

# L' EVALUATION OBJECTIVE DE LA DYSPHONIE: UNE METHODE MULTIPARAMETRIQUE.

Antoine GIOVANNI<sup>1</sup>, Valérie MOLINES<sup>1</sup>, Noel NGUYEN<sup>2</sup> et Bernard TESTON<sup>2</sup>.

<sup>1</sup> Centre Hospitalier Universitaire de la Timone, Marseille.  
<sup>2</sup> URA 261 du CNRS, Institut de Phonétique, Aix-en-provence.  
FRANCE.

## ABSTRACT

The aim of this study is to validate an aid for the evaluation of dysphonia with objective measurements.

We recorded exhaled airflow, fundamental frequency and sound level pressure, for a sustained vowel "a", with 51 dysphonic subjects and 15 normal subjects. The following measurements are made on these three parameters: mean value, standard deviation and coefficient of variation. The exhaled airflow volume was also computed for a duration of 2 seconds. A principal components analysis of the measurements indicated that it is possible to recognize the classes of vocal evaluation and vocal pathology. These findings reinforce on objective aid for the vocal evaluation of dysphonia.

## 1. INTRODUCTION

Dans un exposé fort célèbre, M. HIRANO [1] détaille remarquablement les différentes méthodes d'évaluation objectives de la production vocale : "1- to diagnose the etiologic disease, 2- to determine the degree and the extent of the etiologic diseases, 3- to evaluate the degree and the nature of dysphonia, 4- to determine the prognosis and, 5- to monitor changes". Son étude, basée sur plusieurs centaines de cas, décrit statistiquement les méthodes d'évaluation objectives au moyen généralement d'histogrammes, en fonction de différents paramètres tels que: le débit d'air buccal et le volume d'air associés au temps maximal de phonation, la fréquence fondamentale usuelle (Fo), celle de l'intensité acoustique de l'émission vocale (SPL), de la pression sous glottique, et d'autre paramètres moins communs. Cependant, HIRANO n'est pas allé plus loin dans ses investigations statistiques, en essayant par exemple de faire émerger des classes pathologiques ou différents degrés de dysphonie en fonction de plusieurs paramètres. C'est ce que nous cherchons à faire dans la ligne du travail de DEJONCKERE [2] dont nous apprécions la méthode. Notre démarche n'est pas innocente. En effet, le but de cette étude est de valider ou invalider une aide objective à l'éva-

luation vocale pour des applications cliniques et de rééducation dans le domaine de la Phoniâtrie, sans être pour autant aussi optimiste que HIRANO dans son introduction..

## 2. PROTOCOLE EXPERIMENTAL

### 2.1. Le dispositif

Il est schématisé dans la figure 1; L'expérience consiste à mesurer 3 paramètres de l'émission vocale. - Le débit d'air buccal en cc/s. - L'intensité acoustique (spl) en dB. - La fréquence fondamentale (Fo) en Hertz. Le débit d'air buccal est mesuré au moyen d'un aérophonomètre [3] développé dans le cadre d'un contrat de faisabilité de l'INSERM. Il se caractérise par une grande dynamique associée à une bonne linéarité. Le détecteur de mélodie est un fréquencesmètre instantané numérique de grande précision (1 Hz à 1000 Hz), développé pour l'étude des phénomènes microprosodiques [4]. Il est associé à un intensimètre qui mesure le logarithme de la valeur efficace du signal de parole avec une intégration pondérée exponentielle de 10 ms. Sa dynamique est de 80 dB avec une précision de + ou - 1 dB. Le signal de parole est enregistré sur un magnétophone REVOX B77 à partir d'un microphone BEYER B75 placé à 20 cm de la bouche du sujet, nous ne faisons pas de calibration au niveau, les mesures sont donc en dB relatifs (à quelques dB de la référence absolue) mais constantes pour tous les locuteurs. Les signaux de parole, de débit, de mélodie et d'intensité sont acquis simultanément sur un micro ordinateur PC vectra ES12 au moyen d'une carte de conversion DATA TRANSLATION 2801-A. Le signal de parole est échantillonné à 4 kHz, les autres signaux à 1 kHz. Ils sont codés sur 12 bits. Les acquisitions sont gérées par le logiciel PHYSIOLOGIA [5]. L'enregistrement sur l'ordinateur et le magnétophone est télécommandé par le locuteur. Il est demandé à chaque locuteur de réaliser, dans la mesure du possible, 4 enregistrements. 1- La voyelle «A» tenue pendant un temps de phonation de 4 à 5 secondes, à la hauteur et intensité la plus naturelle au sujet. 2- La même voyelle «plus haute». 3- La

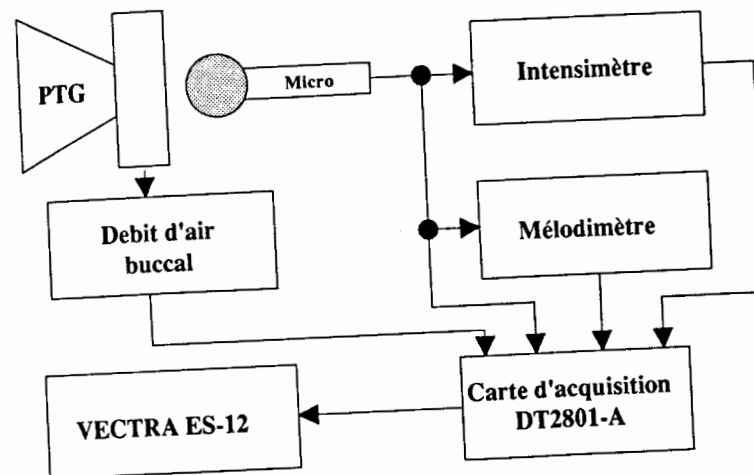


Figure 1: Schéma de principe de la chaîne d'acquisition.

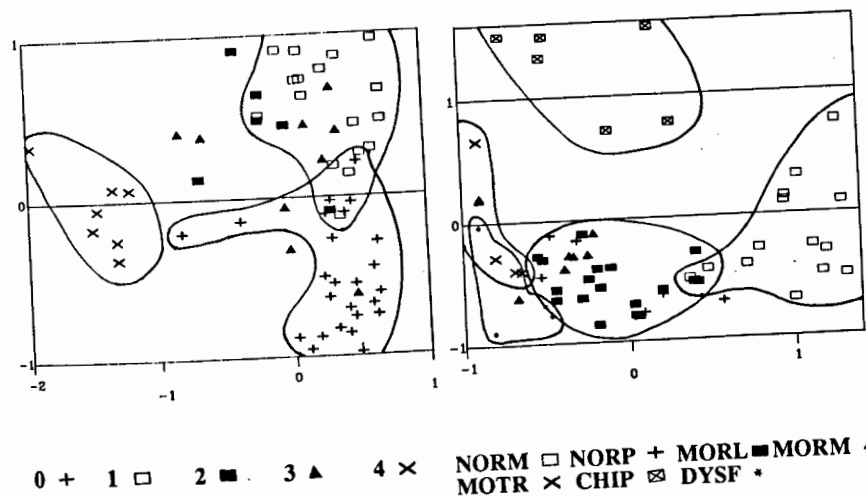


Figure 2: Tableau factoriel des sujets en fonction de l'échelle de qualité de voix

Figure 3: Tableau factoriel des sujets en fonction de l'évaluation pathologique.

même voyelle «encore plus haute». 4- La phrase «elle n'aimait ni maman ni mamie», à la hauteur usuelle.

## 2.2. Le matériel

Matériel pathologique: 50 patients consultants pour dysphonie dans le service ORL du centre hospitalier universitaire de la Timone à Marseille, dont 24 femmes et 26 hommes, de 6 à 62 ans. On distingue deux catégories de patients. D'une part, ceux atteints par une dysphonie jamais explorée auparavant, d'autre part, ceux qui consultent pour un contrôle après traitement (chirurgical ou médical) ou après une rééducation orthophonique, ou les deux associées. Dans ce groupe, se trouve un sous groupe de patients guéris que l'on peut identifier au groupe des normaux.

Pour tester la validité du classement multiparamétrique nous avons décidé de mélanger au matériel pathologique un matériel de normalité, constitué par un échantillon de la population étudiante de l'Université de Provence comprenant 14 sujets, 7 femmes et 7 hommes de 20 à 45 ans, choisi dans l'environnement du laboratoire pour la stabilité de la qualité vocale.

## 3. LA METHODOLOGIE

Pour chaque patient, la dysphonie est évaluée et la pathologie diagnostiquée. L'évaluation acoustique est faite au moyen de 5 facteurs bipolaires déduits de la méthode de HAMMARBERG [6]. Elle aboutit à une échelle de dysphonie de 4 degrés: légère, moyenne, sévère et très sévère. Le degré 0 représente les voix normales. L'évaluation est faite par un phoniatre à partir des enregistrements effectués pendant l'expérience, et de textes lus avant cette dernière. L'examen clinique laryngé est pratiqué par un ORL au moyen d'un laryngoscope rigide avec étude vidéolaryngostroboscopique pour chaque sujet du groupe pathologique. Les déficiences observées sont classées en fonction de la nature physiologique du mécanisme du vibreur laryngien. On distingue ainsi 5 classes de pathologies.

- 1- Altérations morphologiques des cordes vocales:
  - Légère (nodule) codée MORL.
  - Modérée (inflammation) codée MORM.
  - Sévère (Reinke) codée MORS.
- 2- Altération de la motricité codée MOTR.
- 3- Séquelles chirurgicales:
  - Cordectomie codée CHIC.
  - Laryngectomie partielle (Tucker) codée CHIP.
  - Laryngectomie totale codée CHIT.
- 4- Dysfonctionnement avec larynx normal codée DYSF.
- 5- Patients guéris (voix et larynx normaux) codée NORP.

Les sujets normaux sont codés NORM.

Sur chaque paramètre débit, intensité, mélodie, les

mesures suivantes sont effectuées dans une durée normalisée de 2 secondes où les attaques ne sont pas prises en compte: Valeur moyenne, écart type et coefficient de variation, ainsi que le débit intégré pour avoir le volume d'air expiré. Les calculs sont effectués au moyen de l'éditeur de signal de PHY-SIOLOGIA [5].

## 4. ETUDE STATISTIQUE

Les traitements statistiques sont effectués au moyen du logiciel ADDAD [7]. On ne retient de ces mesures que 9 variables sur 10, on supprime la moyenne du débit d'air expiré car elle est parfaitement corrélée avec le volume d'air. On ajoute une variable ordinaire codée sur 5 niveaux pour la qualité de la voix, ainsi qu'une variable qualitative correspondant aux 10 classes pathologiques. Les données représentent un tableau de 174 lignes sur 11 colonnes: 65 productions du «A» usuel, 60 du «plus aigu» et 49 de «l'encore plus aigu». Un premier traitement de statistique élémentaire nous permet d'étudier la répartition des différentes variables sous la forme d'histogrammes.

### 4.1 Classification des sujets en fonction de l'échelle de qualité de voix

Il est difficile de traiter le tableau des données dans sa forme initiale par une analyse factorielle, car les variables descriptives sont hétérogènes (écart type, coefficient de variation, moyenne pour des signaux de nature très différente). L'analyse en composantes principales normée, offre la possibilité d'homogénéiser un tableau en divisant chaque valeur par l'écart type de la variable correspondante. Ainsi, les différentes variables présentent toutes la même variance (=1). Dans notre cas, on hésite cependant à accomplir cette normalisation pour des raisons de nature «sémantique». Comment interpréter un coefficient de variation rapporté à son écart type? La procédure suivie par DEJONCKERE [2] n'est peut être pas à cet égard sans inconvénients. On a donc recodé le tableau sous une forme disjonctive complète. Cela consiste à délimiter sur chaque dimension une dizaine d'intervalles caractérisés par des effectifs égaux. Les intervalles sont considérés comme des classes à l'intérieur desquelles se répartissent les sujets. On a donc transformé une variable quantitative en variable ordinaire. On «écrite» alors cette variable ordinaire, qui comporte les modalités possibles, en une nouvelle variable indicatrice de modalités binaires. Ce nouveau codage permet de rendre les données homogènes et de les soumettre à une analyse des correspondances [8]. Pratiquement, les 9 variables initiales sont découpées en 11 intervalles d'effectifs égaux puis recodées sous forme disjonctive complète. On s'est limité, dans la présente étude, uniquement aux voyelles de hauteur usuelle (n=65). On croise la variable disjonctive de la

qualité de la voix avec les 99 indicatrices de modalité. Ce nouveau tableau est alors soumis à une analyse des correspondances. Les individus n'entrent pas en ligne de compte dans le calcul des facteurs. Cette procédure peut se comparer à une régression multiple au moyen de laquelle la variable «qualité de voix» est expliquée, au sens statistique du terme, par l'ensemble des mesures. Les résultats sont donnés dans le plan factoriel de la figure 2. La répartition des 5 niveaux de qualité de voix est bien différenciée si ce n'est pour les niveaux moyens et sévères qui s'interpénètrent et se répartissent plus largement. Ceci s'explique par le fait que ces niveaux sont plus difficiles à différencier auditivement par un seul auditeur. Les deux exceptions sur la normalité s'expliquent par le fait qu'il s'agit de deux sujets masculins à voix très grave (basse profonde). Le sujet «sévère» perdu au milieu de la normalité est un cas très particulier qui présente une voix au timbre dégradé mais avec une bonne stabilité de la production vocale. Cela montre la nécessité d'une analyse harmonique complémentaire.

### 4.2. Classification des sujets en fonction de l'évaluation pathologique

On cherche par ce traitement à savoir si la variable pathologique recodée sous forme disjonctive, est corrélée avec les variables de mesure. La procédure statistique est la même que précédemment. Cependant, faute d'un nombre suffisant de sujets, nous supprimons les modalités MORS, CHIC, et CHIT. Les résultats sont donnés dans le plan factoriel de la figure 3. Il apparaît que la répartition est encore mieux marquée que pour la qualité de voix. On constate la différence entre les sujets normaux avec les «normaux guéris», qui s'apparentent pathologiquement à de légères lésions morphologiques.

## CONCLUSION

Cette étude fait apparaître qu'il est possible de différencier objectivement au moyen du débit d'air expiré, de l'intensité et de la mélodie, des patients dont l'évaluation du degré de dysphonie et la classification pathologique sont différentes. Nous considérons que ces résultats apportent un argument solide à l'utilisation d'une évaluation vocale assistée par des mesures objectives. Nos résultats ne sont que partiels et portent sur un échantillon de patients assez faible. Cependant, nous les jugeons très encourageants. Le dépouillement continue avec les productions de différentes hauteurs mélodiques. Nous allons ajouter l'analyse harmonique qui nous semble susceptible d'apporter des renseignements complémentaires et approfondir l'importance de l'information contenue dans les différents paramètres.

## BIBLIOGRAPHIE

- [1] HIRANO, M. (1989), "Objective evaluation of the human voice: Clinical aspects.", *Folia Phoniatrica*, Vol 41, 89-144.
- [2] DEJONCKERE, P. H. (1990), "Bruits de turbulence et aperiodicité dans la voix pathologique. Une approche multifactorielle.", *Revue de Laryngologie*, Vol 3, No 4, 353-357.
- [3] TESTON, B. (1983), "A system for the analysis of the aerodynamics parameters of speech.", 11 th ICPS, Utrecht, Aug 1983, Sec 5, 457.
- [4] TESTON, B. et ROSSI, M. (1977), "Un système de détection automatique de la fréquence fondamentale et de l'intensité de la parole.", 8 emes JEP, Groupement des Acousticiens de Langue Française, Aix, 111-117.
- [5] GALINDO, B. et TESTON, B. (1990), "PHY-SIOLOGIA: Un logiciel d'analyse des paramètres physiologique de la parole.", *TIPA*, Vol 13, 197-217.
- [6] HAMMARBERG, B., FRITZELL, B., GAUFFIN, J., SUNDBERG J. and WEDIN, L. (1980), "Perceptual and acoustic correlates of abnormal voice qualities.", *Acta Otolaryngologica*, Vol 90, 441-451.
- [7] LEBART, L., MORINEAU, A. et FENELON, J.P. (1979), "Traitement des données statistiques. Méthodes et programmes", Dunod, Paris, 254 p.
- [8] CAZES, P. (1976), "Régression par boule et par l'analyse des correspondances", *Revue de Statistiques Appliquées*, Vol 24, No 4, 5-22.