

ARTICLE ORIGINAL

Analyse objective des dysphonies avec l'appareillage EVA[®]. Etat des lieux

Antoine Giovanni¹, Ping Yu², Joana Révis¹, Marie-Dominique Guarella¹, Bernard Teston³, Maurice Ouaknine¹

¹ Laboratoire d'Audio-Phonologie Expérimentale et Clinique de l'Université de la Méditerranée - CHU Timone Marseille

² Service ORL - Hôpital de l'Armée de Libération Populaire, Fu-Xing Lu Beijing, Chine Populaire

³ Laboratoire Parole et Langage - CNRS-Université de Provence, Aix-en-Provence

RÉSUMÉ

Hypothèse : L'évaluation objective des dysphonies a pour objectif d'analyser la qualité de la voix au même titre que l'audiogramme par exemple.

But de l'étude : Les progrès des méthodes instrumentales autorisent à proposer un protocole multiparamétrique incluant des mesures acoustiques et aérodynamiques pour classer les échantillons de voix de manière aussi pertinente qu'un jury d'écoute.

Méthode : La voix de 270 sujets féminins dysphoniques et de 38 sujets contrôle a été analysée de même que la voix de 121 sujets masculins et de 20 sujets contrôle. Elle a été jugée globalement de façon perceptive par un jury d'experts. Le protocole instrumental combinait des mesures acoustiques (Jitter, Coefficient de Lyapunov, Rapport signal/bruit, étendue vocale) et des mesures aérodynamiques (Débit d'air buccal, Pression sous-glottique estimée).

Résultats : A l'aide d'une analyse statistique discriminante, les paramètres objectifs ont permis de prédire à 81% chez les femmes et 84% chez les hommes la qualité de la voix telle qu'elle avait été établie par le jury.

Conclusion : Malgré les difficultés de comparaison entre la parole analysée par le jury et les voyelles tenues sur lesquelles portent les mesures instrumentales, l'accent est mis sur les avancées permises par les mesures aérodynamiques et les mesures non-linéaires.

(Fr ORL - 2006 ; 90 : 183-192)

Mots clés : Analyse instrumentale, Voix, Dysphonies, Jury d'écoute.

Date de soumission : novembre 2005

Date d'acceptation : mai 2006

Auteur correspondant Antoine Giovanni

Laboratoire d'Audio-Phonologie Expérimentale et Clinique de l'Université de la Méditerranée
CHU Timone Marseille - 13385 Marseille cedex 5
e-mail : antoine.giovanni@ap-hm.fr

ABSTRACT

Hypothesis: Objective dysphonia assessment can be used to analyze voice quality, similarly to audiograms for analyzing hearing.

Study objective: Advances in tools for objective voice assessment can be used to develop multiparametric voice-analysis protocols based on acoustic and aerodynamic measurements. Our objective was to evaluate a multiparametric protocol comparatively with perceptual analysis by a panel of listeners.

Methods: Voice samples from 270 women with dysphonia and 38 female controls and from 121 men with dysphonia and 20 male controls were subjected to perceptual analysis by a panel of experts and to objective evaluation using acoustic measurements (Jitter, Lyapunov exponents, Signal-to-noise ratio, and Vocal range) and aerodynamic measurements (Oral air flow, Estimated subglottic pressure).

Results: Discriminant analysis showed concordance between the objective assessment and the perceptual analysis in 81% of women and 84% of men.

Conclusion: Despite difficulties in comparing the continuous speech evaluated by listeners and the sustained vowel sounds used for objective measurements, the use of aerodynamic variables and of nonlinear dynamics has improved the concordance between perceptual and objective evaluations.

Key words : Objective voice assessment, Voice, Dysphonia, Panel of listeners.

Analyse objective des dysphonies

INTRODUCTION

Dans le cadre de la prise en charge des dysphonies, l'étape de l'évaluation a pour objectif de permettre des comparaisons entre plusieurs pathologies, plusieurs patients ou plusieurs approches thérapeutiques. Les moyens de cette évaluation sont à l'heure actuelle nombreux, depuis les scores perceptifs et les questionnaires d'auto-évaluation jusqu'aux méthodes instrumentales objectives [revue in 1].

Le travail que nous présentons a pour but de présenter une méthodologie de l'analyse instrumentale de la voix combinant des mesures acoustiques et aérodynamiques [2]

D'innombrables méthodes de mesure [revue in 3 et 4] et de nombreux protocoles combinant diverses mesures ont été présentés dans la littérature [5-6] dans ce sens et, même s'il n'y a pas de protocole unanimement accepté, il existe un certain nombre de points communs entre la plupart des études :

- la validation des mesures instrumentales proposées repose sur l'analyse perceptive par un jury considérée comme la méthode de référence [5],

- plusieurs mesures différentes doivent être combinées pour répondre d'une part à la variabilité des mesures et d'autre part au caractère multidimensionnel de la voix [6].

La validation des mesures instrumentales par comparaison avec la dimension perceptive reste de toute façon nécessaire puisque cette dimension perceptive intervient prioritairement dans la fonction de communication de la voix [7]: c'est parce que la voix est altérée que les patients viennent consulter, c'est donc sur la qualité de la voix que doivent être jugés les résultats des thérapeutiques même si la variabilité des jugements individuels est un écueil majeur qui a été l'objet d'une abondante littérature [8].

Le caractère spécifiquement multidimensionnel de la voix impose à la plupart des études une analyse simultanée de plusieurs paramètres acoustiques et aérodynamiques car les mesures acoustiques isolées restent insuffisantes. Wolfe avait ainsi étudié la combinaison de 4 mesures acoustiques : la Fo moyenne, le Jitter, le Shimmer et le rapport harmonique sur bruit [5]. A l'aide d'une analyse de régression, il avait été possible de montrer que la combinaison de ces paramètres (surtout Fo et Shimmer) permettait de trouver une corrélation de 0.56 entre les mesures perceptives et les mesures acoustiques, ce qui restait faible, bien en deçà de la possibilité d'une réelle utilisation clinique.

Piccirillo avait présenté une étude combinant des

mesures acoustiques et des mesures aérodynamiques [6] chez 97 sujets dysphoniques et 35 sujets témoins. L'analyse objective faisait appel à la mesure de 14 paramètres : pression sous glottique (PSG), débit d'air buccal (DAB), intensité (faible, moyenne, forte), fréquence (la plus basse, moyenne, la plus élevée), étendue vocale, résistance laryngée, efficacité glottique, temps maximum de phonation (TMP), débit d'air transglottique et électroglottographie. A l'aide d'une régression linéaire, 4 variables particulièrement pertinentes avaient émergé parmi les 14 variables objectives mesurées : la pression sous glottique (PSG), le débit d'air buccal (DAB), le temps maximum de phonation (TMP) et l'étendue vocale. Dans le même esprit, une étude multicentrique menée en Belgique par Wuyts a proposé un index de la sévérité de la dysphonie (Dysphonia Severity Index, DSI) [9]. L'étude a été réalisée sur 319 sujets présentant une pathologie variée et 68 sujets nomophoniques. L'analyse objective a été effectuée sur 13 paramètres acoustiques et aérodynamiques. L'analyse perceptive a été réalisée sur une échelle à 4 niveaux entre 0 pour voix normale et 3 pour voix très dysphonique. A l'aide d'une analyse de régression multivariée, il a mis en évidence 4 paramètres pertinents : fréquence la plus élevée (Fo-High in Hz), intensité la plus basse (I-Low in dB), temps maximum de phonation (TMP), et jitter (%). A partir de ces 4 paramètres, il a établi une équation de régression linéaire : $DSI = 0.13 * MPT (s) + 0.0053 * Fo-High (Hz) - 0.26 * I-Low (dB) - 1.18 * Jitter (%) + 12.4$. Le résultat avait montré que plus la valeur de DSI était négative, plus la dysphonie était importante. A l'inverse, plus la valeur de DSI était positive, plus la dysphonie était légère. Mais, par ailleurs, les mêmes auteurs ont réalisé une analyse de type discriminant et la combinaison de ces 4 mêmes variables n'a permis de retrouver que 56% d'adéquation entre les systèmes d'analyse objective et perceptive, ce qui reste également faible pour une véritable utilisation clinique et éventuellement d'expertise.

Depuis 1990, nous avons entrepris une démarche visant à élaborer un appareillage et une méthodologie qui relèverait le défi de mettre au service du clinicien un outil de mesure et d'expertise de la qualité de la voix qui soit susceptible de « remplacer » au quotidien un jury d'experts, impossible à installer pendant une consultation. Ce travail a abouti d'abord à la présentation de l'appareillage EVA conçu avec le Laboratoire Parole et Langage de l'Université de Provence (UMR CNRS 567) permettant de réaliser simultanément, grâce à une embouche spécifique, des mesures acoustiques et des mesures aérodynamiques [2]. Avec cet appareillage,

Analyse objective des dysphonies

nous avons présenté une analyse objective multiparamétrique des dysphonies chez les hommes [10]. L'étude a été menée sur 63 dysphonies et 21 sujets témoins. L'analyse objective avait été réalisée sur 10 paramètres acoustiques et aérodynamiques : la Fo, l'intensité, le jitter, le coefficient de Lyapunov (CL) que nous avons présenté par ailleurs et qui est un indice de stabilité du signal analogue au Jitter par exemple mais ne nécessitant pas de détection préalable de la fréquence [11], le rapport signal sur bruit ($SRf > 0$ et $SRf > 1\text{kHz}$), le débit d'air buccal (DAB), la pression sous-glottique estimée (PSGE), l'étendue vocale et le temps maximum de phonation (TMP). Nous avons obtenu, à l'aide de l'analyse discriminante, 86% de concordance entre l'analyse perceptive et objective. Ces résultats devenaient encourageants et ces chiffres permettaient d'envisager une véritable utilisation clinique.

MATÉRIEL ET MÉTHODES

Sujets étudiés

Nous présentons ici une étude rétrospective sur des échantillons de voix tirés de la banque d'enregistrement de voix de la Fédération ORL du CHU Timone de Marseille. Nous avons retenu tous les enregistrements effectués sur des sujets adultes et dont la qualité technique (absence de saturation notamment) permettait une analyse perceptive correcte.

Le diagnostic avait été posé par laryngoscopie indirecte, réalisée soit au moyen d'une optique rigide (dans la plupart des cas), soit au moyen d'un fibroscope (dans les cas présentant un réflexe nauséux important) lors de l'examen ORL. L'examen stroboscopique avait été utilisé en cas de difficulté diagnostique. Les pathologies rencontrées correspondaient soit à une anomalie morphologique des cordes vocales (lésions d'origine dysfonctionnelle ou lésions d'origine organique) soit à une anomalie de la mobilité des cordes vocales (paralysie laryngée unilatérale). Les enregistrements ont été effectués avant tout traitement médicamenteux ou chirurgical et avant toute prise en charge rééducative.

Analyse perceptive

Les enregistrements vocaux ont été réalisés dans une pièce insonorisée à l'aide d'un magnétophone numérique (Tascam DA-20, Tokyo, Japon). Les sujets étudiés devaient lire les premières phrases de « La chèvre de Monsieur Seguin » d'Alphonse Daudet. Les consignes étaient données à chaque sujet de manière identique : "Parlez à hauteur et intensité confortable, le

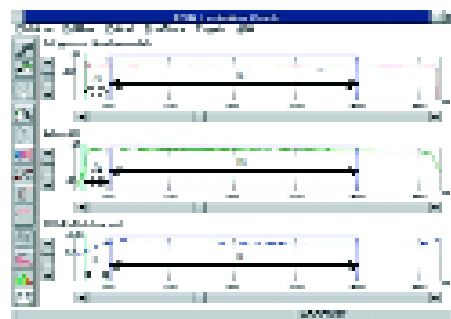
plus naturellement possible". L'analyse de la qualité vocale a été réalisée par le jury d'écoute de façon globale, en utilisant une échelle visuelle analogique de 10 cm de long (0 : voix normale, 10 : dysphonie sévère). Le jury était constitué de 4 auditeurs expérimentés depuis plusieurs années au jugement de la qualité de la voix (2 orthophonistes, 2 médecins phoniatres). Les enregistrements ont été présentés au jury de façon aléatoire et à 3 reprises. Chaque échantillon était donc l'objet de 12 jugements (3×4). Le meilleur jugement et le moins bon étaient ôtés et la moyenne effectuée sur les 10 jugements restants. A partir de cette note moyenne, le grade était établi en utilisant la méthode déjà présentée [12] (le grade 0 voix normale correspondait aux notes de 0 à 1, le grade 1 dysphonie légère aux notes de 1 à 5, le grade 2 dysphonie moyenne aux notes de 5 à 9, le grade 3 dysphonie sévère aux notes de 9 à 10).

Mesures instrumentales

Les mesures instrumentales ont été effectuées avec l'appareillage EVA® (SQ-Lab, Aix-en-Provence, France). L'examineur demandait au sujet d'émettre une voyelle /a/ tenue à une intensité confortable et à une hauteur la plus proche possible de sa hauteur de voix habituelle. Plusieurs essais étaient effectués jusqu'à ce que l'examineur retienne une émission qui lui semblait correspondre à la voix spontanée du sujet, observée pendant l'enregistrement de la voix de lecture. Cette émission était alors enregistrée en tant que fichier informatique.

Figure 1 : Présentation du signal sur l'appareillage EVA.

En haut, courbe de la fréquence en fonction du temps, au milieu, courbe de l'intensité, en bas, courbe du Débit d'Air Buccal. Mise en place de la fenêtre de calcul des principaux paramètres. A : durée de 200ms après le début du signal non analysée ; B fenêtre de 2000ms utilisée pour l'analyse.



Analyse objective des dysphonies

Les données étaient recueillies à partir d'une fenêtre normalisée, d'une durée de deux secondes, placée 200 ms après l'attaque vocale afin d'éliminer les phénomènes liés à celle-ci (Figure 1). Pendant la durée de cette fenêtre les valeurs suivantes étaient mesurées : Fo (en Hz), intensité (en dB), Jitter factor (en %), débit d'air buccal (DAB en cm³/s), rapport signal sur bruit total (SRf>0) et rapport signal sur bruit dont la fréquence est supérieure à 1kHz (SRf>1kHz). Sur la même portion du signal, nous avons effectué la détermination du coefficient de Lyapunov (CL) à l'aide d'un logiciel reproduisant l'algorithme déjà présenté [11] reposant sur l'analyse du portrait de phase du signal. Le CL était exprimé en bits/s. L'ensemble de la procédure était répété trois fois et c'est la valeur moyenne qui a été finalement utilisée.

Nous n'avons pas retenu la mesure du Shimmer, qui semblait peu fiable sur la version 1 de l'appareillage EVA® dont nous disposons. Cette difficulté de mesure a fait par la suite l'objet d'un remaniement logarithmique qui a permis de résoudre le problème et de fiabiliser la mesure du Shimmer sur la version 2 dont nous ne disposons pas au moment de l'étude.

La pression sous-glottique estimée a été mesurée en hectoPascal (hPa), à l'aide du capteur intégré à l'appareillage EVA, par l'intermédiaire du logiciel Phonedit. L'examineur plaçait une sonde dans la bouche du sujet en lui demandant de la maintenir entre les dents sans l'écraser. Le sujet devait émettre une série de 10 /pa/ à hauteur et intensité usuelles à raison de 1 par seconde environ. Les mesures de pression sous glottique estimée (PSGE) étaient relevées à partir de 6 productions de /pa/ en enlevant systématiquement le premier et le dernier. Nous avons calculé pour chaque sujet les moyennes de pression sous-glottique estimée.

Nous avons également demandé à chaque sujet de produire la note la plus grave et la plus aiguë possible. À l'aide de l'appareillage EVA, nous avons déterminé la hauteur en Hz de chaque production. L'étendue vocale rapportée ici correspond à la hauteur de la note la plus aiguë diminuée de la hauteur de la note la plus grave. Puis, nous avons demandé au sujet d'émettre une voyelle le plus longtemps possible à hauteur et intensité confortable. Ce temps était mesuré sur le signal numérisé par l'intermédiaire de curseurs. Nous avons effectué trois essais consécutifs et la valeur la plus élevée est rapportée dans les résultats.

Traitement statistique

Les données ont été analysées à l'aide du logiciel Systat 7.0® for Windows® (Chicago, USA). Nous

avons effectué deux procédures statistiques : l'analyse descriptive et l'analyse comparative.

Dans l'analyse descriptive, nous avons calculé les moyennes et les écart-types (SD) respectivement en fonction des groupes de pathologie et de l'intensité de la dysphonie. Il s'agissait d'évaluer l'influence de la pathologie sur la voix par rapport au groupe témoin. Dans l'analyse comparative, nous voulions savoir si les mesures objectives permettaient de différencier les voix pathologiques des voix normales d'une part, et de déterminer les différentes intensités de dysphonie d'autre part. En raison du caractère non "gaussien" de certaines variables et de la disproportion observée au sein des différentes pathologies, la comparaison des groupes pathologiques et du groupe témoin ainsi que la comparaison entre les valeurs obtenues pour les grades de dysphonie ont été réalisées grâce au test non paramétrique de Mann-Whitney. Le niveau de signification était fixé à $p < 0.05$.

L'analyse discriminante « post hoc » a été réalisée à l'aide d'un algorithme identique à celui qui avait été présenté précédemment [8].

RÉSULTATS

Nous avons analysé 449 enregistrements de voix provenant de 391 patients avec une voix pathologique (270 femmes et 121 hommes) et de 58 sujets contrôle avec une voix normale (38 femmes, 20 hommes). La moyenne d'âge des patients était de 42 ans (19 - 76) et celle des sujets contrôle de 38 ans (20 - 62).

La répartition des pathologies rencontrées est donnée dans le Tableau I. On peut constater que les pathologies rencontrées dans la population féminine sont différentes des pathologies rencontrées dans la population masculine ce qui interdit toute comparaison entre les deux populations. Les mesures concernant les paramètres objectifs sont données dans les Tableaux IIa et IIb. Pour chaque paramètre, la validité était testée par comparaison avec les mesures obtenues dans le groupe contrôle. La comparaison des paramètres mesurés avec l'analyse perceptive est donnée dans le Tableau III. On peut noter que 49 femmes étaient jugées grade 0 (alors que seulement 38 sujets témoins avaient été introduits dans l'étude) et que 25 hommes étaient jugés grade 0 (contre 20 sujets témoins).

L'analyse discriminante rétrospective reposait sur la combinaison de mesures objectives de manière à retrouver un classement des voix le plus proche possible du classement permis par l'analyse perceptive. Nous avons

Analyse objective des dysphonies

Tableau I : Distribution des pathologies rencontrées.

Noter la différence entre les pathologies rencontrées chez les femmes et les hommes par exemple concernant les nodules : 89 chez les femmes soit 29%, 8 chez les hommes soit 5%).

	Femmes	Hommes
Sujets contrôle	38	20
Dysphonie fonctionnelle	15	9
Nodules	89	7
Polypes	50	41
Oedème de Reinke	47	8
Kystes	23	4
Sulcus	10	4
Paralysie Unilatérale	36	29
Leucoplasie	-	19
Total	308	141

Tableau II : Mesures obtenues classées par pathologies.

On note que la plupart des paramètres sont statistiquement différents des valeurs obtenues chez les sujets normaux tandis que les sujets présentant une dysphonie dysfonctionnelle non compliquée présentent quant à eux des différences non significatives.

IIa Femmes

	n	Fo (Hz)	Jitter (%)	Etendue (Hz)	CL (bits/s)	Int (dB)	DAB (cm3/s)	PSGE (hPa)	S/N R (dB)	S/N R>1khz (dB)	TMP (s)
Controles	38	215 (23)	0,51 (0,41)	418 (124)	112 (68)	87,7 (3,3)	136 (56)	6,6 (1,5)	60,6 (10,9)	23,9 (7,5)	13,4 (5,7)
Dysphonie fonctionnelle	15	198 (67) <i>0,313</i>	0,51 (0,10) <i>0,525</i>	405 (142) <i>0,230</i>	157 (94) <i>0,095</i>	86,6 (5,3) <i>0,230</i>	140 (61) <i>0,747</i>	6,8 (1,8) <i>0,960</i>	53,9 (15,7) <i>0,247</i>	18,9 (7,2) <i>0,040</i>	11,9 (3,3) <i>0,565</i>
Nodules	89	207 (27) <i>0,086</i>	0,94 (1,66) <i>0,004</i>	275 (145) <i>0,001</i>	206 (329) <i>0,005</i>	88,8 (5,5) <i>0,353</i>	186 (84) <i>0,001</i>	8,4 (2,4) <i>0,001</i>	54,9 (16,6) <i>0,083</i>	18,4 (8,8) <i>0,001</i>	10,0 (4,0) <i>0,001</i>
Polypes	50	196 (33) <i>0,005</i>	1,15 (1,07) <i>0,001</i>	202 (95) <i>0,001</i>	220 (183) <i>0,001</i>	89,3 (4,1) <i>0,145</i>	213 (109) <i>0,001</i>	12,1 (12,0) <i>0,001</i>	54,2 (14,0) <i>0,032</i>	15,7 (6,1) <i>0,001</i>	7,7 (2,8) <i>0,001</i>
(Edème de Reinke	47	163 (39) <i>0,001</i>	2,77 (4,73) <i>0,001</i>	166 (57) <i>0,001</i>	351 (360) <i>0,001</i>	89,7 (5,6) <i>0,113</i>	236 (106) <i>0,001</i>	10,2 (3,4) <i>0,001</i>	47,6 (17,2) <i>0,001</i>	12,3 (7,2) <i>0,001</i>	7,5 (3,0) <i>0,001</i>
Kystes	23	208 (27) <i>0,382</i>	1,33 (2,388) <i>0,001</i>	251 (102) <i>0,001</i>	278 (387) <i>0,002</i>	89,7 (6,6) <i>0,047</i>	191 (88) <i>0,011</i>	10,4 (3,9) <i>0,001</i>	52,2 (13,6) <i>0,035</i>	15,7 (7,0) <i>0,001</i>	7,3 (2,2) <i>0,001</i>
Sulcus	10	229 (36) <i>0,297</i>	0,82 (0,50) <i>0,002</i>	226 (87) <i>0,001</i>	179 (79) <i>0,006</i>	90,7 (5,4) <i>0,147</i>	216 (28) <i>0,041</i>	10,2 (3,9) <i>0,004</i>	53,6 (12,4) <i>0,094</i>	14,7 (6,3) <i>0,002</i>	9,3 (5,3) <i>0,041</i>
Paralysie unilatérale	36	220 (43) <i>0,758</i>	1,78 (3,27) <i>0,001</i>	204 (94) <i>0,001</i>	334 (421) <i>0,001</i>	87,1 (5,4) <i>0,435</i>	244 (147) <i>0,001</i>	8,7 (2,9) <i>0,001</i>	50,8 (17,7) <i>0,016</i>	15,1 (6,6) <i>0,001</i>	7,0 (5,7) <i>0,001</i>

La valeur moyenne est accompagnée entre parenthèses de l'écart-type et, en italiques, de la valeur du p pour la comparaison entre la valeur du groupe pathologique et la valeur du groupe témoin.

Analyse objective des dysphonies

IIb Hommes											
	n	Fo (Hz)	Jitter (%)	Etendue (Hz)	CL (bits/s)	Int (dB)	DAB (cm3/s)	PSGE (hPa)	S/N R (dB)	S/N R >1kHz (dB)	TMP (s)
Controles	20	124 (27)	0,55 (0,15)	320 (120)	151 (75)	91,8 (5,0)	156 (61)	6,3 (1,7)	62,1 (16, 2)	20,1 (10, 9)	21,7 (7,3)
Dysphonie fonctionnelle	9	130 (21) <u>0,528</u>	0,71 (0,37) <u>0,370</u>	355 (135) <u>0,524</u>	209 (153) <u>0,579</u>	94,1 (7,3) <u>0,164</u>	163 (90) <u>0,850</u>	7,7 (3,0) <u>0,069</u>	45,1 (22,8) <u>0,085</u>	17,5 (12,5) <u>0,505</u>	19,0 (7,0) <u>0,493</u>
Nodules	7	127 (41) <u>0,839</u>	1,6 (1,1) <u>0,049</u>	301 (178) <u>0,700</u>	404 (546) <u>0,297</u>	93,4 (5,1) <u>0,518</u>	206 (87) <u>0,174</u>	10,9 (4,1) <u>0,013</u>	49,4 (22,4) <u>0,205</u>	13,0 (6,1) <u>0,264</u>	14,9 (13,3) <u>0,393</u>
Polypes	41	127 (26) <u>0,510</u>	3,1 (4,7) <u>0,001</u>	138 (92) <u>0,001</u>	387 (427) <u>0,014</u>	93,2 (6,4) <u>0,259</u>	294 (145) <u>0,001</u>	10,4 (3,7) <u>0,001</u>	53,2 (14,1) <u>0,054</u>	14,1 (8,0) <u>0,046</u>	9,7 (5,5) <u>0,001</u>
(Edème de Reinke	8	120 (30) <u>0,740</u>	2,32 (3,12) <u>0,319</u>	125 (76) <u>0,001</u>	271 (113) <u>0,014</u>	94,0 (4,1) <u>0,280</u>	237 (60) <u>0,008</u>	12,4 (5,0) <u>0,002</u>	60,0 (9,1) <u>0,689</u>	18,5 (5,7) <u>0,689</u>	8,0 (4,4) <u>0,001</u>
Kystes	4	133 (37) <u>0,699</u>	1,13 (1,18) <u>0,486</u>	150 (72) <u>0,045</u>	204 (125) <u>0,268</u>	93,3 (6,7) <u>0,727</u>	271 (150) <u>0,075</u>	14,2 (3,2) <u>0,008</u>	56,6 (21,2) <u>0,551</u>	18,7 (12,7) <u>0,670</u>	10,5 (8,4) <u>0,067</u>
Sulcus	4	171 (20) <u>0,028</u>	1,60 (1,74) <u>0,293</u>	141 (115) <u>0,005</u>	323 (122) <u>0,021</u>	93,7 (7,7) <u>0,749</u>	318 (156) <u>0,083</u>	11,9 (4,9) <u>0,018</u>	55,0 (15,8) <u>0,482</u>	13,3 (7,4) <u>0,421</u>	6,9 (1,7) <u>0,006</u>
Paralysie unilatérale	29	151 (33) <u>0,005</u>	6,75 (8,53) <u>0,001</u>	153 (78) <u>0,001</u>	582 (623) <u>0,001</u>	90,6 (5,1) <u>0,555</u>	474 (287) <u>0,001</u>	24,4 (7,5) <u>0,001</u>	53,1 (15,6) <u>0,07</u>	11,5 (7,3) <u>0,004</u>	4,2 (2,7) <u>0,001</u>
Leucoplasic	19	147 (43) <u>0,143</u>	4,97 (3,83) <u>0,001</u>	134 (51) <u>0,001</u>	1002 (767) <u>0,001</u>	87,5 (4,9) <u>0,041</u>	232 (118) <u>0,075</u>	16,5 (9,5) <u>0,001</u>	43,7 (15,0) <u>0,040</u>	7,7 (4,4) <u>0,002</u>	10,6 (5,4) <u>0,005</u>

Tableau III : Résultats bruts de la moyenne des paramètres mesurés chez les hommes H et les femmes F en fonction des différents grades de l'analyse perceptive (écart-type entre parenthèses).

Noter l'importance des écarts-types témoignant de la variabilité des résultats intra-groupes.

		Grade 0 - F/M (49/25)	Grade 1 - F/M (83/26)	Grade 2 - F/M (141/57)	Grade 3 - F/M (35/33)
Fo (Hz)	F	218 (28)	204 (34)	195 (37)	208 (57)
	M	122 (24)	131 (27)	130 (33)	167 (32)
Int (dB)	F	88 (4)	88 (5)	89 (5)	89 (5)
	M	92 (5)	92 (7)	93 (5)	89 (9)
Jitter(%)	F	0,49 (0,13)	0,60 (0,28)	1,13 (1,53)	4,51 (5,54)
	M	0,55 (0,16)	1,02 (0,65)	2,02 (2,75)	10,58 (13,21)
CL (bits/s)	F	117 (74)	133(76)	249 (291)	622 (536)
	M	156 (83)	217 (141)	323 (251)	950 (1000)
SN-R	F	58,4 (13,8)	56,7 (14,2)	52,2 (15,4)	40,6 (19,2)
f>0 (%)	M	62,1 (14,1)	49,2 (19,2)	58,7 (11,1)	40,7 (14,4)
SN-R (%)	F	23,1 (7,8)	19,2 (7,4)	15,0 (7,2)	10,1 (6,4)
f>1kHz	M	21,7 (9,9)	15,5 (10,3)	15,1 (6,5)	7,0 (4,3)
DAB (cm3/s)	F	136 (58)	168 (59)	210 (102)	305 (151)
	M	153 (66)	191 (69)	322 (162)	446 (287)
PSGE (hPa)	F	6,7 (1,5)	7,7 (2,0)	9,9 (3,2)	11,2 (4,0)
	M	6,7 (1,6)	9,7 (3,4)	10,8 (3,3)	14,2 (4,2)
Etendue (Hz)	F	410 (127)	316 (130)	199 (87)	168 (81)
	M	294 (133)	215 (136)	141 (86)	141 (81)
TMP (s)	F	13,6 (5,1)	10,5 (4,0)	7,6 (2,9)	5,3 (3,0)
	M	24,0 (8,8)	13,4 (5,0)	9,3 (4,8)	4,7 (2,7)

Analyse objective des dysphonies

utilisé pour cela l'algorithme non linéaire de discrimination que nous avons déjà présenté [8]. Les paramètres qui sont apparus les plus « utiles » dans la première phase de l'étude pas à pas ont été chez les femmes l'étendue vocale, les coefficients de Lyapunov, la PSGE, le TMP, le DAB, le SRf<1kHz et la Fo. Chez les hommes il s'est agi de l'étendue vocale, du coefficient de Lyapunov, du TMP, de la PSGE, de la Fo et du SNR. Les tables de discrimination utilisant ces variables avec l'algorithme sont présentées dans les Tableaux IVa et IVb. On peut voir que la discrimination est tout à fait satisfaisante avec 81% de fiabilité chez les femmes et 84% chez les hommes. En ce qui concerne le diagnostic différentiel entre voix normale et dysphonie légère, le chiffre est de 117 sujets sur 126 chez les femmes (soit 93%) et de 44 sur 47 chez l'homme (93%). Nous n'avons pas testé cette discrimination en fonction des différentes pathologies.

DISCUSSION

Notre objectif était de valider sur une importante série le protocole de mesures acoustiques et aérodynamiques que nous réalisons avec l'appareillage EVA. Cet objectif est atteint avec 81% de corrélations exactes entre les mesures objectives et perceptives chez les femmes et 84% chez les hommes. La méthode peut donc être proposée comme une méthode fiable d'analyse et d'expertise des dysphonies. Cependant certains points méritent discussion.

Variabilité des jugements perceptifs

Chaque auditeur a sa propre échelle de sévérité des problèmes de voix qui dépend des conditions de son apprentissage (ancrage auditif) ainsi que des conditions de l'écoute (le jugement d'une voix est influencé par exemple par la voix du patient précédent dans une consultation ou par un certain nombre d'autres facteurs). Cette variabilité est d'autant plus importante que les auditeurs sont moins expérimentés dans le domaine de l'analyse perceptive [12]. La mise en place d'un jury d'écoute constitué d'experts permet donc de contourner cette difficulté de la variabilité et représente la modalité de référence perceptive la plus communément admise. Nous avons par ailleurs proposé d'utiliser, à la place du classique jugement de 0 à 3, une échelle visuelle analogique secondairement transformée en échelle de classe de 0 à 3 pour contourner la difficulté habituelle pour les jurés de coter les dysphonies de

Tableau IV : Tables de discrimination entre les Grades (perceptifs) et les groupes (paramètres mesurés).

IVa Femmes (n : 308)

	Groupe 0	Groupe 1	Groupe 2	Groupe 3	% correct
Grade 0 (n : 49)	45	4	0	0	92%
Grade 1 (n : 83)	5	72	6	0	87%
Grade 2 (n : 141)	2	19	104	16	74%
Grade 3 (n : 35)	0	0	6	29	85%
Total	52	95	116	45	81%

IVb Hommes (n : 141)

	Groupe 0	Groupe 1	Groupe 2	Groupe 3	% correct
Grade 0 (n : 23)	22	1	0	0	96%
Grade 1 (n : 26)	2	22	2	0	84%
Grade 2 (n : 57)	0	10	42	5	74%
Grade 3 (n : 33)	0	0	1	32	97%
Total	24	26	22	12	84%

Les sujets bien classés sont en gras sur la diagonale

grades intermédiaires [13]. Sur le plan clinique, cette méthode est cependant extrêmement lourde et il reste nécessaire de chercher à proposer des alternatives instrumentales à l'expertise par un jury. Le but des méthodes instrumentales d'analyse de la voix est donc de contribuer à donner au clinicien une information fournie par un outil préalablement validé par la comparaison avec une analyse perceptive par un jury d'experts.

Différence de matériaux sonores entre analyse objective et analyse perceptive

Même après validation statistique, il reste difficile pour un clinicien de comparer des mesures qui n'ont pas été réalisées sur le même échantillon sonore puisqu'une voyelle tenue est nécessaire pour les mesures objectives tandis que l'analyse perceptive porte le plus sou-

Analyse objective des dysphonies

vent sur un fragment de parole. Cette différence reste conceptuellement difficile à admettre et entache toujours les résultats d'une analyse objective de la voix. En réalité, il apparaît que le cadre théorique de l'analyse perceptive telle qu'elle est réalisée par le jury d'experts reste très vague : il s'agit de juger d'une « qualité » comme la raucité ou le souffle par exemple sur un échantillon de voix. Or la dysphonie n'est pas un phénomène constant affectant toutes les voyelles de façon identique dans un discours ou une simple phrase. Dans le cas d'une paralysie laryngée unilatérale par exemple, l'accolement des cordes vocales et la mise en vibration peuvent être obtenus de façon variable au long de la phrase test et pendant le discours. Des lors, la dysphonie perçue par le jury sera nécessairement différente pour chaque phonème (grade 0 à certains moments, grade 2 sur d'autres phonèmes par exemple) et il existe un processus de décision chez l'auditeur qui lui fait finalement affecter tel ou tel grade de dysphonie. Nous ignorons actuellement presque tout de ce processus qui fait partie du domaine cognitif. Une analyse spécifique de cet aspect de la dysphonie est en cours par Révis [14] mais de nombreuses études restent à développer. Cette étape nous semble indispensable avant de pouvoir envisager une automatisation des méthodes d'analyse du signal visant à reproduire ces processus cognitifs et à enfin pouvoir commencer à réaliser des études objectives sur des phénomènes de parole ou au moins sur des voyelles faisant partie d'un processus de parole et non plus sur des voyelles isolées.

Validité des mesures instrumentales

La mesure instrumentale pour être fiable doit répondre à trois critères : (1) les mesures doivent avoir une signification physiopathologique ; (2) elles doivent être capables de différencier les voix normales des voix pathologiques ; (3) leur valeur doit être logique, c'est à dire manifester une cohérence avec l'intensité de dysphonie [3].

Les résultats obtenus par ce travail confirment la fiabilité des mesures objectives fournies par l'appareillage EVA® : toutes les mesures peuvent être interprétées sur le plan physiologique et sont interprétables d'un point de vue statistique. Nous reviendrons plus loin sur le cas particulier des mesures issues des méthodes non linéaires.

Il est possible de considérer la fiabilité des paramètres objectifs sous l'angle de la variabilité. Cette variabilité augmente avec la sévérité de la dysphonie. Conformément aux données de la littérature [3-4], nos

résultats montrent que cette variabilité peut rendre suspects notamment les résultats obtenus pour les mesures de l'instabilité du signal. Titze a abordé ce problème lors du Workshop on Voice Analysis tenu à Denver en février 1994. Il a proposé de classer les signaux vocaux en trois types qui pouvaient relever de méthodologies d'analyse objective différentes [non publié]. Les signaux de type 1 sont des signaux de type périodique ou quasi périodique (voix normale ou dysphonie légère). Ces signaux ne montrent pas de changements pendant la durée de la fenêtre d'analyse. Dans ces cas, la mesure statistique de la périodicité est une approche satisfaisante. Dans cette logique, seule une valeur de Jitter inférieure à 5% peut être considérée comme significative. Les signaux de type 2, comme dans le cas d'une dysphonie moyenne, contiennent des sub-harmoniques, des modulations et des bifurcations et ne peuvent pas être définis par un F_0 moyenne. Les outils de type graphique comme les spectrogrammes, les portraits de phase ou les courbes de F_0 sont plus adaptés mais leur quantification est difficile. Les signaux de type 3, en cas de dysphonie sévère, sont chaotiques et ne contiennent pas de structure périodique. Les mesures perceptives de la raucité sont les seules méthodes utilisables. Il apparaît cependant difficile de proposer en expertise un telle approche puisque précisément elle consiste à définir d'abord le niveau de dysphonie puis à lui appliquer telle ou telle mesure. Si on compare à un audiogramme, cela reviendrait à mesurer les surdités légères moyennes et sévères avec des méthodes et des audiomètres différents. On comprend que cette solution ne peut pas être retenue dans une perspective clinique et surtout dans une perspective d'expertise. L'analyse multiparamétrique qui permet de combiner les renseignements venant de paramètres différents est certainement à l'heure actuelle le meilleur moyen de contourner ces difficultés. Ici encore, on comprend que le véritable problème n'est pas forcément la méthodologie de mesure instrumentale mais l'adéquation des matériaux phonétiques. En théorie, il serait souhaitable d'effectuer tout l'ensemble des mesures instrumentales sur chaque voyelle, puis d'affecter à l'ensemble des mesures effectuées un algorithme de décision inspiré des processus utilisés par les auditeurs et qui restent à élucider.

En pratique, l'ensemble de nos résultats est cohérent avec les données de la littérature et notamment, Picciello, Wuyts et Dejonckere [6, 9, 15]. Deux des mesures les plus fiables dans toutes ces études semblent finalement être l'étendue vocale (avec différentes variantes) et le temps maximum de phonation qui sont des mesures extrêmement simples à mettre en œuvre y

Analyse objective des dysphonies

compris dans une approche clinique quotidienne. Par ailleurs, toutes ces études mettent l'accent sur de nombreuses incertitudes concernant les mesures acoustiques mais insistent sur la fiabilité des mesures aérodynamiques telles que le DAB ou la PSGE. Inversement, cette étude confirme qu'il n'y a guère de place dans l'état actuel des indices proposés pour une expertise strictement acoustique reposant sur l'analyse de la voix par un logiciel ou par un autre. Il n'est cependant pas impossible que des progrès soient à attendre en particulier du point de vue de l'analyse spectrale et notamment du rapport signal/bruit.

Les mesures issues des mathématiques non linéaire s

L'intérêt des méthodes non linéaires d'analyse du signal dans le domaine de l'analyse des dysphonies a été mis en avant depuis une dizaine d'années notamment par Titze et Herzel [16-17] ainsi que Berhman [18]. Nous y avons-nous même contribué avec la présentation d'un algorithme de calcul des coefficients de Lyapunov [11], algorithme qui est au cœur du travail présenté ici.

L'exposé du cadre théorique et des méthodes « non linéaires » et issues des théories du chaos est en dehors du cadre du travail que nous présentons ici mais l'idée générale qui sous-tend la plupart de ces travaux est que le fonctionnement parfaitement symétrique et synchrone des cordes vocales est un fonctionnement théorique et qu'en pratique, il existe toujours des asymétries, des irrégularités, des « imperfections » des cordes vocales qui tendent à s'exprimer comme autant de degrés de liberté [17-18]. Les mesures non linéaires ont pour objectif de rechercher cette expression, parfois partielle, des degrés de liberté latents sous la forme d'indices nouveaux et dont la signification physiopathologique nous échappe encore un peu : les dimensions fractales d'une part [19] et les coefficients de Lyapunov d'autre part [11]. Ces indices sont approximativement des indices de la stabilité du signal comme pourraient l'être le Jitter et le shimmer par exemple mais ne nécessitent pas de calcul ou de détection préalable de la fréquence dont on connaît précisément la variabilité dans le domaine des dysphonies. Les années qui viennent vont voir de nouveaux concepts émerger à la lumière des théories non linéaires du signal et permettront de valider ces méthodes et de mieux en saisir la signification et la portée clinique.

CONCLUSION

Réaliser un ensemble de mesures objectives permettant d'expertiser une pathologie vocale en complétant la subjectivité du clinicien devant le jugement perceptif de ses propres patients et de ses propres résultats est maintenant accessible.

Notre étude confirme d'abord que les mesures du Temps Maximum de Phonation et d'étendue vocale qui sont des outils phoniatriques classiques sont des mesures extrêmement pertinentes et fiables.

Elle confirme ensuite que les mesures de débit et de pression sont des compléments appréciables même si les appareils qui les autorisent restent coûteux.

Enfin, elle démontre l'intérêt des méthodes d'analyse non linéaire du signal qui seront certainement de plus en plus présentes dans les années qui viennent.

RÉFÉRENCES

1. Guerrier B, Giovanni A, Remacle M. La corde vocale. Editions de la Société Française d'ORL et Chirurgie Cervico-Faciale 2004.
2. Giovanni A, Robert D, Estublier N, Teston B, Zanaret M, Cannoni M. Objective evaluation of dysphonia: Preliminary results of a device allowing simultaneous acoustic and aerodynamic measures. *Folia Phoniatri Logop.* 1996; 48: 175-185.
3. Baken R, Orlikoff R. Clinical measurement of speech and voice. San Diego: Singular Publishing group 2000, pp 603.
4. Kent RD, Ball MJ. Voice Quality Measurement. San Diego: Singular Publishing group 2000, pp 492.
5. Wolfe V, Fitch J, Cornell R. Acoustic prediction of severity in commonly occurring voice problems *J Speech Hear Res.* 1995; 38: 273-279.
6. Piccirillo J, Colin P, Dennis F, Frederickson J. Multivariate analysis of objective vocal function. *Ann Otol Rhinol Laryngol.* 1998; 107: 107-112.
7. Titze I. Principles of voice production. Englewood Cliffs : Prentice Hall 1994.
8. 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: 21-40.
9. Wuyts FL, De Bodt MS, Molenberghs G, 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.

Analyse objective des dysphonies

10. Yu P, Ouakine M, Révis J, Giovanni A. Objective voice analysis for dysphonic patients: a multiparametric protocol including acoustic and aerodynamic measurements. *J Voice* 2001; 15: 529-542.
11. Giovanni A, Ouaknine M, Triglia JM. Determination of the largest Lyapunov exponents of vocal signal. Applications to unilateral laryngeal paralysis. *J Voice* 1999; 13: 341-354.
12. Dejonckere PH, Obbens C, De Moor GM, Wieneke GH. Perceptual evaluation of dysphonia: Reliability and relevance. *Folia Phoniatica* 1993; 1008-1019.
13. Yu P, Revis J, Wuyts FL, Zanaret M, Giovanni A. Correlation of instrumental voice evaluation with perceptual analysis using a modified visual analogic scale. *Folia Phoniatr Logop.* 2002; 54: 271-281.
14. Révis J. Contribution à l'étude de l'analyse perceptive des dysphonies. Thèse de 3^e cycle Université de la Méditerranée 2004
15. Dejonckere P, Bradley P, Clemente P, et al. A basic protocol for functional assessment of voice pathology, especially for investigating the efficacy of (phonosurgical) treatments and evaluating new assessment techniques. *Eur Arch Otorhinolaryngol.* 2001; 258: 77-82.
16. Titze I, Baken R, Herzel H. Evidence of chaos in vocal fold vibration. In Titze I ed *Frontiers in Basic Sciences*. San Diego: Singular Publishing Group 1993. pp143-188.
17. Herzel HP, Berry D, Titze I, Saleh M. Analysis of vocal disorders with methods of non linear dynamics. *J Speech Hear Res.* 1994; 37: 1008-1019.
18. Behman A, Agresti C, Blumstein E, Lee N. Microphone and electroglottographic data from dysphonic patients: type 1,2 and 3 signals. *J Voice* 1998; 12: 249-260.
19. Boek W, Wieneke GH, Dejonckere P. Clinical relevance of the fractal dimension of F_0 perturbations computed by the box counting method. *J Voice* 1997; 11: 437-442.