, désagréable, attirante	6	21 %	Gaie, attirante, apaisante	Effrayante	4	29 %
Sensations	Sucrée, ronde	Moelleuse, râpeuse	4	9 %	Moelleuse, chaude, caressante, belle, douce	âpre	6	43 %
Autre	Nuancée, animée, chantée, énergique		4	9 %			0	0 %

Tableau XXVII : changements positifs et négatifs dans les caractéristiques citées, classées selon leur nature

Le qualificatif « énergique » est le seul à faire l'objet d'un relatif consensus : il est cité après les séances comme caractérisant leur propre voix par 3 patients sur les 4, alors qu'aucun ne l'avait cité avant les séances, ni comme caractérisant sa propre voix ni comme caractérisant la voix qu'il aimerait avoir. On peut faire l'hypothèse que la visualisation du spectrogramme comme répartition de l'énergie acoustique selon les fréquences leur a donné une sensibilité plus forte à l'adjectif « énergique ».

On peut classer les différents adjectifs selon leur nature, comme l'indique le Tableau XXVII.

On constate que 8 adjectifs positifs concernant les émotions et les sensations ont été perdus. En outre, 6 adjectifs négatifs concernant les émotions et les sensations n'ont pas été repris lors de la deuxième évaluation : le pôle « adjectifs perdus concernant les sensations et les émotions » explique 57 % des changements négatifs et 18 % des changements positifs.

A l'inverse, les 6 adjectifs concernant la richesse du timbre font toujours l'objet d'un changement positif, et ils en expliquent 18 %. On peut faire l'hypothèse, sous réserve d'une étude plus poussée patiente par patiente, que les séances proposées aux patients ont pour effet :

- d'améliorer l'image qu'ont les patients de leur propre voix sur le plan de la richesse du timbre,
- de faire perdre aux patients une vision de leur voix axée sur les sensations et les émotions,
- de détériorer l'image qu'ont les patients de leur propre voix en termes d'adjectifs se rapportant aux émotions et aux sensations.

Le nombre de changements positifs est calculé comme étant la somme du nombre d'adjectifs positifs non cités avant les séances et qui sont cités après les séances d'une part, et d'autre part du nombre d'adjectifs négatifs cités avant les séances et qui ne sont plus cités après les séances. Les changements négatifs sont constitués des modifications inverses. Ils sont répertoriés dans le Tableau XXVIII et illustrés par la Figure XI et la Figure XI.

	Nombre d'adjectifs cités pour « ma voix telle qu'elle est » (adjectifs « caractérisants »)				Nombre d'adjectifs cités pour « la voix que j'aimerais avoir » (adjectifs souhaités)		Changements positifs	Changements négatifs
	Positifs		Négatifs		Avant	Après		
	Avant	Après	Avant	Après				
Patiente A	26	46	6	5	23	34	9	1
Patiente B	30	22	16	9	19	10	17	9
Patiente C	10	6	7	3	9	5	5	1
Patiente D	30	36	1	1	15	17	11	8
total							42	19

Tableau XXVIII : nombre d'adjectifs cités par les patientes avant et après protocole, et nombre de changements positifs et négatifs

On constate :

- une diminution du rapport entre le nombre d'adjectifs cités pour « ma voix telle qu'elle est » négatifs et le nombre d'adjectifs cités pour « ma voix telle qu'elle est » positifs : tous les patients jugent leur voix globalement plus positivement après les séances,
- une diminution du rapport entre le nombre d'adjectifs cités pour « la voix que j'aimerais avoir » et le nombre d'adjectifs cités pour « ma voix telle qu'elle est » positifs : les patients expriment tous moins de « frustration » face à leur voix après les séances.

La patiente D est celle qui se montre la plus satisfaite de sa voix, suivie dans l'ordre (rapports négatifs / positifs croissants, voir Figure X) par la patiente A, la patiente B et enfin la patiente C qui

est celle qui est la plus insatisfaite de sa voix : elle est la seule à donner plus d'adjectifs négatifs que positifs pour qualifier sa voix. Cette hiérarchie entre les patientes est celle qu'on retrouve aux scores au VHI-10. On constate également une tendance à une corrélation entre l'insatisfaction des patientes et leur évolution plus forte vers plus de satisfaction.

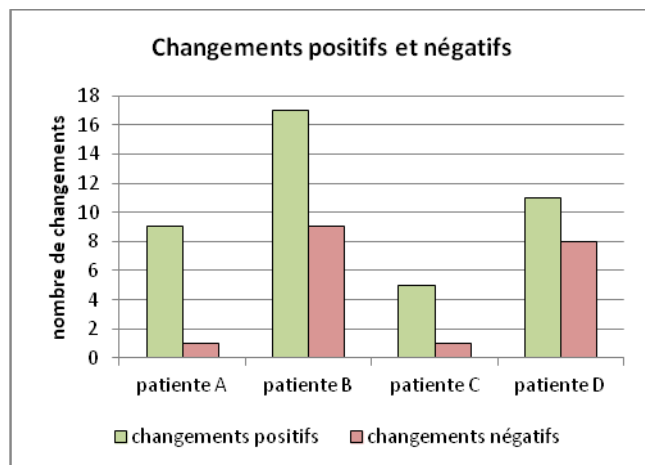
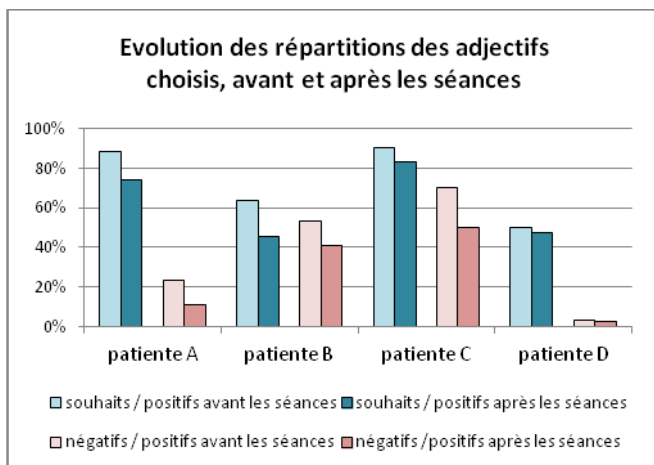


Figure X : Evolution des répartitions des adjectifs choisis Figure XI : changements positifs et négatifs dans les adjectifs choisis

5.2.3.3 Conclusion

Les questionnaires d'auto-évaluation révèlent de la part des patientes une meilleure image de leur voix en fin de protocole.

5.3 Commentaires autour du spectrogramme

5.3.1 Impact sur le geste vocal

5.3.1.1 Sensations acoustiques et proprioceptives

Réponses au questionnaire en ligne

Les 10 personnes ayant répondu au questionnaire répondent de façon plutôt homogène (écart type 0,7) à l'effet de l'image spectrographique sur l'affinage de leurs sensations auditives et proprioceptives (entre « plutôt d'accord » et « tout à fait d'accord »), sensiblement plus pour les sensations auditives. Elles sont moins catégoriques pour dire que le spectrogramme ne les coupe pas de leurs sensations auditives et proprioceptives (elles répondent en moyenne entre « je ne sais pas » et « plutôt pas d'accord » aux questions « le spectrogramme vous coupe de vos sensations auditives / sensation internes »). Au total, pour les personnes répondant au questionnaire, le spectrogramme n'a pas d'influence négative sur les sensations auditives et proprioceptives et permet de les affiner.

Possible focalisation sur l'image au détriment des sensations

Les patients ayant répondu au questionnaire précisent parfois qu'ils ont été effectivement coupés de leurs sensations : « au départ mes yeux étant rivés sur le spectrogramme tout en monitorant ma voix, il m'était difficile de me concentrer sur mes sensations, cela m'a même valu quelques fatigues vocales. Il y a eu un temps d'adaptation » ; « je trouve qu'effectivement, lorsque j'utilise le spectrogramme toute seule, je peux avoir tendance à me focaliser sur le tracé produit, et [non] plus sur ce que j'entends ». Cette patiente sous-entend que le thérapeute l'aiderait à revenir à ses sensations auditives. La patiente A est plus pessimiste sur le rôle potentiel du thérapeute. Elle pense que parfois l'outil

spectrographique induit pour le thérapeute « *quelque chose qui vient s'interposer par rapport à une attention à ce que je fais corporellement* ».

La patiente A n'arrive pas « *être dans les deux à la fois* » et « *préfère privilégier les sensations* ». Pour les cas où elle se laisse déconcentrer par l'image spectrographique elle rapporte : « *je ne savais pas ce que j'avais fait, donc [l'image] m'intéressait moins.* »

La patiente D relève aussi le cas des transitions de voyelles où « *[elle] arrivai[t] mieux en fermant les yeux en sentant ce qui se passe chez [elle] plutôt qu'en regardant* ». Toutefois, cela ne l'empêche pas de s'intéresser après coup à l'image : « *ça m'a aidée de voir l'histoire des voyelles, des passages é e u i etc, d'associer l'image sur le spectrogramme avec le son que je faisais* ».

Ouverture des sensations auditives

Pour la patiente C : « *visualiser ça donne une troisième dimension, ça ouvre les oreilles, c'est vraiment intéressant* » ; « *on entend différemment ce qu'on produit comme son, par exemple sur la qualité d'une attaque, et du coup on peut rectifier et on voit tout de suite ce que ça produit, sur ça et sur les harmoniques* ».

Une patiente écrit dans un témoignage écrit spontané : « *VV [le spectrogramme de VoceVista] m'aide à affiner mon écoute* ».

Ouverture des sensations proprioceptives

Certains patients répondant au questionnaire Internet font directement le lien entre ce qu'ils voient sur le spectrogramme et des paramètres internes : « *le spectrogramme m'est utile pour m'autocorriger. J'ai compris que ma mâchoire n'était pas assez détendue et que j'actionnais [davantage] de muscles que nécessaires pour l'articulation* » ; « *je trouve que [le spectrogramme] aide à ressentir les sensations internes, car c'est en changeant ces dernières qu'on modifie ce qui se passe sur l'écran* ».

Pour la patiente C : « *ça répond bien au geste, on voit bien le geste qu'on fait, c'est bien* » « *on voit directement à l'écran* ». Cependant la patiente A relativise : « *globalement ça correspond à peu près à la sensation que j'en ai, pas toujours exactement. Des fois j'ai l'impression que ce que je fais n'est pas propre et à l'écran vous me dites « mais si ça l'était » et des fois c'est l'inverse ou alors y'a dans l'intention ça va marcher, mais je vois bien que pas tant que ça.* ». Une patiente hors protocole ajoute, pour le niveau de souffle dans la paille : « *ça montre que ce qu'on a ressenti et ce qui est réel n'est pas forcément la même chose* ».

Aide lorsque les sensations et l'audition interne sont perturbées

La patiente D pense que « *ça aide pour les gens qui n'ont peut-être pas une bonne proprioception* ». La patiente B estime avoir un ressenti perturbé par la douleur qu'elle ressent lors de ses émissions vocales : « *le spectrogramme me permet de rationaliser mes sensations et mon audition interne qui sont en décalage avec ce qui est perçu par l'auditeur* » ; « *ça permet d'avoir une expérience scientifique de sa voix, ne plus être dans le ressenti, parce que mon ressenti est perturbé, il n'est pas fiable* ». Effectivement, pour elle, « *le son ne reflète pas la douleur ou le mécanisme qui se bloque* ».

Pour la patiente B, le feedback auditif externe est nécessaire pour une prise de conscience gratifiante : « *ce que j'entends n'est pas ce que les autres entendent* » ; « *ça a fait prendre conscience aussi que l'effort de lutte que j'avais en interne ne s'entendait pas du tout dans le son (chanté en tout cas)* ».

Pour la patiente C, sur la consonne [s] qu'elle fait un peu serrée : *« de la voir c'est vraiment plus intéressant que de juste l'entendre parce qu'on n'entend pas bien »*

5.3.1.2 Temps réel

Résultats du questionnaire

Les résultats sont disparates pour la question du temps réel. En effet, chaque patient a des réactions différentes devant le fait que l'image s'affiche en même temps que l'on fait l'exercice. De plus, nous faisons l'hypothèse que la réponse peut être différente suivant les exercices. Les réponses ne sont pas toujours consistantes (certaines préfèrent ne regarder le spectrogramme ni en temps réel, ni après coup). On obtient tout de même un très léger avantage pour le temps réel.

Besoin de se concentrer sur le geste vocal

Ce besoin avait été pointé en séance et lors de l'entretien de fin de protocole, dans le cadre de la gêne qui pouvait survenir du fait de l'affichage en temps réel du spectrogramme. A propos de l'exercice de la sirène, la patiente C rapporte : *« ça me gênait de le voir, je ne pouvais plus me concentrer sur le geste vocal »*.

Besoin de liberté, pas de supervision

En essayant de réaliser la sirène en regardant le spectrogramme, la patiente C nous dit : *« ça me fait peur de le voir, je me lâche beaucoup moins [...] j'aime pas trop [...] ça me perturbe »*. Elle rapporte ensuite : *« le spectrogramme te met dans une forme de contrôle : tu contrôles toi, « contrôle » dans le sens « supervision/superviseur ». Pour la sirène, et pour d'autres exercices, le ouhouh par exemple, je regardais pas le truc. Là pour le coup j'ai besoin de liberté, de me lâcher, de me mettre en légèreté, de me mettre « ailleurs » ; la sirène, ou les ouhouh, du coup fallait pas que je le regarde, j'ai besoin de liberté »*.

Au contraire, une patiente hors protocole pense justement que c'est le spectrogramme qui *« permet d'acquérir un peu plus de liberté, de spontanéité, parce qu'au début le face à face [avec le thérapeute] justement, y'a un je-ne-sais-quoi d'intimidant ; de nous concentrer sur un autre objet, ça détourne un petit peu la préoccupation de ce face à face. »*

Temps réel plus utile sur des phénomènes longs

Pour la patiente A, le temps réel serait utile sur des sons tenus ou des phrases musicales, mais pas sur l'attaque : *« pour le problème de l'attaque, quand on le voit il est trop tard »*. *« Ça me déconcentrait de le regarder à l'écran parce que de toutes façons c'est trop tard »*.

Intérêt du spectrogramme même de façon différée

Même lorsque regarder l'image en temps réel déconcentre les patients, le spectrogramme garde son intérêt car il permet de voir ce qui a été émis après coup. Tous les patients du groupe protocole le disent : pour la patiente A, pour l'attaque, il est tout de même *« intéressant après-coup d'avoir confirmation qu'il y a un problème, mais pas en temps réel »*. Pour la patiente C, sur la sirène qu'elle ne parvenait pas à effectuer en regardant le spectrogramme : *« ça restait intéressant de le voir après »*. Pour la patiente D, les difficultés dues au temps réel sont apparues dans l'exercice des transitions entre

les voyelles. Après coup « *ça m'a conforté dans le fait qu'il y avait quelque chose de spécial dans le son* ». On note que cette patiente a justement cité en premier cet exercice de transition entre les voyelles comme un des exercices les plus intéressants. Cela montre que le spectrogramme peut tout à fait avoir un intérêt même lorsqu'il n'est pas utilisé en temps réel.

Un patient précise dans le questionnaire en ligne : « *je préfère regarder en même temps si je veux agir sur ma voix et après si c'est seulement pour me fier à mes sensations et vérifier si elles sont justes* ».

On note que la patiente B ne rapporte aucune gêne et n'a jamais eu besoin de quitter le spectrogramme des yeux pour réussir un exercice. On suppose que cela est lié au fait que cette patiente a un très gros contrôle sur elle-même et une longue expérience d'analyse de son geste vocal.

5.3.1.3 Comportement vocal

Résultats du questionnaire

En moyenne, les patientes pensent que l'utilisation du spectrogramme est gênante pour prendre une posture adaptée à l'émission vocale. Cependant cette opinion moyenne est due à une minorité de patientes (1 « tout à fait d'accord » et 3 « plutôt d'accord » sur 10). La disparité observée pour la posture est encore plus marquée pour le geste vocal : une patiente pense que le spectrogramme ne la gêne pas du tout (« pas du tout d'accord »), et 4 pensent qu'il gêne un petit peu (« plutôt d'accord »). Les patientes sont plus unanimes pour la respiration, disant en moyenne qu'elles sont « plutôt pas d'accord » avec le fait que le spectrogramme gêne la respiration. Une seule personne (la patiente A) pense que le spectrogramme peut gêner la respiration.

Enfin, le fait de travailler devant un écran est gênant pour cette même personne seulement, les autres étant en majorité « pas du tout d'accord » (4 patients) ou « plutôt pas d'accord » (2 patients) avec le fait que l'écran soit gênant.

Ordinateur et posture

Problème du micro

Un patient répondant au questionnaire en ligne précise : « *le seul souci principal sur l'ordinateur, c'est la fixité du micro, qui oblige à ne pas trop bouger parce que sinon, je n'ai plus la même référence par rapport à ce que j'ai pu faire juste avant. Pour éviter ça, je pense qu'il suffit juste d'avoir un micro mobile, qui permettrait plus de souplesse* ».

Problème de la visualisation sur un écran

Durant les entretiens post-protocole, la patiente A a été la seule à pointer que « *d'être là devant l'écran et éventuellement tordu, penché, c'est contradictoire* ». Cette patiente confirme cette impression lors du questionnaire écrit, quelques semaines plus tard : « *j'ai été très gênée par le fait d'utiliser le spectrogramme sur un petit écran de portable, ce qui induisait une posture médiocre, peu propice à une bonne phonation. Il me semble que ce travail serait plus judicieux devant un écran de vidéoprojection* ».

La patiente C la rejoint, dans ce même questionnaire : « *difficile de projeter la voix au loin lorsqu'on a un micro sur la bouche et un écran d'ordinateur juste devant soi, besoin plutôt de regarder au loin pour accompagner certains gestes vocaux de projection de la voix* ».

Prise de conscience d'un comportement qui modifie la phonation

Une patiente vue en dehors du protocole dit pendant un exercice d'auto-analyse : « *en réécoutant, j'entends que le début du deuxième est plus clair que le premier. C'est en entendant que je m'en rends compte* ». Cependant elle pense que c'est uniquement le feedback auditif qui lui permet cette prise de conscience.

Pour la patiente B, le constat est le même mais elle l'attribue directement au spectrogramme : « *on l'a vu avec le spectrogramme, parce que quand je me préparais à faire le son de façon soutenue, il y avait beaucoup moins de spasmes [sous-entendu : sur le spectrogramme], donc dans la vie de tous les jours il faut que je soutienne beaucoup plus la voix pour minimiser la dysphonie* ». Lorsqu'on lui demande si elle en avait déjà conscience avant, elle répond « *ça m'en fait prendre de plus en plus conscience* » et que cela pourra la motiver davantage pour « *changer ma façon de parler pour minimiser la dysphonie* ».

5.3.2 Intérêts particuliers de la visualisation du spectrogramme

5.3.2.1 Réactions et émotions face à l'image

Nous relevons ici les réactions des patientes qui ont découvert le spectrogramme avec nous-mêmes uniquement (les 4 patientes du groupe protocole et une du groupe hors protocole)

Inquiétude

Lorsque la patiente C voit le spectrogramme sur la première fois, elle demande, inquiète « *ouh là qu'est-ce que c'est ?* » « *C'est du Picasso ?* ». Lorsque nous lui donnons les explications, elle est enthousiasmée : « *c'est ouf le son !* ». Une patiente hors protocole s'inquiète : « *j'suis à combien ?* » lorsqu'elle peut vérifier pour la première fois la hauteur de sa fréquence fondamentale.

Étonnement

Le premier commentaire de la patiente D lorsqu'on introduit le spectrogramme est « *c'est drôle ! [...] moi je l'ai pas entendu comme ça en fait, du coup de le voir c'est marrant* ». Sur l'exercice [s z i], la patiente B est intriguée : « *je suis ni contente ni pas contente, ça m'intrigue* », « *je suis un peu perturbée, par les sensations, et l'image... par rapport à ce que j'entends, ça a l'air plus facile que ce que je ressens* ».

Agacement

Selon la patiente D, « *le fait de le faire en le voyant c'est génial et en même temps c'est très frustrant quand on voit qu'il ne se passe pas ce qu'on attend* ». Elle est agacée de ne pas faire apparaître un [s] comme celui de l'expérimentateur. « *le vilain ! [...] coquinou on le voit pas beaucoup [...] c'est fou quand même !* », « *c'est fou, c'est super frustrant [le [s] qui ne se voit pas]* ». Nous avons accompagné la patiente en relativisant le rôle du modèle du thérapeute (« *comme moi c'est pas forcément mieux* »), et en valorisant l'image de sa voix par rapport à la nôtre, notamment sur un autre phonème.

Insatisfaction

La patiente D n'est pas tout à fait satisfaite de ses attaques, leur trouvant toujours un petit coup de glotte. Nous la rassurons sur ses capacités, qui sont déjà très bonnes, notamment en comparant avec nos productions qui ne sont pas parfaites non plus. La patiente A a la même problématique : « *je n'arrive pas à faire l'attaque nette* ». Nous pointons avec elle sur l'image les moments où elle a réussi (après les attaques soufflées).

Lors que la patiente D n'est pas contente de sa production : « *on voit rien* », nous lui faisons préciser : « *j'ai l'impression que chez toi on voit mieux* ». Après que nous avons de nouveau relativisé le poids du modèle du thérapeute, elle parvient à analyser finement l'image et à trouver qu'on y voit aussi bien sur sa production que sur la nôtre.

Déception

La patiente D voulait tester un changement d'émission vocale en voix chantée, elle est déçue qu'on ne puisse pas voir de façon nette ce changement sur le spectrogramme, ni ses inspirations bruyantes.

Enthousiasme

La patiente C, lorsqu'elle émet un [f] précis et mieux dirigé, s'exclame : « *c'est beau, ça, ça me plait !* » (voir Figure XXXIV en annexe). La patiente D est très intéressée par les formants vocaliques : « *c'est énorme ! [la différence entre le [a] et le [i]]* » (voir Figure XL en annexe). Elle est également très impressionnée par l'image de la sirène : « *c'est chan-mé !* » « *cool !* » « *sympa !* » (voir Figure XXXIII en annexe).

Satisfaction

La patiente C est heureuse de constater qu'elle réussit à enrichir ses harmoniques. Elle commente immédiatement après sa réalisation, parlant des harmoniques : « *j'ai réussi à les prendre, à les garder* ». Nous pointons volontiers la grande richesse harmonique de la patiente D y compris dans le fry. Elle en est très satisfaite (« *joie* »). Elle est fière de sa production également sur les formants vocaliques (« *Yes !!* »).

Prise de conscience

La patiente C pense que son vibrato est exagéré. Lorsqu'elle peut le voir et l'entendre, elle dit : « *j'ai eu la sensation d'en faire des caisses et le vibrato est plus petit que ce que je pensais* ».

Soulagement

La patiente C est soulagée de voir les harmoniques du formant du chanteur dans sa voix. La patiente D se trouve également rassurée de les voir.

Plaisir et jeu

Sur l'exercice des [s] colorés par les voyelles [i] et [u], la patiente C commente « *tu fais de la peinture* », puis « *j'ai fait une mammographie de trois seins* ». Sur ce même exercice, la patiente D s'amuse : « *c'est rigolo parce qu'on voit le « u » visuellement* ».

L'exercice des formants sur le fry fait dire à la patiente D : « *on dirait un diplodocus sur une colline, c'est très beau* ». Nous lui rappelons qu'il ne faut tout de même pas forcer la voix pour faire apparaître

quelque chose. Sur les alternances [ɛ] / [ɛ̃] la patiente voit « *de vraies têtes de mort* ». La patiente B voit « *une tour Eiffel* » dans les transitions de voyelles.

5.3.2.2 Paramètres vocaux que le spectrogramme permet d'observer et d'améliorer

Résultats du questionnaire

Les personnes ayant répondu au questionnaire sont unanimes pour dire que le spectrogramme permet d'observer des paramètres vocaux (8 « tout à fait d'accord » sur 10 personnes). Cela est d'autant plus vrai pour la parole⁶³ (9 « tout à fait d'accord ») et l'attaque (8 « tout à fait d'accord »), et reste tout à fait vrai pour le timbre (8 « tout à fait d'accord ») pour lequel on note tout de même que la patiente A répond « je ne sais pas ».

Les réponses sont beaucoup plus partagées concernant la possibilité d'améliorer les différents paramètres. C'est le timbre qui paraît le plus difficile à améliorer, puis l'attaque, puis les différents sons émis par la parole, mais celle-ci avec une grande disparité, la patiente C étant la seule à avoir répondu « pas du tout d'accord ».

En résumé, les personnes ayant répondu au questionnaire pensent que le spectrogramme peut permettre d'observer des caractéristiques de leur voix ; ils sont partagés sur le fait que cela peut permettre de les améliorer, tout en restant globalement plutôt d'accord avec cette affirmation. C'est le timbre de la voix qui leur est le plus difficile à observer et à améliorer.

Impact de la visualisation d'un timbre altéré

Possible image « violente » de la dysphonie

Une patiente donne spontanément ce témoignage écrit : « *le spectrogramme a été pour moi une révélation en ce sens qu'il permet de tracer objectivement les défauts à combattre* ».

Pour les patientes dysphoniques du protocole, le fait d'identifier sur l'image les manifestations de la dysphonie a été un événement marquant. La patiente B rapporte : « *la première chose c'était d'identifier le spasme* ». « *Au tout début c'était un peu pénible, notamment quand j'avais des spasmes, même si je le sais qu'il y a une dysphonie spasmodique, mais c'était un peu douloureux au début d'être encore dans la prise de conscience que « oui, c'est là »* ».

La patiente C rapporte également que le spectrogramme lui a permis de voir « *les ruptures de son dont [elle est] victime* » ainsi que « *le souffle qui vient se greffer dans la voix* ». Cette patiente n'hésite pas à mentionner que cela l'a parfois « *chagrinée* » : « *déjà qu'on l'entend, mais en plus on le voit, ça marque bien le truc, je pense qu'il faut être un petit peu blindé* ». Elle pose la question des réactions qu'elle aurait pu avoir lorsqu'elle avait la voix « *très très cassée* » : « *ça aurait été peut-être un peu violent, au premier abord* ». Cette agressivité de l'image de soi-même renvoyée par le spectrogramme est évidemment à prendre en compte dans les thérapies vocales. Lorsque nous lui suggérons qu'une voix même cassée peut toujours émettre des sons du type « ff » par exemple ou observer des harmoniques sur du souffle comme dans l'exercice « sss-i sss-ou », elle acquiesce : « *Ça peut être une entrée en matière, pour quelqu'un qui a la voix très très cassée, très éraillée* ».

⁶³ introduite comme les « différents sons émis par la parole (voyelles, consonnes) »

Intérêt majeur : pouvoir travailler dessus

Cette même patiente C, à la question « qu'est-ce que ça fait de voir ces choses un peu négatives ? », répond sans hésitation qu' « *on peut les corriger. On voit la rupture du son donc, par exemple [sur les sirènes], ça permet de rectifier le geste* » (voir Figure XXXVIII en annexe).

La visualisation a permis à la patiente B d'identifier des situations où il y avait plus ou moins de spasmes : « *sur la voix parlée, ce que ça m'a montré, c'est que quand je prends [une voix] avec une intensité normale, il y a beaucoup plus de spasmes, et dès que je prends une voix un peu plus projetée [ça va mieux]* ».

Attaques facilement visualisées

Dans les paramètres de la voix que le spectrogramme leur a permis de visualiser, toutes les patientes du protocole citent les différents types d'attaques. La patiente D cite notamment l'exercice d'auto-analyse pour la visualisation d'attaque en coup de glotte : « *ça m'a aidé à prendre conscience des attaques* » ; « *ça le fait plus que ce que j'aurais pensé, [je m'en rends compte] parce que je le vois et aussi parce que tu as attiré mon attention dessus* ». Elle ajoute : « *je trouve que pour les attaques c'est super clair, pour les gens qui pourraient mal sentir ou mal doser ça se voit* ». Une patiente du protocole trouve les exemples sur l'attaque « *très clairs* », « *c'est marrant comme ça se voit !* ».

Phonétique

Seule la patiente A rappelle cet aspect du travail dans l'entretien post-protocole : elle « *voi[t] bien que ça prend une autre réalité, c'est un peu plus palpable aussi* » avec le spectrogramme.

Harmoniques

La patiente C rapporte : « *[on a travaillé] la prise de conscience de la **qualité du son** sur les harmoniques, j'ai vraiment découvert des trucs* ». La patiente D pense que « *prendre conscience des résonateurs aussi c'est chouette* ».

Dans les exercices d'auto-analyse, pour l'échelle de la richesse du timbre, il est fréquent que les patients disent « *je ne sais pas juger de ça* » avant de voir le feedback visuel. Une patiente hors protocole rapporte : « *c'est vraiment quelque chose que je ne sais pas sentir, mais tout compte fait il y en a [...] il y a en a même pas mal* ».

Le **timbre vocalique** et les possibilités d'accorder les formants avec une note chantée ont été cités par les patientes C et D comme des paramètres de leur voix observés grâce au spectrogramme, comme l'évoque la patiente C : « *[le fait qu'une note va sonner plus ou moins bien selon la voyelle], ça on le voit sur le spectrogramme : je l'entendais mais sans vraiment comprendre pourquoi* ». A ce propos, la patiente D déclare que : « *je me rappelle que sur le moment j'avais trouvé ça génial* ».

Pour la patiente B, le **singing formant**, « *c'est le truc qui intéresse le plus les chanteurs, c'est très narcissique : avoir des harmoniques, une voix timbrée, vibrante, résonante* ». La patiente D exprime à sa façon que cette notion paraît un peu magique : « *pour moi c'était même un truc d'alien que moi j'avais pas, un truc de chanteur lyrique* ».

La patiente A relie directement l'image et la sensation auditive : « *c'est confus, instinctif, c'est pas un truc que j'entends clairement, donc c'est un peu mystérieux pour moi le fait qu'il y ait ces harmoniques qui apparaissent à l'image, par contre je peux le relier assez facilement au fait que je chante bien,*

qu'il y a quelque chose dans le timbre qui me plaît et que comme par hasard, c'est à ce moment-là qu'il y a des harmoniques ». La patiente C a le même ressenti : « quand les harmoniques sont apparues, pour moi je sentais que le son était bien accroché [...] ça correspond exactement dans la sensation ».

5.3.2.3 Moyen d'analyse scientifique d'étude et de connaissance

Moyen de satisfaire une soif de connaissance

La patiente D rapporte : *« moi je suis très curieuse de ce qui se passe, je suis friande de « comment ça marche », de où ça va, de « qu'est-ce qui se passe », donc ça c'est vrai que pour ça c'était vraiment excellent de voir tout ça » ; « c'est génial parce que c'est le support de plein d'explications ».*

La patiente B s'intéresse elle-aussi très précisément aux mécanismes de la voix chantée : *« ça a été un vrai plus, parce qu'aussi je fais partie des chanteurs avancés, et ça me permet de mieux comprendre les mécanismes que j'étudie ».*

Possibilité d'adopter un protocole de rééducation encourageant

La patiente B rapporte : *« si je comprends, ça me permet de mieux adapter mon protocole de rééducation ».* Elle ajoute que le spectrogramme permet de *« voir scientifiquement les différentes configurations vocales pour le prof de chant ».*

Un patient répondant au questionnaire internet précise : *« le spectrogramme est un instrument très utile; je regrette, dans mon cas, ne pas avoir connu plus tôt ce support de travail. Cela m'aurait motivé à travailler plus scientifiquement donc encouragé ».*

Attirer l'attention sur une partie du son

Pour la patiente A, cette connaissance peut être intéressante, mais pas toujours forcément utile. *« Ça m'intéresse intellectuellement, c'est intéressant d'avoir une image de choses qu'on n'entend pas toujours ou qui sont un petit peu fines : voir qu'il s'est bien passé un truc » ; « ça peut attirer l'attention sur un truc bizarre, tiens à quoi ça correspond au niveau auditif ? »*

Pour la patiente D, pour pouvoir améliorer un paramètre, il a d'abord fallu que le thérapeute attire l'attention dessus : *« c'est parce que je le vois et aussi parce que tu as attiré mon attention dessus » ; « tes explications ont attiré mon attention sur des trucs et du coup je les ai comprises avec tes explications, et de voir aussi ça m'a bien aidée »* et le spectrogramme permet de pointer précisément ce dont on veut parler dans le son. Elle ajoute : *« les feedbacks auditifs, ça m'apprend beaucoup, mais là le fait que ça soit visuel, c'est vraiment complet, parce que tu peux pas mettre ton attention pour tout »* et que l'image permet au thérapeute de pointer, en différé notamment, ce qui reste à travailler.

Intérêt intellectuel au détriment d'une utilité pour la voix ?

L'intérêt intellectuel pour le fonctionnement de la voix n'est pas toujours partagé par les patients. La patiente A l'exprime ainsi : *« des fois vous êtes beaucoup rentrée dans les détails et je sais pas si j'en avais tant besoin que ça, je ne comprenais pas toujours pourquoi vous m'expliquiez tout ça alors que je pensais que j'étais là plutôt pour que vous fassiez vos mesures, vos trucs, mais au bout d'un moment j'étais un peu perdue. »*

Elle pense, en outre, que cette analyse scientifique du son émis n'aide pas à travailler sa voix : « *si on essaye d'améliorer le son mais sans avoir conscience que ça vient du souffle on va être assez vite limité, on va être dans un critère un peu esthétique [pas adapté]* ».

A ce sujet, une patiente hors protocole fait un témoignage écrit spontané : « *dans ce travail, je ne souhaite pas rester au niveau « scientifique » (« spectrogramme », « formant » etc ... tous ces mots techniques que je découvre). Le dessin, comme la voix, le chant, utilisent certes la technique, mais la dépassent* ». Pour elle, le spectrogramme permet de « *découvrir le dessin des voyelles, des consonnes, des syllabes, des mots chantés ou parlés, chargés de diverses émotions* » ; « *ils parlent de moi, de mon humanité* ». On voit ici qu'on peut aller au-delà de l'aspect scientifique, selon la sensibilité des patients.

5.3.2.4 Reflet objectif rendu positif par le thérapeute

Reflète objectif

Tous les patients qui ont répondu au questionnaire pensent que le spectrogramme fournit la preuve objective de ce qu'ils avaient remarqué sur leur voix. La patiente B précise : « *c'est plus palpable, parce que la voix c'est quelque chose d'immatériel, d'impalpable.* ». Pour elle l'image permet d'être en simultané et apporte plus d'objectivité, ce que l'enregistrement seul ne réussit pas : « *j'ai une grosse expérience des enregistrements, de m'entendre, mais l'image apporte peut-être encore plus : les deux combinés font que j'ai plus tendance à être objective par rapport à moi-même, que même d'avoir qu'un enregistrement, du fait aussi qu'on voit en simultané* ».

Possibilité de mettre son problème de voix à distance, et dédramatiser

La patiente B pense qu'« *on n'est pas toujours objectif, on a tendance, surtout quand on est face à une pathologie, à dramatiser, comme tous les patients dans le monde entier. Même si on le dit pas* ». Elle pense que « *c'est moins dramatisant d'être objectif sur le souci* ». Elle a ressenti qu'« *une partie des difficultés, c'est pas qu'elle ne s'entend pas, mais qu'elle est moins dramatique que ce qu'[elle] sen[t]* ».

La patiente A la rejoint, en rapportant que l'image spectrographique « *permet de se dire : tiens finalement c'est pas si grave, c'est pas mal : vérifier qu'on n'a pas la berlué, qu'on n'a pas rêvé. C'est intéressant d'avoir confirmation de quelque chose qu'on entend ou pas* ».

La patiente D a pu s'affranchir d'un problème mineur qui avait pris de l'importance pour elle parce qu'elle ne comprenait pas ce qui pouvait arriver. Le fait de voir en image ce point précis l'a rassurée, sans pourtant qu'on ait réglé le problème. « *J'avais une copine qui m'avait dit t'as les [s] qui sifflent c'est la catastrophe. Quand on a regardé, justement j'étais très étonnée parce qu'ils n'apparaissent pas je pensais que ça ferait krkrkrkr ou j'sais pas trop quoi, et donc là effectivement, j'ai vu qu'il n'y avait pas de problème dans le sens où je pensais. Je pensais que c'était trop fort ou trop soufflé trop j'en sais rien et en fait finalement on les voit moins mais j'ai pas pris ça comme un gros problème je ne me suis pas attardée là-dessus.* » Apparemment, c'est surtout son ignorance du problème exact et de la raison pour laquelle une personne avait pointé cela chez elle qui la perturbait. Objectiver une différence sur le spectrogramme a permis de ne plus considérer cette différence comme un problème qu'il faudrait régler.

Une patiente hors protocole écrit ce témoignage spontané: « VV [le spectrogramme de VoceVista] me permet de m'apprivoiser, de me regarder, de m'écouter sans peur, avec humour quand je trouve que c'est moche ! » ; « avant, je n'aimais pas écouter ma voix ; VV me permet de l'apprivoiser ». Elle ajoute : « devant le spectrogramme là j'étais en état de liberté subjective devant un matériel objectif. Liberté intérieure de ne pas être accusée que c'est génial ou pas génial, bien ou moins bien » ; « à travers cet écran, je me fais moins de cinéma intérieur ».

Possibilité d'observer l'évolution d'un travail

La patiente A ajoute : « y'a quelque chose d'assez rassurant aussi, y'a des gens qui peut-être ne se rendent pas forcément compte qu'ils ont progressé, que oui, il y a des harmoniques dans leur voix, c'est vrai que ça fait plaisir : « la preuve, c'est là », et on peut voir aussi l'évolution d'un travail ».

Le formant du chanteur est quelque chose de rassurant à objectiver pour la patiente D, surtout dans le cadre de sa reconversion dans les métiers du chant : « le formant du chanteur j'étais bien contente de le voir, genre, on te donne le tampon : « c'est bon, vous l'êtes, pffiou [au suivant] », ça c'est bien rassurant ».

Possibilité d'une relation apaisée avec le thérapeute

La patiente B pense que le spectrogramme « permet une relation plus objective [avec le thérapeute], parce qu'il n'y a pas de ressenti, y'a pas ce que vous avez entendu et ce que moi j'entends, il y a vraiment une objectivité scientifique ».

Pour elle, le spectrogramme permet d'accorder les visions qu'ont les thérapeutes et les patients de la voix du patient, en objectivant le phénomène physique qui gêne effectivement le patient, ou en permettant au contraire au patient de s'affranchir d'un ressenti négatif en lui montrant qu'il n'y a objectivement pas d'impact acoustique à sa gêne: « même si c'est psychologique, il y a quand même un phénomène physique, je pense que le spectrogramme ça permet ça (dans les deux sens: à la fois « vous voyez bien que ça va pas » et « vous voyez bien que ça va bien ») ». Le spectrogramme permet à la relation thérapeutique de s'établir sur une base commune objective, qui a parfois manqué à la patiente B dans son parcours thérapeutique.

La patiente C nous dit : « j'ai bien aimé être à deux à regarder le troisième truc. J'ai vraiment bien aimé. On produit pour la machine avec l'œil de l'expert. Y'a une machine qui va vraiment voir ce que je produis et qu'on va décrypter ». Elle a été par le passé mal à l'aise avec des approches thérapeutiques manquant, à son avis, de méthode formalisée et de feedback. L'utilisation d'un outil comme le spectrogramme est pour elle un gage de qualité pour une prise en charge.

Un patient répondant au questionnaire en ligne précise : « je vois un vrai intérêt à l'utilisation du spectrogramme par des professionnels. En effet, cela met en évidence les explications parfois imagées des professeurs de chant ou orthophonistes, qui peuvent induire en erreur. Sans le spectrogramme, l'élève ou le patient ne peut faire confiance qu'au sens de l'observation (développé ou non), du praticien ».

Impact positif de l'interprétation par le thérapeute

La patiente B est sûre qu'elle est « beaucoup plus positive (depuis ce travail avec le spectrogramme) ». Elle rapporte que des personnes extérieures ont également reconnu les progrès

réalisés. Elle pense que « *les échanges, l'opinion de l'expérimentateur, ses remarques* » l'ont aidée autant que l'outil spectrogramme. On ressent que pour elle, le côté positif du spectrogramme est beaucoup induit par l'attitude du thérapeute face à l'image.

5.3.3 Utilisations préconisées et possibilité d'autonomie

5.3.3.1 *Public visé*

Résultats du questionnaire

Les utilisations souhaitées sont plus avec un professeur de chant qu'avec un orthophoniste. De plus les patientes répondent de façon disparate lorsqu'il s'agit d'une utilisation avec un orthophoniste et de façon plus unanime pour le professeur de chant (écart type 0,6), entre « tout à fait d'accord » et « plutôt d'accord ».

2 patientes sur les 10 qui ont répondu n'ont pas envie d'utiliser le spectrogramme (1 « pas du tout » et 1 « plutôt pas »), les autres ayant majoritairement « tout à fait » envie, ce qui conduit à une grande disparité des réponses (écart type 1,5).

Quel type de public

La patiente D pense que « *c'est pas du tout une méthode invasive, douloureuse, pénible* ». Elle compare notamment aux difficultés rencontrées classiquement en laryngoscopie avec le réflexe nauséux notamment. Dans ce cadre, elle pense que l'outil peut facilement être utilisé pour tout type de public. Elle pense en particulier au côté ludique pour les enfants et très intéressant pour les chanteurs.

De façon générale, les patientes pensent que le spectrogramme est un outil « *un peu avancé* » (patiente C) et trouvent a priori qu'il est plus adapté pour les personnes qui sont « *à l'écoute de leur voix* » et qui « *ont des notions* ».

La patiente D pense que « *ça pourrait mettre mal à l'aise quelqu'un qui est mal à l'aise avec les machines* » puis se ravise en disant que « *même les personnes âgées, quelqu'un qui est un peu allergique à l'intervention d'une machine, si ça lui permet de voir l'objectif à atteindre et à régler les problèmes, je vois pas du tout en quoi ça peut être gênant* ».

Pour une patiente hors protocole : « *on peut se dire ah non ça peut être gênant, et personnellement, avec les limites vocales que j'ai, je trouve que c'est un outil intéressant, qui ne fait pas peur* ».

Utilisation pour les chanteurs avancés

La patiente B est professeur de chant. Elle a envie d'utiliser le spectrogramme « *plutôt [avec] des élèves avancés, pour objectiver des choses notamment pour le singing formant, mais notamment aussi pour tout ce qui est la voix mixte* ».

Alors qu'elle cherchait à produire un son chanté un peu en gorge ou non, afin de voir les différences visibles au spectrogramme, la patiente D remarque tout de même les limites de l'outil : « *cette machine-là ne mesure pas tout.[...] : on entend mais on ne voit pas vraiment, il y a pas un avant / après qui soit clair, c'est pas clair de savoir ce qui se passe : c'est la limite de l'outil* ». En réalité, sur son exemple, on voit des différences minimes dans la répartition de l'intensité dans les harmoniques, mais il reste difficile de savoir ce qui est le mieux, ce qu'on cherche.

L'oreille experte d'un professeur de chant reste indispensable.

Utilisation par les professionnels de la voix

La patiente C, journaliste, pense que le spectrogramme est indiqué pour « *tous ceux qui utilisent leur voix* » ; « *un journaliste radio, il a besoin de poser sa voix, et ça peut être un bon complément de voir comme ça, il va entendre que si il fait son intonation comme ça, il voit que quand il fait sa chute de phrase, il arrivera mieux à poser sa voix* ».

Utilisation pour les dysphoniques

La patiente B pense que « *dans le cas de dysphonie je pense que ça peut vraiment être un bon outil* ». La patiente D imagine que l'image spectrographique permet aux patients dysphoniques d'être plus motivés pour arriver à un objectif : « *j'imagine que [du fait que l'orthophoniste] leur [donne] l'exercice qui fait que ça va mieux, en voyant le résultat, ils vont être plus motivés, avoir plus envie de le faire* ».

Cependant, la patiente A rappelle que la phonation reste un « *travail corporel* ». Elle ne peut s'appuyer sur l'image ni sur ce qu'on a découvert ensemble pour améliorer son geste vocal. Elle ajoute que le fait que le spectrogramme soit sur un petit écran constitue un obstacle pour « *produire du son, à chercher des petites variantes, à voir ce que ça change sur l'écran* », qui est ce qui l'aurait intéressé, mais qui, selon elle, n'a pas été fait lors des séances.

5.3.3.2 Compréhension de l'image spectrographique

Résultats du questionnaire

7 personnes sur 10 sont « plutôt d'accord » (5 personnes) ou « tout à fait d'accord » (2 personnes) avec le fait que l'image du spectrogramme soit déroutante. L'image du spectrogramme est en moyenne plutôt déroutante, mais avec une forte disparité d'opinion suivant les personnes (écart type 1,3).

L'interprétation de l'image est facile pour la majorité des personnes ayant répondu au questionnaire (6 « plutôt d'accord » et 1 « tout à fait d'accord »). On note que l'image peut être déroutante sans forcément gêner l'interprétation. Nous pensons que c'est le fait que le spectrogramme fait parfois émerger des émotions diverses (voir paragraphe 5.3.2.1) qui peut faire dire aux patients que l'image est déroutante. En moyenne, les personnes ayant répondu au questionnaire interprètent plutôt facilement l'image spectrographique.

Une petite majorité des personnes ayant répondu au questionnaire pensent qu'ils comprennent suffisamment le spectrogramme pour pouvoir l'utiliser seules (6 sur 10, 5 « plutôt d'accord » et 1 « tout à fait d'accord »). La disparité des réponses est ici maximale (écart type 1,4).

Remarques qualitatives

Un patient écrit : « *le spectrogramme représente pour moi « l'électrocardiogramme » de la voix* ». Il sait reconnaître les signes de raucité dans « *le bruit entre les harmoniques* ».

On se rend compte tout de même en séance que la compréhension est souvent partielle. De plus, les patients en autonomie interprètent parfois l'image de façon assez fantaisiste. Un patient notamment était très ennuyé par le « trou » lorsqu'il enregistrait des [i], alors que cette configuration des

harmoniques provient directement de la nature de la voyelle. De même, un autre avait pour objectif de remplir tout l'espace entre 0 et 4 kHz d'harmoniques, indépendamment des formants constitutifs de la voyelle émise.

Selon un exemple de patient répondant au questionnaire en ligne : « *dans mon cas, je ne parviens pas à distinguer mes harmoniques au niveau sonore, je vois en revanche sur le spectrogramme que j'en perds certaines en montant dans les fréquences.* », nous ne sommes pas certains ici de ce qu'a compris le patient. Même s'il interprète bien l'image, ici, on ne sait pas s'il a bien compris l'objectif : en montant la fréquence fondamentale, il est inévitable de faire diminuer le nombre d'harmoniques dans une bande de fréquences donnée.

Enfin, nous avons constaté que même après les 3 séances du protocole consacrées uniquement à la découverte du spectrogramme, les patients pouvaient avoir oublié ce que représentent les traits horizontaux, à quoi correspond la fréquence fondamentale, comment se matérialise le bruit, etc.

On en déduit que parfois les patients peuvent surestimer leur connaissance de l'outil, ce qui risque alors de les mener à de fausses interprétations et à des fausses pistes d'exploration et de rééducation de leur voix, si leur erreur n'est pas repérée et corrigée par le thérapeute.

La patiente C, en fin de protocole, n'a d'abord rien reconnu dans un extrait de voix chantée. Nous l'avons encouragée à chanter à nouveau l'extrait vocal en regardant le spectrogramme s'afficher en temps réel. Elle a été alors tout à fait capable de reconnaître ce qu'elle avait produit de façon pertinente.

5.3.3.3 Utilisation en autonomie

Résultats du questionnaire

Tous les patients qui ont répondu au questionnaire sont d'accord avec le fait que le spectrogramme permet de s'amuser, et de visualiser des caractéristiques objectives de la voix, mais ils sont unanimes surtout pour dire que le spectrogramme permet d'analyser la voix (entre « plutôt d'accord » et « tout à fait d'accord »), sauf la patiente A qui répond « ne sais pas ».

5 patients sur 10 ont téléchargé le logiciel, mais sur les 5, deux n'ont pas pu l'utiliser pour des raisons techniques (ils utilisent des ordinateurs avec le système d'exploitation Macintosh, pour lequel le logiciel proposé au téléchargement ne fonctionne pas ; pour ce système d'exploitation il semble que les spectrogrammes suffisamment fonctionnels soient tous payants).

Page		Nombre de visites	Durée moyenne
Mode d'emploi de vocevista		96	1 min 34
Exercices	attaque (début de son)	85	2 min 08
	timbre (richesse harmonique)	80	2 min 36
	hauteur et intensité (sirènes)	61	1 min 55
Lecture de spectrogramme	attaque (début de son)	60	2 min 20
	timbre (richesse harmonique)	53	2 min 24
	curiosités	23	16 s

Tableau XXIX : statistiques de visites du site internet pour les patients en autonomie

4 patients sur les 10 ont utilisé le site internet proposé. Les pages préférées sont celles sur la lecture d'intensité et de hauteur, et celle sur la lecture du timbre. C'est la page « banque de son » qui a été la moins utilisée, mais il est difficile de faire des statistiques sur 3 patients uniquement. Selon les

statistiques proposées par le fournisseur d'accès free (phpMyVisites : Tableau XXIX), en 3 mois (février à avril 2013) 939 pages ont été vues pour 142 visiteurs et 187 visites. Ces statistiques montrent que les pages d'exercices sont plus visitées que celles sur la lecture de spectrogramme.

Problèmes techniques et convivialité

L'ordinateur est parfois un obstacle pour les patients, soit parce que le logiciel est perçu comme instable : « *je voyais ton ordi bugger, et mon ordi maison rame beaucoup, j'ai peur qu'il ne suive pas* » (patiente C), soit par manque d'envie de passer du temps encore sur un ordinateur : « *je passe déjà tellement de temps devant l'ordinateur que j'ai pas envie de m'installer devant encore* » (patiente A).

Dans un témoignage spontané, un patient hors protocole raconte : « *je maîtrise mal « la technique » et le simple fait d'appuyer sur un me stresse* ». Certains témoignages rapportent également les conditions matérielles d'enregistrement (bruit du PC, qualité du micro et des haut-parleurs) ne donnant pas toujours des résultats probants (voix perçue comme « *métallique* » par exemple).

Nécessité d'un accompagnement pour interpréter l'image

Pour la patiente C, « *c'est le spectrogramme avec l'explication qui décrypte ce qu'on voit à l'écran sinon ça sert à rien.* »

La nécessité d'être initié est pointée par une patiente qui répond au questionnaire en ligne : « *c'est un outil ludique et instructif, à condition d'y avoir été initié auparavant, pour bien comprendre comment ça marche et ce qu'on peut faire avec* ». Une patiente du protocole précise que dans une utilisation en autonomie: « *il manqu[e] la relecture avec quelqu'un qui a l'expérience et la main sur le logiciel et qui sait interpréter d'une manière qu'il puisse m'expliquer avec des mots* ». Elle recherche pourtant cette utilisation en autonomie, un peu comme des « *Travaux Pratiques* » : « *le spectrogramme permet au thérapeute de mettre le patient sur les rails pour faire un travail individuel à la maison* ».

Pour certains patients cependant, qui prennent ici une attitude de chercheur, l'apprentissage en autonomie des correspondances entre l'auditif et le visuel est possible : « *progressivement je vais apprendre à lire mes dessins en les regardant, les comparant, les écoutant* ».

Nécessité d'un accompagnement pour guider le travail

Le travail en autonomie avec le spectrogramme est freiné par le fait que les patients ont besoin d'accompagnement : comme le dit la patiente B : « *j'ai besoin que de temps en temps il y ait des gens qui me disent « mais si, c'est la bonne voie/x, c'est ça... »* » ; et comme l'évoque la patiente C : « *j'ai besoin d'un contrôle [extérieur]. [...] Besoin de quelqu'un qui dit « exactement c'était ça » ou alors « tiens si tu faisais ça »* ». Cet accompagnement permet d'encourager mais également de proposer des idées, des pistes.

Une patiente précise dans le questionnaire en ligne : « *pour de l'autocorrection, je pense qu'un praticien doit aiguiller son patient dans des exercices spécifiques à sa pathologie. Ou simplement pour lui indiquer la tonalité à partir de laquelle travailler ; en effet, ne connaissant pas mon registre, j'ai eu tendance à forcer dans le grave ou l'aigu avec des sons ouverts* ». Une autre patiente évoque : « *peut-être à ne pas utiliser uniquement seul, parce que grâce à lui on peut voir si ce qu'on fait est*

"correct" ou non, mais on n'a pas forcément les moyens de comprendre pourquoi ou comment y remédier si besoin ».

Les personnes utilisant le spectrogramme seules ont parfois trouvé les consignes des exercices « *un peu floues* » sur le site Internet. Ils évoquent la nécessité d'avoir un « *journal de bord* » où noter les exercices qui ont été faits et comment ils ont été réussis, ainsi que l'envie de faire le lien avec « *des techniques vocales avec des exercices pas à pas* » : « *en fait, un lien direct avec les exercices que proposent tous les profs de chant (même sur le web), afin d'avoir un didacticiel qui nous met davantage dans une dynamique de progrès que de constat de sa qualité vocale* ».

Pour une patiente hors protocole, le spectrogramme est « *utile mais pas suffisant* ». Elle précise : « *j'attends que ça débouche sur un complément de dialogue avec une personne* »

Nécessité d'un objectif

Pour les patients qui se sentent prêts à se servir seuls du spectrogramme, un objectif précis est nécessaire : la patiente D fournit un exemple avec une interrogation qu'elle a depuis un cours de chant précis « *je pense que j'irai voir ce qui se passe dans les différents [i] [selon] les différentes positions de langues et d'ouverture de bouche* ». La patiente A pense s'en servir uniquement dans le cadre d'un travail précis sur les harmoniques⁶⁴.

Cependant, pour la patiente D, le protocole a été plus une exploration qu'un travail sur un objectif en particulier (sauf la dernière séance où l'on a travaillé sur un extrait chanté et amélioré les aigus). Elle se demande comment on pourrait travailler sur un objectif en particulier, surtout si la voix est pathologique. De même, pour la patiente A : « *dans ce qu'on fait ensemble, il n'y a pas de soin, c'est plus un constat. Je le vois plus comme un truc de vérification, de confirmation* ».

Utilisation autonome ludique

Lorsqu'elles ne se sentent pas prêtes à travailler leur voix avec l'aide du spectrogramme, les patientes du protocole évoquent la possibilité d'utiliser le spectrogramme de façon ludique (« *plutôt un gadget* », dit la patiente A), parfois pour dédramatiser la dimension d'analyse que le spectrogramme comporte « *peut-être si mon neveu vient à la maison je le mettrai et on va jouer avec ; une utilisation ludique. J'ai toujours du mal à chanter chez moi, faut être capable de se lâcher un peu pour s'amuser devant le truc, c'est de l'auto-exploration et j'ai du mal à faire ça* » (patiente C).

5.3.4 Conclusion

Les différents témoignages sur le spectrogramme confirment les avantages, mais également les limites pointées par les auteurs. Il est remarquable cependant que, quels que soient les sensibilités proprioceptive et auditive des patients, leur niveau de connaissance des mécanismes de production de la voix, et leur niveau d'atteinte vocale, ils ont été très majoritairement intéressés par l'usage du spectrogramme. En outre, pour chacune des patientes du protocole, il a été l'occasion d'une expérience positive autour du travail et de la découverte de leur voix.

⁶⁴ Choque D. (2004) Perception et émission des harmoniques. In *3^{ème} séminaire Voix, approches corporelles et sensorielles*. Musicothérapie Expressions 45

6 Discussion

6.1 Validation des hypothèses de travail

6.1.1 Hypothèse 1 : prise de conscience des paramètres vocaux

Dans notre étude, la conduite des exercices d'auto-analyse permet aux patients de distinguer les différents paramètres à évaluer (intensité, hauteur, attaque, richesse, clarté du timbre et geste vocal), avec l'aide d'un feedback visuel constitué par le spectrogramme, et un apprentissage rapide des différents paramètres à évaluer par des ancrages externes pour l'attaque, la richesse et le timbre, présentés systématiquement avant les auto-analyses.

6.1.1.1 Fiabilité inter-juges

En accord avec les conclusions de Dejonckere (2006) [22], la fiabilité inter-juges des auditeurs du jury d'écoute est améliorée par l'adjonction d'une image spectrographique. Cela est particulièrement vrai pour la clarté, la richesse et le geste vocal.

6.1.1.2 Observation des paramètres

Les auteurs pensent que le spectrogramme peut être un outil d'aide à la prise de conscience de certains paramètres vocaux et des progrès effectués. Ils citent la hauteur (Coudière, 2003, p. 146) [19], le timbre, l'attaque, la respiration, la précision des phonèmes, le soutien en fin de phrase, la justesse (Nair, 1999) [78].

Les patientes et les personnes ayant répondu au questionnaire en ligne sont unanimement d'accord avec le fait que le spectrogramme peut permettre d'observer des caractéristiques de leur voix. Dans le questionnaire post-protocole, les patientes du protocole citent en particulier la prise de conscience des éléments précis de la dysphonie (spasme, ruptures de son..), les attaques dures et surtout les harmoniques (timbre vocalique, singing formant).

6.1.1.3 Auto-évaluation de ses propres qualités vocales : exercice d'auto-analyse

Une bonne prise de conscience de ses paramètres vocaux se traduit notamment par une auto-évaluation fiable de ses propres qualités vocales. Il existe effectivement peu d'études s'intéressant à l'auto-évaluation par les patients de leur propre voix. (Eadie *et al.*, 2010, p. 564b) [25]. Les locuteurs baseraient leurs jugements sur des indices différents de ceux des auditeurs.

Meilleure distinction des différents paramètres

D'après Hammarberg et Kreiman (Kreiman *et al.*, 1992 ; in Revis, 2004, p. 81) [55] [89], « *les auditeurs candides tendent à fonder leurs impressions sur la hauteur et la puissance davantage que sur les notions de souffle et de raucité* ». Notre étude permet d'observer les effets d'un apprentissage des différents paramètres vocaux en s'appuyant sur l'image spectrographique. Si les patients ont au départ une vision « en bloc » de leur dysphonie, ils sont en revanche capables, après feedback auditif et visuel, de dissocier les paramètres de hauteur, d'intensité, de qualité d'attaque, de clarté et de richesse et de les apprécier chacun indépendamment.

Meilleure objectivité

Pour Kasama et Braslotto⁶⁵, « *les patients jugent leurs troubles vocaux par rapport aux retentissements de ceux-ci sur leur vie quotidienne et non par rapport à l'altération de leurs qualités acoustiques* » (Lescarmontier, 2012, p. 86) [63]. D'autres études montrent cependant que la corrélation entre l'auto-évaluation par les patients de la sévérité globale de leur dysphonie et les résultats à un test de qualité de vie (comme le VHI par exemple) reste modérée (Behrman et al., 2004) [8], même si elle semble plus forte que la corrélation avec les tests perceptifs faits par un jury d'écoute (Ma & Yiu, 2001) [64].

Dans notre étude, les modifications apportées par le spectrogramme vont dans le sens d'une plus grande objectivité. Les feedback auditif et visuel permettent effectivement aux patients d'affiner leur évaluation de l'intensité, de la hauteur, de l'attaque. Les différents paramètres sont mieux distingués (moins corrélés entre eux) du fait de ces feedback additionnels. Les corrélations observées chez les patients s'approchent de celles observées chez les auditeurs du jury d'écoute lorsqu'on ajoute les feedbacks externes. Pour la clarté et la richesse, les corrélations avec certaines mesures objectives s'améliorent pour les patients du groupe protocole.

Moindre corrélation entre clarté et geste vocal

Dans l'étude de Eadie et ses collègues (2010) [25], les auteurs ont notamment voulu déterminer les différences entre les auto-évaluations de 20 locuteurs dysphoniques et 4 locuteurs contrôle, et les évaluations d'auditeurs expérimentés et inexpérimentés en terme de « sévérité globale » et d' « effort vocal »⁶⁶ sur des extraits de voix parlée en lecture, entendus au casque. La sévérité globale et l'effort vocal étaient évalués sur des échelles visuelles analogiques de 100 mm. On peut rapprocher ces échelles de nos propres échelles « clarté du timbre : altéré / clair » et « comportement vocal : adapté / non adapté ». Les auteurs y font la même observation que dans une étude précédente sur des patients dysphoniques : la corrélation entre les deux échelles « sévérité globale » et « effort vocal » est bien moindre lors des auto-évaluations que lors des évaluations par un jury d'écoute. Cela rejoint également nos conclusions concernant les corrélations entre « clarté du timbre » et « geste vocal » : ces deux échelles sont corrélées plus chez les auditeurs du jury d'écoute (**0,85**** à **0,78****), mais moins pour les patients (**0,71**** à **0,41****). On note d'ailleurs que cet effet est encore plus marqué après les feedbacks externes additionnels. Eadie et ses collègues font l'hypothèse que les patients, même lorsqu'ils jugent leur effort vocal par un extrait réentendu, ne tiennent pas uniquement compte des qualités acoustiques du signal audio. Ils citent Brandt et ses collègues⁶⁷, qui pensent que certains auditeurs peuvent également faire la relation entre ce qu'ils entendent et leur perception de l'effort dans la production de la voix. Cela expliquerait que les corrélations entre les jugements d'effort vocal par les patients sont plus fortes avec les auditeurs plus expérimentés qu'avec les auditeurs moins expérimentés (Eadie *et al.*, 2010, p. 571) [25]. Nous ne retrouvons pas cet effet de corrélation plus forte pour les auditeurs plus expérimentés dans notre étude, quelle que soit la condition du feedback.

⁶⁵ : « vocal perception and life quality » faculdade de medicina de ribeirao prteo, universidade de sao paulo

⁶⁶ “overall severity (OS)” and “vocal effort (VE)”

⁶⁷ Brandt J.F., Ruder K.F., Shipp T. Vocal loudness and effort in continuous speech. J Acoust Soc Am. 1969; 46:1543–1548.

Pour les auditeurs du jury d'écoute, le geste vocal reste corrélé à la clarté et un peu à la qualité de l'attaque. Pour les patients, il semble que le feedback externe modifie leurs évaluations de la clarté et de la richesse mais pas du geste, ce qui aboutit à cette dé-corrélation. Avant le feedback, les patients, en moyenne, jugent la richesse, la clarté et le geste d'un seul bloc (bloc comprenant, dans une moindre mesure, la qualité de l'attaque). Le feedback externe améliore la prise de conscience des différences entre qualité d'attaque, clarté et richesse, de façon indépendante du geste vocal : lorsque des modifications sont apportées à la clarté et à la richesse du timbre, les auditeurs ne modifient pas systématiquement leur évaluation du geste, comme si les auditeurs et les patients, ne « voyant » pas le geste vocal sur l'image, ne pensaient pas à le modifier par la suite pour le mettre en cohérence avec la clarté du timbre. Cet effet est particulièrement prégnant pour les patients, qui ont moins besoin que les auditeurs du jury d'écoute de se baser sur des feedbacks externes pour juger ce paramètre.

Impact sur les sensations proprioceptives

Pour Sapir et ses collègues (in Lee *et al.*, 2005, p. 360) [61], les patients se basent sur ce qu'ils entendent et sur ce qu'ils ressentent, tandis que les thérapeutes se basent sur ce qu'ils entendent et ce qu'ils voient du geste vocal. Behrman et ses collègues (2004, p. 1698) [8] soulignent également que les sensations d'effort vocal, d'inconfort musculaire conduisent les patients à évaluer leur voix différemment des thérapeutes de la voix, qui, eux, ne jugent du comportement vocal qu'à partir de ce qu'ils peuvent voir. Dans notre étude, le feedback visuel semble éloigner l'attention des patients de leur geste vocal, et dans le même temps, leur permet d'avoir des évaluations plus proches de celles des auditeurs du jury d'écoute. Dans le cadre d'une thérapie vocale, on pourrait se demander si l'évaluation première des patients, tenant apparemment plus fortement compte de leur geste vocal, n'est pas justement à valoriser, au contraire de l'évaluation perceptive et acoustique externe. Le thérapeute aura à choisir entre favoriser le caractère valorisant du feedback visuel et favoriser les sensations proprioceptives et la conscience du geste vocal. Les indications sont évidemment diverses selon les patients et selon leurs évolutions dans leur rééducation, sur le plan de la qualité subjective, objective, perceptive de leur voix, et sur le plan motivationnel et relationnel.

Enfin, on note que les patients, surtout ceux répondant au questionnaire en ligne, pensent que le spectrogramme affine leurs sensations proprioceptives, ce qui n'est pas confirmé par les résultats de l'analyse des exercices d'auto-analyse. Des études supplémentaires seraient nécessaires pour préciser cette question. On peut penser que suivant l'utilisation qui est faite du spectrogramme (en temps réel ou non, et selon les types d'exercices, la durée d'entraînement et d'autonomie par exemple), des réafférences peuvent s'établir, ou non, entre les sensations proprioceptives et les sensations auditives. Notre étude confirme que le feedback visuel a un grand intérêt quand les sensations proprioceptives sont perturbées (notamment pour la patiente B hypersensible à ses sensations internes). Nous pensons que le lien entre sensations proprioceptives et auditives est favorisé par une visualisation en temps réel, ce qui n'est pas le cas lors de notre exercice d'auto-analyse.

6.1.1.4 Conclusion sur les hypothèses de travail

Notre hypothèse 1 est validée : l'utilisation du spectrogramme en rééducation vocale permet aux patients une meilleure prise de conscience de leurs paramètres vocaux tels que l'intensité, la hauteur,

le caractère de l'attaque pour notre population. Ces paramètres sont, semble-t-il, influencés autant par le feedback auditif que par le feedback visuel.

Les feedbacks visuels et auditifs permettent également de distinguer dans la voix ce qui relève de la clarté et ce qui relève de la richesse du timbre, de façon indépendante de la qualité de l'attaque. Cette distinction est favorisée par l'ajout du feedback visuel constitué par le spectrogramme.

On note en revanche au travers de notre exercice d'auto-évaluation une absence de corrélation pour les patientes entre l'évaluation de la clarté et celle du geste vocal. Or notre exercice fait intervenir le feedback visuel comme un feedback externe après-coup. Nous sommes enclins à considérer que la prise de conscience des liens entre feedback visuel et feedback proprioceptif est favorisée par le feedback visuel en temps réel.

6.1.2 Hypothèse 2 : auto-contrôle des paramètres vocaux

Dans notre étude, malgré le faible nombre de séances de travail (3 séances) et une population réduite (4 patientes), on mesure une amélioration des productions des patientes du protocole en termes d'augmentation du rapport signal sur bruit et de diminution du shimmer, mais cependant d'autres paramètres objectifs (le jitter par exemple et le rapport des TMP z/s) se détériorent dans le même temps.

Les personnes répondant au questionnaire en ligne sont aussi partagées en ce qui concerne la possibilité d'améliorer les paramètres de la voix observés sur le spectrogramme. Le feedback auditif et visuel est souvent ressenti comme insuffisant pour améliorer leurs productions vocales. Il est considéré plutôt comme un moyen d'affiner leurs sensations auditives et proprioceptives, qui permettent, à leur tour, d'adapter leur geste vocal.

Bastian et Nagorsky (1987, p. 1347) [6] indiquent que le feedback visuel laryngoscopique avec des sujets dysphoniques doit venir en complément des perceptions habituelles, qui restent les principaux moyens de parvenir à un changement dans la production de la voix.

Comme Anna-Maria Laukkanen et ses collègues dans leur étude, nous constatons que certains sujets peuvent être plus à l'aise avec le feedback auditif, et que dans ce cas ils auront de moins bons résultats avec un feedback visuel, tandis que certains autres sujets, ne disposant pas d'une perception auditive suffisamment fine, auront au contraire absolument besoin du feedback visuel (Laukkanen *et al.*, 2004, p. 73) [58]. Ces auteurs notent que les sujets disposant du feedback visuel semblent avoir fait plus d'effort pour arriver au but, et que cela peut nuire à la qualité vocale que l'on recherche. Les auteurs préconisent de ne pas négliger les feedbacks sensoriels habituels et de toujours porter attention à la qualité de la voix (Laukkanen *et al.*, 2004, p. 74) [58]. Nous notons également des cas d'utilisation du spectrogramme en autonomie, où des comportements excessifs de patients qui voudraient « *faire une belle image* » pourraient nuire à l'établissement d'un comportement vocal adapté (Coudière, 2003, p. 134,147; Manteau, 2004, p. 66; Maulet, 1997, p. 86; Nair, 1999, p. 17; Roch, 2001, pp. 62–63; Verdolini & Krebs, 1999, pp. 236–238) [19] [65] [69] [78] [94] [107].

On note également que l'utilisation du spectrogramme peut induire des comportements vocaux inadaptés du fait du matériel lui-même : induisant une posture inadéquate, gênant à cause de la nécessité du micro, et du stress des dysfonctionnements informatiques. Ces inconvénients sont notés également dans le mémoire de Caroline Coudière (2003, p. 62,143) [19].

Conclusion : notre hypothèse 2 n'est pas validée. L'exercice d'auto-analyse était axé sur le feedback externe différé et s'appuyait sur 4 patientes seulement. Nous faisons l'hypothèse que le feedback en temps réel (plutôt que différé) favorise un meilleur auto-contrôle de ses paramètres vocaux par le patient et que nous ne disposions pas d'une population suffisamment nombreuse pour répondre à cette question.

6.1.3 Hypothèse 3 : construction d'une image plus positive de sa voix

6.1.3.1 Effets observés dans l'exercice d'auto-analyse

Les feedbacks auditif et visuel ont le plus souvent un effet gratifiant sur l'auto-évaluation de sa propre qualité d'attaque et la richesse de son timbre. Pour les patientes hors protocole, l'effet est visible aussi pour la clarté et le geste vocal.

Dans l'étude de Lee *et al.* (2005) [61], les auteurs concluent que le paramètre « Grade » du GRBAS était jugé (en moyenne) moins sévèrement par les cliniciens que par les patients (respectivement 1 et 1,4). Cet effet est également cité par Sapir *et al.*⁶⁸ (in Eadie *et al.*, 2010, p. 569) [25]. Nous ne retrouvons pas cet effet : en moyenne les patients sur-évaluent leur clarté. Les patients du groupe protocole se surévaluent surtout avant le feedback, et les patients hors protocole davantage après le feedback.

6.1.3.2 Observation des réactions face à l'image

Les réactions des patients face à l'image spectrographique sont diverses. Les réactions négatives, comme l'inquiétude, la déception, la frustration, doivent être contournées par le thérapeute. On observe alors de nombreuses réactions enthousiastes et narcissisantes comme le soulagement, la satisfaction, le plaisir, le jeu.

6.1.3.3 Effets observés par l'évolution des patients

Les pourcentages de satisfaction de chacun des patients du groupe protocole augmentent entre le début et la fin du protocole. Le VHI-10 diminue de 3,5 points en moyenne (sur 40) pour la voix parlée. L'examen détaillé des adjectifs utilisés pour caractériser leur voix et la voix qu'ils voudraient avoir montre que les patients ont tous une meilleure image de leur voix à la fin du protocole : ils lui donnent plus de qualificatifs positifs, et sont moins frustrés par rapport à des souhaits non réalisés. En détail, nous réalisons également que les patients sont moins influencés par les qualificatifs se rapportant aux émotions et aux sensations, et plus par les qualificatifs se rapportant aux paramètres acoustiques auxquels on s'est intéressé durant l'étude : intensité, clarté, richesse du timbre. On observe ici, comme dans l'exercice d'auto-analyse, une plus grande objectivité, bénéfique à l'image qu'ont les patients de leur voix. Cette façon de caractériser sa voix correspond plus à la façon de caractériser le timbre des auditeurs experts que des auditeurs naïfs (Ribaute, 2011, p. 148)⁶⁹ [93].

⁶⁸ Sapir S., Aronson A.E., Thomas J.E. (1986) Judgment of voice improvement after recurrent laryngeal nerve section for spastic dysphonia: clinicians versus patients. *Ann Otol Rhinol Laryngol.*95:137-141.

⁶⁹ Garnier M., Henrich N., Dubois D., Castellengo M., Poitevineau J., Sotiropoulos D. (2005) Etude de la qualité vocale dans le chant lyrique, *SCOLIA* n°20, pp.151-169.

6.1.3.4 Effets observés dans les commentaires en fin de protocole et le questionnaire en ligne

Les entretiens de fin de protocole montrent que le spectrogramme, plus qu'une simple réécoute de sa voix, permet une mise à distance du problème de voix (Cazade, 2000, p. 14) [13] et une dédramatisation.

L'attitude du thérapeute face à l'image et son interprétation positive des phénomènes observés y est pour beaucoup, mais également le fait que mettre à plat les difficultés ou les objectifs visés rend la relation à sa voix plus facile. Rendre objectifs les progrès effectués et les possibilités de reproduire avec sa voix le modèle du thérapeute, voire de faire mieux que le thérapeute, semble redonner aux patients confiance en leurs possibilités et leurs qualités vocales.

Nous avons même observé que lorsque les patients pouvaient être confrontés directement à leurs difficultés vocales, cela leur était bénéfique. Les patients expriment que le fait de pouvoir observer les difficultés donnait la possibilité de travailler dessus. La possibilité offerte par le spectrogramme de se référer à un modèle visuel positif plutôt qu'à un feedback verbal négatif (Pearson *et al.*, 2011, p. 174) [80] peut être également un facteur permettant la restauration d'une meilleure image de lui-même du patient.

Nous pensons que les mêmes mécanismes psychologiques que ceux décrits par Valérie Peloux peuvent être en jeu. Valérie Peloux explique en effet cet type d'effet pour les enfants malentendants qui utilisent le SpeechViewer : *« le rééducateur n'est plus celui qui corrige, juge, ou évalue les productions, Il devient celui qui "aide" l'enfant face à la machine, qui conseille et surveille les manipulations.[...] C'est l'ordinateur qui est investi des éléments de jugement, non plus l'adulte »* (Peloux, 1993, p. 61) [81].

Nous expliquons donc mal pourquoi les patientes semblent toutes moins sûres de leur voix, en répondant plus souvent avoir « peur que leur voix les trahissent » dans le questionnaire post-protocole. Cela peut être un effet des divers biais liés à ce questionnaire.

6.1.3.5 Conclusion

Notre hypothèse 3 est validée. Pour la construction d'une meilleure image de sa voix de la part du patient, il semble que l'image mise à distance et objectivée joue un rôle plus important que le feedback auditif concernant notre population. Cependant nous attirons l'attention sur le fait que la technique spectrographique seule ne suffit pas à produire cet effet. L'attitude positive du thérapeute est nécessaire.

6.1.4 Hypothèse 4 : impact positif sur la relation thérapeutique

6.1.4.1 Modification de la position du thérapeute

Caroline Coudière exprime le changement de la position du thérapeute lorsqu'un outil informatique est utilisé : *« l'ordinateur change la place du thérapeute qui va devenir un guide référent de l'utilisation de la machine. Cela entraîne une modification par rapport à la loi : la relation duelle enseignant/ enseigné disparaît, le rééducateur n'est plus dans la position d'un détenteur exclusif du savoir. La démarche se situe dans une optique de transfert de connaissances. L'ordinateur est aussi l'objet de phénomènes transférentiels. Il va devenir pour le thérapeute un partenaire et médiateur de la relation thérapeutique. La triade constituée peut permettre de débloquer une situation conflictuelle*

avec le rééducateur ou de surmonter un refus initial de prise en charge par le patient. Ce déplacement d'affects vers la machine permet aussi parfois d'éviter une approche trop directe du symptôme et de placer le patient dans une situation plus spontanée. [...] La relation praticien / patient existe aussi réellement quand on peut introduire des tours de rôle, stopper l'application pour faire des commentaires ou poser des questions » (Coudière, 2003, p. 61) [19].

Anne Menin-Sicard et Etienne Sicard (2006, p. 23) [71] pensent également que l'outil spectrographique constitue un « médiateur utile » : « *la relation devient alors triangulaire. La charge affective diminue grâce à la mise à distance au profit de la force de conviction du thérapeute* »

On note effectivement que la relation à trois avec l'ordinateur a pu être très appréciée (patiente B) comme elle a pu également être plus difficile (patiente A) selon les personnalités, les attentes et les antécédents parfois malheureux des patientes. La majorité des patientes du protocole ont cependant beaucoup apprécié l'emploi du spectrogramme et les séances du protocole.

6.1.4.2 Base de travail stable et partagée

Pour certaines, et notamment les patientes B et C, le spectrogramme permet l'instauration d'une relation apaisée avec le thérapeute, du fait de son côté (plus ou moins) objectif et indiscutable. Si l'interprétation de l'image spectrographique est subjective, en tout cas cet outil rend possible le fait de s'entendre sur ce dont on parle, en le montrant sur l'écran (Miller & Schutte, 1999a, p. 210) [72].

Harmoniser les perceptions et les attentes du patient et du thérapeute peut être considéré comme un objectif en rééducation vocale (Lee *et al.*, 2005, p. 360) [61]. En ce sens, l'utilisation du spectrogramme peut-être très bénéfique.

6.1.4.3 Nécessité de rester présent au patient

Caroline Coudière (2003, p. 64) [19] cite le dossier de N. Chaze⁷⁰ : « *60% des orthophonistes non informatisés trouvent que l'emploi de l'informatique en rééducation orthophonique est une contrainte nuisant gravement à la qualité et à la richesse de la relation ; alors que 97% des orthophonistes informatisés pensent que l'ordinateur a permis d'enrichir la relation et de compléter le travail.* » En ce sens, il semble que ce soit la non familiarité avec un outil qui fait craindre pour la qualité de la relation. Tout l'art de la rééducation orthophonique consiste à intégrer les différents outils à disposition sans cesser de porter une attention particulière à la relation thérapeutique.

Dans leur étude avec deux professeurs de chant, Welch et ses collègues observent que les professeurs de chant n'ont pas de difficulté à rester attentifs à l'élève, même lorsque les professeurs n'étaient pas familiers avec l'outil informatique (Welch *et al.*, 2005, p. 241) [110].

6.1.4.4 Conclusion

Notre hypothèse 4 est validée. Nous pensons que, pour notre population, l'image spectrographique, à laquelle on peut se reporter objectivement, est à l'origine des modifications bénéfiques de la relation thérapeutique plutôt que le feedback acoustique. De nouveau, nous précisons que c'est l'utilisation de l'outil à bon escient qui fait la relation thérapeutique, et non l'outil lui-même.

⁷⁰ N. Chaze, Apports et limites de l'utilisation de l'informatique en rééducation orthophonique, dossier paru sur <http://www.orthomalin.com>, septembre 2002

6.2 Limites

6.2.1 Biais relatif à la population

6.2.1.1 Biais de recrutement

Nous n'avons pu réaliser de recrutement systématique pour inclure les patients à notre travail. Dans ces conditions, le recrutement était à base de volontariat et les personnes qui étaient volontaires pour le protocole ont une relation à leur voix différente de la population tout-venant.

La seule personne qui a été recrutée par le biais d'une orthophoniste (Patiente A) est celle qui a le moins compris la finalité des séances du protocole, celle qui a été la plus critique vis-à-vis du spectrogramme et des exercices d'auto-analyse. Cela laisse penser que la population tout-venant d'un cabinet d'orthophonie pourrait plus souvent avoir cette attitude réservée face à cette approche que ce que nos résultats laissent entendre.

On peut citer également le peu d'enthousiasme des orthophonistes, et même parfois une franche incompréhension, face au protocole que nous proposons de mettre en place avec certains de leurs patients. C'est justement cet état de fait qui nous a contraints à modifier le protocole et à recruter des personnes en dehors des cabinets d'orthophonie.

De même, les personnes qui ont répondu au questionnaire en ligne sont très certainement celles qui sont les plus intéressées par notre étude.

Malgré le faible nombre de personnes questionnées par notre étude et le fort biais de recrutement, les avis et les expériences récoltées restent intéressants à prendre en compte.

6.2.1.2 Disparité patients hors protocole / patients protocole

Les patientes du groupe protocole étant toutes chanteuses, elles ont une connaissance et une pratique de leur voix bien différente de certains patients hors protocole, dont certains seulement sont chanteurs ou chanteuses.

Eadie et ses collègues ont étudié la différence entre musiciens et non musiciens dans leur fiabilité à évaluer les caractères rauques et soufflés de voix dysphoniques, en rapport avec leur perception de la hauteur tonale. Colton et Estill⁷¹ avaient montré que les musiciens reconnaissaient mieux les timbres de voix chantées non dysphoniques. A leur suite, Eadie et ses collègues concluent que les musiciens ont une meilleure fiabilité inter-juges que les non-musiciens, pour le caractère soufflé (et rauque dans une moindre mesure) de voix dysphoniques, sur des voyelles tenues.

Nous retrouvons cet effet, car l'erreur sur la clarté est moindre pour les patients du groupe protocole (12,4 mm) que pour ceux hors protocole (22,7 mm). De même, les corrélations entre richesse et clarté et certains paramètres acoustiques objectifs sont significatives dans le groupe protocole mais pas pour les patients hors protocole.

D'autre part le groupe hors protocole comprend des hommes et des femmes alors que le groupe protocole est entièrement constitué de femmes. Nous aurions pu tenter de discriminer entre les hommes et les femmes du groupe hors protocole afin d'étudier l'influence de la variable sexe. Le groupe hors protocole est constitué de 11 personnes alors qu'il n'y a que 4 patientes dans le groupe

⁷¹ Colton RH, Estill J. Elements of voice quality: perceptual, acoustic&physiologic aspects. In: *Lass, ed. Speech and language: advances in basic research and practice*. New York: Academic Press; 1981:311-403.

protocole. Les comparaisons statistiques entre ces deux groupes ne sont pas entièrement statistiquement valides. Enfin, les patients du groupe hors protocole ont des pathologies variées, concernant parfois l'audition plutôt que la production vocale.

6.2.1.3 Biais relatifs au questionnaire en ligne

Les réponses au questionnaire en ligne sont à prendre avec du recul : elles ne sont pas toujours cohérentes (par exemple la question sur le temps réel). Il semble que la règle des 3 « oui » (« *quand on a répondu oui à la première question, oui à la question suivante, on a de fortes probabilités de répondre également oui à la troisième* » (Py, 2010, p. 23) [84]) ait sensiblement faussé les réponses au questionnaire.

6.2.1.4 Biais relatifs au bilan subjectif

On note que la relation entre les patientes et l'expérimentateur n'est pas la même au début du protocole qu'à la fin. Cela peut impliquer des modifications dans chacun des questionnaires (même la liste d'adjectifs qui est pourtant remplie par le patient lorsqu'il est seul : en fin d'expérimentation, les patientes savent pour qui elles le remplissent, et la personnalité de l'expérimentateur influence leurs réponses, ce qui n'était pas le cas au début de l'expérimentation). On devra garder à l'esprit que les résultats de notre étude sont bien sûr liés à la façon dont l'expérimentateur a mené les séances et à la relation particulière tissée avec les patientes.

6.2.1.5 Biais relatifs au bilan perceptif et au jury d'écoute

Nous avons suivi au maximum les recommandations de Joana Revis (2004, pp. 90–94) [89]. En revanche, pour des raisons de disponibilité des auditeurs, une seule session d'écoute a été effectuée. De plus la durée de la session d'écoute était trop grande (près d'une heure et demie) pour que toutes les évaluations puissent être effectuées à la suite dans des conditions optimales.

D'autre part, malgré le soin apporté au matériel d'enregistrement, le micro serre-tête n'a pas été positionné de la même façon pour les évaluations avant et après le protocole, ce qui a pu gêner les évaluations perceptives du jury d'écoute. D'autre part, pour des raisons de disponibilité, 2 des auditeurs du jury d'écoute étaient également expérimentateurs, les 2 autres connaissant chacun une patiente du protocole. L'évaluation perceptive n'est pas effectuée rigoureusement en aveugle (Dejonckere *et al.*, 2001, p. 81) [23].

6.2.2 Évaluation difficile de l'apport du feedback visuel

6.2.2.1 Apport spécifique du feedback visuel par rapport au feedback auditif dans l'exercice d'auto-analyse

Notre étude distingue les apports du feedback visuel (spectrogramme) par rapport au feedback auditif (réécoute de l'extrait) de façon claire pour les auditeurs du jury d'écoute, pour lesquels la première condition d'évaluation est une écoute des deux sons, la deuxième condition étant l'écoute des deux sons avec la visualisation du spectrogramme.

En revanche, pour les patients, nous n'avons pas respecté ces deux conditions. En effet, nous voulions connaître leur évaluation immédiate, celle correspondant à ce qu'ils peuvent expérimenter dans leur vie quotidienne. Nous devons donc leur demander leur évaluation directement après avoir émis un son. Afin de ne pas multiplier les évaluations, l'évaluation avec uniquement feedback auditif n'a pas

été retenue. De plus cette décision a été prise avant de renoncer à constituer une population témoin de patients qui auraient travaillé avec un feedback auditif mais sans feedback visuel.

Cependant l'apport spécifique du feedback visuel a pu être appréhendé de différentes manières :

- par l'observation des réactions des patients en séance lorsqu'ils regardent l'image spectrographique sans réécouter leur voix,
- par l'observation des réactions des patients lors de la passation de l'exercice, et en leur demandant « c'est plutôt en entendant ou en voyant que vous changez d'avis sur votre production ? »,
- par l'introduction ponctuelle du feedback auditif seul entre la condition sans feedback externe et la condition avec feedback auditif et visuel. Lorsqu'elle modifiait les évaluations, cette condition avait pour résultat une évaluation intermédiaire entre les deux autres conditions : ce que les patients entendent est en général confirmé par la visualisation du spectrogramme, qui leur permet dans certains cas d'aller encore plus loin dans la modification de leur jugement. La distinction claire de ses deux feedbacks n'ayant été effectuée que sur 2 passations, de la même patiente, elle ne nous autorise pas à conclure,
- par comparaison des résultats moyens des patients du groupe protocole entre leur première passation (où l'on ajoute uniquement le feedback auditif) et les deuxièmes et troisièmes passations (où l'on ajoute en même temps les feedbacks auditif et visuel) de l'exercice d'auto-analyse. Les résultats de cette comparaison sont à interpréter avec réserve du fait du faible nombre de données (4 passations seulement), et du biais introduit par le fait que cette passation sans spectrogramme est systématiquement la première occurrence de l'exercice d'auto-analyse (les patients doivent effectuer l'exercice et analyser leur voix en détail pour la première fois),
- en considérant, avec la réserve que le jury d'écoute est plus expérimenté que les patients, que lorsque le feedback visuel ne modifie pas beaucoup les évaluations du jury d'écoute pour un paramètre, c'est que l'évaluation de ce paramètre dépend surtout du feedback auditif (comme c'est le cas pour l'intensité et surtout la hauteur).

6.2.2.2 Apport spécifique du feedback visuel par rapport aux interventions verbales de l'expérimentateur dans l'exercice d'auto-analyse

Lorsque les patients s'évaluent pour la deuxième fois, les expérimentateurs prennent en général plus de temps (en moyenne environ 10 minutes contre environ 4 min)⁷² pour questionner le patient et le guider, notamment afin de réinterpréter avec lui l'image spectrographique lorsqu'elle est mal comprise, et parfois pour faire référence à nouveaux aux extraits vocaux constituant les ancrages pour la richesse et la clarté du timbre. Dans ces conditions, il devient plus difficile de faire la part des effets induits par l'image spectrographique de ceux induits par les interventions des expérimentateurs. Cependant nous considérons le spectrogramme comme un outil mis à la disposition des thérapeutes. L'effet de cet outil est indéniablement lié à la façon dont les thérapeutes l'utilisent.

On note que la passation des patients hors protocole (hormis une patiente pour 2 passations) était faite en présence d'Isabelle Marié-Bailly, tandis que la passation des patients protocole était toujours faite

⁷² Calculé sur 19 passations mesurées sur les 23 (les 4 non mesurées concernent toutes les patients hors protocole)

par nous-même. La présence d'Isabelle Marié-Bailly et son approche spécifique induit éventuellement des attitudes différentes par rapport à l'exercice et par rapport au spectrogramme.

6.2.2.3 Rigueur du conditionnement

Il est clair que la complexité des tâches demandées lors de productions vocales variées ne se prêtait pas à un conditionnement type « tout ou rien » ou même sur une échelle, comme préconisé dans les méthodes de conditionnement strict, dans les méthodes de biofeedback (André *et al.*, 1986, pp. 295–300) [4]. De même, on n'a pas pu établir un protocole rigoureux avec un conditionnement avec feedback externe pendant une période, une évaluation intermédiaire, puis l'arrêt de ce feedback externe, et une deuxième évaluation à distance avec reprise du conditionnement instrumental si nécessaire (Volin, 1999, pp. 254–256) [108]. Nous n'avons donc notamment pas pu étudier si les acquis dus au spectrogramme se généralisaient sans spectrogramme. Pour cela, le protocole était trop court et la population trop peu nombreuse.

Un programme aussi strict nous semblait par ailleurs nous éloigner du sens de notre étude, qui se voulait proche des conditions de rééducation orthophonique en cabinet libéral, et qui se devait donc de s'adapter aux besoins des patients. En ce sens cependant, les patientes du protocole ont pointé que les séances proposées lors du protocole ne constituaient pas un soin puisqu'elles n'avaient pas l'objectif de réduire un trouble. Les séances consistaient plus en une découverte de leur voix et de l'apport du spectrogramme. Nous n'avons pas reproduit des séances de rééducation classiques où l'on aurait pointé ponctuellement des caractéristiques intéressantes à visualiser sur le spectrogramme, comme dans la leçon de chant décrite par Garyth Nair : « *L'image en temps réel est toujours affichée lorsque l'élève chante, mais elle est utilisée seulement pour des temps courts pendant la session pour répondre à des événements spécifiques et courts* »⁷³, le reste du temps la leçon est conduite de façon traditionnelle (Nair, 1999, p. 69) [78].

Il est toutefois notable que les découvertes et les jeux effectués avec le spectrogramme, bien que ne demandant pas un investissement considéré comme un travail, ont été relevés comme induisant des progrès et des prises de conscience. Or la prise de conscience est bien la première étape du soin (Amy de la Bretèque, 2012, p. 132; Estienne, 1998, p. 92; Howard *et al.*, 2012, p. 9; Roch, 2001, p. 61; Sarfati, 1998, pp. 53–54) [3] [27] [51] [94] [99].

6.2.3 Approche incomplète de la voix

Notre approche de la rééducation dans le protocole n'a jamais considéré l'articulation, le débit, l'intonation, le rythme de la parole. Ces paramètres ont cependant parfois été travaillés avec le spectrogramme, pour les patients hors protocole dans le cadre de leur rééducation phoniatrique classique.

Pour l'attaque, l'échelle bipolaire proposée est réductrice : une attaque peut être jugée à la fois soufflée et dure, ou de mauvaise qualité d'une autre manière (avec un port de voix par exemple). La même simplification est observée pour le timbre, évalué sur seulement deux échelles (clarté et richesse). En particulier, nous n'avons pas distingué souffle et raucité pour les altérations du timbre.

⁷³ “the VRTF image is running the entire time a student is singing, but it is used only briefly during the session to address specific short time events”

Cette donnée, ainsi que la couleur du timbre, pouvaient cependant trouver leur place en commentaire libre sur la feuille d'évaluation.

6.2.4 Complexité de l'exercice d'auto-analyse

L'exercice d'auto-analyse s'est avéré un peu trop complexe, cette complexité se traduisant souvent par un grand nombre de ratures sur les feuilles de passation :

- les extrémités à droite des échelles ne correspondent pas toujours à l'objectif : le son le plus clair, le plus riche, le geste le plus adapté sont des objectifs de l'exercice et sont, tous, tout à fait à droite. Cependant, pour l'attaque, le point le plus à droite correspond à une attaque très dure, ce qui n'est pas l'objectif de l'exercice. Cela a parfois conduit à des inversions d'échelles ;
- pour certains patients et pour certains auditeurs du jury d'écoute, placer deux signes l'un à côté de l'autre, séparés de 2 à 5 mm, signifie que la personne pensait que les deux sons étaient similaires pour ce paramètre,
- les signes sont parfois placés en dehors de la ligne, ce qui complique la mesure des évaluations,
- lorsqu'un son se modifie, il serait utile de pouvoir juger de ses paramètres avant et après le changement. Le cas s'est présenté deux fois avec la même patiente, pour laquelle le son plus fort commençait clair et finissait très altéré,
- un auditeur du jury d'écoute n'a pas utilisé l'échelle « clarté » à bon escient, s'intéressant plutôt à un timbre « sombre » plutôt qu'à un timbre « altéré » dans le sens du grade G du GRBASI.

6.3 Perspectives

6.3.1 Faciliter la pratique des orthophonistes

Notre travail pourrait être complété par la formalisation d'une liste plus exhaustive des exercices qui peuvent être proposés en rééducation vocale et pour lesquels la visualisation du spectrogramme apporte un feedback intéressant, comme nous avons commencé à le faire au paragraphe 10.3 en annexe. En outre il nous semblerait particulièrement judicieux de donner aux orthophonistes désirant utiliser cet outil des clés permettant de renvoyer au patient une image positive de lui-même et de dédramatiser les différentes caractéristiques de la dysphonie mises en évidence par l'image spectrographique.

6.3.2 Étude d'une utilisation en autonomie

Le questionnaire en ligne reflète que les personnes voulant utiliser le spectrogramme ressentent parfois que les exercices ne sont pas assez détaillés sur le site internet construit dans le cadre de ce mémoire, et qu'ils ont besoin d'un cadre de travail plus suivi. Il pourrait être intéressant d'étudier à quel point une autonomisation du patient est possible, souhaitable ou non.

6.3.3 Impact du temps réel notamment sur les sensations proprioceptives

Notre étude a montré que le spectrogramme garde un grand intérêt même lorsqu'il n'est pas utilisé en temps réel, mais juste après la phonation. Pour certains exercices et pour certains patients, c'est même cette utilisation qui est la plus profitable. Cette utilisation est peu décrite dans la littérature. Des études complémentaires sur le sujet sont nécessaires pour affiner ces affirmations.

La même remarque est également valable en ce qui concerne les sensations proprioceptives, qui, lors de l'utilisation du spectrogramme, peuvent être considérées comme volontairement mises de côté afin d'obtenir plus d'objectivité ou au contraire considérées comme grandement renforcées par un feedback visuel. Les effets semblent dépendre de la population, du thérapeute et des exercices proposés, des différentes utilisations (en temps réel ou non, de façon intensive ou non) du spectrogramme.

6.3.4 Réduction des biais observés dans notre étude

Chacun des biais observés dans notre étude gagnerait à être mieux contrôlé, notamment par des populations plus nombreuses et la présence d'une population témoin. Notre travail pourrait être le point de départ par exemple de l'observation de l'utilisation du spectrogramme au sein de séances de rééducations orthophoniques classiques, ce qui était notre projet initial.

6.3.5 Étude sémantique et linguistique des commentaires autour du spectrogramme

Notre étude comprend deux parties bien distinctes (exercice d'auto-analyse et recueil de commentaires autour de l'utilisation du spectrogramme), qui auraient mérité chacune des études plus approfondies, notamment dans l'analyse linguistique et sémantique de ce qui se dit autour du spectrogramme et autour de sa propre voix, pendant et après les séances.

6.3.6 Applications à d'autres pathologies de la parole

Le spectrogramme est également utilisé pour traiter la parole dans les surdités et les dysarthries. Pour aborder ce thème, la bibliographie sur l'utilisation du spectrogramme pour la parole, riche essentiellement dans l'apprentissage des langues étrangères, pourrait être explorée.

7 Conclusion

L'étude des exercices d'auto-analyse couplée à celle des commentaires autour du spectrogramme nous a permis de dégager plusieurs conclusions, pour notre population.

L'analyse de la voix par l'image spectrographique permet une meilleure prise de conscience et une meilleure connaissance des diverses composantes de la voix.

Cependant l'observation de ces caractéristiques ne débouche pas systématiquement vers une meilleure réalisation du geste vocal.

La visualisation du spectrogramme a un effet bénéfique sur l'image qu'a le patient de sa propre voix, et gratifiant pour ce qui est de la richesse de son timbre et sa qualité d'attaque. Cet effet est principalement dû à l'attitude positive du thérapeute face à l'image spectrographique.

Enfin, le travail avec un outil tiers objectif comme le spectrogramme peut être vécu comme un apaisement de la relation thérapeutique dans la mesure où le thérapeute devient un allié du patient davantage que son juge.

Nous préconisons l'emploi du spectrogramme en rééducation vocale, tout en veillant à respecter des éléments importants.

Il est nécessaire pour le thérapeute de valoriser l'image spectrographique de la voix du patient, et de rectifier, avec celui-ci, les idées faussées sur sa voix et/ou une image de sa voix dévalorisante.

Le spectrogramme s'intègre comme un outil supplémentaire au sein d'une rééducation vocale classique, mais ne se substitue pas à l'expertise des thérapeutes. Cet outil doit perturber au minimum le déroulement d'une séance habituelle : le thérapeute doit rester présent au patient et attentif à son geste vocal.

Une projection de l'image spectrographique sur grand écran et l'utilisation d'un micro qui puisse se faire oublier permettraient de limiter les inconvénients de l'utilisation de l'informatique sur le comportement vocal du patient et de son thérapeute, notamment en termes de posture et de sensation de liberté.

Le thérapeute de la voix doit garder à l'esprit que cet outil ne conviendra pas nécessairement à tous les patients, et son utilisation devra être adaptée aux besoins et à la personnalité de chacun. L'outil spectrographique s'avère cependant un allié considérable, notamment pour les personnes qui sont à la recherche de données scientifiques objectives, ainsi que pour celles qui ont une mauvaise image de leur propre voix.

Enfin il est nécessaire pour le thérapeute d'avoir la possibilité de proposer lui-même des exemples vocaux explicites, mettant en corrélation geste vocal et image spectrographique.

Nous espérons que la construction du site internet à l'occasion de ce mémoire soit une opportunité supplémentaire de démocratisation de la connaissance de la voix et l'occasion d'heureuses découvertes pour de nombreux patients, orthophonistes et professionnels de la voix.

8 Table des matières

Remerciements	2
Sommaire	3
Index des tableaux.....	1
Index des figures.....	2
Introduction	1
1 Assises théoriques.....	2
1.1 Dysphonie dysfonctionnelle et rééducation vocale	2
1.1.1 Dysphonie dysfonctionnelle et image de soi	2
1.1.1.1 Définition	2
1.1.1.2 Forçage vocal et dérèglement fonctionnel	2
1.1.1.3 Dysphonie et mauvaise image de soi	3
1.1.2 Rééducation vocale.....	3
1.1.2.1 Objectifs.....	3
1.1.2.2 Complémentarité des conceptions scientifiques et traditionnelles	4
L'orthophoniste partenaire du médecin et du chirurgien	4
La dimension psychologique.....	4
Technique de rééducation orthophonique	5
1.1.2.3 Conclusion	5
1.1.3 Rétro-contrôle (feedback) en rééducation vocale.....	5
1.1.3.1 Définition de l'apprentissage par feedback.....	5
1.1.3.2 Trois phases de l'apprentissage moteur	6
1.1.3.3 Particularité du rétro-contrôle pour la production vocale	6
Rétro-contrôles habituels pour la voix	6
Sensoriel, proprioceptif, postural	6
Auditif : boucle audio-phonatoire	7
Feedback verbal.....	7
Imitation	8
Suppléer les feedbacks auditifs, proprioceptifs et verbaux	8
Particularité de l'apprentissage vocal en tant que comportement moteur.....	9
1.1.3.4 Feedback acoustique en rééducation vocale	9
1.1.3.5 Conclusion	10
1.2 Le spectrogramme.....	10
1.2.1 Présentation du spectrogramme	10
1.2.1.1 Notions d'acoustique	10
1.2.1.2 Définition	11
1.2.1.3 Lecture de spectrogramme : la parole.....	11
1.2.1.4 Lecture de spectrogramme : le timbre	11
Définition	11
Timbre vocalique.....	12

Les harmoniques	12
Le bruit entre les harmoniques	12
Classifications	13
Exemples en voix chantée	14
Singing formant.....	14
Voix soufflée	14
Voix serrée	14
1.2.1.5 Lecture de spectrogramme : l'attaque.....	14
1.2.1.6 Conclusion	15
1.2.2 Utilisation du spectrogramme en temps réel	15
1.2.2.1 Formation vocale des chanteurs.....	15
1.2.2.2 Langue étrangères et accent, prosodie	16
1.2.2.3 Parole dans le cas de surdités.....	17
1.2.2.4 Dysarthrie.....	17
1.2.2.5 Rééducation vocale	18
1.2.2.6 Conclusion	18
1.3 Spectrogramme et rééducation vocale : avantages et limites.....	18
1.3.1 Avantages	19
1.3.1.1 Prise de conscience	19
Importance de la prise de conscience en rééducation vocale	19
Apport du spectrogramme	19
1.3.1.2 Motivation.....	20
1.3.1.3 Autonomisation.....	21
1.3.1.4 Enrichissement de compétences, diversification du travail orthophonique.....	21
1.3.1.5 Conclusion	21
1.3.2 Limites de l'analyse spectrographique.....	22
1.3.2.1 Fiabilité de l'image	22
1.3.2.2 Complémentarité des approches classiques et du feedback visuel	22
1.3.2.3 Pertinence de l'information donnée par le spectrogramme	23
1.3.2.4 Nécessité d'une formation spécifique.....	24
1.3.2.5 Conclusion	25
1.3.3 Aspects ambivalents	25
1.3.3.1 Concentration et charge cognitive	25
1.3.3.2 Nécessité du temps réel.....	26
1.3.3.3 Analyse fine de sa production vocale	27
1.3.3.4 Lien avec les feedbacks auditifs et proprioceptifs	27
1.3.3.5 Nécessité de retirer progressivement le feedback.....	28
1.3.3.6 Objectivité / subjectivité	28
Evaluation et confirmation de l'intérêt de la rééducation	28
Feedback verbal facilité	29
L'interprétation reste subjective.....	29
1.3.3.7 Conclusion	30
2 Problématique.....	31

3	Hypothèses	32
4	Méthodologie.....	33
4.1	Matériel	33
4.1.1	Présentation générale.....	33
4.1.2	Description des populations.....	33
4.1.2.1	Patients protocole.....	33
	Généralités.....	33
	La patiente A.....	33
	La patiente B.....	34
	La patiente C.....	34
	La patiente D.....	34
4.1.2.2	Patients hors protocole.....	34
4.1.2.3	Personnes répondant au questionnaire en ligne.....	34
4.1.2.4	Jury d'écoute.....	34
4.2	Méthode.....	35
4.2.1	Protocole.....	36
4.2.1.1	Déroulement général.....	36
4.2.1.2	Evaluation pré- et post- protocole.....	36
	Auto-évaluation du handicap ressenti.....	36
	Bilan vocal.....	36
	Court questionnaire sur la relation à sa voix.....	37
	Entretien semi-dirigé.....	38
	Qualification de sa propre voix telle qu'elle est et telle qu'on voudrait qu'elle soit.....	38
4.2.1.3	Utilisation du spectrogramme.....	38
4.2.1.4	Exercices proposés.....	38
4.2.1.5	Ressenti face au spectrogramme.....	40
4.2.2	Patients en autonomie	41
4.2.2.1	Site internet.....	42
4.2.2.2	Questionnaire en ligne.....	42
4.2.3	Exercice d'auto-analyse	43
4.2.3.1	Description de l'exercice d'auto-analyse.....	43
	Principe de l'exercice.....	43
	Intérêt de l'exercice.....	43
	Échelles visuelles analogiques.....	44
	Banque de sons.....	45
	Conditions d'enregistrement et d'écoute.....	46
	Conditions de passation.....	46
	Condition de l'évaluation perceptive.....	46
4.2.3.2	Méthode d'analyse des résultats de l'exercice d'auto-analyse.....	47
	Mesures / vérifications / recherche des valeurs aberrantes.....	47
	Calcul du paramètre « qualité d'attaque ».....	47
	Mesures objectives.....	47
	Homogénéisation des évaluations du jury.....	47

Choix des paramètres étudiés	48
Variables gratification / aide / erreur / corrélation	48
Gratification	48
Erreur / aide	48
Résumé utilisation / gratification / aide / erreur	48
Corrélations	49
Nature du feedback additionnel et nombre de données.....	49
Fiabilité des évaluations du jury d'écoute pour l'exercice d'auto-analyse	50
5 Analyse des résultats	51
5.1 Auto-analyse : échelles visuelles analogiques.....	51
5.1.1 Utilisation des échelles	51
5.1.1.1 Distribution des marques sur les échelles visuelles analogiques	51
Remarque préliminaire :	51
5.1.1.2 Corrélations entre les différentes échelles (tous groupes confondus et tous feedbacks confondus).....	53
Pôle clarté / geste / qualité d'attaque.....	53
Pôle richesse /intensité / caractère de l'attaque (soufflée à dure).....	53
Geste.....	54
Hauteur	54
5.1.1.3 Synthèse des corrélations	54
5.1.2 Fiabilité de l'évaluation subjective	54
5.1.2.1 Fiabilité intra-juge.....	54
Conclusion sur la fiabilité intra-juge	54
Effet d'apprentissage supposé : test/retest.....	55
Effet d'apprentissage	55
5.1.2.2 Fiabilité inter-juges	56
Conclusion sur la fiabilité inter-juges de la mesure subjective par des auditeurs du jury d'écoute	56
5.1.3 Corrélation avec les critères acoustiques	57
5.1.4 Influence de l'image sur la perception de la qualité de la réalisation	58
5.1.4.1 Utilisation du feedback additionnel	58
Généralités.....	58
Pour le feedback auditif et visuel, effet de l'appartenance à un groupe (jury d'écoute, patients hors protocole, patients protocole).....	59
5.1.4.2 Effet gratifiant du feedback additionnel	59
Prise de conscience d'un défaut	59
Prise de conscience d'une qualité, confiance en soi :	60
Pour le feedback auditif et visuel, effet de l'appartenance à un groupe (jury d'écoute, patients hors protocole, patients protocole).....	60
Différence pour les patients du groupe protocole, entre la première passation sans spectrogramme, et celles avec spectrogramme	60
5.1.4.3 Aide du feedback additionnel	60
5.1.4.4 Mise en perspective erreurs / utilisation / aide / gratification	61

Richesse, clarté et qualité d'attaque	61
Richesse du timbre	61
Clarté du timbre.....	62
Qualité d'attaque	62
Différence d'intensité	62
Geste.....	63
Différence de hauteur	63
5.1.4.5 Corrélations entre les évaluations des différents paramètres	64
Richesse et intensité	64
Intensité et attaque.....	64
Intensité et clarté	65
Clarté et hauteur	65
Isolation de l'évaluation du geste.....	65
5.1.4.6 Conclusion	65
5.2 Évolution des patients du groupe « protocole ».....	66
5.2.1 Bilan vocal.....	66
5.2.1.1 Évaluation objective.....	66
5.2.1.2 Évaluation perceptive.....	66
5.2.2 Évolution de l'exercice d'auto-analyse	67
5.2.3 Modification de la façon de voir sa voix.....	67
5.2.3.1 Évolution des questionnaires d'auto-évaluation	67
5.2.3.2 Liste d'adjectifs.....	68
5.2.3.3 Conclusion	71
5.3 Commentaires autour du spectrogramme	71
5.3.1 Impact sur le geste vocal.....	71
5.3.1.1 Sensations acoustiques et proprioceptives	71
Réponses au questionnaire en ligne	71
Possible focalisation sur l'image au détriment des sensations.....	71
Ouverture des sensations auditives	72
Ouverture des sensations proprioceptives	72
Aide lorsque les sensations et l'audition interne sont perturbées.....	72
5.3.1.2 Temps réel.....	73
Résultats du questionnaire.....	73
Besoin de se concentrer sur le geste vocal	73
Besoin de liberté, pas de supervision	73
Temps réel plus utile sur des phénomènes longs	73
Intérêt du spectrogramme même de façon différée	73
5.3.1.3 Comportement vocal.....	74
Résultats du questionnaire.....	74
Ordinateur et posture	74
Problème du micro	74
Problème de la visualisation sur un écran	74
Prise de conscience d'un comportement qui modifie la phonation.....	75

5.3.2	Intérêts particuliers de la visualisation du spectrogramme	75
5.3.2.1	Réactions et émotions face à l'image.....	75
	Inquiétude.....	75
	Étonnement.....	75
	Agacement.....	75
	Insatisfaction	76
	Déception	76
	Enthousiasme	76
	Satisfaction	76
	Prise de conscience.....	76
	Soulagement	76
	Plaisir et jeu.....	76
5.3.2.2	Paramètres vocaux que le spectrogramme permet d'observer et d'améliorer	77
	Résultats du questionnaire.....	77
	Impact de la visualisation d'un timbre altéré	77
	Possible image « violente » de la dysphonie.....	77
	Intérêt majeur : pouvoir travailler dessus	78
	Attaques facilement visualisées	78
	Phonétique	78
	Harmoniques	78
5.3.2.3	Moyen d'analyse scientifique d'étude et de connaissance	79
	Moyen de satisfaire une soif de connaissance.....	79
	Possibilité d'adopter un protocole de rééducation encourageant	79
	Attirer l'attention sur une partie du son	79
	Intérêt intellectuel au détriment d'une utilité pour la voix ?	79
5.3.2.4	Reflet objectif rendu positif par le thérapeute.....	80
	Reflet objectif.....	80
	Possibilité de mettre son problème de voix à distance, et dédramatiser	80
	Possibilité d'observer l'évolution d'un travail	81
	Possibilité d'une relation apaisée avec le thérapeute	81
	Impact positif de l'interprétation par le thérapeute	81
5.3.3	Utilisations préconisées et possibilité d'autonomie	82
5.3.3.1	Public visé.....	82
	Résultats du questionnaire.....	82
	Quel type de public	82
	Utilisation pour les chanteurs avancés	82
	Utilisation par les professionnels de la voix.....	83
	Utilisation pour les dysphoniques	83
5.3.3.2	Compréhension de l'image spectrographique.....	83
	Résultats du questionnaire.....	83
	Remarques qualitatives	83
5.3.3.3	Utilisation en autonomie	84
	Résultats du questionnaire.....	84

Problèmes techniques et convivialité	85
Nécessité d'un accompagnement pour interpréter l'image	85
Nécessité d'un accompagnement pour guider le travail.....	85
Nécessité d'un objectif	86
Utilisation autonome ludique	86
5.3.4 Conclusion.....	86
6 Discussion.....	87
6.1 Validation des hypothèses de travail	87
6.1.1 Hypothèse 1 : prise de conscience des paramètres vocaux.....	87
6.1.1.1 Fiabilité inter-juges	87
6.1.1.2 Observation des paramètres	87
6.1.1.3 Auto-évaluation de ses propres qualités vocales : exercice d'auto-analyse	87
Meilleure distinction des différents paramètres	87
Meilleure objectivité	88
Moindre corrélation entre clarté et geste vocal	88
Impact sur les sensations proprioceptives	89
6.1.1.4 Conclusion sur les hypothèses de travail	89
6.1.2 Hypothèse 2 : auto-contrôle des paramètres vocaux.....	90
6.1.3 Hypothèse 3 : construction d'une image plus positive de sa voix	91
6.1.3.1 Effets observés dans l'exercice d'auto-analyse	91
6.1.3.2 Observation des réactions face à l'image.....	91
6.1.3.3 Effets observés par l'évolution des patients.....	91
6.1.3.4 Effets observés dans les commentaires en fin de protocole et le questionnaire en ligne	92
6.1.3.5 Conclusion	92
6.1.4 Hypothèse 4 : impact positif sur la relation thérapeutique	92
6.1.4.1 Modification de la position du thérapeute	92
6.1.4.2 Base de travail stable et partagée	93
6.1.4.3 Nécessité de rester présent au patient	93
6.1.4.4 Conclusion	93
6.2 Limites.....	94
6.2.1 Biais relatif à la population	94
6.2.1.1 Biais de recrutement	94
6.2.1.2 Disparité patients hors protocole / patients protocole.....	94
6.2.1.3 Biais relatifs au questionnaire en ligne	95
6.2.1.4 Biais relatifs au bilan subjectif.....	95
6.2.1.5 Biais relatifs au bilan perceptif et au jury d'écoute	95
6.2.2 Évaluation difficile de l'apport du feedback visuel.....	95
6.2.2.1 Apport spécifique du feedback visuel par rapport au feedback auditif dans l'exercice d'auto-analyse.....	95
6.2.2.2 Apport spécifique du feedback visuel par rapport aux interventions verbales de l'expérimentateur dans l'exercice d'auto-analyse.....	96
6.2.2.3 Rigueur du conditionnement.....	97

6.2.3	Approche incomplète de la voix	97
6.2.4	Complexité de l'exercice d'auto-analyse	98
6.3	Perspectives.....	98
6.3.1	Faciliter la pratique des orthophonistes.....	98
6.3.2	Étude d'une utilisation en autonomie.....	98
6.3.3	Impact du temps réel notamment sur les sensations proprioceptives	99
6.3.4	Réduction des biais observés dans notre étude.....	99
6.3.5	Étude sémantique et linguistique des commentaires autour du spectrogramme	99
6.3.6	Applications à d'autres pathologies de la parole.....	99
7	Conclusion.....	100
8	Table des matières	101
9	Bibliographie.....	109
10	Annexes	117
10.1	Classifications des images spectrographiques caractérisant la qualité vocale	117
10.2	Évolution objective des patients du groupe protocole	118
10.3	Le spectrogramme en pratique	119
10.3.1	Introduction du spectrogramme	119
10.3.1.1	Temps qui passe	119
10.3.1.2	Hauteur.....	119
10.3.1.3	Intensité.....	120
10.3.1.4	Bruit, bruit entre les harmoniques.....	120
10.3.2	Travail avec les consonnes fricatives et la paille.....	120
10.3.2.1	Paille	120
10.3.2.2	Consonnes fricatives	121
	Respiration et posture.....	121
	Différenciation des consonnes.....	121
	Alternance voisé / non voisé	121
	Vers les voyelles.....	121
10.3.3	Travail des attaques et du mécanisme 0 (fry)	122
10.3.3.1	Trois types d'attaque.....	122
10.3.3.2	Mécanisme 0 (fry).....	122
10.3.3.3	Pouff d'air	122
10.3.4	Exploration des formants vocaliques.....	122
10.3.4.1	Sur des consonnes	122
10.3.4.2	A partir du fry	122
10.3.4.3	Sur des voyelles (« château des voyelles »).....	123
10.3.4.4	Transitions entre voyelles	123
10.3.5	Exploration de la voix chantée	123
10.4	Documents du protocole	124
10.5	Questionnaire en ligne sur le spectrogramme	129
10.6	Exemples d'images spectrographiques.....	130

9 Bibliographie

- [1] Aimadeddine, S., Kerlan, M. (2003). *Le rôle de la rééducation orthophonique dans la prise de conscience de la voix et l'interprétation du trouble vocal, en cas de dysphonie avec composante fonctionnelle : chez deux populations culturellement différentes, française et marocaine : essai d'analyse de quelques cas*. Mémoire d'orthophonie de l'Université de Besançon, Besançon.
- [2] Amy de la Bretèque, B. (2004). *L'équilibre et le rayonnement de la voix*. Marseille: Solal., 122 p.
- [3] Amy de la Bretèque, B. (2012). Règles générales de la rééducation vocale. In *La voix parlée et la voix chantée* (pp. 130–147). Montpellier: Sauramps médical.
- [4] André, Brugerolle de Fraisinette, B., Chellig, L. (1986). le biofeedback en rééducation motrice. *Annales de réadaptation et de médecine physique*, (29), 238–310.
- [5] Barnes-Burroughs, K., Lan, W.Y., Edwards, E., Archambeault, N. (2008). Current attitudes toward voice studio teaching technology: a bicoastal survey of classical singing pedagogues. *Journal of voice*, 22(5), 590–602.
- [6] Bastian, R.W., Nagorsky, M.J. (1987). Laryngeal image biofeedback. *Laryngoscope*, 97(11), 1346–1349.
- [7] Beauchene, P., Bouquay, A. (1994). *Approche rééducative de la dysarthrie post-traumatique et de la dysprosodie sur le logiciel Speech Viewer II*. Mémoire d'orthophonie de l'Université de Nantes, Nantes
- [8] Behrman, A., Sulica, L., He, T. (2004). Factors predicting patient perception of dysphonia caused by benign vocal fold lesions. *Laryngoscope*, 114(10), 1693–1700.
- [9] Bele, I.V. (2005). Reliability in Perceptual Analysis of Voice Quality. *Journal of Voice*, 19(4), 555–573.
- [10] Bergeras, M., Tain, L. (2007). Des orthophonistes dans le monde de la voix : savoir, savoir-faire et légitimité. In *Le métier d'orthophoniste : langage, genre et profession* (pp. 173–180). Rennes: Éd. de l'École nationale de la santé publique.
- [11] Blayac, J.P., Bensoussan, P.-L., Marotin, O. (2012). La voix dans le soin, la voix en musicothérapie. In *La voix parlée et la voix chantée* (pp. 148–155). Montpellier: Sauramps médical.
- [12] Borel, S. (2002). *Evaluation de l'impact de la paralysie récurrentielle sur la qualité de vie à l'aide du "Voice Handicap Index" : étude de 6 cas*. Mémoire d'orthophonie de l'Université Pierre et Marie Curie, Paris.
- [13] Cazade, A. (2000). De l'usage des courbes sonores et autres supports graphiques pour aider l'apprenant en langues. *Alsic. Apprentissage des Langues et Systèmes d'Information et de Communication*, 2(2), 3–32.
- [14] Chan, K.M.K., Yiu, E.M.-L. (2002). The effect of anchors and training on the reliability of perceptual voice evaluation. *Journal of speech, language, and hearing research: JSLHR*, 45(1), 111–126.

- [15] Chevré-Muller, C. (1990). Ecouter, mesurer la voix. In *L'Esprit des voix : études sur la fonction vocale* (pp. 53–67). Grenoble: la Pensée sauvage Ed.
- [16] Cooper, M. (1974). Spectrographic Analysis of Fundamental Frequency and Hoarseness before and after Vocal Rehabilitation. *Journal of Speech and Hearing Disorders*, 39(3), 286–297.
- [17] Cornut, G. (2009). *La voix*. Paris: Presses universitaires de France. 127p.
- [18] Cotellet, S. (1993). *Etude de la relation patient-thérapeute dans la rééducation vocale : Echo du féminin et du masculin*. Mémoire d'orthophonie de l'Université Pierre et Marie Curie, Paris.
- [19] Coudière, C. (2003). *De l'utilité des logiciels pour la voix en rééducation de dysphonies dysfonctionnelles*. Mémoire d'orthophonie de l'Université Paul Sabatier, Toulouse.
- [20] Deary, I.J., Wilson, J.A., Carding, P N., Mackenzie, K. (2003). The dysphonic voice heard by me, you and it: differential associations with personality and psychological distress. *Clinical Otolaryngology & Allied Sciences*, 28(4), 374–378.
- [21] Dejonckere, Ph. (2001). La systématisation du traitement orthophonique en éléments : une méthode pour catégoriser et quantifier les composantes d'un traitement fonctionnel. In *Dysphonies et rééducations vocales de l'adulte* (pp. 349–356). Marseille: Solal.
- [22] Dejonckere, Ph. (2006). améliorer la fiabilité de l'analyse perceptive de la voix pathologique par un support visuel sonographique. In *Voix parlée et chantée* (pp. 35–42). Ville-d'Avray: C. Klein-Dallant.
- [23] Dejonckere, Ph., Bradley, P., Clemente, P., Cornut, G., Crevier-Buchman, L., Friedrich, G., Woisard, V. (2001). A basic protocol for functional assessment of voice pathology, especially for investigating the efficacy of (phonosurgical) treatments and evaluating new assessment techniques. Guideline elaborated by the Committee on Phoniatics of the European Laryngological Society (ELS). *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*, 258(2), 77–82.
- [24] Delacroix, L., Morsomme, D., Remacle, M. (2009). Rééducation logopédique des paralysies cordales. In *Précis d'audiophonologie et de déglutition. Tome II, Les voies aéro-digestives supérieures* (pp. 423–430). Marseille: Solal.
- [25] Eadie, T.L., Kapsner, M., Rosenzweig, J., Waugh, P., Hillel, A., Merati, A. (2010). The role of experience on judgments of dysphonia. *Journal of Voice*, 24(5), 564–573.
- [26] Espanol, C. (2011). *Etude de différents facteurs pouvant favoriser la récupération dans les paralysies récurrentielles unilatérales post-chirurgicales*. Mémoire d'orthophonie de l'Université Pierre et Marie Curie, Paris.
- [27] Estienne, F. (1998). *Voix parlée, voix chantée : examen et thérapie*. Paris: Masson, 194 p.
- [28] Estienne, F. (2001). Examen et thérapie de la voix, réflexions et actions. In *Dysphonies et rééducations vocales de l'adulte* (pp. 27–49). Marseille: Solal.

- [29] Estienne, F., Morsomme, D. (2009). Traitement orthophonique des pathologies vocales. In *Précis d'audiophonologie et de déglutition. Tome II, Les voies aéro-digestives supérieures* (pp. 257–268). Marseille: Solal.
- [30] Estienne, F., Piérart, B. (2006). Les bilans de langage et de voix : fondements théoriques et pratiques. Paris: Masson, 312 p.
- [31] Ferrand, C.T. (1995). Effects of practice with and without knowledge of results on jitter and shimmer levels in normally speaking women. *Journal of Voice*, 9(4), 419–423.
- [32] Fisher, J., Kayes, G., & Reeve, M. (2003). Advanced Vocal Training with Vocal Process: Computer Voice Analysis. Poster presented at the *Fifth Pan European Voice Conference (PEVoC 5)*, Graz.
- [33] Foisneau, S., Riedel, S., Chevalier, D., & Decorte, D. (2003). *La voix dans tous ses états*. Isbergues: Ortho éd., 197 p.
- [34] Fortin, E., Guth, P., Tain, L. (2007). Entre “parler droit” et “sonner juste” : les arrangements entre orthophonistes et patientèle au cours d’une rééducation vocale. In *Le métier d’orthophoniste : langage, genre et profession* (pp. 235–243). Rennes: Éd. de l’École nationale de la santé publique.
- [35] Fournier, C. (1989). *La voix, un art et un métier : anatomie, physiologie, acoustique, phonétique : technique de la voix professionnelle*. Seyssel, France: CCL editions, 241 p.
- [36] Fournier, C., Dupessey, M. (2001). Education et rééducation vocales : les bases d’une approche parallèle. Cas particulier des comédiens et des chanteurs. In *Dysphonies et rééducations vocales de l’adulte* (pp. 221–245). Marseille: Solal.
- [37] Fugain, C. (1990). La dysphonie fonctionnelle. In *L’Esprit des voix : études sur la fonction vocale* (pp. 99–106). Grenoble: la Pensée sauvage Ed.
- [38] Ghio, A. (2012). bilan instrumental de la dysphonie. In *La voix parlée et la voix chantée* (pp. 69–104). Montpellier: Sauramps médical.
- [39] Gillie, C. (2007). *La Voix au Risque de la Perte ; de l’aphonie à “l’a” phonie*. Paris 7, Ecole Doctorale de Recherches en Psychanalyse.
- [40] Giovanni, A., Aumelas, E. (2004). Anatomie et physiologie de la phonation. In *Le bilan d’une dysphonie : état actuel et perspectives* (pp. 7–43). Marseille: Solal.
- [41] Giovanni, A., Fenollar, C. (2004). Analyse subjective par le patient de son handicap vocal. In *Le Bilan D’une Dysphonie : État Actuel Et Perspectives* (pp. 45–66). Marseille: Solal.
- [42] Giovanni, A., Robieux, C., Santini, L. (2012). Physiologie de la phonation. In *La voix parlée et la voix chantée* (pp. 9–16). Montpellier: Sauramps médical.
- [43] Hart, M. (2002a). L’analyse sonagraphique : un nouvel outil pédagogique ? *Le journal de l’AFPC*, (9), 25–31.

- [44] Hart, M. (2002b). Utilisation de l'analyse sonographique de la voix chantée dans les écrits de trois grands pédagogues américains : William Vennard, D. Ralph Appelman, Berton Coffin. *Le journal de l'AFPC*, (9), 52–54.
- [45] Heinrich, N. (2012). Physiologie de la voix chantée : vibrations laryngées et adaptations phonorésonnantes. In *La voix parlée et la voix chantée* (pp. 17–32). Montpellier: Sauramps médical.
- [46] Hoppe, D., Sadakata, M., Desain, P. (2006). Development of real-time visual feedback assistance in singing training: a review. *Journal of computer assisted learning*, 22(4), 308–316.
- [47] Howard, D.M. (2002). The real and the non-real in speech measurements. *Medical engineering & physics*, 24(7), 493–500.
- [48] Howard, D.M., Welch, G.F., Brereton, J., Himonides, E., DeCosta, M., Williams, J., Howard, A. W. (2004). WinSingad: A real-time display for the singing studio. *Logopedics Phoniatrics Vocology*, 29(3), 135–144.
- [49] Howard, D.M. (2005). Human hearing modelling real-time spectrography for visual feedback in singing training. *Folia phoniatica et logopaedica*, 57(5-6), 328–341.
- [50] Howard, D.M, Brereton, J., Welch, G.F., Himonides, E., Decosta, M., Williams, J., Howard, A. W. (2007). Are real-time displays of benefit in the singing studio? An exploratory study. *Journal of voice*, 21(1), 20–34.
- [51] Howard, D.M., Abberton, E., Fourcin, A. (2012). Disordered voice measurement and auditory analysis. *Speech Communication*, 54(5), 611–621
- [52] Julien, P. (2009). *Perception de la musique et dysphonie*. Mémoire d'orthophonie de l'Université d'Aix-Marseille, Marseille.
- [53] Klein-Dallant, C. (2001). Le travail de l'espace et de la résonance. In *Dysphonies et rééducations vocales de l'adulte* (pp. 119–125). Marseille: Solal.
- [54] Kreiman, J., Gerratt, B.R., Kempster, G.B., Erman, A., Berke, G S. (1993). Perceptual evaluation of voice quality: review, tutorial, and a framework for future research. *Journal of speech and hearing research*, 36(1), 21–40.
- [55] Kreiman, J., Gerratt, B.R., Precoda, K., Berke, G.S. (1992). Individual differences in voice quality perception. *Journal of speech and hearing research*, 35(3), 512–520.
- [56] Lamesch, S. (2012). Les ressources du numérique dans l'enseignement du chant. *Journal de l'AFPC*, 9, 61–65.
- [57] Larson, C.R., Altman, K.W., Liu, H., Hain, T.C. (2008). Interactions between auditory and somatosensory feedback for voice F₀ control. *Experimental Brain Research*, 187(4), 613–621.
- [58] Laukkanen, A.-M., Syrjä, T., Laitala, M., Leino, T. (2004). Effects of two-month vocal exercising with and without spectral biofeedback on student actors' speaking voice. *Logopedics, phoniatrics, vocology*, 29(2), 66–76.

- [59] Le Huche, F., Allali, A. (1989). *La voix. Tome 4*. Paris: Masson, 211 p.
- [60] Leclerc, C. (2007). “Spécialistes” de la rééducation vocale : des parcours entre deux mondes. In *Le métier d’orthophoniste : langage, genre et profession* (pp. 105–116). Rennes: Éd. de l’École nationale de la santé publique.
- [61] Lee, M., Drinnan, M., Carding, P. (2005). The reliability and validity of patient self-rating of their own voice quality. *Clinical Otolaryngology*, 30(4), 357–361.
- [62] Leipp, E. (2010). *Acoustique et musique : données physiques et technologiques, problèmes de l’audition des sons musicaux, principes de fonctionnement et signification acoustique des principaux archétypes d’instruments de musique, les musiques expérimentales, l’acoustique des salles*. Paris: Presses des Mines-Mines ParisTech., 380 p.
- [63] Lescarmontier, E. (2012). *Exclusion du facteur personnel dans le jugement des dysphonies*. Mémoire d’orthophonie de l’Université François Rabelais, Tours.
- [64] Ma, E.P., Yiu, E.M. (2001). Voice activity and participation profile: assessing the impact of voice disorders on daily activities. *Journal of speech, language, and hearing research*, 44(3), 511–524.
- [65] Manteau, E. (2004). Rééducation ou conservation du langage oral et de la parole dans les surdités appareillées ou non, y compris en cas d’implantation cochléaire. In *Les approches thérapeutiques en orthophonie* (pp. 29–79). Isbergues: Ortho édition.
- [66] Marié-Bailly, I. (2004). *La voix sensorielle - percevoir pour émettre*. Mémoire de Phoniatrie de l’Université de Paris 13, Bobigny.
- [67] Martens, J.W., Versnel, H., Dejonckere, Ph. (2007). The effect of visible speech in the perceptual rating of pathological voices. *Archives of Otolaryngology—Head & Neck Surgery*, 133(2), 178–185.
- [68] Martin, P. (2008). *Phonétique acoustique*. Paris: A. Colin, 163 p.
- [69] Maulet, M. (1997). Expériences d’utilisation de SpeechViewer auprès d’enfants présentant des troubles de la voix et de la parole. Analyse, applications et limites. En particulier, chez l’enfant atteint de déficience auditive. *Entretiens d’orthophonie*. (pp. 74-87). Paris
- [70] McCoy, S.J. (2004). *Your voice, an inside view : multimedia voice science and pedagogy*. [Princeton, N.J.]: Inside View Press, 181 p.
- [71] Menin-Sicard, A., Sicard, É. (2006). *Vocalab: aide à l’évaluation et à la rééducation de la voix et de la parole : de la théorie à la pratique*. Isbergues (76-78 rue Jean-Jaurès, 62330): Ortho éd. 183p
- [72] Miller, D.G., Schutte, H.K. (1999a). The use of spectrum analysis in the voice studio. In *Voice-tradition and technology : the use of computer feedback in voice training* (pp. 189–210). San Diego, Calif.: Singular Publishing.
- [73] Miller, D.G., Schutte, H.K. (1999b). The use of the electroglottograph in the voice studio. In *Voice-tradition and technology : the use of computer feedback in voice training* (pp. 211–225). San Diego, Calif.: Singular Publishing.

- [74] Miller, D.G., Schutte, H.K. (1990). Feedback from spectrum analysis applied to the singing voice. *Journal of Voice*, 4(4), 329–334.
- [75] Miller, D.G. (2008). *Resonance in singing : voice building through acoustic feedback*. Princeton, NJ: Inside View Press, 130 p.
- [76] Miller, R., Gouëlou, J.-M. (1993). *La Structure du chant pédagogie systématique de l'art du chant*. Paris: Cité de la musique-Centre de ressources musique et danse, 395 p.
- [77] Morsomme, D., Estienne, F. (2006). Bilan vocal. In *Les bilans de langage et de voix : fondements théoriques et pratiques* (pp. 223–256). Paris: Masson.
- [78] Nair, G. (1999). *Voice-tradition and technology : the use of computer feedback in voice training*. San Diego, Calif.: Singular Publishing, 326 p.
- [79] Ormezzano, Y. (2000). *Le Guide de la voix*. Paris: O. Jacob, 432 p.
- [80] Pearson, P., Pickering, L., Da Silva, R. (2011). The impact of computer assisted pronunciation training on the improvement of Vietnamese learner production of English syllable margins. In *Proceedings of the 2nd Pronunciation in Second Language Learning and Teaching Conference*, (pp. 169–180). Presented at the Pronunciation in Second Language Learning and Teaching Conference, Ames, IA: Iowa State University: J. Levis & K. LeVelle.
- [81] Peloux, V. (1993). *Outil informatique et rééducation : intérêt de Speech Viewer*. Mémoire d'orthophonie de l'Université de Montpellier I, Montpellier.
- [82] Pierce, J. R. (1984). *Le son musical musique, acoustique et informatique*. Paris: Pour la science, 242 p.
- [83] Pillot-Loiseau, C., Vaissière, J. (2009). La portée de la voix parlée et chantée : aspects scientifiques et rééducatifs. In *La voix dans tous ses maux*, (pp 243–249). Isbergues: Ortho édition.
- [84] Py, B. (2010). *La statistique sans formule mathématique avec 150 questions et exercices corrigés d'entraînement aux examens*. Paris: Pearson education, 306 p.
- [85] Remacle, M. (2009). Les laryngopathies dysfonctionnelles. In *Précis d'audiophonologie et de déglutition. Tome II, Les voies aéro-digestives supérieures* (pp. 243–255). Marseille: Solal.
- [86] Remacle, M., Morsomme, D. (2009). Evaluation objective de la voix. In *Précis d'audiophonologie et de déglutition. Tome II, Les voies aéro-digestives supérieures* (pp. 121–134). Marseille: Solal.
- [87] Rémond, A., Rémond, A., Flor-Henry, P. (1994). *Biofeedback: principes et applications*. Paris; Milan; Barcelone: Masson, 242 p.
- [88] Revis, J., Barberis, S., Giovanni, A. (2000). Definition of a new temporal voice onset measurement. *Revue de laryngologie - otologie - rhinologie*, 121(5), 291–296.
- [89] Revis, J. (2004). L'analyse perceptive des dysphonies. In *Le bilan d'une dysphonie : état actuel et perspectives* (pp. 67-104). Marseille: Solal.

- [90] Revis, J., Cayreyre, F. (2004). Rééducation des troubles de la voix d'origine organique ou fonctionnelle. In *Les approches thérapeutiques en orthophonie. Tome 3* (pp. 91–103). Isbergues: Ortho édition.
- [91] Revis, J., Galant, C., Minghelli, E. (2012). Analyse perceptive de la dysphonie, historique et perspective. In *La voix parlée et la voix chantée* (pp. 57–68). Montpellier: Sauramps médical.
- [92] Revis, J., Giovanni, A., Triglia, J.-M. (2002). Influence de l'attaque sur l'analyse perceptive des dysphonies. *Folia Phoniatica et Logopaedica*, 54(1), 19–25.
- [93] Ribaute, J. (2011). *Evaluation perceptive des dysphonies: étude des stratégies sous-jacentes à la perception de la qualité vocale par l'analyse de catégorisations libres d'auditeurs avertis, experts et dysphoniques*. Mémoire d'orthophonie de l'Université Paul Sabatier, Toulouse.
- [94] Roch, J.B. (2001). Douze clefs pour la rééducation de la voix. In *Dysphonies et rééducations vocales de l'adulte* (pp. 51–65). Marseille: Solal.
- [95] Rodríguez, W. ., Saz, O., Lleida, E. (2012). A prelingual tool for the education of altered voices. *Speech Communication*, 54(5), 583–600.
- [96] Romand, C. (1999). Utilisation de Speechviewer dans le traitement des aspects segmentaux et suprasegmentaux de la parole chez des enfants sourds profonds. *Glossa*, (68), 28–35.
- [97] Rosen, C.A., Lee, A.S., Osborne, J., Zullo, T., Murry, T. (2004). Development and validation of the voice handicap index-10. *Laryngoscope*, 114(9), 1549–1556.
- [98] Rousseaux, F. (1989). *De l'usage de la relaxation dans les dysphonies dysfonctionnelles: le point du vue du patient* Mémoire d'orthophonie de l'Université Pierre et Marie Curie, Paris
- [99] Sarfati, J. (1998). *Soigner la voix*. Marseille, France: Solal, 127 p.
- [100] Singly, F. de. (2008). *Le questionnaire*. Paris: A. Colin, 128 p.
- [101] Siupsinskiene, N., Razbadauskas, A., Dubosas, L. (2011). Psychological distress in patients with benign voice disorders. *Folia phoniatica et logopaedica*, 63(6), 281–288.
- [102] Steinhauer, K., Grayhack, J.P. (2000). The role of knowledge of results in performance and learning of a voice motor task. *Journal of voice*, 14(2), 137–145.
- [103] Steinhauer, K., Grayhack, J.P., Smiley-Oyen, A.L., Shaiman, S., McNeil, M.R. (2004). The relationship among voice onset, voice quality, and fundamental frequency: a dynamical perspective. *Journal of Voice*, 18(4), 432–442.
- [104] Sundberg, J. (2001). Level and Center Frequency of the Singer's Formant. *Journal of Voice*, 15(2), 176–186.
- [105] Vaissière, J. (2011). *La phonétique*. Paris: Presses universitaires de France, 127 p.
- [106] Van Leer, E., Connor, N.P. (2010). Patient Perceptions of Voice Therapy Adherence. *Journal of Voice*, 24(4), 458–469.

- [107] Verdolini, K., Krebs, D.E. (1999). Some considerations on the science of special challenges in voice training. In *Voice-tradition and technology : the use of computer feedback in voice training* (pp. 227–239). San Diego, Calif.: Singular Publishing.
- [108] Volin, R. A. (1999). Applying gram to common speech pathology and voice therapy techniques. In *Voice-tradition and technology : the use of computer feedback in voice training* (pp. 241–258). San Diego, Calif.: Singular Publishing.
- [109] Welch, G. F. (1985). A Schema Theory of How Children Learn to Sing in Tune. *Psychology of Music*, 13(1), 3–18.
- [110] Welch, G.F., Howard, D.M., Himonides, E., Brereton, J. (2005). Real-time feedback in the singing studio: an innovatory action-research project using new voice technology. *Music Education Research*, 7(2), 225–249.
- [111] Wilson, P.H., Lee, K., Callaghan, J., Thorpe, C.W., Maimets-Volk, K., Parncutt, R., Ross, J. (2007). Learning to sing in tune: Does real-time visual feedback help? In *Proceedings of the Conference on Interdisciplinary Musicology (CIM07)* (p. 9), Tallinn, Estonia: R. Parncutt, A. Kessler & F. Zimmer.
- [112] Wilson, P.H., Thorpe, C.W., Callaghan, J. (2005). Looking at singing: does real-time visual feedback improve the way we learn to sing? In *2nd APSCOM Conference: Asia-Pacific Society for the Cognitive Sciences of Music* (p. 10), Séoul, Corée du Sud.
- [113] Woisard, V. (2004). Le bilan vocal en pratique. In *Le bilan d'une dysphonie : Etat actuel et perspectives* (pp. 197–207). Solal.
- [114] Wolfe, V. I., Steinfatt, T.M. (1987). Prediction of vocal severity within and across voice types. *Journal of speech and hearing research*, 30(2), 230–240.
- [115] Wuyts, F.L., De Bodt, M.S., Van de Heyning, P.H. (1999). Is the reliability of a visual analog scale higher than an ordinal scale? An experiment with the GRBAS scale for the perceptual evaluation of dysphonia. *Journal of voice*, 13(4), 508–517.
- [116] Yanagihara, N. (1967). Significance of Harmonic Changes and Noise Components in Hoarseness. *Journal of Speech and Hearing Research*, 10(3), 531–541.
- [117] Yasui, A. (2004). Objective study of vocal disorders uses spectrograph'. *Bulletin of Universities and Institutes*, 2, 35–46.
- [118] Yumoto, E., Sasaki, Y., Okamura, H. (1984). Harmonics-to-Noise Ratio and Psychophysical Measurement of the Degree of Hoarseness. *Journal of Speech and Hearing Research*, 27(1), 2–6.
- [119] Zonder, A. (2002). *Le biofeedback électromyographique, outil d'évaluation en orthophonie : application aux paralysies faciales, infirmités motrices cérébrales et troubles de la cinématique temporo-mandibulaire*. Mémoire d'orthophonie de l'Université Pierre et Marie Curie, Paris.

10 Annexes

10.1 Classifications des images spectrographiques caractérisant la qualité vocale

<p>Type 0 et 0+ : pas de raucité</p>	<p>Composantes harmoniques normales sur toutes les fréquences, sans composante de bruit.</p>
<p>Type I : les harmoniques sont mélangées avec du bruit, principalement dans la zone formantique des voyelles.⁷⁴</p>	<p>Composantes harmoniques normales mixées avec des composantes de bruit principalement dans la région des formants des voyelles.</p>
<p>Type II : le bruit domine sur les harmoniques au niveau du second formant du [ε] et du [i], et on observe aussi un peu de bruit dans les hautes fréquences au-delà de 3000 Hz pour les voyelles [ε] et [i].⁷⁵</p>	<p>Composantes de bruit prédominant sur les composantes harmoniques pour le second formant, et légères composantes de bruit dans la région des hautes fréquences (environ 3 kHz).</p>
<p>Type III : les seconds formants du [ε] et du [i] sont totalement remplacés par du bruit. La composante de bruit au-delà de 3000Hz augmente en intensité et en étendue fréquentielle.⁷⁶</p>	<p>Le second formant est complètement remplacé par les composantes de bruit, et des composantes de bruit additionnel supérieures à 3 kHz s'intensifient et élargissent leur champ.</p>
<p>Type IV : les seconds formants de [ə] et [ε] et du [i] sont remplacés par du bruit, et même les premiers formants de toutes les voyelles perdent leurs harmoniques au profit du bruit. De plus le bruit est intensifié au niveau des hautes fréquences.⁷⁷</p>	<p>Les premiers et deuxièmes harmoniques sont remplacés par du bruit, et les composantes de bruit sont intensifiées dans les hautes fréquences.</p>

Tableau XXX : classification de Yanagihara et ses adaptations ultérieures (Yanagihara, 1967, pp.533-534, (Remacle & Morsomme, 2009, p.129, Wolfe & Steinfatt, 1987, p.233, Cooper, 1974, p.291) [116] [86] [114] [16]

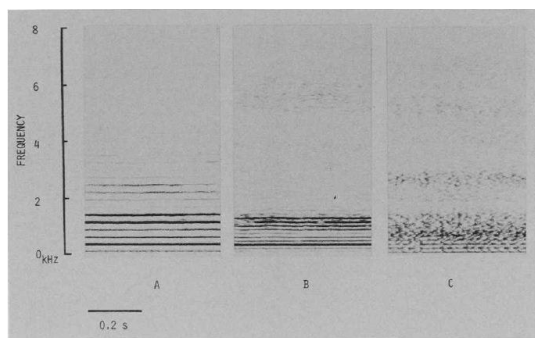


FIGURE 2. Spectrograms both judges rated as type 1 (A), 3 (B), and 4 (C). The H/N ratios of these phonatory samples were 9.1 dB, 0 dB, and -7.6 dB, respectively.

Figure XII : spectrogrammes jugés comme type I, III et IV par les deux juges (Yumoto *et al.*, 1984, p. 4) [118]

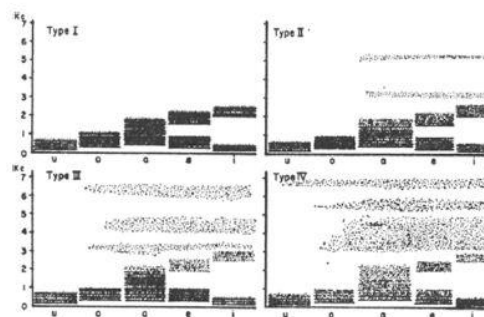


Figure XIII : représentation schématique des 4 degrés de raucité sur le spectrogramme (Yanagihara, 1967, p.535) [116]

⁷⁴ The regular harmonic components are mixed with the noise component chiefly in the formant region of the vowels

⁷⁵ “The noise components in the second formants of [ε] and [i] predominate over the harmonic components, and slight additional noise components appear in the high frequency region above 3000 Hz in the vowels [ε] and [i]” .

⁷⁶ “The second formants of [ε] and [i] are totally replaced by noise components, and the additional noise components above 3000 Hz further intensify their energy and expand their range” .

⁷⁷ “The second formants of [ə], [ε] and [i] are replaced by noise components, and even the first formants of all vowels often lose their periodic components which are supplemented by noise components. In addition, more intensified high frequency additional noise components are seen” .

<p>Analyse de SMLT (Spectre Moyenné à Long Terme) en bande étroite, de 0 à 1000 Hz d'un [a:] tenu. Classification utilisée « <i>en clinique de manière quotidienne</i> ».</p>	<p>Classification à cinq niveaux de la qualité des tracés d'analyse spectrale.</p>
<p>Classe 4 : présence de la fréquence fondamentale et des harmoniques. Absence de bruit. Classe 3 : présence de la fréquence fondamentale et d'harmoniques. Le bruit est encore présent. Classe 2 : présence de la fréquence fondamentale. Présence de bruit. La présence d'harmoniques est possible mais de faibles intensités. Classe 1 : présence de bruit mais absence de signal. Classe 0 : le tracé ne comporte pas d'harmonique identifiable. Le tracé est pratiquement plat signant la faiblesse de l'énergie acoustique et l'absence de sonorisation.</p>	<p>Tracé 4 : fondamental et harmoniques bien présents, absence de bruit, la voix est normale ». Tracé 3 : fondamental plus distinct, harmoniques présents mais persistance du bruit. Tracé 2 : apparition de la fréquence fondamentale, mais toujours pauvreté en harmoniques et présence de bruit. Tracé 1 ; présence de bruit et absence de son. Tracé 0 : plat, sans harmoniques ni sonorisation.</p>

Tableau XXXI : classification de Remacle & Morsomme (2009, p.130) [86] à gauche, et son pendant à droite par Delacroix et al. (2009, p.427) [24], des tracés spectrographiques caractérisant la raucité vocale

10.2 Évolution objective des patients du groupe protocole

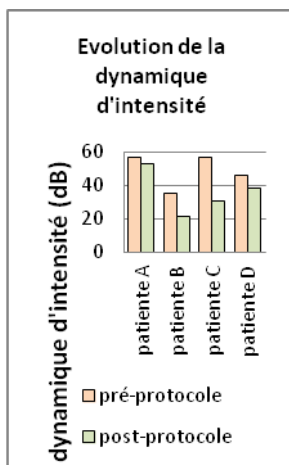


Figure XIV : évolution de la dynamique d'intensité

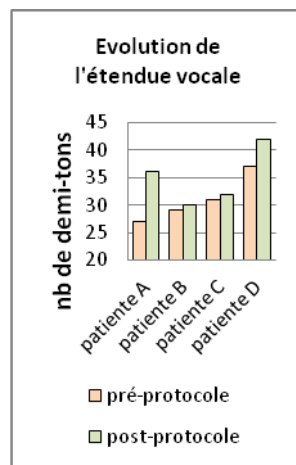


Figure XV : évolution de l'étendue vocale

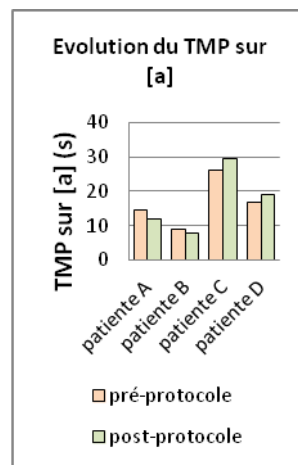


Figure XVI : évolution du TMP sur [a]

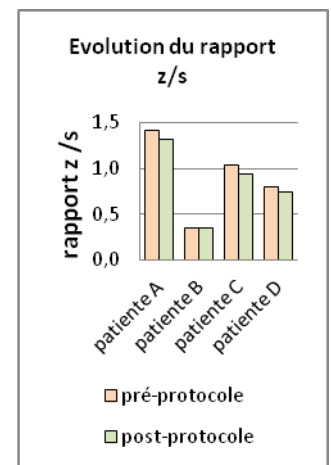


Figure XVII : évolution du rapport z/s

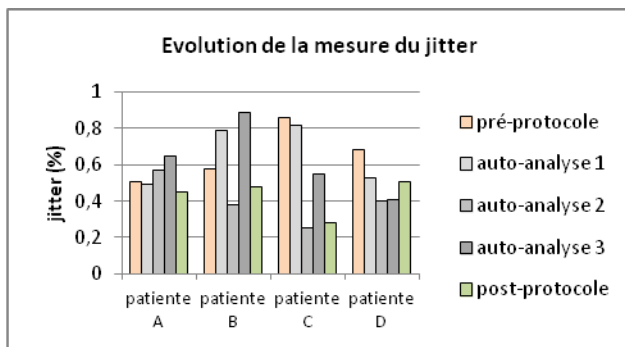


Figure XVIII : Evolution de la mesure du jitter

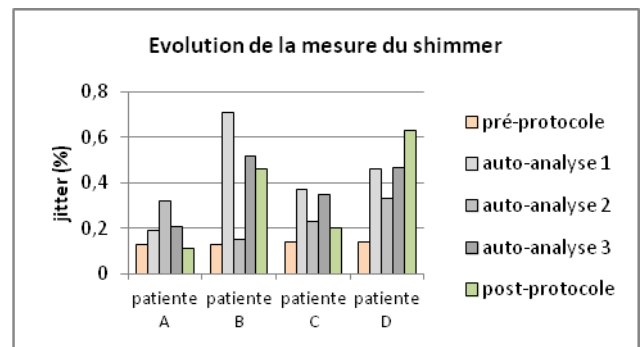


Figure XIX : évolution de la mesure du shimmer

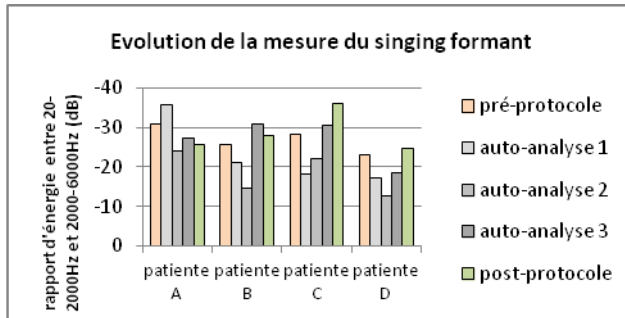


Figure XX : évolution de la mesure du singing formant

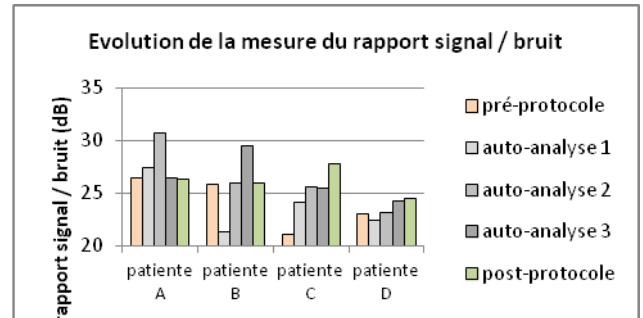


Figure XXI : évolution de la mesure du rapport signal / bruit

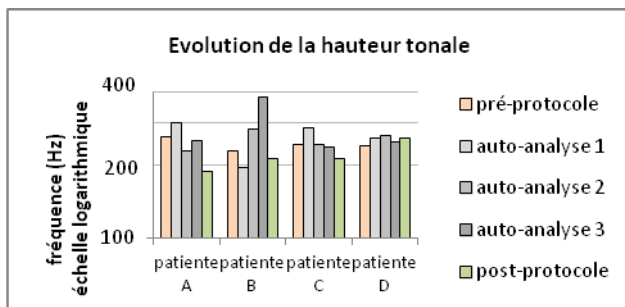


Figure XXII : évolution de la hauteur tonale

10.3 Le spectrogramme en pratique

10.3.1 Introduction du spectrogramme

Elle se fait grâce à la visualisation de l'exercice d'auto-analyse effectué juste avant, au début de la séance n°2.

10.3.1.1 Temps qui passe

Le temps qui passe est facilement visualisé lorsqu'on fait rejouer le son, ce qui fait apparaître progressivement le spectrogramme.

10.3.1.2 Hauteur

On présente les harmoniques : la première est la fréquence fondamentale, qui donne la hauteur du son, les autres sont les fréquences supérieures, multiples de la fréquence fondamentale. La notion de hauteur est particulièrement visible sur le spectrogramme d'une sirène sur un [a]. On peut également avoir une idée de la hauteur par la distance entre les harmoniques. Les hommes ont une fréquence fondamentale en moyenne une octave plus basse que les femmes (fréquence divisée par deux), ce qui se traduit par des harmoniques deux fois plus rapprochées.

Pour certains patients, il est nécessaire de clarifier la notion de « fréquence fondamentale » (ou « fondamentale ») : c'est la note musicale entendue, à distinguer de la notion de « fondamentale »

dans un accord. On fait également la distinction entre premier harmonique (note fondamentale) et premier formant vocalique (voir les exercices sur formants).

10.3.1.3 Intensité

En comparant le son faible et le son fort de l'exercice d'auto-analyse, on explique que l'intensité globale (visible sur la forme d'onde au-dessus du spectrogramme) est la somme de toutes les harmoniques. On voit plus d'harmoniques sur un son fort que sur un son faible, même si la répartition des harmoniques peut être la même (en particulier, il peut y avoir des harmoniques aiguës, dans la zone du formant du chanteur, mais que l'on ne voit pas, parce que d'intensité trop faible).

10.3.1.4 Bruit, bruit entre les harmoniques

On introduit la notion de bruit avec des consonnes fricatives non voisées comme le [ʃ] et le [f]. Il n'y a pas d'harmoniques car ces sons n'ont pas de hauteur tonale. L'exemple de consonnes explosives non voisées [p] [t] [k] ou le bruit d'un claquement de mains montre que ce bruit n'est pas forcément continu.

On fait alors reconnaître ce genre de « nuages » dans un son voisé, entre les harmoniques. Certains patients peuvent penser que « *le bruit c'est le noir* », il est alors important de préciser que le bruit, comme les harmoniques, est en couleur (en général bleu pour le bruit, et de bleu à rouge en passant par le jaune pour les harmoniques).

La présence de bruit dans un son voisé signe un timbre altéré, un manque de clarté, par la présence d'un souffle ou d'une raucité. Les patients n'entendent pas toujours à quoi cela correspond dans le son.

Le bruit correspond aussi parfois à un artefact non vocal dû à la prise de son (micro, bruit ambiant..), surtout lorsqu'il est visible dans les basses fréquences.

10.3.2 Travail avec les consonnes fricatives et la paille

La coordination pneumophonique peut être travaillée par la « sonorisation continue de consonnes fricatives », ou l'expiration fluide sonorisée dans une paille (Amy de la Bretèque, 2004; Revis & Cayreyre, 2004, p. 101; Sarfati, 1998, pp. 63–64) [2] [90] [99]. Lors de ces exercices, il est important de ne pas aller au bout du souffle. Ces exercices permettent d'équilibrer les pressions sus- et sous-glottiques.

10.3.2.1 Paille

On commence par visualiser sur le spectrogramme le bruit du souffle dans la paille. Il faut être extrêmement attentif à la posture du patient tout en gardant l'extrémité de la paille proche du micro pour que celui-ci capte le son du souffle. Lorsqu'on ajoute le son dans la paille, celui-ci doit être intérieur, très doux. On ne doit pas chercher à le faire résonner à l'extérieur. L'important est de toujours visualiser le souffle de la paille. La visualisation du son dans la paille est parfois noyée dans l'affichage du bruit du souffle. Le son devient visible lorsqu'il est émis plus aigu (recto tono ou avec une sirène).

10.3.2.2 Consonnes fricatives

Respiration et posture

On peut commencer par un exercice de respiration, surtout chez certains patients qui n'ont pas automatisé la respiration costo-diaphragmatique et l'entrée d'air silencieuse par la bouche. On demande alors uniquement des [ʃ ʃ ʃ] en portant attention à la posture et à la respiration.

Différenciation des consonnes

On différencie ensuite les consonnes fricatives sourdes [f] [s] [ʃ] entre elles. Cela nécessite de changer l'échelle du spectrogramme pour visualiser au moins jusqu'à 8kHz (pour la consonne [s]). Le [f] est plus diffus, parfois il apparaît peu sur le spectrogramme. Le [s] est dans les aigus. Le modèle donné par le thérapeute n'a pas valeur stricte : il est important que le patient ne cherche pas à reproduire exactement le modèle du [s] ou du [ʃ], mais qu'il s'intéresse plutôt aux différences entre ces deux consonnes, dans son émission personnelle.

Ces exercices peuvent être une entrée en matière dédramatisante pour les personnes dont la voix est très altérée car ils ne nécessitent pas de voisement et intéressante par la visualisation de l'émission de leurs harmoniques aiguës.

Alternance voisé / non voisé

On visualise chaque paire de consonnes ayant le même lieu d'articulation, en passant en alternance de l'une à l'autre : [ʃ ʒ] [s z] [f v]. Les sons voisés font apparaître des harmoniques (traits horizontaux) correspondant à une hauteur tonale : on peut « chanter » sur les fricatives voisées.

On est toujours attentif à ce que le patient n'aille pas au bout de son souffle, en donnant des modèles courts. Le bruit du [v] est peu visible lorsqu'on l'émet naturellement. De même, on met naturellement moins d'air dans le [ʒ] que dans le [ʃ]. L'exercice consiste à ajouter du voisement sur les consonnes non voisées, en gardant la même sensation proprioceptive au point articulaire. Le spectrogramme permet de caractériser la qualité de transition entre un son non voisé et un son voisé, par l'apparition du voisement, sous la forme d'harmoniques et la continuité du souffle. L'objectif est de mettre en vibration les plis vocaux dans le souffle installé préalablement, comme pour l'exercice de la paille, afin d'avoir un début de son non glottique.

Vers les voyelles

On demande au patient de faire suivre les constrictives par une voyelle : [f v u] [s z i] [ʃ ʒ u]. On observe en temps réel et de façon différée les trois temps de l'exercice : bruit seul sur la consonne non voisée, bruit et harmoniques sur la consonne voisée, harmoniques seules sur la voyelle. L'objectif est d'obtenir des transitions douces entre chaque phonème, visibles sur le spectrogramme. Il est possible de faire l'exercice sur des tonalités progressivement plus aiguës, demi-ton par demi-ton, afin de travailler toute la tessiture, et notamment la zone de passage qui peut poser problème en voix chantée. On note que certains patients peuvent se crisper sur l'image à obtenir. Il convient dans ce cas de changer d'exercice pour quelque chose de plus libre, afin de ne pas induire de forçage ou de mauvais geste vocal.

10.3.3 Travail des attaques et du mécanisme 0 (fry)

Le travail de l'attaque est préconisé par différents auteurs (Miller & Gouëlou, 1993; Revis et al., 2002, p. 24 ; Ormezzano, 2000 ; Cornut, 2009 ; Fournier, 1989, p. 130; Amy de la Bretèque, 2012, p. 132; Revis & Cayreyre, 2004, pp. 100–101) [76] [92] [79] [17] [35] [3] [90].

10.3.3.1 Trois types d'attaque

Après avoir expliqué la physiologie des trois types d'attaques, on visualise leur traduction sur un spectrogramme. Il est important d'attirer l'attention sur le début du son : le bruit que l'on peut voir ensuite entre les harmoniques ne fait pas partie de l'attaque en elle-même. Le coup de glotte forme un trait vertical plus ou moins net au début du son. Il n'est pas toujours évident pour le thérapeute de produire des exemples parfaits, c'est pourquoi on peut s'aider de l'exemple enregistré en ligne.

10.3.3.2 Mécanisme 0 (fry)

La visualisation du spectrogramme permet de faire la distinction entre une émission saine du mécanisme 0 et un serrage : dans le serrage, on entend une hauteur qui se matérialise par des harmoniques, tandis que dans le mécanisme 0, on ne distingue pas d'harmoniques (traits horizontaux) (Figure XXXVI). On peut s'exercer à alterner le fry sur un [a] avec un [a] voisé à différentes hauteurs.

10.3.3.3 Pouff d'air

L'objectif est de sentir des « bulles d'air » au niveau du larynx. Pour cela, on peut partir du fry et le ralentir. Le patient cherche à obtenir un seul accolement des cordes vocales, et pour cela il doit reconnaître proprioceptivement la juste pression nécessaire. On observe sur le spectrogramme de petits traits verticaux. Parfois le son n'est pas assez fort pour être visualisé, mais le patient peut sentir qu'il fait une seule « bulle ». Il peut être avantageux dans ce cas de placer un micro très proche de la source sonore. Le son peut également être avantageusement amplifié lorsque le patient modifie son espace de résonance.

10.3.4 Exploration des formants vocaliques

Prendre conscience des formants vocaliques permet au patient de préciser ses voyelles et de gagner en efficacité vocale. L'exploration des harmoniques aiguës lors des transitions entre voyelles ouvre en outre des possibilités nouvelles d'écoute et de prise de conscience du timbre extra-vocalique.

10.3.4.1 Sur des consonnes

Lorsqu'on émet des consonnes non voisées (comme [f] [s] [ʃ] ou [t] [p] [k]) en mettant son conduit vocal dans la forme voulue pour une certaine voyelle, les formants de ces voyelles sont accentués dans le bruit des consonnes, et cela est visible sur le spectrogramme (Figure XXXV).

10.3.4.2 A partir du fry

La source glottique, pour le fry, fournit un large spectre d'harmoniques qui permet d'observer les premiers formants des voyelles. Cependant on voit parfois mal le deuxième formant lorsqu'il est aigu (sur le [i] par exemple, Figure XL) On explore le triangle vocalique, par exemple avec la séquence [i e ε a] ou [u o ɔ a]. Il peut être utile d'écrire les voyelles sur un papier « aide-mémoire » pour le patient.

10.3.4.3 Sur des voyelles (« château des voyelles »)

Le même travail est effectué sur une voyelle chantée. Il est nécessaire de prendre la note assez bas, surtout pour les femmes, de façon à émettre des harmoniques assez proches l'une de l'autre pour que les formants soient visibles. Une note aigüe, même riche en harmoniques, ne comprendra pas suffisamment d'harmoniques pour que l'on observe leur renforcement sur certaines zones de fréquences, notamment dans les fréquences basses. Il s'agit de trouver un compromis entre la possibilité pour le patient d'émettre correctement la note (il faut qu'elle soit dans sa tessiture) et qu'elle soit suffisamment grave.

Les patients ne comprennent pas facilement où se trouvent les formants et ce qu'ils deviennent lorsqu'on change de note. Il est important de leur montrer précisément, pour les deux premiers formants, chaque transition de note, et de leur montrer que les changements sont parfois minimes, et qu'il ne faut pas s'attendre à des changements plus marqués.

Les patients ont tendance à vouloir afficher le plus d'harmoniques possible sur le spectrogramme. Or, ce travail est l'occasion de la prise de conscience de la différence entre des voyelles postérieures [o] et [u], dont les deux premiers formants sont graves (en-dessous de 1000 Hz) et les voyelles antérieures comme le [e] et le [i], dont les deux premiers formants sont très éloignés l'un de l'autre, et où on observera un « trou » fréquentiel (entre 500 Hz et 3000 Hz pour le [i]).

10.3.4.4 Transitions entre voyelles

L'objectif de cet exercice est de passer très progressivement d'une voyelle à l'autre, par exemple en ouvrant la bouche de [i] vers [a] sur [i e ε a], en laissant s'ouvrir lentement la mâchoire, sur un son recto-tono. Lorsque l'on entend des formants inhabituels, ne correspondant pas à une voyelle connue du français, c'est l'occasion d'être attentif au son, non plus comme une note (la fréquence fondamentale) et une voyelle (les formants), mais comme une matière sonore résonante, avec toutes ses harmoniques, qui deviennent précisément audibles, lorsqu'on passe de l'une à l'autre très lentement et très progressivement (par exemple du [i] au [a] ou du [i] ou [u]). Certaines personnes peuvent être plus ou moins sensibles à certaines plages de fréquences. On peut également au début se concentrer sur ces sensations en fermant les yeux, puis regarder le spectrogramme de manière différée.

10.3.5 Exploration de la voix chantée

Le singing formant est facilement observable sur le spectrogramme. Certaines techniques de chant (le « twang » notamment) permettent de renforcer ces harmoniques aiguës. Avec ces techniques de chant, on peut observer des harmoniques jusqu'à 10 kHz

On observe également le vibrato : on le voit peu sur la première harmonique, mais surtout sur les harmoniques plus aiguës.

Lorsqu'on travaille un extrait chanté (Figure XLI), l'image spectrographique est complexe, mais elle peut être décodée plus facilement en regardant l'image qui se construit, en chantant ou en réécoutant l'extrait. Le travail avec le spectrogramme peut permettre de modifier des voyelles (par exemple sur un fa# un [oe] peut être remplacé par un [a] pour améliorer la justesse, sans que cette modification n'entrave la compréhension du texte). Modifier les voyelles permet également une correspondance des formants avec les harmoniques qui font mieux résonner la voix. Ce travail peut être effectué à partir

du fry afin de bien visualiser à quelle hauteur se situe le formant d'une voyelle afin de trouver la couleur de la voyelle qui correspond à la hauteur des harmoniques de la note tonale que l'on doit chanter (Miller, 2008, p.96) [75].

On travaille également le legato et la prise de conscience des différentes consonnes non voisées sur lesquelles il n'y a pas de hauteur, et les consonnes explosives, qui coupent le son : on voit explicitement un trou dans l'image spectrographique. On peut observer également les effets de micro-prosodie sur [d] qui peut jouer sur la justesse du son. On travaille la différenciation [d] / [t] et le voisement du [d] à la hauteur de la voyelle qui va suivre.

10.4 Documents du protocole

Ortho : _____ Patient : _____ Date : _____

auto-analyse des réalisations

marquer d'un signe (x ou o) chaque caractéristique des sons réalisés, sur chacune des lignes suivantes

x : premier son (le moins fort) **o** : deuxième son (le plus fort)

Couleur utilisée pour la première marque, juste après la réalisation :

Couleur utilisée pour la deuxième marque, après feedback :

Impression générale (qualification libre) :

• Intensité	Faible _____	Forte _____
• Hauteur	Grave _____	Aigu _____
• Attaque	Soufflée _____	Dure _____
	Equilibrée _____	

Qualification libre de l'attaque :

• Clarté du timbre	Timbre altéré _____	Timbre clair _____
• Richesse du timbre	Peu d'harmoniques _____	Beaucoup d'harmoniques _____

Qualification libre du timbre :

• Comportement vocal	Inadapté _____	Adapté _____
----------------------	----------------	--------------

Qualification libre du comportement vocal :

Figure XXIII : feuille de passation de l'auto-analyse de 2 [a]

Formulaire de Consentement libre et éclairé

Enregistrement d'échantillons vocaux pour un mémoire d'orthophonie portant sur :

« A l'écoute de sa voix en rééducation vocale : apports du spectrogramme ; étude de cas »

Sous la direction de Isabelle Maréchal-Bailly

Mémoire de fin d'études du Département Universitaire d'Enseignement et de Formation en Orthophonie
 Université Paris VI - Faculté de Médecine Pierre et Marie Curie
 91, boulevard de l'Hôpital - 75013 Paris

Je certifie avoir donné mon accord pour participer à une étude d'orthophonie sur la voix. J'accepte volontairement de participer à cette étude non-remunérée. Je comprends que ma participation n'est pas obligatoire et que je peux la stopper à tout moment sans avoir à me justifier ni à encourir aucune responsabilité (cf. article L.1122-1). Mon consentement ne décharge pas les organisateurs de la recherche de leurs responsabilités et je conserve tous mes droits garantis par la loi.

Au cours de cette étude, j'accepte que soient recueillis certaines informations personnelles et un enregistrement audio d'échantillon vocal. Je comprends que ces données recueillies seront strictement confidentielles et à usage exclusif des investigateurs concernés.

J'ai été informé que mon identité n'apparaîtra dans aucun rapport ou publication et que toute information me concernant sera traitée de façon confidentielle. Si je le souhaite, je serai informé(e) sur les résultats globaux à l'issue de la recherche (cf. article L.1122-1).

Date :/...../.....

Nom du volontaire : Mail :

Signature du volontaire (précédée de la mention « lu et approuvé ») :

Nom de l'expérimentateur : Florence PARMENTIER

Signature de l'expérimentateur :

Figure XXIV : formulaire de consentement libre et éclairé

Périmètre de l'exercice	
<p>« nous allons faire maintenant dans le temps qui nous reste un exercice spécifique qui fait partie du protocole de l'étude : est-ce que nous pouvons enregistrer sur un dictaphone à partir de maintenant ? »</p> <p>Consigne : « vous allez faire 2 /a/, à la même hauteur, un faible (piano) et l'autre fort. Vous allez ensuite marquer sur ces échelles la où vous pensez que les sons que vous avez faits se situent, en fonction des divers critères nommés sur la feuille : Intensité, Hauteur, Qualité de l'attaque, Clarté et Richesse du timbre, Geste vocal. Il n'y a pas de bonnes réponses, c'est un exercice pour prêter attention à chacun de ses critères dans votre voix. Dans un deuxième temps, nous réécouterons l'enregistrement en regardant le spectrogramme et vous pourrez peut-être modifier ce que vous avez noté en premier. Est-ce que vous avez bien compris ? »</p> <p>« les 2/a/ doivent durer chacun 2 à 3s sans attendre trop entre les deux »</p>	<p>Enregistrer avec un système indépendant de l'ordinateur qui prend les /a/ à auto-évaluer (dictaphone)</p>
<p>Familiarisation avec les termes (la première fois notamment) : en montant sur les échelles où se situeraient chacun des sons exemple entendus</p>	<p>Intensité : Fort / pas fort</p>
<p>Hauteur : Grave / aigu (en cas d'incompréhension donner à entendre des notes graves et aiguës au piano)</p>	<p>Qualité de l'attaque : Faire écouter et voir si besoin les 3 types d'attaque de la banque de son</p>
<p>Clarté du timbre : Faire écouter et voir 3 types de timbre de la banque de son</p>	<p>Richesse du timbre : Faire écouter et voir 3 types de timbre de la banque de son</p>
<p>Geste vocal : Posture / respiration / serrage / forçage / hypotonie...</p>	<p>Donner la feuille d'auto-analyse au patient.</p>
<p>Le patient effectue l'exercice</p>	<p>enregistrement + spectrogramme (avec VOCEVISTA) avec un micro de bonne qualité et en contrôlant la distance à la bouche</p>
<p>« qu'est-ce que vous pensez de ce que vous venez de faire ? »</p>	<p>Encourager les commentaires</p>
<p>« pouvez-vous maintenant évaluer votre réalisation : une première marque (x) pour le son faible, et une deuxième (o) pour le son fort, pour chacun des critères, selon ce dont vous vous souvenez »</p>	<p>Cette phase risque de poser problème au patient, on ne doit pourtant pas redonner les exemples sonores.</p>
<p>« nous allons maintenant réécouter en regardant le spectrogramme afin de voir si cela vous aide pour cet exercice »</p>	<p>Donner un crayon d'une autre couleur</p>
<p>« qu'est-ce que vous en pensez ? »</p>	<p>On écoute une première fois</p>
<p>« nous allons réécouter critère par critère »</p>	<p>Encourager les commentaires</p>
<p>« par rapport à ce que vous entendez dans les exemples, où vous situez-vous ? »</p>	<p>Les mêmes aides que lors de la familiarisation sont données, avec autant d'écoute avec visualisation que nécessaire.</p>
<p>« si vous comparez l'image de votre son avec les images, où vous situez-vous ? »</p>	<p>orienter l'écoute / orienter la lecture du spectrogramme (en fonction des fiches d'interprétation du spectrogramme, que le patient peut consulter) en fonction de chaque critère.</p>
<p>« est-ce que vous avez l'impression que la visualisation apporte quelque chose ? quoi ? Est-ce plutôt d'avoir réentendu votre voix qui modifie votre perception ou d'avoir vu le spectrogramme ? »</p>	<p>Archiver les enregistrements : les 2/a/ tenus, et l'enregistrement indépendant (dictaphone) de l'exercice en entier.</p>
<p>« merci ! »</p>	

Tableau XXXII : déroulement de l'exercice d'auto-analyse

Figure XXVI : Evaluation Perceptive : GRBAS(I)

Ortho : _____ Patient : _____ Date : _____

- comment qualifieriez-vous votre voix parlée actuelle ?
- que reprochez-vous à votre voix ?.....
- qu'aimez-vous dans votre voix ?
- Quel est votre pourcentage de satisfaction actuel ?
- Qualifiez (par un ou plusieurs adjectifs) chacune des caractéristiques de votre voix :
 - hauteur :
 - intensité :
 - timbre (qualité vocale) :
 - attaque (début de son) :
 - geste vocal (votre manière physique d'émettre des sons) :
- Lorsque vous entendez un enregistrement de votre voix, qu'en pensez-vous ?
- Quels sont les objectifs que vous estimez avoir travaillé ?
- quels sont ceux que vous estimez n'avoir pas pu réaliser ?
- Quels sont Ceux qui n'ont pas été travaillé :
- Avez-vous encore besoin d'aide pour votre voix ?
Pourquoi ?
- qu'est-ce qui vous a le plus aidé dans votre travail sur votre voix ?
- vos échanges avec l'expérimentateur, précisez :
- les techniques ou les outils utilisés, précisez :

Figure XXVII : entretien semi-dirigé pré-protocole

Date : _____ Auditeur : _____
Evaluation Perceptive : comparaison extraits A et extraits B.
 (1) : Extrait A meilleur (4) : Extrait B légèrement meilleur
 (2) : Extrait A légèrement meilleur (5) : Extrait B meilleur
 (3) : A et B semblables

Voix parlée

	(1)	(2)	(3)	(4)	(5)	
1A						1B
2A						2B
3A						3B
4A						4B

Remarques facultatives

Voix d'appel

	(1)	(2)	(3)	(4)	(5)	
1A						1B
2A						2B
3A						3B
4A						4B

Remarques facultatives

3 attaques

	(1)	(2)	(3)	(4)	(5)	
1A						1B
2A						2B
3A						3B
4A						4B

Remarques facultatives

Figure XXV : évaluation perceptive : comparaison extraits A et extraits B

Date : _____ Auditeur : _____
Evaluation Perceptive : GRBAS(I)
 Attention si vous écoutez les extraits des bilans vocaux, vous ne devez plus modifier les comparaisons faites précédemment.
 Pour chaque sujet 4 extraits vocaux sont disponibles: voix parlée, voix d'appel, sirène et 3 attaques.

Sujet 1

	0	0,5	1	1,5	2	2,5	3	
G								Remarques facultatives
R								
B								
A								
S								

Sujet 2

	0	0,5	1	1,5	2	2,5	3	
G								Remarques facultatives
R								
B								
A								
S								

Sujet 3

	0	0,5	1	1,5	2	2,5	3	
G								Remarques facultatives
R								
B								
A								
S								

Sujet 4

	0	0,5	1	1,5	2	2,5	3	
G								Remarques facultatives
R								
B								
A								
S								

1. comment qualifieriez-vous votre voix parlée actuelle ?
2. que reprochez-vous à votre voix ?.....
3. qu'aimez-vous dans votre voix ?
4. Quel est votre pourcentage de satisfaction actuel ?
5. Qualifiez (par un ou plusieurs adjectifs) chacune des caractéristiques de votre voix :
 - hauteur :
 - intensité :
 - timbre (qualité vocale) :.....
 - attaque (début de son) :
 - geste vocal (votre manière physique d'émettre des sons) :
6. Lorsque vous entendez un enregistrement de votre voix, qu'en pensez-vous ?
7. Quels sont vos objectifs en ce qui concerne votre voix ?.....
8. Quelle aide attendez-vous du thérapeute de votre voix ?
9. Prenez-vous soin de votre voix ?
- Si oui comment ?
10. Avez-vous reçu des informations sur la manière dont on émet la voix parlée ?
- Si oui lesquelles ?
11. Chantez-vous ?
- Si oui, dans quelles circonstances ?
- Avez-vous pris des cours de chant ?.....

Figure XXVIII: entretien semi-dirigé post-protocole (première partie)

12. Avez-vous apprécié l'emploi du spectrogramme ?
Précisez :
13. Pensez-vous que cet outil a été utile dans votre travail sur votre voix ?
Précisez :
14. Le spectrogramme a-t-il permis d'observer des caractéristiques de votre voix ?
Précisez :
15. Le spectrogramme a-t-il été un obstacle pour le travail sur votre voix ?
Précisez :
16. avez-vous eu envie de l'utiliser également en autonomie ?
- l'avez-vous fait ?
17. pensez-vous que l'utilisation du spectrogramme a modifié la relation entre l'expérimentateur et vous ?
.....
18. lors de la première présentation du spectrogramme, avez-vous été
- intéressé / amusé / indifférent / gêné ?
.....
- autre :
19. A votre avis, pour quel usage le spectrogramme est-il le mieux adapté ?
.....

Figure XXIX : entretien semi-dirigé post-protocole (deuxième partie : évaluation du spectrogramme)

Questionnaire permettant de caractériser votre voix

Vous devez répondre à chacune des propositions en ne cochant qu'une seule case à chaque fois.

Notations utilisées : J=jamais ; PJ=presque jamais ; P=parfois ; PT=presque toujours ; T=toujours

	J	PJ	P	PT	T
On m'entend difficilement à cause de ma voix					
On me comprend difficilement dans un milieu bruyant					
Mes difficultés de voix limitent ma vie personnelle et sociale					
Je me sens écarté(e) des conversations à cause de ma voix					
Mes problèmes de voix entraînent des pertes de revenus					
J'ai l'impression que je dois forcer pour produire la voix					
La clarté de ma voix est imprévisible					
Mon problème de voix me tracasse					
Je me sens handicapé(e) à cause de ma voix					
On me demande : « Qu'est-ce qu'il va pas avec ta voix ? »					

Comment est votre voix actuellement :

Comme avant	<input type="checkbox"/>
Moins bonne qu'avant	<input type="checkbox"/>
Meilleure qu'avant	<input type="checkbox"/>

Figure XXX : VHI-10

	J	PJ	P	PT	T
J'ai peur que ma voix me trahisse					
Je trouve du plaisir à utiliser ma voix					
Ma voix correspond à ce que je suis aujourd'hui					
J'accorde de l'importance à la voix des autres					
La voix est un sujet qui m'intéresse					
J'accorde de l'importance à ma propre voix					
J'aime ma voix					

Figure XXXI : questionnaire relatif à sa relation à sa voix

Ortho :

Patient :

Date :

Voici une liste d'adjectifs sur la voix, entourez ceux qui semblent s'appliquer à votre voix telle qu'elle est, et soulignez ceux qui pourraient qualifier la voix que vous aimeriez avoir.

acide, accusatrice, affreuse, agréable, agitée, aguicheuse, aigre, aiguë, amère, amortie, ample, angoissée, animée, apaisante, apaisée, âpre, apprêtée, artificielle, assurée, atone, attirante, autoritaire, bâillée, basse, de basse chantante, de basse profonde, belle, bien posée, blanche, bourdonnante, brillante, brutale, bruyante, calme, de camelot, caressante, cassante, caverneuse, chaleureuse, chantante, chantée, charmeuse, de châtré, chaleureuse, chatoyante, chaude, chuchotée, claire, criée, de colorature, de commandement, contenue, de couloir d'hôpital, couverte, creuse, crispante, cristalline, décalée, décidée, désagréable, discordante, douce, douée, dure, éclatante, effrayante, émouvante, emphatique, endurente, énergique, énervante, enjôleuse, ennuyeuse, enrouée, envoûtante, éraillée, étouffée, étranglée, fabriquée, faible, fausse, ferme, flagorneuse, forcée, forte, gaie, geignarde, gouailleuse, grasse, grave, grossière, grumeleuse, gutturale, haute, humble, hypocrite, implorante, incisive, infantile, insipide, irrégulière, irritante, jeune, lasse, lumineuse, mal posée, mélancolique, mélodieuse, métallique, de micro, mielleuse, moelleuse, molle, monotone, mordante, morte, moyenne, musicale, naturelle, nasale, nasillarde, nuancée, d'or, pâle, parlée, passive, pauvre, pénétrante, perçante, pincée, plaintive, plate, pleine, pleureuse, pleurnicharde, poignante, pointue, prenante, de pucelle, puissante, râpeuse, rauque, régulière, résonnante, riche, rocailleuse, ronde, rude, serrée, sexy, saccadée, sarcastique, sévère, sifflée, sirupeuse, sombre, sonore, de soprano, soufflée, spasmée, spontanée, stridente, sucrée, suffisante, suppliante, tendue, de ténor, terne, terrorisante, timbrée, timide, traînante, tranchante, triste, de vamp, veloutée, vieille, voilée, vulgaire.

Figure XXXII : liste d'adjectifs sur la voix

10.5 Questionnaire en ligne sur le spectrogramme

Bonjour,
Dans le cadre de mon mémoire d'orthophonie sur l'utilisation du spectrogramme, je vous invite à répondre à quelques questions ci-dessous.
Vos réponses sont strictement confidentielles et seront utilisées uniquement dans le cadre de cette étude.

*Obligatoire

Votre email

cette donnée est facultative, mais elle permet notamment de ne pas vous renvoyer de mails supplémentaires de rappel !

Comment avez-vous utilisé le spectrogramme (jusqu'à maintenant) ? *

- Je ne l'ai jamais utilisé
- Je l'ai entre-aperçu rapidement
- Je l'ai utilisé une ou deux fois avec ma voix
- Je l'ai utilisé plusieurs fois avec ma voix
- Je l'ai utilisé régulièrement, y compris en autonomie

Comment est votre voix habituellement ? *

- Ma voix est enrrouée et/ou difficile à produire
- Ma voix reste fragile
- Je n'ai pas de problème particulier avec ma voix

Etes-vous chanteur / chanteuse ? *

- Je ne chante jamais
- Je chante occasionnellement
- Je chante régulièrement et/ou je fais partie d'un groupe de chant
- Je chante souvent et/ou je prends des cours de chant
- Je suis chanteur / chanteuse professionnel/le

Commentaires libres

L'image spectrographique et les autres sensations *

Les sensations internes sont les sensations d'ouverture ou de serrage, d'appui, de vibration... Les sensations auditives sont des impressions de qualité du son, de hauteur, d'intensité, de couleur du timbre, etc. Pour chacune des affirmations suivantes, dites si vous êtes tout à fait d'accord, plutôt d'accord, plutôt pas d'accord, pas du tout d'accord.

- L'image proposée par le spectrogramme est un reflet objectif de ma voix
- L'image proposée par le spectrogramme est parfois déroutante
- J'interprète facilement l'image proposée par le spectrogramme
- Le spectrogramme m'a permis d'affiner mes sensations auditives par rapport à ma voix
- Le spectrogramme m'a permis d'affiner mes sensations internes par rapport à ma voix
- Utiliser le spectrogramme me coupe de mes sensations auditives
- Utiliser le spectrogramme me coupe de mes sensations internes

Commentaires libres

Utilisation du spectrogramme *

Pour chacune des affirmations suivantes, dites si vous êtes tout à fait d'accord, plutôt d'accord, plutôt pas d'accord, pas du tout d'accord

- Le spectrogramme m'a permis d'observer des caractéristiques de ma voix
- Le spectrogramme m'a permis d'observer ma qualité d'attaque vocale
- Le spectrogramme m'a permis d'améliorer ma qualité d'attaque vocale
- Le spectrogramme m'a permis d'observer ma qualité de timbre vocal (clarté, richesse)
- Le spectrogramme m'a permis d'améliorer ma qualité d timbre vocal (clarté, richesse)

- Le spectrogramme m'a permis d'observer les différents sons émis par la parole (voyelles, consonnes)
- Le spectrogramme m'a permis de jouer avec les différents sons émis par la parole (voyelles, consonnes)

Technique vocale *

Pour chacune des affirmations suivantes, dites si vous êtes tout à fait d'accord, plutôt d'accord, plutôt pas d'accord, pas du tout d'accord

- L'utilisation du spectrogramme m'empêche d'avoir une bonne posture
- L'utilisation du spectrogramme m'empêche d'avoir une bonne respiration
- L'utilisation du spectrogramme m'empêche d'avoir un bon geste vocal
- Pour moi, c'est gênant de produire la voix pour un écran d'ordinateur
- Souvent, je préfère regarder l'image seulement après avoir produit la voix
- Souvent, je préfère regarder l'image en même temps que je produis la voix

Commentaires libres

En particulier, qu'est-ce qui vous a semblé le plus utile ? *

Vous pouvez préciser quel(s) exercice(s) en particulier vous a plu

En particulier, qu'est-ce qui vous a semblé le moins approprié ? *

Vous pouvez préciser quel(s) exercice(s) en particulier vous a déplu.

Intérêt pour le spectrogramme *

- Le spectrogramme me permet de m'amuser
- Le spectrogramme me permet d'analyser ma voix
- Le spectrogramme fournit la preuve objective de ce que j'avais remarqué sur ma voix
- J'aimerais utiliser le spectrogramme avec un orthophoniste
- J'aimerais utiliser le spectrogramme avec un professeur de chant
- Je n'ai pas envie d'utiliser le spectrogramme en ce moment

Commentaires libres

Utilisation du spectrogramme en autonomie *

Pour chacune des affirmations suivantes, dites si vous êtes tout à fait d'accord, plutôt d'accord, plutôt pas d'accord, pas du tout d'accord

- Je comprends assez bien le spectrogramme pour pouvoir l'utiliser seul(e)
- J'ai téléchargé un logiciel de spectrogramme en temps réel
- Je n'ai pas utilisé le logiciel pour des raisons techniques

Commentaires libres

Utilisation du site <http://florence.parmetier.free.fr/acoustique/spectrogramme/> *

Réponses possibles : « je ne l'ai pas du tout utilisé », « je l'ai regardé rapidement », « je l'ai utilisé », « c'était très intéressant pour moi »

- Partie tutoriel d'utilisation de vocevista
- Partie « lecture de spectrogramme » attaque (début de son)
- Partie « lecture de spectrogramme » : intensité, hauteur (sirènes)
- Partie « lecture de spectrogramme » : timbre (richesse harmonique)
- Partie « idées d'exercices » : attaque (début de son)
- Partie « idées d'exercices » timbre (richesse harmonique)
- Partie « banque de son »

Commentaires sur l'utilisation du site

Qu'avez-vous trouvé le plus intéressant, le moins intéressant ?

Quel autre contenu auriez-vous souhaité trouver sur ce site ?

Commentaires libres sur votre participation à cette étude, etc.

10.6 Exemples d'images spectrographiques

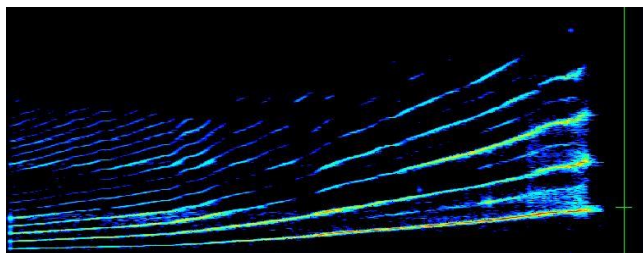


Figure XXXIII : sirène sur [a] (patiente D) ; échelle 8kHz



Figure XXXIV : f non efficace puis f efficace (patiente C) ; échelle 8kHz

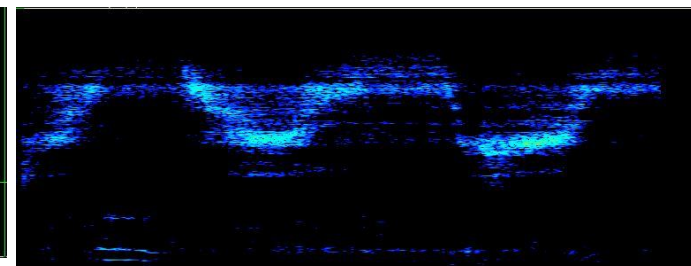


Figure XXXV : "mammographie" sur l'exercice de coloration du [s] (patiente B) ; échelle 8kHz

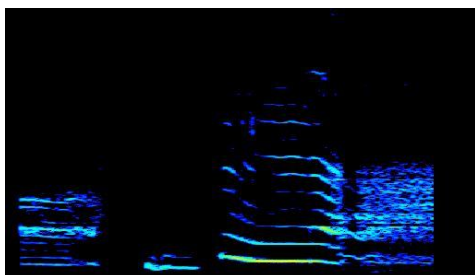


Figure XXXVI : essai infructueux de fry puis fry réussi après une sirènes descendante (patiente C), échelle 8kHz

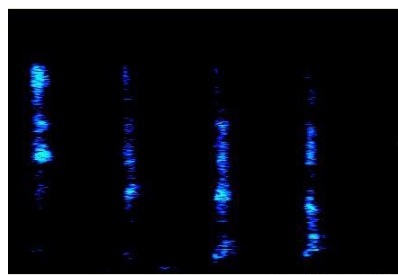


Figure XXXVII : "ti té tè ta" toniques (patiente B), échelle 8kHz

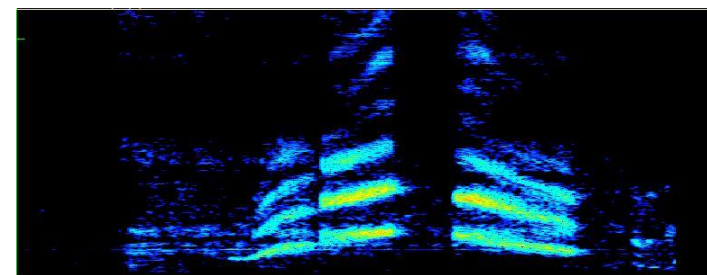


Figure XXXVIII : sirène en battements de lèvres, présentant des ruptures (patiente C) ; échelle 4 kHz

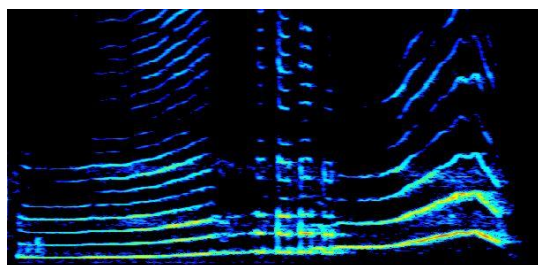


Figure XXXIX : visualisation du spasme sur une sirène sur [a] (patiente B) ; échelle 4kHz

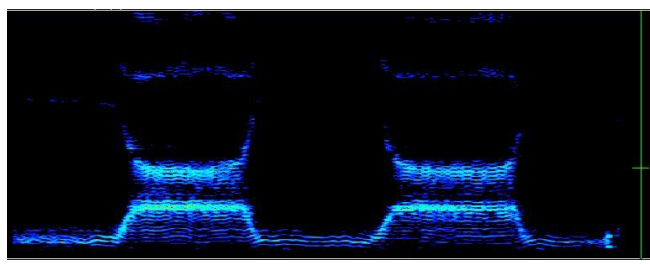


Figure XL : fry sur [i] et [a] alternés (patiente D) ; échelle 4kHz

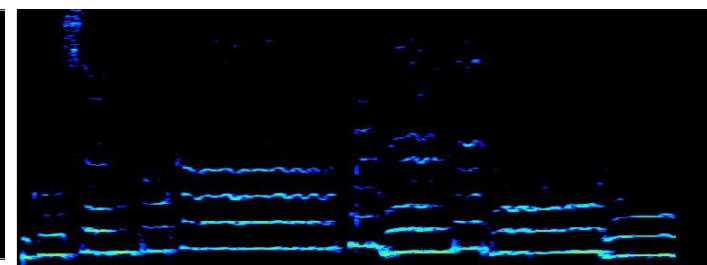


Figure XLI : extrait chanté (patiente A) ; échelle 4 kHz

D'autres exemples sont à voir et à écouter sur le site internet <http://florence.parmentier.free.fr/acoustique/spectrogramme/>

Je soussignée Florence Parmentier déclare être pleinement consciente que le plagiat de documents ou d'une partie d'un document publiés sur toutes formes de support, y compris l'Internet, constitue une violation des droits d'auteur ainsi qu'une fraude caractérisée. En conséquence, je m'engage à citer toutes les sources que j'ai utilisées pour écrire ce mémoire.

Apprécier sa voix en rééducation vocale : apport du spectrogramme. Etude de cas.

Le spectrogramme matérialise par un feedback visuel de nombreux paramètres acoustiques de la voix, notamment le timbre et l'attaque. Nous avons mené une étude incluant 4 patientes travaillant leur voix avec le spectrogramme, et une analyse statistique d'exercices d'auto-analyse et un recueil de témoignages chez 19 sujets. Nous concluons que la visualisation du spectrogramme, couplée à l'attitude positive du thérapeute, peut avoir un impact favorable sur la relation thérapeutique, sur l'image qu'a le patient de sa propre voix ; elle permet une meilleure reconnaissance de diverses composantes acoustiques de la voix.

mots-clés : rééducation vocale, spectrogramme, feedback visuel, auto-évaluation, relation à sa voix

The perception of one's voice: the impact of spectrogram visual feedback in voice therapy.

With a spectrogram, many acoustical parameters, such as voice onset and voice quality, may be visualized on a screen. We studied 4 patients training their voice with a spectrogram. We also performed a statistical analysis of self-assessment exercises and reports on the use of the spectrogram with 19 subjects. Our conclusion is that spectrogram visual feedback associated with the voice therapists' comforting attitude can have a positive impact on the therapeutic relationship and on the image the patient has of his voice and permits a better comprehension of the numerous acoustic parameters of the voice.

key words: voice therapy, spectrogram, visual feedback, self-awareness, self-perception

130 pages : 116 pages, table des matières et bibliographie incluses, 14 pages d'annexes

119 références bibliographiques