



HAL
open science

Bilan instrumental de la dysphonie

Alain Ghio

► **To cite this version:**

Alain Ghio. Bilan instrumental de la dysphonie. *Rééducation orthophonique*, 2013, 254, pp.9-29.
hal-01486662

HAL Id: hal-01486662

<https://hal.science/hal-01486662v1>

Submitted on 19 Apr 2018

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.

Bilan instrumental de la dysphonie

Alain Ghio

Résumé

Bien que peu répandue en pratique orthophonique, l'utilisation de méthodes d'analyse instrumentale des troubles de la voix et de la parole apparaît nécessaire en complément de toutes les autres formes d'évaluation (imagerie, perception, questionnaires) à l'instar de la démarche suivie en phonétique expérimentale et clinique. Le développement de logiciels d'analyse acoustique contribue à la diffusion d'une pratique des mesures objectives en milieu orthophonique. Si l'intérêt réel en termes de bilan reste sujet à discussion, l'apport de ces dispositifs dans la compréhension des mécanismes complexes de production vocale, notamment en cas de troubles, est indiscutable.

Mots clés : dysphonie, analyse acoustique, phonétique clinique.

Instrumental assessment of dysphonia

Abstract

Instrumental assessment of voice/speech disorders is not a widespread practice among speech therapists. Nevertheless, instrumental methods are needed in addition to all other forms of evaluation (imaging, perception, questionnaires), as is the case for experimental and clinical phonetics. The development of acoustic analysis software contributes to the use of objective measures in speech therapy. Although the true value of these methods regarding assessment remains a subject of debate, their contribution to the understanding of the complex mechanisms involved in vocal production is indisputable, particularly in cases of pathology.

Key Words : dysphonia, acoustic analysis, clinical phonetics.

Alain GHIO
Ingénieur de Recherche au CNRS
UMR7309 CNRS - Université d'Aix-
Marseille
Laboratoire Parole & Langage
5 avenue Pasteur
BP 80975
13604 Aix en Provence Cedex 1
Courriel : alain.ghio@lpl-aix.fr

◆ Pourquoi recourir aux mesures instrumentales dans un bilan vocal ?

Dans le cadre de la prise en charge de la dysphonie, l'évaluation perceptive de la qualité de la voix est un outil de dépistage, de diagnostic et d'évaluation des thérapies, qu'elles soient chirurgicales ou rééducatives. Elle vient en complément d'autres examens considérés comme un « tronc commun » pour l'évaluation de la plupart des dysphonies (De Jonckere et al, 2001) : vidéo-strobo-endoscopie du larynx, analyses acoustiques, examens aérodynamiques, auto-évaluation... En pratique orthophonique, l'évaluation perceptive est considérée comme le « gold standard ». Cependant, même si elle est très largement utilisée et si une écoute attentive contribue de façon indéniable à dresser le tableau clinique complet du patient, son importance réelle dans un procédé fiable d'évaluation peut être discutée. En effet, de nombreuses études ont mis en évidence une variabilité certaine, observée dans le jugement porté par différents auditeurs sur une même voix ou pour un même auditeur entre diverses séances d'écoute (Kreiman et al., 1992 & 1993). Une explication partielle de ces phénomènes perceptifs indésirables en pratique clinique a été proposée (Giovanni et al., 2009). Dans la mesure où une fiabilité suffisante ne peut être obtenue que par le recrutement de plusieurs auditeurs experts, la nécessité de proposer plusieurs sessions d'écoute, celle de mettre en place des séances d'apprentissage consensuel (Ghio & Dufour, 2011), l'évaluation perceptive s'avère finalement consommatrice en temps et en ressources humaines, ne permettant pas finalement une utilisation optimale en pratique clinique.

En complément de l'évaluation perceptive, l'analyse instrumentale est conçue pour qualifier et surtout quantifier les dysfonctionnements vocaux à partir de mesures acoustiques, aérodynamiques (débit d'air, pression intra-orale, pression sous-glottique) ou électrophysiologiques (EGG, EMG...) Ces mesures sont réalisées sur le patient en cours de production vocale par le biais de capteurs conçus pour enregistrer et étudier de multiples paramètres de la production de parole. La majorité des études portant sur ces procédés font apparaître la

nécessité de combiner différentes mesures complémentaires afin de tenir compte du caractère multidimensionnel de la production vocale.

La littérature abonde de comparaisons entre les méthodes perceptives et instrumentales. Baken (2000, p.146) affirme que les « *méthodes d'évaluation (perceptives) sont inadéquates... incapables de mettre en évidence de petites différences* ». De même, Morsomme et Estienne (2006) confirment que si « *notre premier outil d'évaluation acoustique est l'oreille humaine...l'audition est insuffisante...(car) elle traite le signal acoustique dans sa globalité. L'oreille peut donc être dupée* ». Enfin, Kent et al. (2000) participent à ce tableau comparatif en signalant :

p1. « *Certains avancent que la qualité de la voix peut être décrite et évaluée uniquement par des moyens auditivo-perceptifs* »

p1. « *D'autres affirment que les méthodes auditivo-perceptives sont enclines à diverses sources d'erreurs...* »

p117. « *Il est généralement admis dans la recherche actuelle que la nature impressionniste de la description perceptive devrait être augmentée par la nature plus objective de l'analyse instrumentale quand cela est possible* »

p117. « *C'est important cependant de ne pas supposer que des approches instrumentales sont d'une façon ou d'une autre les seules correctes et infaillibles. Des limites sur la précision existent avec toutes les techniques et l'interprétation de résultats n'est pas toujours directe* ».

Ce débat entre mesure physique et perception de la parole, récurrent dans la pratique clinique, n'est pas sans rappeler un épisode fondateur de la phonétique expérimentale. Dès 1875, la Société de Linguistique de Paris envoie une délégation pour consulter Etienne-Jules Marey, un physiologiste français travaillant sur l'étude des mouvements des organismes vivants et ayant mis au point une méthode graphique permettant de visualiser lesdits phénomènes (Teston, 2004a). Les linguistes de l'époque font appel à E.J. Marey dans le but « *d'appliquer la méthode graphique à l'étude des mouvements si complexes et si variés qui se produisent dans la parole [...] pour obtenir une trace objective des mouvements des organes articulateurs, cage thoracique, larynx, langue, lèvres, voile du palais, au cours de l'articulation des différentes unités phonétiques* ». Cette démarche est la conséquence d'un constat posé par certains linguistes qui témoignent que « *l'oreille, la vue et la palpation ne sont pas suffisants pour appréhender les mouvements successifs et simultanés nécessaires à la réalisation d'un phonème* ». De cette rencontre finissent par émerger divers instruments dont celui de Rosapelly (Figure 1, gauche), et surtout le kymographe de Roussetot (Figure 1 droite), qui va perdurer longtemps dans les études de phonétique articulaire jusque dans les années 1970.

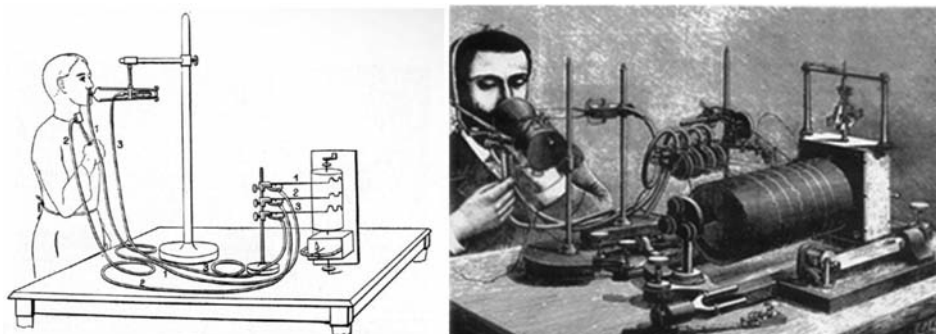


Figure 1 : Premiers dispositifs d'évaluation instrumentale de la parole.
A gauche, montage expérimental de Rosapelly permettant l'enregistrement de la pression nasale, des vibrations laryngées et des mouvements des lèvres (1875).
A droite, le kymographe de Rousselot (1892).

La phonétique expérimentale avec notamment la description des mécanismes de production de la parole est directement fondée sur les mesures obtenues avec ces dispositifs qui ont permis de s'extraire de spéculations théoriques. L'apparition, par la suite, de procédés permettant l'enregistrement et la restitution sonore puis de techniques d'analyse acoustique a définitivement ancré les trois composantes principales de la phonétique moderne : phonétique articulaire, phonétique acoustique et phonétique perceptive. Par la suite, l'avancée des connaissances dans ces trois domaines s'est faite en parallèle, les progrès résidant dans la mise en relation des faits expérimentaux observés avec notamment la mise en relation entre phono-articulation et résultats acoustiques (quels sont les liens entre le fonctionnement des organes de la parole et le signal sonore produit ?) ainsi que la mise en relation entre acoustique et perception (quel est l'impact perceptif de variations acoustiques ?). Très vite, les phonéticiens ont donc intégré le fait que l'étude des mécanismes de production de la parole nécessite une approche instrumentale héritée de Marey, s'affranchissant de la composante perceptive appréhendée comme un autre phénomène. Utiliser la perception auditive pour observer la production de la parole d'un (inter)locuteur reste très contestable en phonétique sauf si l'objectif est d'étudier la perception en tant que telle ou l'intégration des deux mécanismes appréhendés dans un même phénomène de production/perception.

La phonétique clinique, qui intègre une approche phonétique sur des phénomènes cliniques tels que la dysphonie, s'inscrit dans cette démarche dans laquelle les instruments, et notamment l'analyse acoustique est le moyen privilégié d'observer et mesurer les phénomènes de production de la parole, la perception s'inscrivant elle comme un phénomène distinct.

◆ Précautions techniques nécessaires à l'évaluation acoustique

Les conditions d'enregistrements nécessaires à l'évaluation acoustique

L'analyse acoustique de la voix et de la parole nécessite une bonne qualité du signal sonore, particulièrement dans l'étude des dysfonctionnements où il s'avère parfois nécessaire de mesurer de fines nuances qui ne doivent pas être entachées de bruit. Actuellement, la diffusion très large d'ordinateurs équipés en série d'entrées audio autorisant l'enregistrement sonore sans difficulté ne doit pas occulter le fait que certaines précautions doivent être prises. Typiquement, il est déconseillé d'utiliser un microphone bas de gamme connecté directement sur l'entrée micro de l'ordinateur : le résultat pourra être convenable à l'oreille mais les logiciels d'analyse auront des difficultés à fournir des mesures correctes sur des données altérées (Figure 2). Il est préférable, d'une part d'utiliser des microphones de qualité, le plus souvent électrostatiques. D'autre part, il est conseillé de connecter le microphone sur un préamplificateur externe et non directement à l'ordinateur. En effet, ces dispositifs périphériques (ex : Roland TM Quad Capture) génèrent beaucoup moins de bruit et, comme la plupart intègrent à présent les convertisseurs analogique-numérique. L'expérimentateur récupère alors directement le signal numérisé, évitant ainsi de multiples manipulations génératrices de problèmes.

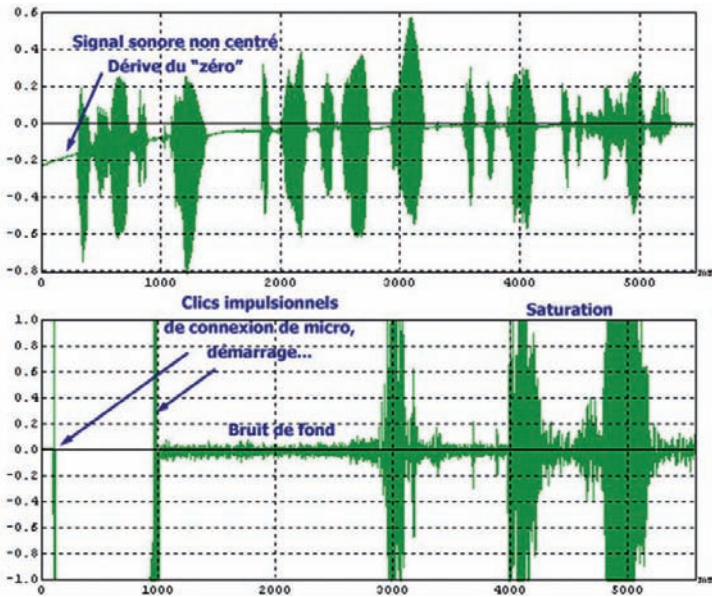


Figure 2 : Signaux de parole enregistrés sans précaution sur des dispositifs de mauvaise qualité. Une analyse acoustique fiable ne sera pas possible sur de telles données. En haut, on remarque un offset initial (le signal n'est pas centré sur zéro) et une dérive pro-

gressive. En bas, on observe des clics au démarrage, un bruit de fond résiduel important et la saturation du signal dans les parties énergétiques.

Le problème du calibrage de l'intensité

Le sonomètre reste l'instrument idéal pour mesurer l'intensité sonore car il fournit une mesure SPL calibrée. Il permet de faire des études comparatives fiables sur le niveau sonore. Ainsi, par exemple, il est possible de mesurer le fait que tel locuteur a augmenté de 6 dB sa puissance sonore en 3 mois de rééducation. A partir du moment où le signal sonore est enregistré avec un microphone dont on ne connaît pas la sensibilité, ce type d'analyse comparative devient délicat. En effet, si les caractéristiques de la chaîne d'enregistrement constituée par le microphone, le préamplificateur et le convertisseur numérique ne sont pas parfaitement connues, il est impossible de connaître la valeur SPL finale de l'enregistrement, en particulier s'il est stocké sous forme numérique en fichier (ex : fichier wav). En effet, l'amplitude observée du signal dépend du microphone utilisé, du volume d'enregistrement utilisé et de divers autres réglages. Autrement dit, on peut obtenir pour un locuteur à voix forte un fichier audio avec de faibles valeurs d'amplitude tout simplement parce qu'on a utilisé un microphone à faible sensibilité, placé loin du locuteur, avec un faible volume d'enregistrement. Inversement, un locuteur hypophonique peut fournir un fichier audio avec de fortes valeurs d'amplitude si un microphone sensible placé à 10 cm du locuteur avec un enregistreur à fort gain a servi pour l'enregistrement.

Hormis le sonomètre, il existe deux moyens d'obtenir des données sonores permettant la mesure calibrée en intensité. La première consiste à maîtriser parfaitement toute la chaîne d'acquisition. C'est le cas dans le dispositif EVA2 fabriqué par SQLab-LPL (www.sqlab.fr) où un microphone précis est connecté à un préamplificateur réglé pour fournir 10 V crête à 110 dB SPL à 30 cm. La deuxième solution consiste à utiliser une source sonore étalon, ceci avant chaque enregistrement. On peut utiliser un « buzzer » dont l'intensité est connue (ex : 83 dB à 30 cm). En émettant un bip en début d'enregistrement, on peut alors en déduire, par comparaison, l'intensité du signal de parole (ex: -10 dB par rapport au buzzer signifiera 73 dB). Si aucune précaution de calibrage n'a été prévue par l'opérateur, il n'est pas possible de comparer diverses intensités sonores entre divers enregistrements.

La délicate tâche de détection de la fréquence fondamentale

La majeure partie des indices exploités dans l'évaluation acoustique de la voix ou de la mélodie dérive de la mesure de la fréquence fondamentale, qui s'avère cruciale. La plupart des problèmes de détection de F0 porte sur 2 catégo-

ries : le voisement et les sauts d'octave. La décision du voisé/non voisé est un compromis délicat : soit le détecteur élimine trop de parties voisées en les considérant comme non voisées et par conséquent, la F0 n'est pas calculée sur ces parties intéressantes ; soit le détecteur favorise trop le voisement et des parties non voisées sont soumises à un calcul de F0 qui se solde par des valeurs aberrantes car portant sur du bruit a périodique. Il peut être intéressant d'introduire une contre-réaction sur la décision voisé/non voisé suite au calcul de F0 en cas de valeurs aberrantes et déclarer non voisées des suites anarchiques de F0 dues, par exemple, à un calcul sur du bruit. Toutefois, cette solution peut s'avérer malvenue notamment dans le cas de la parole pathologique. Dans ce cas précis, il est préférable d'introduire un réglage semi-manuel laissé à l'expertise de l'observateur.

Le deuxième problème fréquent dans la mesure de F0 est le saut d'octave. En effet, dans les techniques temporelles de détection telles que l'AMDF (Hess, 1983), si une périodicité est observée pour un temps t , on considérera qu'il s'agit alors de la période fondamentale T_0 , inverse de la F0. Mais $2T_0$, $3T_0$, $4T_0$ constituent aussi des candidats à la périodicité. Si l'algorithme, pour diverses raisons, sélectionne comme résultat $T=2T_0$, cela entraîne une erreur de calcul de fréquence fondamentale égale à $F_0/2$. Inversement, dans les techniques spectrales de type analyse harmonique, méthode du peigne (Hess, 1983), si un espacement f constant est observé entre deux raies spectrales, on considérera qu'il s'agit alors de la fréquence fondamentale F0, cette valeur étant l'écart systématique entre les harmoniques. Mais $2F_0$, constitue aussi un candidat à l'intervalle spectral. Si l'algorithme, pour diverses raisons, sélectionne comme résultat $F=2F_0$, cela entraîne une erreur de calcul de fréquence fondamentale (doublement).

La détection de F0 n'est donc pas un processus sans faille comme peut l'être le calcul de l'intensité ou du spectrogramme. Il existe des seuils, des limites, des choix qui ont une incidence sur le résultat. Aussi, face à une courbe de F0, il faut toujours garder un œil critique. Tout saut brutal ou toute montée/descente tronquée doivent être vus avec circonspection. En même temps, dans le cas de parole pathologique, ces « sauts » de F0 particulièrement visibles, par exemple, dans les diplophonies, font partie entière de la production vocale. L'œil expert reste donc important.

◆ Evaluation acoustique de l'altération vocale

Les travaux portant sur l'évaluation acoustique des troubles vocaux sont innombrables. Nous ne présenterons que les principaux paramètres. Des détails

peuvent être consultés dans (Baken et al., 2000). Nous nous focaliserons dans ce chapitre sur l'utilisation de l'analyse acoustique pour l'exploration de l'altération laryngée, terminologie empruntée à la classification de l'Organisation Mondiale de la Santé¹. Dans ce modèle, régulièrement adapté, l'altération désigne (Rossignol, 2007) « une perte de substance ou l'altération d'une structure organique comme par exemple, la perte d'un membre ou la détérioration de certains tissus... L'invalidité désigne une restriction objective (du fait d'une altération corporelle) d'une capacité à réaliser certains actes... Le handicap désigne les difficultés qu'est susceptible de rencontrer un individu dans sa vie sociale ou personnelle du fait d'une altération corporelle ou d'une invalidité. L'importance de ces difficultés étant fonction de la situation particulière, matérielle, familiale et sociale, dans laquelle il se trouve ». Adaptée aux troubles de la voix, l'apparition de nodules ou une cordectomie sont considérées comme des altérations. Ces altérations induisent des mécanismes de production vocale anormaux potentiellement mesurables par des techniques instrumentales mis en œuvre, la plupart du temps, sur des voyelles tenues. A ce stade, l'objectif est de faire un contrôle technique de l'appareil phonatoire.

Les mesures de vibration laryngée

La fréquence fondamentale moyenne

La fréquence fondamentale F0 reste un bon indicateur des caractéristiques biomécaniques des cordes vocales. La F0 moyenne apporte une mesure globale de la hauteur de la voix du sujet (voix aiguë, grave...). La Table 1 synthétise diverses études qui fournissent des valeurs moyennes sur des populations normophoniques.

Table 1 : moyennes et écart-type de F0 par groupe de locuteurs selon diverses sources

Source	Peterson et Barney, 1952			Fitch & Holbrook, 1970		Yu et al., 2007	
Age (ans)	Inconnu			18-25		19-59 (moy=38)	
Groupe	Hommes	Femmes	Enfants	Hommes	Femmes	Hommes	Femmes
Langue	Americ.	Americ.	Americ.	Americ.	Americ.	Français	Français
Effectif	33	28	15	100	100	38	20
F ₀ moyenne	131 Hz	223 Hz	264 Hz	116 Hz	217 Hz	122 Hz	218 Hz
Ecart-type F ₀	19 Hz	27 Hz	35 Hz	34 Hz	45 Hz	24 Hz	28 Hz

1 « International classification of Impairments, Disabilities and Handicaps, a manual of classification relating to the consequences of disease » Geneva : World Health Organisation, 1980, dont l'acronyme est ICIDH-80.

Sauf écart très important (ex : voix d'homme à 250 Hz) par rapport à ces données « normatives », il est difficile de corréliser un dysfonctionnement et une simple valeur de F0 moyenne. Ceci est d'autant plus vrai qu'il existe en fonction de l'âge un glissement différent de la F0 moyenne pour les hommes et les femmes, la F0 augmentant chez les hommes, diminuant chez les femmes (Figure 3a). (Yu et al., 2007) ont retrouvé aussi cette tendance en fonction du degré de dysphonie (Figure 3b).

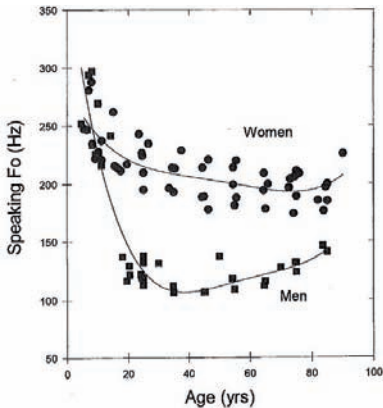


Figure 3a : évolution de la F0 moyenne chez les hommes et les femmes en fonction de l'âge (Baken et al., 2000, p.174)

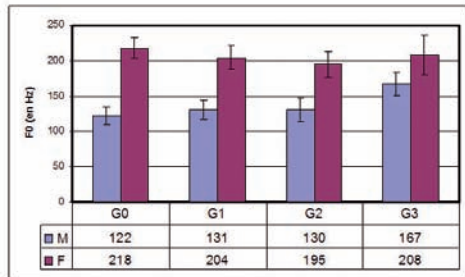


Figure 3b : évolution de la F0 moyenne chez les hommes et les femmes en fonction du grade G de dysphonie (G0 = normal, G3 = dysphonie sévère) (Yu et al., 2007)

L'ambiguïté contradictoire des variations de Fo

Les variations de F0 peuvent rendre compte du double phénomène de contrôle et de régulation du vibrateur laryngé. Dans un fonctionnement correct, la capacité de pouvoir faire varier la fréquence de vibration des cordes vocales est révélatrice d'une bonne maîtrise vocale. C'est particulièrement le cas chez le chanteur qui possède une importante étendue tonale. On mesure alors une adéquation entre variation importante de F0 et bon fonctionnement. Inversement, de fortes perturbations instantanées du cycle vibratoire se mesurent par des variations de F0 révélatrices d'une dérégulation du système phonatoire. On retrouve alors une corrélation opposée à la précédente entre variations et dysfonctionnement. En fait, seules les variations contrôlées sont synonymes de bonne maîtrise. D'où l'importance cruciale du contexte de production vocale et des consignes.

Une consigne de production de voyelle tenue permet de focaliser la mesure sur la stabilité et régulation du vibrateur laryngé et d'écarter les variations contrôlées. C'est d'ailleurs l'un des intérêts de l'utilisation de voyelles tenues. Dans ce cas-là, où la stabilité est demandée, comme une sorte de posture statique avec recherche d'immobilité, toute variation est synonyme de dysfonctionnement. Inversement, dans de la lecture de texte offrant une possibilité d'expressivité, tâche nécessitant un dynamisme s'apparentant à de la marche, les changements lents traduisent une bonne capacité vocale alors que les variations à court terme, voire les dévoisements inattendus, révèlent le dysfonctionnement. D'où la nécessité d'introduire plusieurs paramètres.

Instabilité de la vibration

La dynamique ou l'instabilité laryngée se traduit par des variations de F0 au cours du temps. Elle peut se mesurer par l'**écart type de la F0**, qui correspond à l'ampleur en Hz des variations de F0 autour de la moyenne. Le **coefficient de variation de F0** correspond à l'ampleur en % des variations de F0 par rapport à la F0 moyenne. Il est donc un bon indice pour explorer globalement la dynamique ou l'instabilité du vibrateur laryngé (Figure 4). Sur des voyelles tenues, il devient important en présence de vibrato pathologique, chevrottement, instabilité d'origine neurologique.

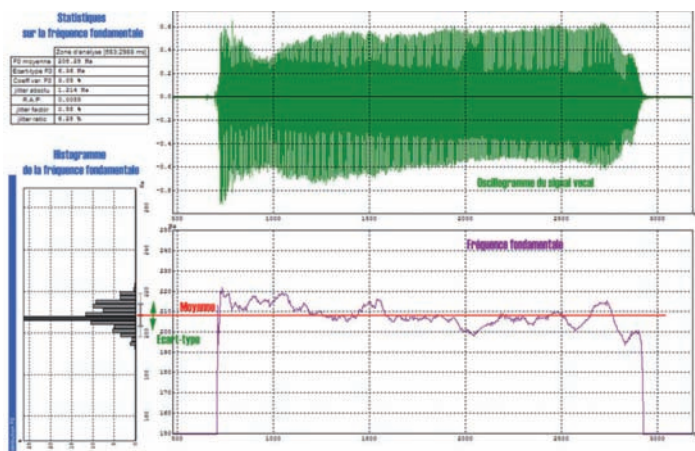


Figure 4 : signal vocal sur un /a/ tenu d'un adolescent en cours de mue « difficile ». Décours temporel, histogramme et statistiques de la fréquence fondamentale. On observe les variations importantes de F0 alors que la consigne est à la stabilité. Le coefficient de variation F0 de 3% mesure cette instabilité globale. En revanche, le faible jitter de 0.58 % traduit une relativement bonne stabilité instantanée.

L'instabilité à court terme de la F0 se traduit par des variations de fréquence entre chaque cycle d'oscillation (Figure 6). Elle se calcule par le **JITTER absolu moyen** qui est la moyenne de la différence de F0 entre deux cycles de vibration consécutifs. Ces variations sont calculées cycle à cycle, de manière très précise. Le **JITTER factor** permet de relativiser le jitter moyen en le comparant à la F0 moyenne. Ainsi, à un jitter moyen de 0.677 Hz pour une F0 moyenne de 180 Hz correspond un jitter factor de 0.38%, valeur faible. Le même jitter moyen pour une F0 moyenne de 80 Hz fournit un jitter factor de 0.84%, valeur plus problématique. Le jitter factor est donc un meilleur indice pour explorer la stabilité de la fréquence fondamentale à court terme. Il reste le meilleur corrélat acoustique de la raucité, bien que cette corrélation simple ne permette pas de rendre compte complètement de cette dimension vocale. Par consensus, il est admis qu'une valeur du jitter factor supérieure à 1% est une trace d'instabilité laryngée (Figure 5). De même, toute valeur qui dépasse 5% traduit une forte instabilité.

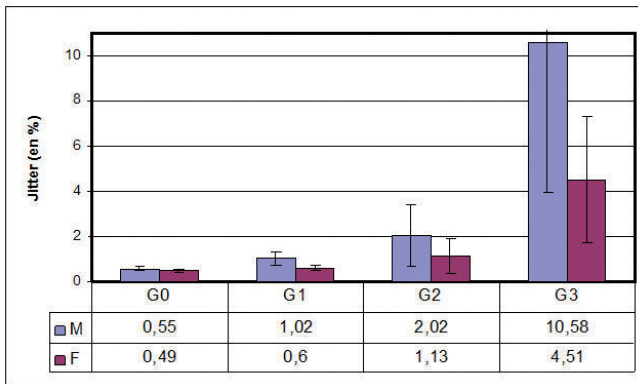


Figure 5 : évolution du Jitter factor chez les hommes et les femmes en fonction du grade G de dysphonie G0 = normal, G3 = dysphonie sévère (Yu et al., 2007)

Le **RAP** (Relative Average Perturbation) mesure la moyenne des variations de trois périodes consécutives rapportée à la période moyenne du signal observé. Le **JITTER ratio** mesure la moyenne des variations de période, inverse de la fréquence, entre deux cycles de vibration consécutifs et relativise cette valeur par la période moyenne du signal observé.

Il faut signaler qu'un jitter élevé entraîne toujours un coefficient de variation important. La réciproque n'est pas vraie. En effet, de petites variations croissantes ou décroissantes de F0 entre chaque cycle ne fourniront pas un jitter important mais pourront donner lieu à des variations globales importantes de F0 (exemple typique : un vibrato).

L'instabilité à court terme de l'amplitude de la vibration se traduit par des variations d'amplitude entre chaque cycle d'oscillation (Figure 6). Elle se calcule par le **SHIMMER moyen** qui est la moyenne des rapports d'amplitudes entre deux cycles de vibration consécutifs. Ce rapport est généralement exprimé en dB. Ces variations sont calculées cycle à cycle, de manière très précise. L'**APQ** (Amplitude Perturbation Quotient) mesure la moyenne des variations d'amplitude en tenant compte de 11 périodes consécutives, le tout rapporté à l'amplitude moyenne du signal observé.

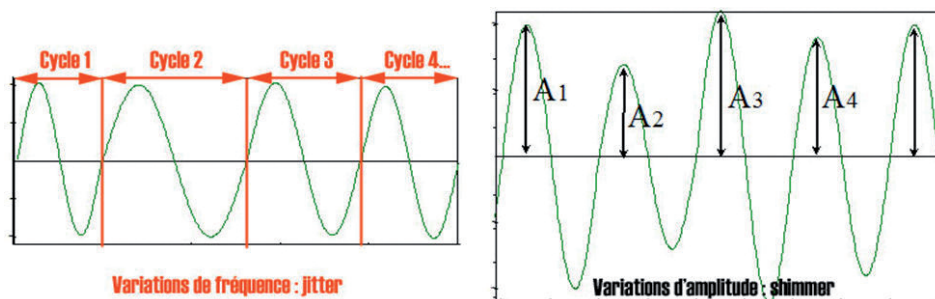


Figure 6 : Variations instantanées des caractéristiques de la vibration. A gauche, l'écart de fréquence entre deux cycles consécutifs entraînera un jitter important ; à droite, l'écart d'amplitude entre deux cycles consécutifs entraînera un shimmer important

Il existe de nombreux paramètres de ce type et il serait fastidieux de les présenter tous. Pour plus de détails, le lecteur pourra se référer à (Baken et al., 2000).

Les mesures de signal/bruit

Les mesure de signal et de bruit explorent la présence (ou absence) de « bruits » au cours de la phonation. Ce bruit peut être dû à un écoulement turbulent causé par un mauvais accolement des cordes vocales. On parle alors de « bruit additif ». Toutefois, il faut signaler qu'une vibration très instable peut être elle aussi mise en évidence par cette analyse : l'écart de la F0 et de ses harmoniques par rapport à la valeur moyenne est détectée comme un « bruit » de modulation.

Le rapport Harmonique/Bruit (Harmonic to Noise Ratio, **HNR**) se calcule à partir d'une méthode temporelle décrite dans (Yumoto et al., 1982). 25 cycles consécutifs du signal voisé permettent d'établir une forme ondulatoire moyenne ; la soustraction entre le signal originel et la forme moyenne permet d'obtenir un résidu considéré comme le bruit. Si H est l'énergie de l'onde

moyennée périodique et N l'énergie du bruit résiduel, le rapport H/N , calculé régulièrement, fournit la valeur du HNR éventuellement converti en dB.

Le Signal Ratio (Sr) se calcule à partir d'une méthode spectrale décrite dans (Hiraoka et al., 1984). On considère le signal vocal comme la somme de deux composantes : une composante périodique (F_0 & harmoniques) et une composante de bruit (le reste). A partir d'un spectre effectué sur une portion de signal, l'énergie présente dans les raies est vue comme la composante périodique du signal. Le calcul de cette énergie dans les raies par rapport à l'énergie totale fournit un rapport appelé Sr . De plus, l'auteur préconise aussi de faire le même type de calcul non pas sur tout le spectre mais uniquement sur des parties moyennes et hautes fréquences comme, par exemple, au dessus de 1 kHz. Dans ce cas-là, le Signal Ratio est particulièrement révélateur de dysphonie car il devient très faible chez ces patients dont le dysfonctionnement entraîne une composition spectrale appauvrie dans les aigus, caractéristique captée par le Sr calculé sans les graves (Figure 7).

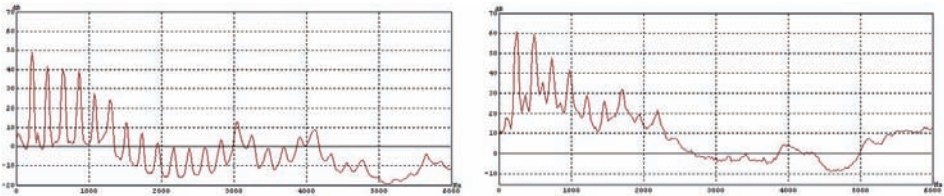


Figure 7 : Spectres d'une voix normale (à gauche) et d'une dysphonie (à droite). Les composantes harmoniques disparaissent clairement au dessus de 2000 Hz sur l'échantillon dysphonique.

L'énergie Normalisée du Bruit (**NNE**) se calcule à partir d'une méthode spectrale décrite dans (Kasuya et al., 1986). Cette méthode est assez proche de la précédente excepté dans la formulation finale qui indique non pas le pourcentage de périodicité mais plutôt la part de bruit exprimée en dB par rapport au signal complet.

Ces indices qui traduisent l'importance de la composante de bruit dans la voix sont reconnus (Yu et al., 2007) pour être des mesures acoustiques corrélées avec le degré de sévérité de la dysphonie (Figure 8).

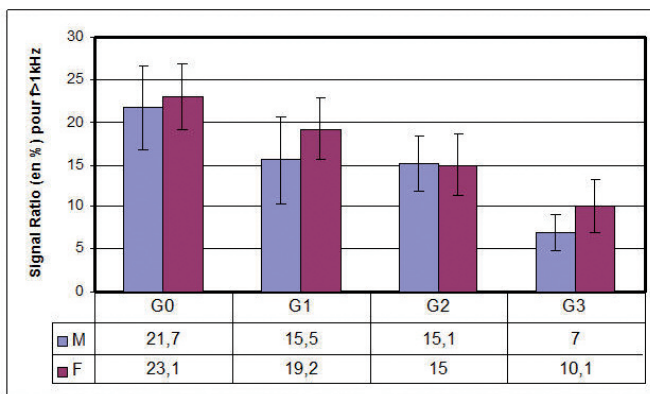


Figure 8 : évolution du Signal Ratio ($f > 1\text{kHz}$) chez les hommes et les femmes en fonction du grade G de dysphonie ; G0 = normal, G3 = dysphonie sévère (Yu et al., 2007)

Les mesures non linéaires

L'analyse quantitative des attracteurs a fait l'objet de travaux dans le domaine. Le calcul des coefficients de Lyapunov a été proposé sur le plan théorique par Titze puis explicitement développé dans l'évaluation vocale par (Giovanni & Ouaknine., 1999). Le principe général est le calcul de la divergence de deux trajectoires voisines d'un portrait de phase. Toutefois, le Plus Grand Exposant de Lyapounov, bien que paramètre d'analyse très prometteur, reste peu utilisé du fait de la difficulté de relier cette mesure à une réalité physiologique.

◆ Les mesures acoustiques pour observer l'invalidité vocale

Dans le cadre des troubles de la voix et de la parole, l'invalidité (ou le dysfonctionnement d'après Rossignol, 2007) pourrait résider dans la restriction des capacités de communication orale, notamment dans la perte de qualité vocale, d'étendue fréquentielle, d'intensité, d'endurance, de coordination... A ce stade, l'objectif est de faire un contrôle de l'usage de l'appareil phonatoire.

Les mesures de performances

Bien que sujets à des problèmes de reproductibilité, les tests de performances tels que l'étendue vocale ou le temps maximal de phonation (TMP) sont des mesures qui s'avèrent particulièrement robustes notamment le TMP (Hirano et al., 1968). Les difficultés liées à ces méthodes reposent essentiellement sur la difficulté d'appliquer la consigne et la certitude d'obtenir la performance de la part du locuteur.

Le phonétogramme (Voice Range Profile) a été recommandé par l'Union Européenne des Phoniâtres et standardisé par Schutte et al. (1983). Il est défini par Dejonckere et al. (2003) comme la représentation graphique du champ dynamique vocal obtenue en quantifiant les intensités sonores maximale et minimale en fonction de la hauteur tonale du son fondamental sur toute l'étendue de la voix. Il permet ainsi de mettre en évidence une réduction de l'espace de fonctionnement laryngé en cas de dysfonctionnement (Figure 9). Si cette méthode est particulièrement adaptée à l'évaluation de la voix des chanteurs, il n'en est pas de même avec des patients qui ont la plupart du temps des difficultés à reproduire précisément la note cible. Il est alors possible d'obtenir une version plus réduite mais plus simple en demandant de produire un glissando (sirène) vers le grave puis vers l'aigu, ce qui permet d'obtenir l'étendue tonale en fréquences (Figure 10b). De même, il peut s'avérer intéressant d'obtenir l'intensité la plus basse et la plus forte en phonation.

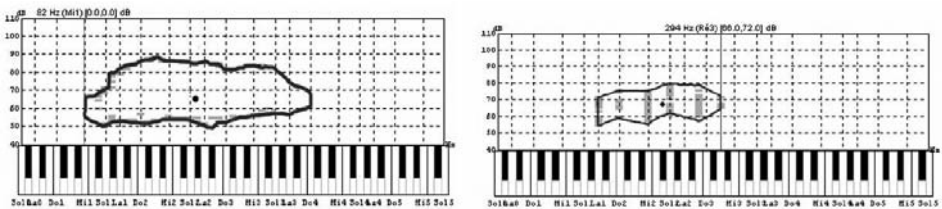


Figure 9 : phonétogramme d'un locuteur normal à gauche et d'un patient parkinsonien à droite (Source : Teston, 2004b)

Le temps maximal de phonation (TMP) est le temps qu'une personne peut tenir, après une profonde inspiration, en produisant une voyelle tenue à intensité et hauteur confortable. Cette méthode est souvent décrite comme un moyen rapide et simple de faire une évaluation aérodynamique (Dejonckere et al., 2001). En effet, cette durée dépend à la fois de la capacité à mobiliser l'air pulmonaire en phonation et de la qualité d'accolement des cordes vocales (plus la fuite glottique est importante, plus le TMP est court). 3 essais sont conseillés pour obtenir une valeur pertinente, ce qui parfois peut poser des problèmes de fatigue selon l'état du patient. La littérature reporte des valeurs normatives de 25 à 35 secondes pour des hommes et de 15 à 25 secondes pour des femmes (Figure 10a). Le fait de ne pas pouvoir distinguer par cette mesure la capacité à mobiliser l'air pulmonaire en phonation de la qualité d'accolement des cordes vocales (mécanismes très différents) peut être problématique en terme interprétatif. L'adjonction d'un capteur de débit d'air donne en revanche une dimension très complète aux mesures observées dans cet examen.

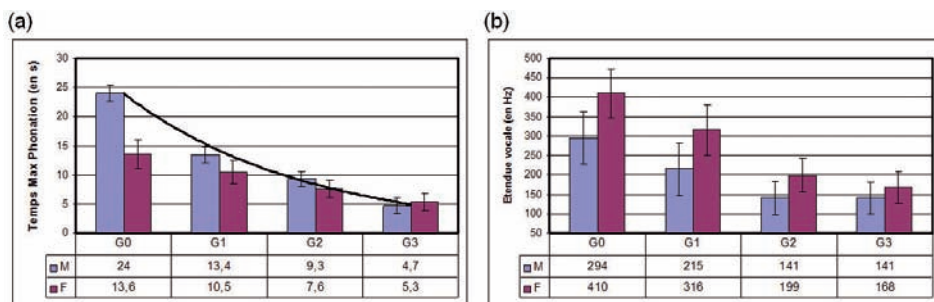


Figure 10 : évolution du Temps Max de Phonation et de l'étendue tonale chez les hommes et les femmes en fonction du grade G de dysphonie ; G0 = normal, G3 = dysphonie sévère (Yu et al., 2007)

La piste du V.O.T.

Le VOT, c'est à dire le Voice Onset Time (Temps d'établissement du voisement) est la durée entre le relâchement d'une occlusive et le début de voisement du phonème suivant. Cette durée est un bon indicateur des capacités du locuteur à coordonner son étage phonatoire et articulaire. Un dysfonctionnement des organes peut entraîner un allongement ou une réduction de cette durée pour des raisons diverses. Des VOT « anormaux » ont été observés dans l'apprentissage d'une langue seconde, dans l'aphasie ou l'apraxie de la parole, la dysarthrie, le bégaiement, la dysphagie, la dysphonie spasmodique. Cette mesure a été testée dans le cadre général de l'évaluation des dysphonies (Freddouille et al., 2009) et les auteurs ont mis en évidence un allongement de ce temps d'établissement du voisement en fonction du grade de dysphonie, ce qui peut être expliqué par la difficulté d'initier correctement la vibration des cordes vocales chez des patients dysphoniques (Figure 11). Cela va dans le même sens que les observations de Révis et al. (2002) qui portaient sur l'attaque de voyelles tenues.

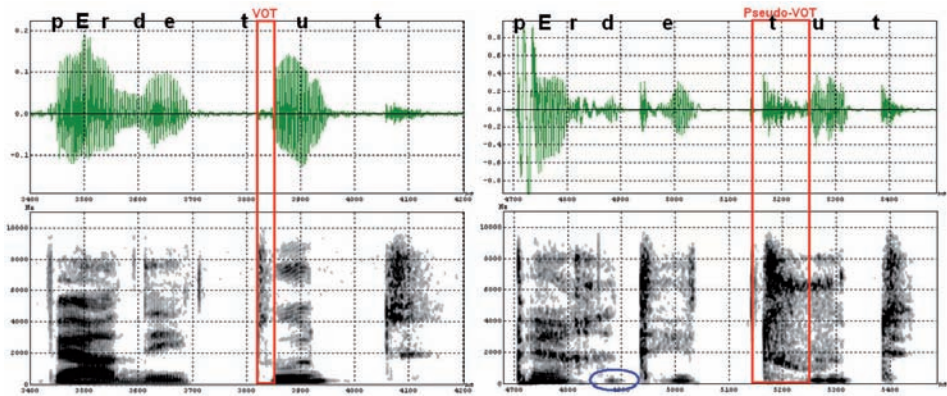


Figure 11 : oscillogramme (en haut) et spectrogramme (en bas) de l'énoncé « il les perdait toutes » pour une voix normale à gauche, une dysphonie sévère à droite. A droite, on note d'une part, l'allongement excessif du VOT sur le 1er /t/ de « toute » et la quasi transformation de cette occlusive en fricative. D'autre part, on observe le dévoisement de la séquence /rd/ de « perdait »

Les tendances high-tech

De nouvelles tendances "high-tech" se dessinent dans le but d'améliorer la pertinence des mesures acoustiques pour l'analyse de troubles vocaux. Ces techniques innovantes ont la vocation à se développer, notamment du fait de leur simplicité d'utilisation et de leur capacité à appréhender de la parole continue.

Alpan et al. (2011) ont proposé une analyse de la dyspériodicité vocale fondée sur la méthode du variogramme généralisé qui permet l'estimation des dyspériodicités inter-cycles sur de la parole continue, quelles que soient leurs causes. Le principe est d'estimer les variations qui apparaissent entre des trames de signaux successifs en tenant compte de divers paramètres. Les auteurs ont montré sur divers corpus de voix dysphoniques que le rapport signal/dyspériodicité semble corrélé fortement avec le degré perçu de la sévérité d'une dysphonie.

La Reconnaissance Automatique du Locuteur (RAL) s'intéresse aux informations extralinguistiques véhiculées par un signal de parole, informations porteuses de renseignements sur les spécificités d'un individu (identité, émotivité, caractéristiques physiques, particularités régionales...). Son objectif est d'identifier une personne à l'aide de sa voix grâce à la variabilité inter-locuteurs qui permet de reconnaître une voix parmi plusieurs possibles. Diverses tentatives ont été effectuées pour adapter ces techniques non pas à l'identification

d'un locuteur mais à un groupe de locuteurs, en l'occurrence des dysphoniques (Fredouille et al., 2009). Le principe est de fournir des données acoustiques de locuteurs atteints de troubles de la voix. Sur le signal sonore sont extraites des informations spectrales qui sont modélisées par une phase d'apprentissage supervisée. Durant cette étape, le système peut être confronté à des dysphonies légères, moyennes ou sévères. A la fin de l'apprentissage, il a établi de façon automatique des caractéristiques statistiques de chaque groupe. La phase d'utilisation proprement dite consiste par la suite à confronter le système à une nouvelle voix qui va être associée à l'un des groupes appris (ex : voix normale ou dysphonie légère ou dysphonie moyenne ou dysphonie sévère).

◆ L'aérophonométrie

L'aérophonométrie est directement héritée des travaux de Marey et Rousset (Figure 1). Le principe est de capter les flux d'air mis en jeu dans la production de parole. Dès (Teston, 1976), des versions modernes d'aérophonètres ont été développées. En France, une première utilisation en milieu clinique de ces techniques a été publiée dans (Giovanni et al., 1991). Les avancées successives ont donné naissance par la suite au dispositif d'Evaluation Vocale Assistée (EVA©), développé par le LPL et SQLab (www.sqlab.fr). Cet appareillage permet d'enregistrer simultanément des signaux acoustiques et aérodynamiques par l'utilisation d'une pièce à main contenant un microphone, un sonomètre, un pneumotachographe mesurant le débit d'air oral et nasal (Ghio & Teston, 2002) ainsi que deux capteurs de pression (intra-orale par exemple). D'autres capteurs, comme l'électroglottographie, peuvent aussi être enregistrés en parallèle.

Ainsi, mesurer le débit d'air oral et mettre en évidence une fuite glottique importante dans une paralysie laryngée est plus direct que de mesurer un rapport signal/bruit faible, révélateur de la fuite. D'autre part, cela permet une certaine sélectivité, par exemple dans la distinction entre mécanisme oral vs nasal. Globalement, le débit d'air oral informe sur la quantité d'air pulsé par les poumons à travers la glotte (phonation) ou la fermeture/ouverture du conduit vocal (articulation). Le débit d'air nasal informe sur la fermeture/ouverture du voile du palais, fonction intéressante à observer dans certaines atteintes neurologiques (Robert et al., 1995). Il permet aussi d'objectiver d'éventuelles fentes palatines ou fistules (Pech et al., 1995). La pression sous-glottique renseigne sur la quantité d'énergie utilisée par le locuteur au niveau de sa source et sur l'éventuel hyper/hypo fonctionnement pneumo-phonatoire. Rappelons que la pression sous-glottique peut être estimée de façon indirecte par la « airway interrupted

method » à l'aide d'un cathéter placé d'un côté dans la cavité orale du locuteur, de l'autre sur un capteur de l'appareillage EVA©. Le locuteur prononce alors une série de /papapa/. Sur la tenue de [p], les lèvres sont fermées. La glotte est ouverte. Un équilibre des pressions s'établit dans le conduit vocal : la pression sous-glottique peut être estimée en mesurant la pression intra-orale, bien plus facile à obtenir. Cette méthode s'avère particulièrement pertinente dans l'évaluation multiparamétrique des dysphonies (Yu et al. 2007). D'autres détails sur l'aérophonométrie peuvent être consultés dans (Teston, 2007).

◆ Conclusion

Bien que peu répandue en pratique orthophonique, l'utilisation de méthodes d'analyse instrumentale des troubles de la voix et de la parole apparaît nécessaire en complément de toutes les autres formes d'évaluation (imagerie, perception, questionnaires) à l'instar de la démarche suivie en phonétique expérimentale et clinique. Le développement de logiciels tels que VOCALAB (Menin-Sicard, 2003) ou de dispositifs plus complets mais moins accessibles tels qu'EVA (§ 5 ci-avant) contribuent à la diffusion d'une pratique des mesures objectives en milieu orthophonique. Si l'intérêt réel en termes de bilan reste sujet à discussion, l'apport de ces dispositifs dans la compréhension des mécanismes complexes de production vocale et de la parole, notamment en cas de troubles, est indiscutable.

REFERENCES

- ALPAN A., MARYN Y., KACHA A., GRENEZ F., SCHOENTGEN J. (2011), Multi-band dysperiodicity analyses of disordered connected speech *Speech Communication*, Vol. 53, 1, pp. 131 - 141, 2011.
- BAKEN R.J., ORLIKOFF R.F. (2000), *Clinical Measurement of Speech and Voice*. San Diego, CA, USA : Singular Publ. Group, Inc., 2e édition, 604 p.
- DEJONCKERE P, BRADLEY P, CLEMENTE P, CORNUT G, CREVIER-BUCHMAN L, FRIEDRICH G, VAN DE HEYNING P, REMACLE M, 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", *Eur Arch Otorhinolar*, 258 : 77-82, 2001.
- DEJONCKERE, P. H., Van WIJCK, I., & SPEYER, R. (2003). Evaluation de l'efficacité de la rééducation vocale par le phonétogramme. *Revue de laryngologie, d'otologie et de rhinologie*, 124(5), 285-289.
- FITCH JL, HOLBROOK. A. (1970). Modal vocal fundamental frequency of young adults. *Archives of Otolaryngology*, 92(4), 379-382.
- FREDOUILLE, C.; POUCHOULIN, G.; GHIO, A.; RÉVIS, J.; BONASTRE, J.-F.; GIOVANNI, A. (2009). Back-and-Forth Methodology for Objective Voice Quality Assessment : From/to Expert Knowledge to/from Automatic Classification of Dysphonia. *EURASIP Journal on Advances in Signal Processing*, vol. 2009. 2009, p. 1-13.

- GHIO A., DUFOUR S., ROUAZE M., BOKANOWSKI V., POUCHOULIN G., REVIS J., GIOVANNI A. (2011), "Mise au point et évaluation d'un protocole d'apprentissage de jugement perceptif de la sévérité de dysphonies sur de la parole naturelle", *Revue de Laryngologie*, volume 132.1, p. 19-27, 2011.
- GHIO, A., & TESTON, B. (2002). Caractéristiques de la dynamique d'un Pneumotachographe pour l'étude de la production de la parole : aspects acoustiques et aérodynamique. In Actes, Journées d'Etude sur la Parole (JEP) (p. 337 340).
- GIOVANNI, A.; RÉVIS, J.; GHIO, A. (2009). Le phoniatre face à l'évaluation instrumentale des dysphonies. *Revue de laryngologie, otologie, rhinologie*, vol. 130, no. 1. 2009, p. 3-10.
- GIOVANNI A, OUAKNINE M, TRIGLIA JM. (1999), Determination of the largest Lyapunov exponents of vocal signal. Applications to unilateral laryngeal paralysis. *J Voice* 1999 ; 13 : 341-354
- GIOVANNI, A.; MOLINES, V.; NGUYEN, N.; TESTON, B. (1991), « L'évaluation objective de la dysphonie : une méthode multiparamétrique », *Proceedings of International Congress of Phonetic Sciences (ICPhS)* (12 : 1991 août 19-24 : Aix-en-Provence, FRANCE). 1991, p. 274-277.
- HESS W. (1983), *Pitch Determination of Speech Signals, Algorithms and Devices*, Springer Series in Information Sciences, 698 p.
- HIRAOKA, N., KITAZOE Y., UETA H., (1984), « Harmonic-intensity analysis of normal and hoarse voices », *J.A.S.A.*, 76(6), 1648-1651.
- HIRANO, M., KOIKE, Y., & Von LEDEN, H. (1968). Maximum phonation time and air usage during phonation. *Clinical study. Folia phoniatrica*, 20(4), 185 201.
- KASUYA H., OGAWA S., MASHIMA K., EBIHARA S., (1986), « Normalized noise energy as an acoustic measure to evaluate pathologic voice », *J.A.S.A.*, 80(5), 1329-1334.
- KENT R., BALL M., 2000, *Voice Quality Measurement*. San Diego, USA : Singular Publ. Group, Inc., 2000, 492 p.
- KREIMAN J., GERRATT BR., PRECODA K., BERKE GS. (1992), Individual differences in voice quality perception. *J Speech Hear Res.* 1992 ; 35(3), 512-520.
- KREIMAN J., GERRATT BR., KEMPSTER GB., ERMAN A., BERKE GS., Perceptual evaluation of voice quality : review, tutorial, and a framework for future research. *J Speech Hear Res.* 1993 ; 36(1), 21-40.
- MENIN-SICARD, A., & SICARD, E. (2003). *Vocalab : a new software for voice evaluation and therapy*. *Revue de laryngologie - otologie - rhinologie*, 124(5), 311 314.
- MORSOMME D., ESTIENNE F. (2006), *Bilan Vocal dans Les Bilans de Langage et de Voix, Fondements théoriques et pratiques*, Masson.
- PECH C, LOGNOS C., CASSASSOLLES S., BARDOT J., MAGALON G. (1995), « Résultats phoniatriques après pharyngoplastie », *Actes du congrès sur le Voile Pathologique, Société Française de phoniatry*, Lyon, p.43-51.
- PETERSON, G. E., BARNEY, H. L. (1952). Control methods used in a study of the vowels. *Journal of the Acoustic Society of America*, 24, 175-184.
- REVIS J., GIOVANNI A., TRIGLIA JM. (2002), « Influence de l'attaque sur l'analyse perceptive des dysphonies », *Folia Phoniatr Logop* 2002 ; 54 : 19-25.
- ROBERT D., SANGLA I., AZULAY J.P., GIOVANNI A., CANNONI M., POUGET J. (1995), « Diagnostic et suivi de l'insuffisance vélaire dans les formes bulbaires des maladies du motoneurone », *Actes du congrès sur le Voile Pathologique, Société Française de phoniatry*, Lyon, p.63-74.
- ROSSIGNOL, C. (2007). Classifications internationales des altérations corporelles, dysfonctionnements et handicaps. Pour une clarification des concepts. *Actes, Entretiens de Bichat, Orthophonie (2007 septembre 28-29 : Paris, FRANCE)*. In *Entretiens de Bichat, Entretiens d'orthophonie*. ISBN 2-7046-1725-2. Paris : Expansion Formation Editions. 2007, p. 62-69.
- SCHUTTE, H. K. & SNEIDER, W. (1983). Recommendation by the union of european phoniatricians (UEP) : Standardizing voice area measurement/phonetography. *Folia Phoniatrians*, (35), 286-288.
- TESTON, B. (1976). « Etude et réalisation d'un système d'analyse des paramètres articulatoires : le polyphonomètre ». *Travaux interdisciplinaires du Laboratoire parole et langage d'Aix-en-Provence (TIPA)*, vol. 3. 1976, p. 151-207. ISSN 1621-0360.

- TESTON, B., (2004a), L'œuvre d'Etienne-Jules Marey et sa contribution à l'émergence de la phonétique dans les sciences du langage, *Travaux Interdisciplinaires du Laboratoire Parole et Langage d'Aix-en-Provence (TIPA)* **23** (2004) 237-266.
- TESTON B. (2004b), « L'évaluation instrumentale des dysphonies : État actuel et perspectives » In Giovanni A., « Le bilan d'une dysphonie : État actuel et perspectives », Solal, Marseille, 244 p.
- TESTON, B. (2007). L'étude instrumentale des gestes dans la production de la parole : Importance de l'aérophonométrie. In P. Auzou ; V. Rolland-Monnoury, S. Pinto, C. Ozsancak (ed.) Les Dysarthries. Marseille : SOLAL. 2007, p. 115-117.
- YU P., GARREL R., NICOLLAS R., OUAKNINE M., GIOVANNI A., (2007), " Objective voice analysis in dysphonic patients. New data including non linear measurements ", *Folia Phoniatica et Logopaedica*, 59 : 20-30.
- E.YUMOTO, W.GOULD, (1982) "Harmonics-to-noise ratio as an index of the degree of hoarseness", *J.A.S.A.*, 71(6), 1544-1550