

NOUVELLE MODELISATION PARAMETRIQUE DE L'ACTIVITE  
GLOTTIQUE " PATHOLOGIQUE"

M.F SMIEJ, M. S. RIFFI TEMSAMANI, M. LAMOTTE, M. KARMOUCHE

CRAN URA 821 Université de NANCY I, Faculté des Sciences  
BP 239, 54506 Vandœuvre-lès-Nancy, France

RÉSUMÉ

Pour l'évaluation de la qualité fonctionnelle de la production de la parole dans le but d'aide au diagnostic, nous proposons un nouveau modèle paramétrique prenant en compte la dissymétrie des cordes vocales. Dans ce modèle cinétique, déduit de celui de Titze, les paramètres sont en liaison directe avec la configuration physiologique pré-phonatoire. Partant des équations de mouvements des deux cordes vocales, nous déduisons d'abord les trois signaux : débit glottique, aire glottique, aire de contact. Ces signaux nous permettent d'avoir une idée sur l'activité glottique pendant un cycle. Une méthode d'extraction simultanée des paramètres glottiques et de l'onde glottique est décrite dans la deuxième partie. Nous validons le modèle par une simulation des signaux glottiques, afin d'étudier l'influence des paramètres de symétrie sur la forme de ces signaux .

ABSTRACT

In order to evaluate the speech production quality in connection with diagnosis, we propose a new parametric model which takes into account vocal cords dissymetry. This model is deduced from the Titze's study. The parameters are directly related to the pre-phonatory configuration. From the vocal cords movement equations three signals are derived : glottal flow, area and contact area. These signals give a good idea about glottal activity during a speech period. Then a simultaneous extraction method of glottal parameters and the glottal wave is described. Finally, glottal signals simulation shows the symmetry effect on the signal shape.

1- INTRODUCTION

En 1984 TITZE propose un modèle cinétique pour l'amélioration de la qualité de la synthèse de la parole; ce modèle ne permet pas d'évaluer un geste phonatoire défectueux. C'est dans ce sens que nous avons introduit de nouveaux paramètres de symétrie pour individualiser le mouvement de chaque corde vocale et ainsi pouvoir décrire certaines pathologies laryngiennes.

2- MODELES DES MOUVEMENTS ASYMETRIQUES  
DES CORDES VOCALES

2-1 Equations de mouvements

A partir de la configuration pré-phonatoire des deux cordes vocales, nous déduisons les équations de mouvements et les différents paramètres glottiques. Nous supposons qu'en phase pré-phonatoire la forme verticale des cordes vocales est trapézoïdale comme l'indique la figure suivante:

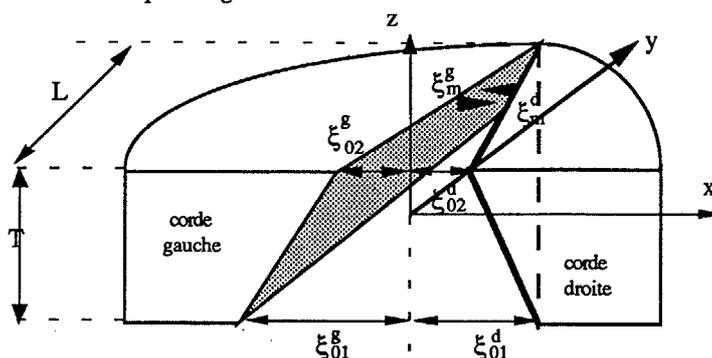


Figure 1 : Configuration pré-phonatoire des deux cordes vocales

Le plan (yoz) est le plan médian, l'axe oy a la direction postéro-antérieure et l'axe des x est parallèle à la direction latérale. La longueur L et la hauteur T est la même pour les deux cordes vocales (aucune dissymétrie de dimensions des cordes vocales). Les indices d et g représentent respectivement la corde vocale droite et la corde vocale gauche (fig.1). Sous ces conditions, les positions pré-phonatoires verticales de chaque corde sont déterminées par les équations suivantes :

$$\xi_0^d(y,z) = \left\{ \xi_{01}^d - \frac{\xi_{02}^d - \xi_{01}^d}{T} z \right\} \left\{ 1 - \frac{y}{L} \right\} \quad (2-1)$$

$$\xi_0^g(y,z) = \left\{ \xi_{01}^g - \frac{\xi_{02}^g - \xi_{01}^g}{T} z \right\} \left\{ 1 - \frac{y}{L} \right\} \quad (2-2)$$

Notons que l'axe (ox) pour chacun des déplacements ou des positions, sera toujours pris dans le sens plan médian vers la corde considérée. Cette convention nous évite des déplacements de signes contraires pour chacune des cordes et nous permet la déduction des signaux glottiques, notamment l'aire glottique.

Les déplacements relatifs des cordes vocales par rapport à leur position initiale sont supposés sinusoïdaux dans le temps et dans l'espace :

$$\xi_1^d(y,z,t) = \xi_m^d \sin\left(\frac{\pi y}{L}\right) \sin\left(2\pi ft - \frac{\phi^d z}{T}\right) \quad (2-3)$$

$$\xi_1^g(y,z,t) = \xi_m^g \sin\left(\frac{\pi y}{L}\right) \sin\left(2\pi ft - \frac{\phi^g z}{T}\right) \quad (2-4)$$

où  $\phi^d$  et  $\phi^g$  sont les déphasages entre les bords supérieurs et inférieurs de chaque corde.



Les équations de mouvements des deux cordes vocales par rapport au plan médian sont donc :

$$\xi^d(y,z,t) = \xi_m^d \left\{ \left[ \frac{\xi_{01}^d}{\xi_m^d} - \frac{\xi_{01}^d - \xi_{02}^d}{T \xi_m^d} z \right] \left[ 1 - \frac{y}{L} \right] + \sin\left(\frac{\pi y}{L}\right) \sin\left(2\pi ft - \frac{\phi^d z}{T}\right) \right\} \quad (2-5)$$

$$\xi^g(y,z,t) = \xi_m^g \left\{ \left[ \frac{\xi_{01}^g}{\xi_m^g} - \frac{\xi_{01}^g - \xi_{02}^g}{T \xi_m^g} z \right] \left[ 1 - \frac{y}{L} \right] + \sin\left(\frac{\pi y}{L}\right) \sin\left(2\pi ft - \frac{\phi^g z}{T}\right) \right\} \quad (2-6)$$

## 2-2 Définition des paramètres

Nous définissons d'abord les différents paramètres pour chacune des cordes, à savoir : quotient d'abduction droite et d'abduction gauche

$$Q_a^d = \frac{\xi_{02}^d}{\xi_m^d}, \quad Q_a^g = \frac{\xi_{02}^g}{\xi_m^g} \quad (2-7 \text{ et } 8)$$

quotient de forme droite et de forme gauche :

$$Q_s^d = \frac{\xi_{01}^d - \xi_{02}^d}{\xi_m^d}, \quad Q_s^g = \frac{\xi_{01}^g - \xi_{02}^g}{\xi_m^g} \quad (2-9 \text{ et } 10)$$

et enfin, quotient de phase droite et de phase gauche :

$$Q_p^d = \frac{\phi^d}{2\pi}, \quad Q_p^g = \frac{\phi^g}{2\pi} \quad (2-11 \text{ et } 12)$$

Ainsi, nous disposons de six paramètres sans dimension et de deux paramètres supplémentaires correspondant à l'amplitude de vibration de chaque corde.

Nous introduisons également des paramètres sans dimension caractérisant la symétrie des cordes vocales du point de vue de l'amplitude de vibration, de la forme et de la phase. Nous donnons les définitions de ces paramètres

$$\text{- Quotient de symétrie d'amplitude : } Q_{sya} = \frac{\xi_m^d}{\xi_m^g} \quad (2-13)$$

$$\text{- Quotient de symétrie de la partie inférieure : } Q_{sinf} = \frac{\xi_{01}^d}{\xi_{01}^g} \quad (2-14)$$

$$\text{- Quotient de symétrie de la partie supérieure : } Q_{syup} = \frac{\xi_{02}^d}{\xi_{02}^g} \quad (2-15)$$

$$\text{- Quotient de symétrie de phase : } Q_{syp} = \frac{\phi^d}{\phi^g} \quad (2-16)$$

Ainsi, l'une des deux cordes (la corde gauche par exemple) étant prise comme référence pour les deux paramètres  $Q_a$  et  $Q_s$ , les mêmes paramètres de l'autre corde s'expriment en fonction des quotients de symétrie de la manière suivante

$$Q_a = Q_a \frac{Q_{syup}^g}{Q_{sya}^g}, \quad Q_s = \frac{Q_{sinf}^g - Q_{syup}^g}{Q_{sya}^g} Q_a + \frac{Q_{sinf}^g}{Q_{sya}^g} Q_s \quad (2-17 \text{ et } 18)$$

ou encore :

$$Q_{sinf} = \frac{Q_{sya}^g Q_s + Q_{syup}^g Q_a}{Q_a + Q_s}, \quad Q_p = Q_{syp} \frac{Q_p^g}{Q_p^d} \quad (2-19 \text{ et } 20)$$

Des dix paramètres définis ci-dessus, nous avons choisi les sept paramètres

- le quotient d'abduction gauche,

- les deux quotients de forme droite et gauche,

- le quotient de phase gauche,

- le quotient de symétrie d'amplitude,

- le quotient de symétrie d'abduction (symétrie de la partie supérieure),

- le quotient de symétrie de phase,

les autres se déduisent des relations (2-18) à (2-20).

Pour simplifier les écritures nous ne précisons pas l'indice gauche et droite sauf pour les deux quotients de forme. Récrivons maintenant les équations de mouvements en fonction des paramètres choisis

$$\xi^g(y,z,t) = \xi_m^g \left\{ \left[ Q_a^g + Q_s^g - Q_s^g \frac{z}{T} \right] \left[ 1 - \frac{y}{L} \right] + \sin\left(\frac{\pi y}{L}\right) \sin 2\pi \left( ft - \frac{Q_p^g z}{T} \right) \right\} \quad (2-21)$$

$$\xi^d(y,z,t) = \xi_m^d \left\{ \left[ Q_a^d + Q_s^d - Q_s^d \frac{z}{T} \right] \left[ 1 - \frac{y}{L} \right] + \sin\left(\frac{\pi y}{L}\right) \sin 2\pi \left( ft - \frac{Q_p^d z}{T} \right) \right\} \quad (2-22)$$

$$\text{Avec : } Q_a' = \frac{Q_{syup}}{Q_{sya}} \text{ et } Q_p' = \frac{Q_{syup} Q_p}{T}$$

## 2-3 Déduction des signaux glottiques

Ce paragraphe consiste à déterminer les trois signaux glottiques discrets à l'instant n. Dans le cas d'une activité glottique " anormale", ces signaux peuvent être affectés par la dissymétrie, comme dans le cas d'une paralysie unilatérale (l'amplitude de vibration de l'une des deux cordes est forcément modifiée) ou bien dans le cas d'une atteinte nerveuse unilatérale " hémiplegie" (le paramètre de symétrie de phase est affecté) ou encore dans le cas d'une atteinte lésionnelle localisée sur la partie supérieure d'une corde (le paramètre de symétrie d'abduction est modifié).

La fonction d'aire glottique, dérivée de celle de Titze, est de la forme:

$$a_n = \min_k \left( \frac{L}{M} \sum_{j=1}^M \xi_{j,k,n}^d + \xi_{j,k,n}^g \right) \quad (2-23)$$

où M est le nombre d'éléments discrets pris sur la longueur L.

L'aire de contact est la somme des aires partielles  $\Delta y \Delta z = (L/M)(T/N)$  pour lesquelles la différence  $\xi_{j,k,n}^d - \xi_{j,k,n}^g$  est négative :

$$c_n = \sum_{j=1}^M \sum_{k=1}^N \Delta c_{j,k,n} \quad (2-24)$$

Le débit glottique est solution de l'équation du second degré :

$$u = \frac{ac}{k} \left\{ - \frac{a}{A^*} \pm \left[ \left( \frac{a}{A^*} \right)^2 + \frac{4k}{c^2 \rho} \left( P_1^+ - P_2^- \right) \right]^{1/2} \right\} \quad (2-25)$$

où  $A^*$  est l'aire effective combinant les aires sous-glottique et sus-glottique, définie par  $\frac{1}{A^*} = \frac{1}{A_1} + \frac{1}{A_2}$ ,  $p_1$  et  $p_2$  sont les pressions incidentes sous et sus-glottiques.

## 2-4 Résultats expérimentaux

Nous avons procédé à des simulations de signaux glottiques, pour estimer le bien-fondé de notre choix des paramètres de symétrie d'une part et pour étudier leurs influences mutuelles. Nous décrivons dans ce paragraphe les différentes conditions de la simulation et nous commentons, à la fin, quelques résultats obtenus.

Pour l'obtention des signaux glottiques, en particulier le débit glottique, nous avons besoin de simuler des conditions de charge de la glotte, et par conséquent nous sommes amenés à simuler le système de propagation de l'onde acoustique dans les tubes sous-

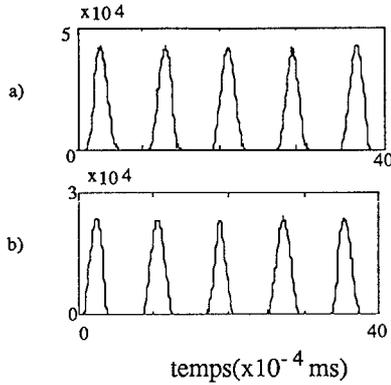
glottiques d'une part, et le conduit vocal d'autre part. Pour l'appareil sous-glottique, nous avons pris quatre tubes acoustiques uniformes dont la section est connue. Pour le conduit vocal, nous sommes servis d'une fonction d'aire correspondant à une voyelle donnée par Mrayati [Mrayati, 1976]. Le système complet débouche sur une pression constante au niveau des lèvres  $P_L$ .

Les conditions de la simulation sont les suivantes :

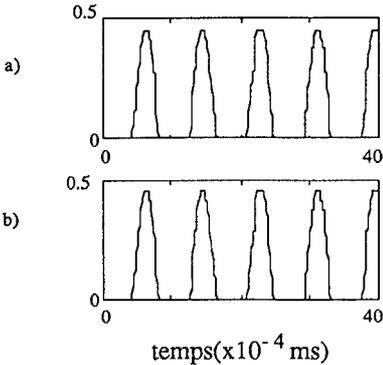
- fréquence d'échantillonnage : 10 kHz,
- fréquence fondamentale : 120 Hz, correspondant à une voix masculine moyenne,
- pression au niveau des poumons : 10 cm H<sub>2</sub>O ou 2,5 cm H<sub>2</sub>O, correspondant à une voix d'intensité respectivement forte et faible,

- amplitude de vibration des cordes vocales : 0,075 ou 0,3 cm. Dans un premier temps, nous avons généré les trois signaux glottiques, le débit, l'aire d'ouverture et l'aire de contact, pour plusieurs valeurs des paramètres pour comparer d'une façon visuelle les deux configurations des cordes vocales : symétrique et asymétrique. Un exemple est donné à la figure 2.

**Débit glottique**



**Aire glottique**



**Aire de contact**

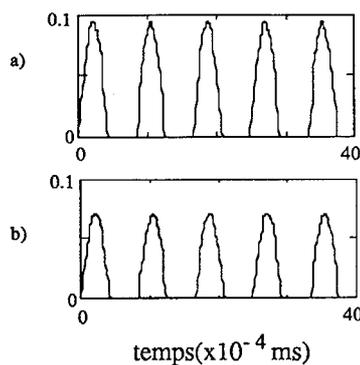


Figure 2 : Signaux glottiques simulés

$\xi_m = 0.075$  cm ;  $PL = 10$  cm H<sub>2</sub>O;  $Q_a = 0$ ;  $Q_s^d = 3$ ;  $Q_s^g = 3$ ;  $Q_p = 0$   
 a) cas d'une configuration symétrique, b) dissymétrie d'amplitude :  
 $Q_{sya} = 0.5$  ( $Q_{sys} = 1, Q_{syg} = 1$ )

Après examen des résultats de la simulation avec différentes valeurs de quotients, nous pouvons tirer les premières conclusions concernant l'introduction des paramètres de symétrie :

- pour le signal débit glottique, la forme de l'onde reste la même dans les configurations symétriques ou asymétriques; seule l'amplitude du signal diminue dans le cas asymétrique,
- pour l'aire glottique, la même constatation peut être faite,
- seule la forme de la fonction aire de contact est affectée.

**3 - EXTRACTION DES PARAMETRES ET DE L'ONDE GLOTTIQUE**

Dans cette partie, nous décrivons une méthode qui sert à la fois à effectuer le filtrage inverse et à extraire les paramètres de l'activité glottique à partir du signal de parole.

Le principe de la méthode est représenté à la figure 3, où  $s(k)$  est le signal de parole,  $u_g(k)$  la sortie du modèle paramétrique,  $1/A(z)$  le modèle tous pôles du conduit vocal et  $L(z)$  la fonction de transfert du rayonnement par les lèvres.

Le critère à minimiser est celui des moindres carrés pondérés

$$V_N(A,P) = \sum_{k=ki}^{ki+N} \lambda_k e_w^2(k) \tag{3-1}$$

où  $e_w(k)$  est l'erreur de prédiction définie par

$$E_w(z) = \left[ S(z) - \frac{L(z) U_g(z)}{A(z)} \right] W(z) \tag{3-2}$$

avec  $W(z)$  un filtre d'ajustement spectral. Le critère est non linéaire par rapport au vecteur  $A$  de  $A(z)$ . Un choix judicieux de la fonction de pondération spectrale  $W(z) = P(z)A(z)$  rend le critère linéaire par rapport au vecteur  $A$ , où  $P(z)$  est un filtre de pré-emphase du signal de parole  $S(z)$

$$E_w(z) = S(z)P(z) - L(z)U_g(z)P(z) \tag{3-3}$$

Pour simplifier les écritures, nous posons

$$S_f(z) = P(z)S(z) \text{ et } U_f(z) = L(z)P(z)U_g(z) \tag{3-4-5}$$

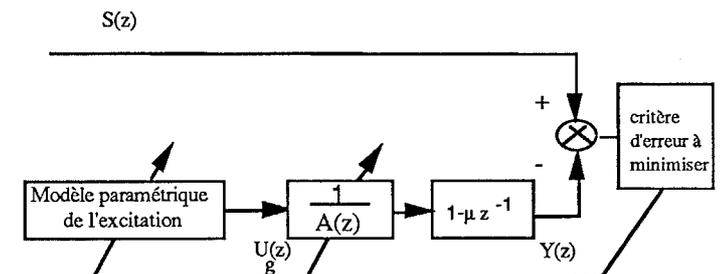


Figure 3 : Schéma-bloc de la détermination simultanée des paramètres de l'activité glottique et du filtre d'analyse  $A(z)$

L'expression de l'erreur de prédiction dans le domaine temporel est

$$e_w(k) = s_f(k) + \sum_{i=1}^p a_i s_f(k-i) - u_f(k) \tag{3-6}$$

En termes vectoriels, si on pose

$$S_N = (s_f(1), \dots, s_f(N))^T \text{ et } U_N(P) = (u_f(1,P), \dots, u_f(N,P))^T$$

où  $P$  est le vecteur paramètres de l'activité glottique, et la matrice  $\Phi_N$  définie par



$$\phi_N = \begin{pmatrix} -s_f(0) & \dots & -s_f(1-p) \\ \vdots & \ddots & \vdots \\ -s_f(N-1) & \dots & -s_f(N-p) \end{pmatrix}$$

le critère défini en (3-1) peut s'écrire

$$V_N(A,P) = (\tilde{S}_N(P) - \phi_N A)^T \Lambda (\tilde{S}_N(P) - \phi_N A) \quad (3-7)$$

$$\text{où } \tilde{S}_N(P) = S_N - U_N(P) \quad (3-8)$$

et  $\Lambda$  une matrice diagonale contenant les facteurs de pondération.

Après avoir écrit le critère sous la forme matricielle, nous décrivons la méthode d'extraction simultanée des vecteurs optimaux  $\hat{P}$  du modèle de la glotte et  $\hat{A}$  du filtre tous pôles  $1/A(z)$ .

Pour un vecteur  $P$  fixé, la méthode des moindres carrés donne comme vecteur optimal  $\hat{A}$

$$\hat{A} = \phi_N^+ \tilde{S}_N(P) \quad (3-9)$$

où  $\phi_N^+ = (\phi_N \phi_N^T)^{-1}$  est la matrice pseudo-inverse de  $\phi_N$ .

Un moyen efficace de calcul numérique de  $\hat{A}$  est la factorisation QR de la matrice  $\phi_N$ , où  $R$  est une matrice triangulaire supérieure et  $Q$  une matrice orthogonale vérifiant  $Q^{-1} = Q^T$ . En utilisant cette factorisation, le critère à minimiser (en prenant  $P(z)=1$  et  $\Lambda = I$ ) devient

$$V_N(A,P) = \|\tilde{S}_N(P) - QRA\|^2 = \|Q^T \tilde{S}_N(P) - RA\|^2 \quad (3-10)$$

La matrice  $R$  étant obtenue par triangularisation de  $\phi_N$ , elle est de la forme

$$R = \begin{bmatrix} R_1 \\ 0 \end{bmatrix} \quad (3-11)$$

où  $R_1$  est une matrice carrée (dimension  $p \times p$ ) triangulaire supérieure.

$$\text{En posant } h(P) = Q^T \tilde{S}_N(P) = \begin{bmatrix} h_1(P) \\ h_2(P) \end{bmatrix} \quad (3-12)$$

où  $h_1(P)$  et  $h_2(P)$  sont deux vecteurs colonnes de dimension respective  $p$  et  $N-p$ , le critère à minimiser devient alors

$$V_N(A,P) = \|h_1(P) - R_1 A\|^2 + \|h_2(P)\|^2 \quad (3-13)$$

Le vecteur  $A$  minimisant  $V_N(A,P)$ , pour un vecteur paramètre  $P$  donné, est donc solution de l'équation

$$R_1 A = h_1(P) \quad (3-14)$$

De même pour le vecteur optimal  $\hat{A}$  (donc solution de (3-14), le vecteur paramètre optimal  $\hat{P}$  est celui qui minimise

$$V_N(\hat{A}, P) = \|h_2(P)\|^2 \quad (3-15)$$

Ainsi, en utilisant la factorisation QR nous avons séparé le problème d'optimisation du critère initial en deux optimisations, chacune par rapport à l'un des deux vecteurs  $A$  et  $P$ .

La procédure d'extraction des paramètres et du filtrage inverse est réalisée selon les étapes suivantes :

- Choix de la fenêtre d'analyse : la fenêtre d'analyse correspond à un cycle glottique. Il est donc nécessaire de commencer par une détection des événements, instants d'ouverture et de fermeture glottique : notons  $N$  la taille de cette fenêtre.

- Pré-emphase du signal sur la fenêtre d'analyse (optionnel).

- Calcul de la matrice  $\phi_N$  (indépendante du vecteur  $P$  des paramètres).

- Factorisation QR de la matrice  $\phi_N$ .

- Choix du vecteur initial  $P_0$  des paramètres.

- Génération sur la fenêtre d'analyse du débit glottique  $u_g(n, P_0)$  et  $u_f(n, P_0)$  partant du modèle paramétrique initial.

- Calcul des vecteurs  $h_1(P)$  et  $h_2(P)$ .

- Recherche du vecteur optimal  $\hat{P}$  des paramètres : recherche du minimum de la fonction  $\|h_2(P)\|^2$  par la méthode d'optimisation exposée plus haut.

- Génération sur la fenêtre d'analyse du débit glottique  $u_g(n, \hat{P})$  partant du modèle paramétrique optimal.

- Calcul des vecteurs  $h_1(\hat{P})$  et  $h_2(\hat{P})$ .

- Détermination du modèle AR optimal : résolution de l'équation matricielle:  $R_1 \hat{A} = h_1(\hat{P})$

- Nous pouvons recommencer la procédure pour le prochain cycle.

Remarques :

- \* la procédure est séduisante puisqu'elle permet d'effectuer le filtrage inverse en même temps que l'extraction des paramètres de l'activité glottique;

- \* les deux points clés de cette méthode sont la factorisation QR de la matrice  $\phi_N$  d'une part, et l'optimisation du critère en fonction des paramètres d'autre part. Quant au temps d'exécution, la procédure est lente et dépend surtout de la convergence une méthode d'optimisation choisie. Nous avons utilisé la méthode de recherche directe de l'optimum. Concernant la décomposition QR, elle ne s'exécute qu'une seule fois par cycle et par conséquent elle a moins d'influence sur le temps d'exécution que la recherche de l'optimum;

- \* outre les paramètres de l'activité glottique, nous pouvons extraire d'autres paramètres, à partir du filtre de synthèse optimal caractérisant le conduit vocal (donc la prononciation); par exemple, la fréquence et la largeur de bande des formants, la fonction d'aire du conduit vocal.

#### 4 - CONCLUSION

Ce travail apporte une nouvelle contribution dans le domaine de la modélisation paramétrique de l'onde glottique, puisque nous avons défini de nouveaux paramètres en étroite liaison avec le comportement différencié des deux cordes vocales. En conséquence, ce modèle peut servir pour effectuer des mesures sur des voix pathologiques dans le but d'étudier la pertinence des paramètres.

#### 5 - REFERENCES BIBLIOGRAPHIQUES

- Titze I.R. Parametrization of the glottal area, glottal flow, and vocal fold contact area, Feb. 1984, *J. Acoust. Soc. Am.*, Vol.75, N° 2.
- Titze I.R. The Human vocal chords : A mathematical model (Part 2), *Phonetica*, Vol.29 pp.1-21.
- Titze I.R. Simulation of vocal fry, 1973, *J. Acoust. Soc. Am.*, Vol.54(A), pp.23
- Titze I.R. Physical and physiological dimensions of intrinsic voice quality, 1979, *ICPh-MIAMI*, pp.217-223.
- Smiej M.F. Estimation et modélisation paramétrique de l'onde glottique en vue de la caractérisation objective de la voix pathologique, 1991, *Thèse d'Université Nancy I*.
- Hess W. Pitch determination of signals, 1983, *Springer Verlag*.