



HAL
open science

Le phoniatre face à l'évaluation instrumentale des dysphonies

Antoine Giovanni, Joana Révis, Alain Ghio

► **To cite this version:**

Antoine Giovanni, Joana Révis, Alain Ghio. Le phoniatre face à l'évaluation instrumentale des dysphonies. *Revue de Laryngologie Otologie Rhinologie*, 2009, 130 (1), pp.3-10. hal-01616007

HAL Id: hal-01616007

<https://hal.science/hal-01616007>

Submitted on 20 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.

Le phoniatre face à l'analyse instrumentale des dysphonies ¹

Use of instrumental vocal analysis in phoniatrics

Giovanni A.
Révis J.
Ghio A.
(Marseille) ²

Résumé

La nécessité croissante d'évaluations quantifiées des actes de soin pousse les cliniciens à s'approprier les différentes méthodes instrumentales d'analyse de la voix proposées incluant les méthodes acoustiques et les méthodes aérodynamiques du fait des difficultés et des incertitudes de l'analyse perceptive. Le paradoxe est que cette même analyse perceptive est souvent présentée comme la méthode de référence contre laquelle les méthodes instrumentales sont évaluées. Les auteurs font le point sur l'intérêt clinique de toutes ces méthodes et notamment sur l'intérêt des mesures de débit d'air phonatoire et de pression sous-glottique.

Mots-clés : Analyse perceptive de la voix, analyse acoustique, débit d'air phonatoire, pression sous-glottique, mesures aérodynamiques.

Summary

The growing need for quantitative evaluation of health care intervention along with the pitfalls of perceptual analysis is leading clinicians to utilise different instrumental vocal analysis methods including acoustic and aerodynamic methods. The fact that perceptual analysis is most often presented as the gold standard for vocal evaluation remains paradoxical. The authors present the use of aerodynamic measurements in particular oral airflow and subglottic pressure.

Key-words: Perceptual voice analysis, acoustic analysis, oral airflow, subglottic pressure, aerodynamic measurements.

INTRODUCTION

La nécessité croissante d'évaluations quantifiées des actes de soin, dans le domaine de la voix comme dans les autres, pousse les cliniciens à s'approprier les différentes méthodes instrumentales proposées par les chercheurs. Mais ces méthodes restent d'interprétation difficile car elles reposent sur des analyses variées, souvent mal maîtrisées par les cliniciens et parfois même peu fondées sur le plan scientifique. Notre démarche au sein du Laboratoire Parole et Langage, initiée par Bernard Teston, consiste à mettre à la disposition du clinicien à la fois l'appareillage EVA qui permet des mesures acoustiques et aérodynamiques simultanées et dans le même temps une réflexion méthodologique pour que les mesures obtenues soient le plus fidèles possibles aux données scientifiques validées. L'association de données acoustiques et aérodynamiques recueillies chez 141 hommes et

308 femmes a ainsi permis de classer 85 % des voix analysées dans le même grade qu'un jury d'experts (de G0 à G3 conformément à la méthode décrite par Hirano). Ces travaux ont été publiés [1, 2] mais la place des méthodes instrumentales d'analyse de la voix reste encore relativement marginale. Il nous a semblé utile de proposer une réflexion d'ensemble sur le sujet, incluant le point de vue du clinicien aussi bien que celui de l'ingénieur phonéticien et celui du spécialiste de l'analyse perceptive des dysphonies.

Dans tous les cas les mesures instrumentales doivent systématiquement pouvoir être mises en relation par le clinicien avec les grandes lignes de la physiopathologie vocale, et on peut proposer une classification dans ce sens des méthodes proposées. Les anomalies de la hauteur de la voix sont bien sûr au mieux appréciées par les méthodes de détermination de la fréquence, tandis que les anomalies de l'intensité peuvent être quantifiées par la mesure instrumentale de l'intensité produite. L'instabilité de la vibration peut être appréciée par les variations de fréquence ou d'amplitude selon des méthodes variées (jitter, shimmer, méthodes non-linéaires) tandis que le caractère voilé ou soufflé sera mieux apprécié par l'analyse spectrale et notamment le rapport signal/bruit. La mesure directe de la fuite glottique peut être réalisée par la mesure du débit d'air buccal en phonation mais

1. Communication présentée au 64^{ème} Congrès de la Société Française de Phoniatry et des Pathologies de la Communication, Paris, 14/10/08.
2. Laboratoire "Parole et Langage" (UMR6057), Université de Provence & CNRS, 29, Avenue Robert Schuman, 13621 Aix-en-Provence cedex 1, France.
Correspondance : Pr Giovanni A, CHU La Timone, Fédération d'ORL et de chirurgie Cervico-faciale, 264 rue St Pierre, 13385 Marseille cedex 05, France. Email: antoine.giovanni@ap-hm.fr

Article reçu : 09/11/08

accepté : 02/03/09

Le phoniatre face à l'analyse instrumentale des dysphonies, Giovanni A, Révis J, Ghio A.

3

elle peut aussi être approchée par la mesure du temps maximum de phonation, même si cette durée inclut aussi des informations sur le contrôle respiratoire. Le forçage vocal est plus complexe à analyser sur le plan objectif et nécessite l'utilisation de méthodes indirectes comme l'étude de la pression sous-glottique estimée. Ces méthodes sont décrites en détail dans deux ouvrages de référence [3, 4].

METHODOLOGIE

Il n'existe pas de dispositif d'analyse de la voix permettant une analyse totalement automatisée et épargnant à l'examineur tout souci méthodologique. Au contraire, la variabilité prodigieuse de la voix normale et pathologique, la complexité des mesures effectuées et de leur interprétation imposent un apprentissage minimum des méthodes utilisées.

Le premier choix à faire est celui du matériau phonétique [5]. Il peut s'agir de parole plus ou moins naturelle ou au contraire de sons voisés soutenus ; ces derniers peuvent à leur tour être subdivisés selon que les attaques et les déclinaisons sont incluses ou non dans la fenêtre d'analyse. Ce type de production vocale permet de focaliser la mesure de façon sélective sur la stabilité et la régulation du vibrateur laryngé et d'écarter les variations contrôlées. 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, ce qui n'est pas le cas dans la parole continue. Cependant il est clair que la parole la plus proche de la parole spontanée est a priori souhaitable. Par ailleurs, une fois choisi le matériau phonétique, la tâche demandée au locuteur peut être de produire une voix proche de sa voix habituelle confortable en voix parlée ou chantée ; mais il peut s'agir au contraire de tester les performances vocales comme la voix la plus forte et la moins forte, la voix la plus aigüe et la plus grave. Dans le cadre de ce type d'examen de performance, il n'est pas aberrant de placer le système pneumophonatoire dans des situations peu naturelles, mais pertinentes pour l'observation. Cette situation se retrouve d'ailleurs dans d'autres domaines de la santé comme en cardiologie, où il peut être demandé au patient d'effectuer une série de flexions, attitude peu usuelle, mais qui permet d'évaluer de façon pertinente le comportement cardiaque.

Le choix suivant est celui du capteur [6]. Le signal microphonique est le plus souvent utilisé ; il capte les variations de pression acoustique de la voix après qu'elle a été transformée en parole dans le conduit et la qualité du microphone influe directement sur la qualité du signal enregistré. Si on souhaite recueillir un signal plus proche du signal-source, il convient d'utiliser un laryngophone, qui peut être considéré comme un microphone de contact placé au niveau du larynx (qui diminue mais qui n'élimine pas les transformations acoustiques du tractus vocal) ou encore d'autres capteurs comme l'électroglottographe qui, sur la base de la mesure de l'impédance électrique à

travers le larynx évolue proportionnellement à la surface de contact entre les deux cordes vocales pendant la vibration ou encore le photoglottographe qui évolue proportionnellement à la surface de la glotte. Il peut être utile d'utiliser des capteurs aérodynamiques permettant d'analyser la pression aérienne ou le débit buccal et/ou nasal. L'intérêt majeur de ces capteurs physiologiques est, d'une part, leur plus grande sélectivité que le signal acoustique, résultat d'un mélange complexe de phonation et articulation. Le second intérêt est leur capacité à mesurer des phénomènes non audibles, comme par exemple la pression sous-glottique informatrice d'hyper/hypo fonctionnement pneumophonatoire, information difficilement appréhendée à travers le signal acoustique.

Après amplification et filtrage correct, la numérisation du signal permet l'analyse par tel ou tel logiciel. La méthode de numérisation doit attirer l'attention du praticien car il est clair que toute compression du signal entraînera, même si elle est inaudible pour l'oreille, des modifications qui peuvent perturber l'analyse spectrale ou encore la détermination précise de la période. De même, tous phénomènes de dérive temporelle, d'offset, de bruit électronique lié à l'alimentation électrique, souvent présents sur du signal numérisé avec du matériel d'enregistrement grand public, restent imperceptibles auditivement mais rendent très difficiles une analyse correcte du signal [7].

Puis l'éditeur de signal doit permettre de visualiser le signal «brut» et ses transformations ainsi que les calculs effectués ou proposés (voir par exemple la figure 1 montrant plusieurs indices simultanément en fonction du temps sur l'appareillage EVA). Les mesures obtenues doivent correspondre à des concepts d'analyse connus par le praticien et les valeurs obtenues doivent avoir un sens clinique clair : le clinicien devra rester vigilant devant des mesures obscures, peu explicites, ainsi que devant une véritable avalanche d'indices redondants comme sur certains dispositifs, ce qui est particulière-

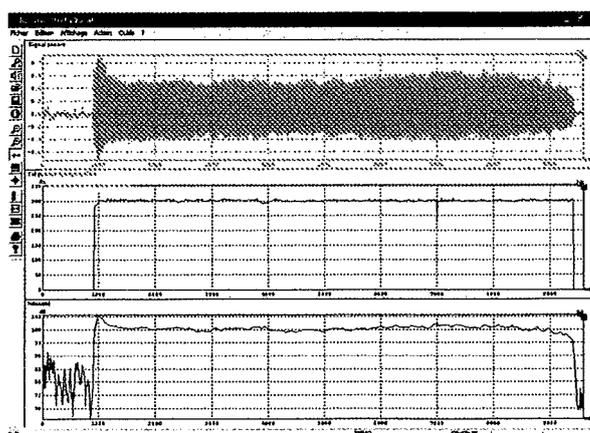


Fig. 1 : Représentation simultanée sur l'écran de l'appareillage EVA des courbes de fréquence instantanée et d'intensité en fonction du temps.

ment retrouvé dans le domaine de l'analyse de la régularité fréquentielle.

Enfin l'interprétation des données obtenues doit faire la part de la variabilité des mesures et on peut dire que globalement, plus la dysphonie est importante, plus le nombre de répétition des mesures doit être important. En pratique, au moins trois mesures doivent être effectuées dans la plupart des cas et la valeur moyenne conservée [8].

La validation des mesures reste, paradoxalement, du domaine des méthodes perceptives car celles-ci restent actuellement la seule évaluation disponible et donc l'unique source de comparaison et validation. Pour l'étalonnage des mesures instrumentales, la plupart des études publiées proposent une évaluation perceptive par un jury. Nous présentons dans le tableau I, les résultats des mesures instrumentales par référence aux résultats de l'analyse perceptive et notamment du critère global G. Le détail des résultats peut être consulté in [2].

L'ANALYSE PERCEPTIVE DES DYSPHONIES

L'analyse perceptive reste donc le standard avec lequel les méthodes d'analyse instrumentales sont évaluées car la voix est un phénomène perceptif par nature aussi bien du point de vue du patient que du point de vue du clinicien. La prééminence de l'analyse perceptive conduit même certains praticiens à considérer que les mesures instrumentales sont d'intérêt modeste au regard de l'écoute. Mais de nombreux arguments scientifiques viennent démontrer que l'analyse perceptive, si elle reste fondée sur le plan pratique, est un instrument d'une fiabilité et d'une pertinence beaucoup plus limitée qu'on ne le croit généralement.

La variabilité des jugements perceptifs est un biais classique et l'essentiel des travaux menés dans ce domaine va dans le sens d'augmenter, directement ou indirectement, la reproductibilité des évaluations [9]. D'une manière générale et quelle que soit la tâche perceptive employée (auditive, visuelle, gustative...) la variabilité s'exprime de deux manières :

1- la variabilité intra-individuelle (un même auditeur évalue le même échantillon différemment d'une fois sur l'autre) et

2- la variabilité inter-individuelle (différents auditeurs donnent une évaluation différente d'un même échantillon). La variabilité intra-individuelle est classiquement plus limitée (environ 72 % de reproductibilité des jugements) que la variabilité inter-individuelle qui dépend de nombreux facteurs liés notamment à la culture et aux préférences individuelles de chaque auditeur, éléments difficiles à maîtriser (environ 68 % de reproductibilité des jugements).

Dans ce contexte, une constatation individuelle est insuffisante pour évaluer une dysphonie. Seul le recrutement d'un jury d'experts permet de limiter les effets de la

variabilité inter-individuelle [9]. Ce jury doit être constitué d'auditeurs spécialistes de la voix, ayant une connaissance homogène des échelles de mesure perceptive et l'habitude d'effectuer la tâche d'évaluation ensemble. Pour des raisons pratiques, l'effectif d'un jury d'experts dépasse difficilement 3 à 7 auditeurs et les rapports dans la littérature de jurys constitués de plusieurs dizaines d'auditeurs sont fortement suspects sur la rigueur de leur recrutement. Par ailleurs, multiplier le nombre d'auditeurs ne permet pas de réduire la variabilité des jugements, mais simplement d'en «lisser» les effets.

Plusieurs facteurs influencent directement la variabilité du jugement perceptif : l'état émotionnel de l'auditeur au moment de l'évaluation, ses valeurs esthétiques, la manière dont il conçoit l'échelle de mesure, etc. La notion de «réfèrent auditif interne» développée par Gerratt et Kreiman [10], fait état d'un prototype mental qui correspondrait à l'idée que l'auditeur se fait de la dysphonie et qui serait instable dans le temps et d'un sujet à l'autre. Dans cette idée, les auteurs proposaient de fournir un exemple fixe et standardisé de voix pathologique (le «standard auditif externe») aux auditeurs, immédiatement avant l'évaluation de chaque échantillon, afin de leur donner une base commune. Ces travaux donnaient des résultats très intéressants pour l'évaluation de voix de synthèse mais ont été abandonnés suite à l'immense difficulté de transposer la méthode aux voix naturelles.

D'autres facteurs, liés à la méthode et à la procédure, influencent la variabilité de manière indirecte. L'échelle mise en place pour l'évaluation de la dysphonie doit tenir compte de deux axes :

1- la qualification, avec le choix des paramètres descriptifs qui doivent être synthétiques et sans ambiguïté car la multiplication des termes utilisés ou la nuance dans leur compréhension («raucité», «érraillement» ou «enrouement» par exemple) peut être responsable de la divergence des jugements ;

2- la quantification qui nécessite un parti pris méthodologique : les échelles de classe (EAI) favorisent la reproductibilité des jugements mais sont relativement grossières, alors que les échelles visuelles analogiques (VAS) permettent une discrimination fine mais induisent une forte variabilité [11]. Au total, l'échelle GRBAS décrite par Hirano en 1981 fait l'objet d'un consensus relatif tant en recherche qu'en clinique, et propose l'évaluation de 5 paramètres qualitatifs (grade global, raucité, souffle, asthénie et forçage) sur une échelle quantitative de 4 classes (0 : normal, 1 : léger, 2 : moyen, 3 : sévère).

La tâche d'évaluation doit être reproduite, au cours de plusieurs sessions d'écoutes (3 le plus souvent), distantes d'au moins une semaine, afin de vérifier et de contrôler la variabilité intra-individuelle. Lors de chaque session, les échantillons vocaux doivent être présentés dans un ordre aléatoire différent afin de limiter l'effet de contexte. En effet, l'ordre de présentation des voix

influence le jugement et un même échantillon sera évalué plus sévèrement s'il apparaît juste après une voix normale que s'il apparaît après une dysphonie avérée. L'ambiguïté de cette démarche réside dans la nécessité de prendre le risque d'augmenter la variabilité intra-individuelle pour être sûr d'obtenir une évaluation représentative de la réalité.

Enfin, l'évaluation doit être réalisée en aveugle, en particulier dans le cas de l'évaluation de l'évolution de la voix après thérapeutique. Au total, les contraintes liées à l'application stricte de ces règles sont donc finalement difficilement compatibles avec la pratique clinique. Mais surtout l'absence de variabilité dans les jugements perceptifs est un idéal dont tout laisse à penser qu'il ne sera jamais atteint. Des lors, il semble licite de remettre en question la légitimité de la perception en tant que standard de référence des mesures instrumentales.

VARIABILITE DE LA PERCEPTION DE LA QUALITE VOCALE

La perception, notamment auditive, peut être considérée, à tort, comme un mécanisme transparent d'enregistrement de la réalité. On peut penser, au premier abord, qu'il existe une correspondance parfaite entre la réalité sonore et ce que l'on en perçoit. On parlerait alors de bi-univocité, c'est-à-dire de relation bijective entre le monde physique sonore et la représentation perceptive que l'on en fait. Cela signifierait qu'à un objet sonore correspond systématiquement une sensation auditive, qu'à une sensation auditive correspond systématiquement un objet sonore, qu'à deux objets sonores distincts correspondent deux sensations auditives distinctes et enfin qu'à deux sensations auditives différentes correspondent deux sons aux caractéristiques distinctes. De nombreux résultats expérimentaux démontrent que ces propriétés ne sont pas respectées.

Prenons l'exemple de la perception de la puissance sonore ("loudness"). Il est bien connu que celle-ci dépend du logarithme de l'amplitude d'un son (notion de décibel). Mais Fletcher et Munson [12] ont montré que cette perception dépend aussi de la fréquence du son (ex: 30 dB à 1000 Hz perçu de façon isosone à 60 dB à 100 Hz). De même, Nishinuma et al [13] ont mis en évidence le rôle de la durée sur la perception de l'intensité des voyelles, un allongement de la durée étant perçu comme une augmentation d'intensité, cette propriété étant d'ailleurs exploitée comme un mode d'accentuation. Au final, une variation de perception de l'intensité peut provenir à la fois d'un changement d'amplitude, de fréquence ou de durée du signal, c'est-à-dire trois primitives physiques indépendantes (1 percept correspond à 3 grandeurs physiques).

La rupture entre le niveau physique et perceptif se retrouve aussi au niveau phonologique. Ainsi, nous pouvons citer le cas des hispanophones incapables de percevoir un /y/ français dans la parole continue [14] ou le cas des langues à tons comme le chinois qui proposent des

distinctions lexicales fondées uniquement sur des variations subtiles de la fréquence fondamentale à l'intérieur de mots possédant exactement les mêmes phonèmes*.

Ces variations tonales à valeur distinctive des mots laissent des auditeurs occidentaux complètement sourds à un tel mécanisme alors que des auditeurs natifs n'ont aucun problème à percevoir ces distinctions. Ces deux exemples de surdité phonologique montrent l'étendue de la variabilité perceptive pour laquelle des caractéristiques physiques peuvent être parfaitement perçues par certains et laisser d'autres individus totalement insensibles.

Nous pouvons aussi mentionner un autre cas de surdité phonologique, celui de l'étude de Dupoux [15] qui montre qu'un japonais soumis à une séquence sonore de type /ebzo/ perçoit ce stimulus comme /ebuzo/. En effet, la langue japonaise possédant une phonotactique de type Consonne Voyelle Consonne Voyelle quasi systématiquement laisse ses auditeurs perturbés lors d'une écoute de type VCCV comme dans /ebzo/. Aussi, le système perceptif d'un japonais va "automatiquement" insérer un phonème vocalique (/u/ en l'occurrence) pour mettre en accord avec son système linguistique ce qui est pour lui un "monstre". On assiste là à l'apparition d'un objet perceptif sans réalité physique.

De la même manière, Meunier et Floccia mettent en évidence l'importance de l'empan temporel lors d'une écoute pour identifier une voyelle [16]. Ainsi, les auteurs montrent que sur un corpus de parole spontanée, l'écoute isolée de voyelles /a/ ne donne que 20 % d'identification correcte. Ce taux atteint 74 % si toute la syllabe est entendue et 92 % dans l'écoute du mot complet. Cela signifie que la relation acoustique/perception peut changer en fonction du zoom temporel utilisé. Enfin, il est important de rappeler des résultats de psycholinguistique relatifs à la cécité et surdité inattentionnelle [17] qui mettent en évidence l'extrême importance du contexte attentionnel dans lequel est plongé un être humain en situation de perception visuelle ou auditive. Il est extrêmement difficile de se placer dans une situation neutre et bien souvent, on finit par entendre ce qu'on veut entendre ou ce qu'on s'attend à entendre et inversement, on peut ne pas entendre ce qui est pourtant évident.

Tous ces exemples de rupture entre la réalité physique et la sensation perceptive doivent nous conduire à une grande prudence quant à l'utilisation d'un procédé perceptif pour étudier précisément un mécanisme de production de la parole, notamment dysfonctionnel comme dans le cas des dysphonies. Pourtant, le processus de communication parlée est un mécanisme robuste qui fonctionne bien, malgré une relation fortement non bi-univoque entre les sons qui nous stimulent et le décodage que l'on opère. Mais ce processus de catégorisation est le

* Ecouter les exemples sonores à l'adresse <http://www.phonetics.ucla.edu/vowels/chapter2/cantonese/si/> ton haut = poème, /si/ ton intermédiaire = essayer, /si/ ton bas = matière, /si/ ton bas montant = temps ...

résultat d'années d'apprentissage de la langue où l'enfant est soumis de façon quasi permanente à cette opération d'association et sa perception catégorielle des phonèmes émerge peu à peu suite à ce processus général d'apprentissage [18]. De plus, les membres d'une même langue subissent globalement le même apprentissage qui autorise alors une bonne reproductibilité et une grande robustesse du processus de décodage de la parole. Malheureusement, la perception de la qualité vocale est loin d'être un phénomène linguistique, contrairement au jugement de l'intelligibilité. L'absence de références partagées entre auditeurs, comme peuvent l'être les phonèmes ou les mots d'une langue, conduit à une inévitable variabilité liée à des référents très personnels et donc très subjectifs. En effet, il s'agit ici d'un auto-apprentissage non supervisé qui conduit chaque individu à des relations associatives acoustique/perception très personnelles. Si l'évaluation perceptive des dysphonies est considérée actuellement comme subjective, ce n'est pas par sa nature perceptive (non subjective dans le décodage d'éléments linguistiques), mais par son procédé actuel déconnecté de référents partagés. Un tel verrou méthodologique a conduit les cliniciens spécialistes de la parole à se tourner donc vers d'autres types d'évaluation comme l'utilisation d'instruments de mesure.

L'APPAREILLAGE EVA ET LES MESURES AERODYNAMIQUES

Le dysfonctionnement vibratoire entraîne dans de nombreux cas une fuite glottique. L'air phonatoire qui n'est pas transformé en son est alors responsable d'un bruit d'écoulement, essentiellement en raison des turbulences aériennes au niveau glottique. Plusieurs approches méthodologiques sont alors envisageables. La première consiste à analyser la voix produite et à mesurer la partie du signal qui correspond au bruit d'écoulement (cf infra méthodes acoustiques). Une autre consiste à mesurer les conséquences cliniques de cet écoulement, en l'occurrence la diminution du temps de phonation. Une autre enfin consiste à directement mesurer le phénomène. Ces différentes méthodes peuvent être mises en perspective avec le mythe de la caverne de Platon : il est possible de choisir de considérer les phénomènes de la vie soit en observant leur reflet sur le mur de la caverne soit en les observant directement [19]

Les méthodes acoustiques

L'analyse spectrale est une méthode instrumentale reposant sur le théorème de Fourier qui stipule que tout son complexe peut être décomposé en une série de sons sinusoidaux (les harmoniques) dont la fréquence est un multiple de la fréquence fondamentale (couramment dénommée F_0). L'interprétation de l'analyse spectrale peut être visuelle sur les graphiques fournis par les dispositifs. Les premières études utilisaient la représentation tri-dimensionnelle classique dite «sonographique» en mode «narrow» en référence à un appareillage déjà ancien (Sonograph® de Kay Elemetrics™) [figure 2]. Le

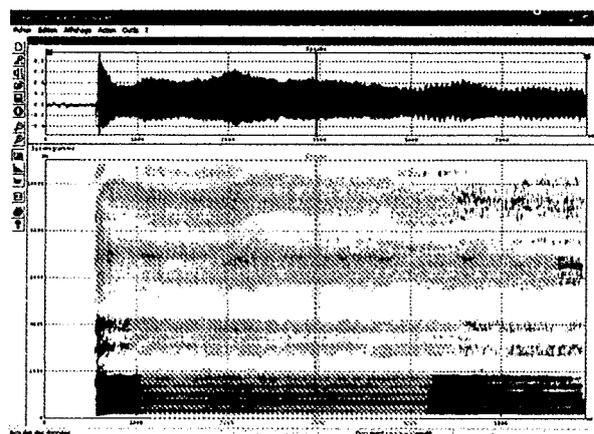


Fig. 2 : Représentation «sonographique» de l'analyse spectrale.

principal intérêt de la méthode est la conservation d'un document graphique relativement aisé à interpréter. Les inconvénients sont cependant nombreux et ont conduit à progressivement abandonner la méthode qui ne survit plus qu'à titre d'illustration : le premier est que la représentation graphique de la voix nécessite toute une technologie et un savoir-faire pour une valeur ajoutée diagnostique modeste, puisqu'il s'agit précisément de représenter sur papier ce que l'auditeur peut entendre. DeJonckere [20] a montré malgré tout l'intérêt de cette visualisation pour "mieux" percevoir auditivement certaines subtilités de la production vocale. Le second est la difficile standardisation de la méthode puisque l'intensité relative de chaque harmonique est très tributaire de l'intensité de la voix, que l'on sait elle-même difficilement reproductible. La plus grande prudence est donc de mise devant cette méthode, malgré ses indéniables qualités visuelles. L'analogie souvent faite avec l'empreinte n'est qu'un leurre : l'analyse spectrale n'est en rien un invariant permettant l'identification d'une personne par sa voix.

La quantification de l'analyse spectrale a fait en revanche l'objet de très nombreux travaux à l'aide d'une des nombreuses variantes (transformée de Fourier, transformée de Hilbert, transformée en ondelettes) [3]. Différentes méthodes se proposent de quantifier 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". Il faut signaler qu'une vibration très instable peut être aussi mise en évidence par cette analyse : l'écart de la F_0 et de ses harmoniques (H_i) par rapport à la valeur moyenne est détectée comme un "bruit de modulation". Il est légitime alors de considérer que le signal de parole est formé de deux composantes : une composante périodique (F_0 & harmoniques) et une composante de bruit (le reste). La somme mathématique de ces deux composantes est considérée comme étant le signal complet. Le Signal Ratio (Sr) [21] est ainsi le quotient $\text{Energie } (F_0 + H_i) / \text{Energie totale}$ alors que

l'énergie Normalisée du Bruit (NNE) [22] se calcule par la formule suivante $NNE = 10 \cdot \log \left[\frac{\text{Energie (bruit)}}{\text{Energie totale}} \right]$. Ces deux paramètres sont fortement redondants. Par contre, il peut être aussi intéressant de calculer ce quotient non pas sur l'ensemble du spectre mais sur sa partie supérieure comme par exemple au dessus de 2 kHz. En effet, il est fréquemment noté que pour des voix dégradées, le nombre d'harmoniques diminuent très vite et on peut assister à l'absence d'harmoniques en hautes fréquences. Nous obtenons ainsi un bon paramètre de discrimination de la sévérité de la dysphonie. Cette diminution du nombre d'harmoniques peut s'expliquer facilement par une perte de dynamique temporelle du vibrateur laryngé (ouverture ou fermeture ralentie de la glotte), qui se traduit par une réduction de la bande passante (moins de fréquences dans le spectre et donc moins d'harmoniques), propriété bien connue en physique sous le terme de dualité temps-fréquence. Cette observation n'est d'ailleurs pas sans rapport avec une méthode proposée par Garrel et al [23] en tussométrie, technique qui consiste à mesurer la dynamique temporelle d'ouverture des cordes vocales lors d'exercice de tout forcée.

Temps Maximum de Phonation (TMP)

En pratique, la détermination du TMP consiste à demander à un patient de produire une voyelle tenue pendant le plus de temps possible. Plusieurs essais sont demandés et le temps maximum est conservé. La méthode la plus simple consiste à utiliser un chronomètre mais il est possible également d'utiliser la mesure du temps par des curseurs sur un éditeur de signal sur un ordinateur.

Il existe des variantes moins utilisées mais qui peuvent être utiles. Le rapport du temps maximum de phonation sur la capacité vitale (mesurée à l'aide d'un spiromètre) donne le quotient phonatoire (qui s'exprime en %) mais rares sont en fait les cabinets médicaux et orthophoniques qui disposent d'un tel spiromètre. La détermination du quotient s/z c'est-à-dire du temps maximum de phonation sur le son s (non voisé) et sur le son z (voisé) donne également une idée indirecte de la fuite glottique.

Les méthodes directes d'étude du débit d'air et de la pression phonatoire

La mesure directe de la quantité d'air traversant les cordes vocales est un moyen conceptuellement simple d'évaluer la fuite glottique même s'il ne s'agit pas stricto sensu d'une mesure de la voix elle-même. Elle permet toutefois d'obtenir des caractéristiques biomécaniques du système pneumophonatoire, autorisant une bonne prédiction du résultat vocal. La méthode utilisée dans l'appareillage EVA est une détermination de la quantité d'air qui traverse un masque en silicone qui est appliqué sur le visage du sujet [figure 3]. Les valeurs mesurées sont données en référence au grade G dans le tableau I.

La mesure directe de la pression est une mesure relativement invasive qui nécessite l'introduction d'un cathéter dans la trachée sous-glottique et son branche-

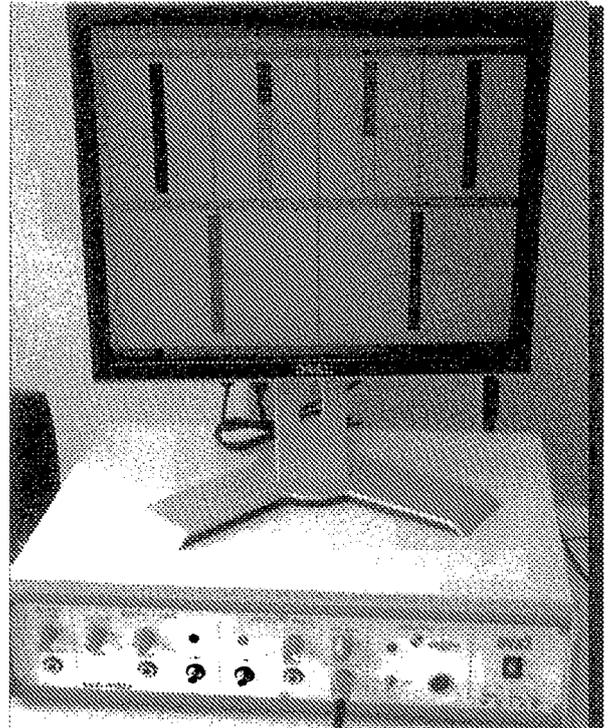


Fig. 3 : Appareillage EVA.

TABLEAU I : Résultats des mesures obtenues [2].

F : Femmes / M : Hommes

| | Grade 0 (F/M) (29/38) | Grade 1 (F/M) (31/28) | Grade 2 (F/M) (14/37) | Grade 3 (F/M) (33/33) |
|--------------------------|--------------------------|--------------------------|--------------------------|--------------------------|
| P_0 (Pa) | | | | |
| F | 218 (38) | 264 (34) | 193 (37) | 203 (37) |
| M | 172 (34) | 131 (37) | 180 (33) | 162 (32) |
| P_1 (dB) | | | | |
| F | 85 (4) | 88 (5) | 89 (5) | 89 (5) |
| M | 92 (5) | 92 (7) | 93 (5) | 89 (9) |
| Time (%) | | | | |
| F | 1.49 (0.13) | 1.60 (0.26) | 1.33 (0.29) | 1.51 (0.28) |
| M | 0.75 (0.16) | 1.03 (0.65) | 2.02 (0.75) | 0.85 (0.30) |
| CL (bits/s) | | | | |
| F | 117 (73) | 130 (76) | 219 (221) | 127 (226) |
| M | 128 (85) | 217 (141) | 323 (251) | 950 (1000) |
| SN R(s) (%) | | | | |
| F | 38.4 (3.8) | 36.7 (4.1) | 32.3 (3.3) | 40.2 (9.2) |
| M | 67.1 (13.1) | 39.2 (19.7) | 38.3 (11.1) | 46.7 (16.4) |
| SN R(z) (dB) | | | | |
| F | 23.1 (7.8) | 15.2 (7.3) | 15.0 (7.2) | 16.1 (6.8) |
| M | 33.7 (9.9) | 15.3 (10.5) | 15.1 (6.5) | 7.0 (3.5) |
| BSX (cm ³ /s) | | | | |
| F | 136 (38) | 168 (89) | 210 (102) | 305 (131) |
| M | 173 (66) | 191 (69) | 323 (163) | 446 (257) |
| PSGR (dbS) | | | | |
| F | 6.3 (3.5) | 7.3 (3.0) | 9.9 (3.7) | 11.2 (4.0) |
| M | 6.7 (3.6) | 6.7 (3.2) | 10.8 (3.5) | 13.5 (4.2) |
| Passive (dB) | | | | |
| F | 410 (277) | 316 (120) | 199 (83) | 168 (81) |
| M | 294 (233) | 215 (136) | 143 (86) | 143 (83) |
| TMP (s) | | | | |
| F | 13.6 (5.0) | 18.5 (4.0) | 7.6 (2.9) | 5.3 (3.6) |
| M | 14.0 (9.8) | 13.4 (8.0) | 6.3 (4.8) | 4.7 (3.7) |

REV LARYNGOL OTOL RHINOL. 2009;130,1:3-10.

ment sur un détecteur de pression tel que disponible dans l'appareillage EVA. C'est pour contourner ce caractère invasif que Smitheran et Hixon ont proposé la "Airway Interrupted method", une méthode d'estimation indirecte par l'intermédiaire de la pression intra-buccale pendant la production d'une série de /papapa/ [24]. Cette méthode très fiable est maintenant répandue et permet une appréciation directe du forçage vocal du sujet. Celle-ci peut être effectuée en conditions usuelles (voix à intensité confortable) ou à la plus faible intensité possible (détermination du seuil phonatoire). Les valeurs mesurées sont données en référence au grade G dans le tableau I.

MESURES MULTIPARAMETRIQUES

La variabilité de chaque mesure et le caractère intrinsèquement multidimensionnel de la voix ont incité de nombreux auteurs à proposer des méthodes multiparamétriques dans lesquelles plusieurs mesures sont combinées. Le choix des paramètres mesurés ne doit cependant pas être laissé au hasard car les dispositifs commerciaux offrent une avalanche de mesures souvent redondantes et qui contribuent encore un peu plus à opacifier le domaine complexe de l'analyse instrumentale. Pour schématiser il existe des logiciels, dont certains en freeware sur internet, qui fournissent des mesures spectrales et temporelles du signal vocal. Il existe par ailleurs des appareillages destinés à des mesures aérodynamiques ou encore des appareillages mixtes destinés à des mesures acoustiques et aérodynamiques simultanées comme l'appareillage EVA que notre équipe a mis au point dans le cadre du CNRS [2]. Cet appareillage est bien sûr beaucoup plus coûteux qu'un simple logiciel du fait de l'industrialisation du «hardware» dont il est constitué.

Sur la base de l'appareillage EVA [tableau I], nous avons élaboré une analyse multiparamétrique avec notamment une étude statistique discriminante destinée à démontrer qu'une combinaison des variables mesurées permet de prédire le jugement perceptif porté par un jury. Nous avons ainsi pu démontrer que chez les hommes comme chez les femmes, cette discrimination est pertinente à plus de 80 % [1, 2]. D'autres tentatives de classement des voix à partir d'un ensemble de mesures notamment acoustiques ont été proposées : signalons notamment le Dysphonia Severity Index [25].

L'introduction de ces méthodes dans la pratique clinique reste encore parcellaire en raison notamment du coût des dispositifs mais aussi, et peut-être surtout en raison des difficultés d'interprétation des résultats, notamment en cas de discordance apparente avec la perception acoustique immédiate de l'auditeur. En réalité, les contraintes méthodologiques ne doivent pas être occultées et aucune machine actuelle ne permet de mesurer totalement automatisée affranchissant le clinicien de toute implication. C'est d'ailleurs le cas dans de nombreuses disciplines de la santé. Insistons notamment sur

les difficultés d'utiliser tel ou tel type de mesure à des fins d'expertise : en pratique, même dans une situation clinique bien connue comme la production vocale avant et après traitement (chirurgical par exemple), il n'est guère possible de considérer les mesures instrumentales comme suffisamment fiables pour avoir valeur de «constante». De ce fait, les mesures instrumentales doivent être incluses dans un protocole de mesures et de méthodes dans lequel la place prééminente reste à l'analyse perceptive et à l'expertise clinique.

La recherche doit cependant être poursuivie car on connaît les limites de l'analyse perceptive dans tous les domaines, qu'il s'agisse de la vision ou de l'audition. On doit se souvenir des bouleversements qu'apportèrent en leur temps les premières «chronophotographies» de Muybridge ou de Marrey. Ces documents sont venus nous démontrer une fois de plus les limites de l'analyse perceptive du mouvement notamment. Depuis, même notre perception visuelle de ces mouvements a été modifiée et il est plus que vraisemblable que les analyses instrumentales nous apporteront dans les prochaines années des renseignements précieux qui enrichiront notre écoute des dysphonies.

Références bibliographiques

1. GIOVANNI A, MOLINES V, NGUYEN N, TESTON B, ROBERT, D, CANNONI M, et al. Une méthode multiparamétrique d'évaluation vocale assistée par ordinateur. ANN OTOLARYNG CHIR CERVICO-FAC. 1992;109:200-206.
2. YU P, GARREL R, NICOLLAS R, OUAKNINE M, GIOVANNI A. Objective voice analysis in dysphonic patients: New data including non linear measurements. FOLIA PHONIATR LOGOP. 2007;59:20-30.
3. BAKEN R, & ORLIKOFF R. Clinical measurement of speech and voice. Singular Publishing, San Diego. 2000.
4. KENT R. The Speech Sciences. Singular Publishing, San Diego. 1997.
5. RÉVIS J, GIOVANNI A, WUYTS F, TRIGLIA JM. Comparison of different voice samples for perceptual analysis. FOLIA PHONIATR LOGOP. 1999;51(3):108-116.
6. TESTON B. L'évaluation instrumentale des dysphonies. Etat actuel et perspectives. In: Giovanni A. Le bilan d'une dysphonie. Solal, Marseille. 2004:107-169.
7. GHIO A. Evaluation acoustique. In: Auzou P, Rolland V, Pinto S, Ozsancak C. Les dysarthries. Solal, Marseille. 2007:236-247.
8. TITZE I. Principles of Voice Production. Prentice Hall, Englewood Cliffs. 1994.
9. 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 Feb;36(1):21-40.
10. GERRATT BR, KREIMAN J, ANTONANZAS-BARROSO N, BERKE GS. Comparing internal and external standards in voice quality judgments. J SPEECH HEAR RES. 1993 Feb;36(1):14-20.
11. REVIS J. Analyse perceptive des dysphonies : approche phonétique de l'évaluation vocale. Thèse de l'Université de la Méditerranée, Marseille 2004.
12. FLETCHER HF & MUNSON WA. Loudness, its definition, measurement and calculation. JOURNAL OF ACOUSTICAL SOCIETY OF AMERICA. 1933;5:82-108.
13. NISHINUMA Y, ESPESSER R, DI CRISTO A. How does vowel duration affect the loudness in a CV syllable? SPEECH COMMUNICATION. 1984;3:39-47.

14. MAGNEN C, BILLIÈRES M, GAILLARD P. Surdit  phonologique et cat gorisation. Perception des voyelles fran aises par les hispanophones. *REVUE PAROLE*. 2005;33:9-30.
15. DUPOUX E, KAKEHI K, HIROSE Y, PALLIER C, MEHLER J. Epenthetic vowels in Japanese: A perceptual illusion? *J. EXP. PSYCHOL. HUM. PERCEPT. PERFORM.* 1999;25:1568-78.
16. MEUNIER C, FLOCCIA C. Syllabe ou mot : quelle unit  permet d'identifier les cat gories phon tiques ? *Actes des 2 mes Journ es d'Etudes Linguistiques, Nantes*. 1999:87-92.
17. SIMONS DJ, CHABRIS CF. Gorillas in our midst: Sustained inattentive blindness for dynamic events. *PERCEPTION*. 1999;28:1059-1074.
18. DAMPER RI, HANARD SR. Neural network models of categorical perception. *PERCEPT. PSYCHOPHYS.* 2000;62(4): 843-867.
19. PLATON. R publique. Livre VII.
20. MARTENS JW, VERSNEL H, DEJONCKERE P. The effect of visible speech in the perceptual rating of pathological voices. *ARCHIVES OF OTOLARYNGOLOGY, HEAD & NECK SURGERY*. 2007;133(2):178-185.
21. HIRAOKA N, KITAZOE Y, UETA H. Harmonic-intensity analysis of normal and hoarse voices. *JOURNAL OF ACOUSTICAL SOCIETY OF AMERICA*. 1984;76(6):1648-1651.
22. KASUYA H, OGAWA S, MASHIMA K, EBIHARA S. Normalized noise energy as an acoustic measure to evaluate pathologic voice. *JOURNAL OF ACOUSTICAL SOCIETY OF AMERICA*. 1986;80(5):1329-1334.
23. GARREL R, AMY DE LA BRET QUE B, GHIO A, GIOVANNI A, GUERRIER B. Int r t de la tussom trie dans les paralysies laryng es unilat rales. *REV LARYNGOL OTOL RHINOL (BORD)*. 2002;123(5):303-6.
24. SMITHERAN J, HIXON T. A clinical method for estimating laryngeal airway resistance during vowel production. *J SPEECH HEAR DISORD*. 1994;1981(46):138-46.
25. WUYTS F, DEBODT M, MOLENBERGHS G, REMACLE M, HEYLEN L, MILLET B, et al. The dysphonia severity index: An objective measure of vocal quality based on a multiparameter approach. *J SPEECH HEAR RES*. 2000;43:796-809.

PRESSE / PRESS

UNE VOIX POUR TOUS - LA VOIX NORMALE ET COMMENT L'OPTIMISER

Genevi ve Heuillet-Martin, H l ne Garson-Bavard et Anne Legr 

2007 - 204 pages - format 21 x 29,7 cm - Prix : 30  

(Tome 1 et 2 pour 57 euros)

ISBN 978-2-35327-023-9

Editions Solal, 111 rue Sainte-C cile, 13005 Marseille, France. Fax: 04 91 80 29 58 - www.editions-solal.fr

Support de la communication orale, outil professionnel, la voix est irrempla able. L'art de la voix, la voix de charme, la voix de la s duction nous troublent. Une voix malade, mutil e, absente, bouleverse la vie d'un individu. La voix qui veut convaincre agit sur autrui.

Les auteurs, phoniatre et orthophonistes, sp cialistes de la voix, font partager leurs connaissances, leur exp rience et leur enthousiasme toujours renouvel s pour servir la cause de la voix. Le premier tome d'«Une Voix pour Tous» pr sente les voix parl es et chant es «normales» et

comment les optimiser, alors que le deuxi me tome traite des pathologies sp cifiques, des anomalies vocales et comment y rem dier.

Ces ouvrages s'adressent aux praticiens, orthophonistes et phoniatres, mais aussi   tous ceux qui font m tier de leur voix ou qui d sirent l'am liorer, enseignants, chanteurs, ainsi qu'  ceux qui rencontrent quelques «pannes vocales». Trouver ou retrouver le plaisir de la voix, en d couvrant ses richesses et en la repla ant dans la globalit  de la personne, am ne un  panouissement de l' tre en chemin vers lui-m me et vers les autres...

Malade et remodel e, banale ou magnifi e par la culture vocale, il existe une voix pour tous !



Genevi ve HEUILLET-MARTIN

M decin ORL, sp cialis e en phoniatre a pendant 35 ann es exerc  en m decine lib rale. Elle  tait attach e de consultation des H pitaux de Marseille et phoniatre de la clinique des Deux Tours, consacr e   la r education ORL. Charg e d'enseignement aupr s des  tudiants en orthophonie, initiatrice des « Carrefours des Voix », responsable du d pistage des troubles de la voix chez les chanteurs du CNIPAL de 1985   1992 et des enseignants, m decin agr e par la DDASS   l'Inspection Acad mique. A pr sent directrice de collections aux  ditions Solal.

H l ne GARSON-BAVARD et Anne LEGR 

Orthophonistes, rattach es   l'Assistance Publique de Marseille ayant travaill  en lib ral, ex-ma tresses de stages   l' cole d'Orthophonie. Animatrices d'ateliers vocaux, dans le cadre de la formation continue aupr s d'enseignants (MAFPEN de l'Acad mie Aix-Marseille), d'orthophonistes (Carrefours des Voix) et de choristes, elles sont des « passionn es » de leur propre voix et celle des autres.