

## Paramètres pour l'évaluation acoustique de la nasalité. Etude sur une population d'enfants

F. PLANTE

*Communication and Neuroscience Department, Keele University, Keele, Staffs ST5 5BG, U.K.*

**Abstract :** This work deals with the performance of some parameters for evaluating nasality in children with velar impairment. French vowels /a/, /i/ and /u/ have been analysed. We used four types of parameters. First cepstral distances compute the global spectral difference with a reference spectrum for each vowel. Secondly, we studied the linear prediction error. It represents the difference between a spectrum and an all pole spectrum (model of normal vocal tract). More in relation to the nasality and its spectral modifications, we computed the ratio between frequency bands for detecting nasal formants. We have also studied a cepstral distance among the three vowels (nasal distance). This parameter shows the convergence of spectral characteristics to nasal target. It analyses a global phonatory behaviour. In order to study the performance of the parameters, we used the ROC curves (Receptor Operational Characteristics). These curves are plotted for 57 velar impairments and 206 control subjects. The nasal distance gives best results. The efficacy of spectral parameters or LP error changes depending the vowel. Differences between parameters could be explained with knowledge about the nasality.

### 1. INTRODUCTION

Les divisions palatines qui sont les malformations congénitales les plus fréquentes, sont à l'origine de l'insuffisance vélaire. Cette pathologie affecte le voile du palais. Celui-ci, soit par manque de mobilité, soit par défaut anatomique ne peut plus séparer correctement la cavité nasale du conduit vocal. Il s'ensuit une déperdition nasale plus ou moins importante. L'insuffisance vélaire peut entraîner des troubles importants au niveau de la phonation. D'une part, tous les phonèmes vont être nasalisés. D'autre part, la fuite nasale rend difficile, voire impossible l'augmentation de pression dans le conduit vocal. La formation des occlusives et fricatives va donc être altérée. L'enfant peut atténuer ces troubles par l'utilisation de phénomènes compensatoires (coup de glotte, fricative laryngée.). Mais ceux-ci découlent d'un apprentissage de mauvais gestes articulatoires. Une fois acquis, ces phénomènes sont difficiles à corriger. Il faut donc que l'enfant retrouve une dynamique du voile et évite ou corrige ces défauts phonatoires. C'est le but de la rééducation.

Celle-ci dure au minimum deux ans. Pour garder la motivation de l'enfant, indispensable à la réussite de la thérapie, il faut pouvoir lui montrer de façon objective et assez précise ses progrès. Dans certain cas de voile trop court, la rééducation est insuffisante. Il est nécessaire de recourir à une opération chirurgicale. Celle-ci doit être réalisée le plus tôt possible afin que l'enfant ne "s'installe" pas dans sa pathologie. Il faut

---

<sup>1</sup>Ce travail a été en partie réalisé au laboratoire "Audition et Voix" URA 1447 de Lyon, avec la collaboration du service de Stomatologie de l'hôpital Debrousse de Lyon.

un compromis entre rapidité et nécessité. Des paramètres précis quantifiant la qualité vocale et la dynamique de la rééducation aideraient le chirurgien et l'orthophoniste dans leur diagnostic.

Les spécialistes de cette pathologie ont donc besoin d'indicateurs précis et objectifs. Malheureusement, l'oreille de l'orthophoniste qui est le plus souvent utilisée, n'a pas toujours ces critères. L'utilisation de méthodes paracliniques pour pallier à ces manques se généralise [1].

Actuellement, la méthode la plus utilisée est la mesure des paramètres aérodynamiques [2]. Cependant, même si cette méthode donne des informations très importantes sur l'état du conduit vocal lors de la phonation, elle ne permet pas de juger de la qualité de la parole. Nous nous sommes donc orientés vers l'analyse acoustique du signal de parole. Celle-ci a été peu étudiée dans le cas de cette pathologie. Il est certain que la variabilité inhérente au signal, accrue sous l'effet de la pathologie et de ses multiples facteurs, apparaît comme un frein dans son utilisation. Cependant, les résultats actuellement obtenus en reconnaissance de la parole ou du locuteur montrent la faisabilité d'une telle analyse. D'autre part, les nombreuses études réalisées sur l'analyse de la parole et de ses modifications, sans lien avec la pathologie, nous permettent de mieux définir les paramètres recherchés.

Dans de précédents travaux, nous avons montré la faisabilité de la détection de la nasalité [3]. Or le but de cette étude est de pouvoir évaluer le degré de nasalité. Il faut donc trouver des paramètres dont les variations soient corrélées avec le degré de nasalité. Ceci implique non plus de regarder le comportement des paramètres à un seuil donné, comme pour une détection, mais d'étudier les potentialités de leur variation. Pour cela, nous avons utilisé les courbes COR (Caractéristiques Opérationnelles du Récepteur).

## 2. METHODES ET PARAMETRES

Dans le travail présenté ici, nous étudions seulement les voyelles /a/, /i/, /u/ dites sans contexte. Elles ont été sélectionnées pour leur position extrême dans l'espace vocalique. La plage de maintien des voyelles a été segmentée sur des critères visuels.

### 2.1 Paramètres acoustiques

La mise en parallèle du conduit nasal avec le conduit vocal entraîne des modifications sur le spectre des voyelles. Nous avons sélectionné quatre types de paramètres qui représentent des approches différentes pour quantifier ces perturbations.

#### 2.1.1 Les distances cepstrales

Nous avons d'abord quantifié l'ensemble des modifications spectrales, en utilisant des distances cepstrales pondérées par l'inverse de la variance. Deux types de représentations spectrales ont été choisies. Les coefficients Mel (MFCC) qui modélisent le fonctionnement logarithmique de l'oreille, et les coefficients issus d'une prédiction linéaire (LPCC) qui modélisent la production de la parole.

Les coefficients cepstraux de référence et leur variance ont été calculés pour chaque voyelle par moyennage des coefficients obtenus pour 120 sujets témoins.

La pondération par l'inverse de la variance pénalise les coefficients ayant une forte variation, donc un pouvoir discriminatif plus faible.

#### 2.1.2 L'erreur de prédiction

Du point de vue spectral, la prédiction linéaire peut être vue comme l'ajustement du spectre réel par un spectre tous pôles [4]. L'erreur de prédiction quantifie cet ajustement. Il peut aussi s'interpréter comme l'aplatissement spectral du signal résiduel. Si pour les sujets témoins la modélisation par un filtre tous pôles n'est pas aberrante, dans le cas des insuffisants vélaire elle est moins optimale. En effet, la mise en parallèle de la cavité nasale avec le conduit vocal entraîne l'apparition de pôles et de zéros. L'erreur de prédiction doit donc être plus élevée dans le cas de voyelles nasalisées.

Contrairement aux distances cepstrales, on ne calcule pas la différence par rapport à un spectre de référence, mais par rapport à un spectre tous pôles.

Les approches précédentes, basées sur des différences spectrales, ne prennent pas en compte le lieu des modifications. Or les plus importantes d'entre elles sont maintenant assez bien connues et localisées.

### 2.1.3 Indices spectraux

Le couplage de la cavité nasale entraîne l'apparition d'une paire de formants nasaux aux alentours des deux premiers formants oraux [5] : le premier vers 200-300 Hz et le deuxième entre F1 et F2. Cependant leur impact sur le spectre varie suivant les voyelles et le degré du couplage.

Pour la voyelle /a/, la modification la plus importante est l'apparition du premier formant nasal (F'1). Dans le cas des enfants, F'1 se confond souvent avec le premier harmonique du fondamental.

Pour le /i/ et le /u/, F1 masque l'apparition de F'1. F2 représente donc la modification la plus importante. Cependant elle est surtout visible pour le /i/ ou F1 et F2 sont éloignés. Dans le cas du /u/, il n'y a pas une apparition nette d'un pic. On a plutôt une sorte de palier en basse fréquence.

### 2.1.4 Distance nasale

Dans ces travaux, G. Feng [6] a montré que les caractéristiques spectrales des voyelles nasalisées convergeaient vers une cible nasale. Pour mettre en évidence ce comportement, nous avons calculé la distance entre les coefficients LPCC des trois voyelles. Cette distance doit diminuer avec l'importance de la nasalité. Contrairement aux paramètres précédents, cette distance donne une indication sur un comportement phonatoire global du sujet.

## 2.2 Les courbes COR

Les courbes COR représentent l'évolution du pourcentage de détection en fonction du pourcentage de fausses alarmes, suivant les variations du seuil.

D'un point de vue théorique, le détecteur optimal est celui qui donne un pourcentage de fausses alarmes nul pour un pourcentage de détection de 100%. Ceci n'est jamais réalisé en pratique. Comme critère de performance, on prend l'éloignement de la courbe par rapport à la deuxième bissectrice.

Il y a deux façons d'utiliser ces courbes. La première consiste à choisir les performances du récepteur (% de détection et de fausses alarmes) et de déterminer la valeur du seuil. C'est l'approche que nous avons utilisé précédemment [3]. La deuxième consiste à regarder les performances du récepteur lorsque le seuil varie. C'est cette approche que nous utilisons ici. En effet, un paramètre sera performant pour l'évaluation si la dispersion des populations témoins et pathologiques sont disjointes, c.à.d si la courbe COR s'éloigne de la deuxième bissectrice. Dans notre cas, la partie de la courbe la plus intéressante est celle se situant entre des taux de fausses alarmes de 0 et 0.4.

## 3. RESULTATS ET DISCUSSION

Les figures 1 à 3 représentent les courbes COR pour les trois voyelles. Dans les trois figures nous avons fait figurer la distance nasale qui correspond à une évaluation sur l'ensemble de ces trois voyelles. Les pourcentages de détection et de fausses alarmes sont calculés respectivement sur 57 et 206 sujets.

Pour les trois voyelles, la distance nasale donne la meilleure courbe sur la plage qui nous intéresse. Il semble donc préférable d'analyser conjointement un ensemble de phonèmes, plutôt que chacun d'eux pris séparément. On minimise ainsi la différence inter-sujets. En effet, si l'espace vocalique d'un sujet est translaté par rapport à son groupe de référence, les paramètres pour chacune des voyelles vont être modifiés, alors que la distance nasale est inchangée.

Pour le /a/, l'indice spectral donne de très mauvais résultats. Ceci est dû au fait que pour les enfants F'1 est masqué par le fondamental. Au début, le comportement de l'erreur LPC est similaire aux autres paramètres, puis il se dégrade. La différence spectrale par rapport à un spectre tous pôles n'est importante que pour un nombre restreint d'insuffisants vélares. La distance cepstrale calculée avec les MFCC donne de meilleurs résultats. Les MFCC décrivent mieux le spectre en basses fréquences que les LPCC.

Pour le /i/, l'erreur de prédiction n'est pas du tout performante. Les autres paramètres ont un comportement similaire. L'apparition du deuxième formant nasal est nette (pour preuve le comportement de l'indice spectral). Dans les deux cas (témoins et pathologiques) les spectres s'identifient bien à des spectres AR. L'erreur LPC est dans ce cas inchangée.

Pour le /u/, l'erreur de prédiction donne de bons résultats par rapport aux autres paramètres. Dans ce cas, contrairement au /i/, l'apparition des formants nasaux crée une sorte de plateau en basses

fréquences et non pas des pics nets. Un spectre AR ne peut donc pas modéliser correctement ce spectre. L'erreur LPC est dans ce cas là un paramètre pertinent.

L'utilisation de différents types de paramètres est importante car cela permet, par leur comparaison, de mieux appréhender les modifications mises en jeu.

#### 4. CONCLUSION

La distance nasale semble le meilleur paramètre pour évaluer la nasalité. Les distances cepstrales donnent de bons résultats, mais il est difficile de les interpréter physiquement. L'efficacité de l'erreur LPC ou des indices spectraux dépend de la voyelle étudiée.

Dans cette étude, les paramètres ont été étudiés séparément. Il est important d'analyser leur complémentarité. Cela pose le problème de la logique à mettre en oeuvre pour leur réunion. Les courbes COR peuvent nous aider à mieux définir les zones où les paramètres sont les plus pertinents.

Les courbes COR ne donnent qu'une idée de la potentialité des paramètres pour l'évaluation de la nasalité. Il faut pouvoir les étudier en suivi de rééducation et regarder leur variation en fonction de l'amélioration d'une part de la qualité de la parole et d'autre part de l'anatomie du conduit vocal. Pour cela il est nécessaire de disposer d'une population assez importante sur une période significative du point de vue de la rééducation.

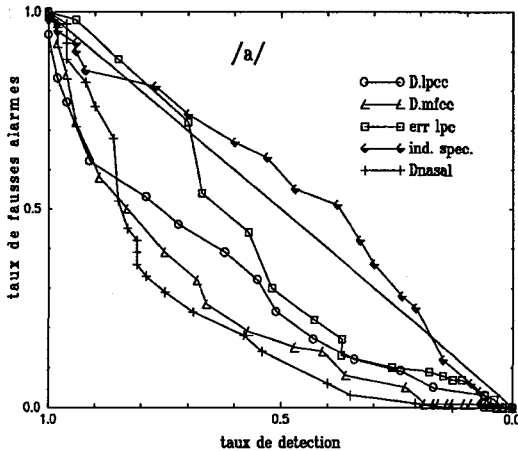


Figure 1 : Courbes COR pour la voyelle /a/.

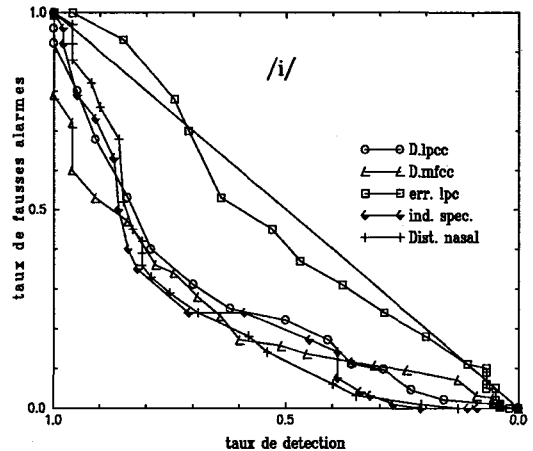


Figure 2 : Courbes COR pour la voyelle /i/.

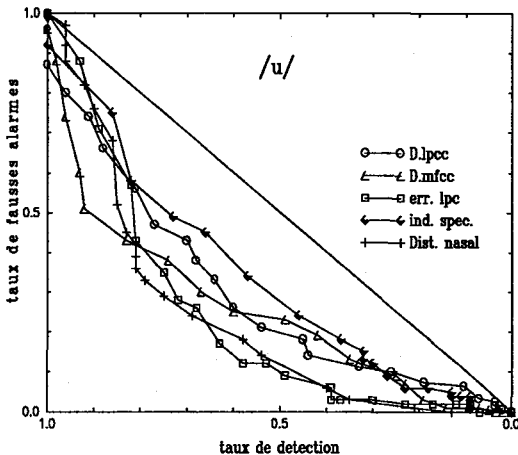


Figure 3 : Courbes COR pour la voyelle /u/.

#### BIBLIOGRAPHIE

- [1] Warren D.W. "The determination of velopharyngeal incompetence by areodynamic and acoustical techniques". *Clinics Plastic Surgery* 2 (1975) 299-304.
- [2] Teston B. "A new air flowmeter design for the investigation of speech production". Eurospeech'93. Berlin 1993 pp.405-408.
- [3] Plante F. "Détection acoustique des pathologies phonatoires chez l'enfant". Thèse INP Grenoble. (1993).
- [4] Makhoul J. "Linear prediction : a tutorial review". *Proc. IEEE* 63 (1975) 561-580.
- [5] Maeda S. "Une paire de pics spectraux comme corrélation acoustique de la nasalisation des voyelles". 13ème J.E.P Bruxelles, (1984) pp.223-224.
- [6] Feng G. "Modélisation acoustique et traitement de la parole. Le cas des voyelles nasales" Thèse INP Grenoble. (1986).