

DES PARAMETRES FORMANTIQUES AU PROFIL ARTICULATOIRE

P. Jospa

Institut de Phonétique, Université Libre de Bruxelles,
50, av Franklin Roosevelt, 1050-Bruxelles,
Belgique.

ABSTRACT

A method for calculating the parameters of an articulatory model [6] is presented. This method takes the formant values and data of labial articulation as its starting point. The method proceeds via the minimisation of the "articulatory effort" under the constraints imposed by articulatory and acoustic data. The expression of these acoustic constraints is based on a variational formulation of the relation linking the resonance modes (formants) to the area function [5]. This formulation takes account of the dissipative effects of the tract and of the lip-spread. Four variants of the procedure are evaluated, with articulatory profiles published in the literature serving as references. In most cases it seems that the simplest version gives the most satisfactory results.

1. INTRODUCTION.

Un modèle articuloire qui rend compte de la position des principaux articulateurs, présente, vis-à-vis de la fonction d'aire, une redondance de degrés de liberté nécessaire pour produire les phénomènes de coarticulation et de compensation entre articulateurs. Un calcul fiable des paramètres du modèle à partir du signal de parole ne peut se concevoir sans disposer d'informations complémentaires de nature acoustique ou articuloire. Dans quelle mesure l'usage d'une stratégie générale de coordination articuloire autorise-il un tel calcul, lorsqu'à côté du signal de parole, la donnée de l'articulation labiale est également fournie?

Nous avons développé une procédure de calcul des paramètres d'un modèle articuloire à partir des trois premières fréquences formantiques, de la donnée optionnelle des largeurs de bande des 2ème et 3ème formants, de la donnée de l'ouverture labiale, ainsi que d'un principe

de moindre "effort articuloire". Le problème est posé en ces termes: satisfaire le principe de coordination articuloire (moindre effort), les contraintes imposées par les données acoustiques et articuloires étant au mieux respectées. Notre méthode de calcul fait un usage direct des données formantiques. Outre la donnée de l'ouverture aux lèvres, la protrusion labiale et l'ouverture de la mâchoire peuvent également être fournies, auquel cas la procédure se limite à la détermination du profil lingual. La distance glotte-incisives est supposée connue. Le modèle articuloire choisi est celui de Maeda [6]. Il comporte 6 paramètres: le degré d'ouverture de la mâchoire, les degrés d'ouverture et de protrusion labiales, et trois paramètres contrôlant l'articulation linguale, à savoir: la position (avant/arrière) du corps de la langue, la forme (arquée/plate) du dos de la langue, et la position (élevée/abaissée) du sommet de la langue.

2. CALCUL DES PARAMETRES FORMANTIQUES.

Le modèle acoustique adopté est celui du conduit vocal dissipatif proposé par Sondhi [10]; il est caractérisé par une admittance des parois proportionnelle à la fonction d'aire et un facteur de forme constant. La distribution d'amplitude ψ et la fréquence ω d'un mode de résonance satisfont l'équation suivante:

$$(1) \quad \partial_x(A\partial_x\psi) + \frac{1}{c^2}(\omega^2 - \omega_p^2)A\psi = 0,$$

avec $A(x)$: la fonction d'aire, c : la vitesse du son dans le conduit vocal, et ω_p : la fréquence de vibration des parois ($2\pi \cdot 200$ hz environ). A l'extrémité glottique, nous imposons la condition:

$$(2) \quad \partial_x\psi = 0, \quad \text{en } x=0,$$

qui correspond à une impédance infinie à la glotte. A l'extrémité labiale, soit en $x=L$, le traitement classique consiste à calculer une correction de longueur ΔL :

$$(3) \quad \Delta L = \frac{3}{8} \sqrt{\pi A(L)}$$

et à imposer à l'amplitude ψ du mode, la condition:

$$(4) \quad \psi(L') = 0, \quad \text{avec } L' = L + \Delta L.$$

Cette condition est bien fondée dans le cas du premier mode. Un traitement différent, fondé jusqu'au troisième mode, consiste à imposer la condition suivante aux lèvres [5]:

$$(5) \quad A\partial_x\psi + q\sqrt{A}\psi = 0, \quad \text{en } x=L,$$

avec q une constante de l'ordre de 2,1.

Le calcul des fréquences ω et des fonctions ψ qui satisfont à l'équation (1), sous les conditions (2) et (4) ou (5), représente un problème aux valeurs propres classique. Ce problème peut recevoir une formulation variationnelle équivalente [4],[5] qui présente certains avantages. Etant de nature intégrale, cette formulation variationnelle conduit à des solutions qui sont liées plus au caractère général qu'au comportement détaillé de la fonction d'aire. Elle permet l'usage des méthodes directes du calcul des variations [3] --telle la méthode de Rayleigh-Ritz utilisée dans ce travail--, qui substituent au problème initial, un problème de valeurs propres de l'algèbre linéaire, de dimension généralement faible, plus simple à résoudre. Enfin, la formulation variationnelle permet d'approcher analytiquement le lien qui uni la fonction d'aire aux fréquences des formants. Dans le cadre de la méthode de Rayleigh-Ritz, et pour un modèle articuloire donné, ce lien prend la forme d'équations algébriques non linéaires reliant les fréquences formantiques f_n aux paramètres articuloires $\{a\}$ [4],[5]:

$$\det\{K(\{a\}) - \lambda_n V(\{a\})\} = 0,$$

$n=1, \dots, N$

$$\text{avec: } \lambda_n = (2L/c)^2 \cdot (f_n^2 - f_p^2)$$

où f_n désigne la fréquence du n ème formant, et où K et V , sont des matrices carrées, symétriques, ne dépendant que

des paramètres articuloires. La définition et le calcul de ces matrices sont exposés ailleurs [5]. L'ordre M de ces matrices peut être relativement faible: $M=8$ dans nos calculs, pour chacun des trois premiers formants. Une fois construites les matrices K et V , la détermination des premières fréquences propres se ramène au calcul des premières racines de l'équation:

$$(6) \quad \det\{K_{l,m} - \lambda V_{l,m}\} = 0.$$

Le calcul des largeurs de bande exige de connaître non seulement la fréquence, mais aussi la distribution spatiale du mode considéré. Une fois calculée la racine λ_n de l'équation (6), le calcul de ψ_n ne fait guère de difficulté [5].

3. CALCUL DES PARAMETRES ARTICULOIRES.

Désignons par a^0_k les premières approximations (ou valeurs initiales) des paramètres articuloires, par $f_n(\{a\})$ et $B_n(\{a\})$ les paramètres formantiques calculés pour un profil articuloire $\{a\}$, et par f_n^* et B_n^* les données formantiques. Désignons encore par K le nombre de paramètres articuloires à calculer, et soit, finalement, p_k et q_k ($k=1, \dots, K$), des pondérations convenablement choisies. La méthode consiste à minimiser une fonctionnelle E des articulateurs ("l'effort articuloire"), de la forme:

$$E(\{a\}) = \sum_{k=1}^K (p_k |a_k - a^0_k| + q_k |a_k|)$$

sous les contraintes acoustiques suivantes:

$$(f_n(\{a\}) - f_n^*) / f_n^* = 0, \quad n=1,2,3$$

$$(B_n(\{a\}) - B_n^*) / B_n^* = 0, \quad n=2,3$$

et sous les contraintes articuloires:

$$-3 \leq a_k \leq 3 \quad k=1, \dots, 4$$

$$0.2 \text{ cm} \leq a_5 \leq 2.4 \text{ cm}$$

$$A_{\text{min}} > 0.3 \text{ cm}^2$$

La procédure adoptée pour résoudre ce problème d'optimisation sous contraintes repose sur une méthode de minimisation séquentielle sans contrainte, faisant usage de fonctions de pénalisation [1]. Soit:

$r_j = (0.1)^j r_0$, ($j=0, 1, \dots, J$), une séquence décroissante de nombres positifs ($r_0 = 5$), telle que $r_j > r_{\text{min}}$ ($r_{\text{min}} = 0.001$).

Nous construisons une suite de $J+1$ systèmes de K équations non linéaires à K inconnues, à résoudre séquentiellement, de

manière optimale au sens des moindres carrés. L'algorithme de Powell [9], utilisé à cet effet, s'est avéré rapide et robuste. La solution obtenue au terme de cette procédure, optimise le respect des contraintes, tout en minimisant l'effort articuloire E . Des exemples de résultats sont illustrés par les figures 1 à 7.

4. EVALUATION.

Une évaluation informelle de la procédure a été réalisée en utilisant les données (fonctions d'aire et fréquences formantiques) publiées par Fant [2] (6 voyelles russes) et par Mrayati [7] (11 voyelles françaises). En l'absence de données fiables relatives aux largeurs de bande, nous utilisons comme données, les largeurs de bande calculées pour les fonctions d'aire publiées. Les données des fréquences formantiques sont celles qui sont publiées par les auteurs: mesurées sur sonogrammes pour Fant, calculées selon le modèle de la ligne de transmission pour Mrayati. Les tableaux I et II résument les résultats obtenus selon que l'une ou l'autre des conditions (4) et (5) est appliquée aux lèvres et selon que les largeurs de bande des formants 2 et 3 sont prises ou non en compte. Quatre versions de la procédure sont ainsi obtenues; leurs résultats sont comparés en fonction de leur degré d'accord entre les fonction d'aire calculées et les fonctions d'aire données. La première colonne indique la version de la procédure, soit:

- Version 1: la condition (4) est adoptée aux lèvres et les largeurs de bande ne sont pas prises en compte.

- Version 2: la condition (4) est adoptée aux lèvres et les largeurs de bande sont prises en compte.

- Version 3: la condition (5) est adoptée aux lèvres et les largeurs de bande ne sont pas prises en compte.

- Version 4: la condition (5) est adoptée aux lèvres et les largeurs de bande sont prises en compte.

La colonne 2 indique la moyenne sur toutes les voyelles v de l'écart logarithmique moyen entre la fonction d'aire calculée A_v et la fonction d'aire de référence A_{vr} :

$$\frac{1}{N_v} \sum_{v=1}^{N_v} \frac{1}{N_p} \sum_{i=1}^{N_p} |\log(A_v(x_i)) - \log(A_{vr}(x_i))|,$$

N_v désignant le nombre de configurations et N_p le nombre de sections par configuration. La colonne 3 indique l'écart absolu moyen, en hertz, pour toutes les voyelles, entre les valeurs calculées et les valeurs données des 3 premières fréquences formantiques. La colonne 4 indique un écart relatif moyen pondéré, en pour-cent, entre les mêmes valeurs formantiques.

Tableau I: données de Fant [2], moyennes sur 6 voyelles russes.

Version:	Ecart log. aires	Fréquences:	
		Eabs(hz)	Erel(%)
1	0.41	87.3	5.30
2	0.46	98.4	7.83
3	0.41	87.1	6.92
4	0.42	95.0	7.40

Tableau II: données de Mrayati [7], moyennes sur 11 voyelles françaises.

Version:	Ecart log. aires	Fréquences:	
		Eabs(hz)	Erel(%)
1	0.44	71.6	4.66
2	0.37	88.1	6.83
3	0.44	59.6	4.95
4	0.47	58.8	4.83

On ne constate pas différences nettement significatives entre les différentes versions. La version 1 réalise l'une des meilleures performances. Etant la moins coûteuse en temps de calcul, et ne demandant pas la connaissance des largeurs de bande, elle semble devoir s'imposer.

Les résultats présentés appartiennent au cas où l'aire d'ouverture et la protrusion labiales sont données; le nombre K de paramètres articulatoires calculés est donc égal à 4, soit l'ouverture de la mâchoire et les trois paramètres de la langue. Des résultats similaires ont été obtenus lorsque seule l'aire de l'ouverture aux lèvres est donnée ($K=5$), et lorsqu'en plus des paramètres labiaux, l'ouverture de la mâchoire est donnée ($K=3$).

L'influence de la première approximation sur la fonction d'aire calculée a été testée, en adoptant, systématiquement, comme première approximation de l'articulation linguale, les valeurs du profil neutre. Les performances moyennes observées ne sont pas significativement modifiées.

5. CONCLUSION.

Nous pensons avoir développé une méthode robuste et relativement économique en temps de calcul pour estimer les paramètres d'un modèle articuloire, à partir des fréquences des trois premiers formants et de la donnée de l'articulation labiale. Appliquée au modèle de Maeda, elle permet une estimation des paramètres de l'articulation linguale. La donnée des largeurs de bandes formantiques ne semble pas utile, et peut, dans certains cas, avoir un effet perturbateur. La condition aux lèvres (5), plus complexe à traiter que la condition (4), ne conduit pas à des performances supérieures; la version 1 de la procédure semble donc pouvoir s'imposer.

REFERENCES:

- [1] ADBY, P.R. et DEMPSTER, M.A.H. (1974), "Introduction to Optimizaton Methods", London: Chapman and Hall.
- [2] FANT, G. (1960), "Acoustic theory of speech production", The Hague: Mouton & Co.
- [3] GOULD, S.H. (1966), "Variational Methods for Eigenvalue Problems", London: Oxford University Press.
- [4] JOSPA, P. (1972) "Forme approchée du conduit vocal déduite des fréquences de résonance. Théorie des perturbations et méthode variationnelle", Lannion: JEP, GALF, 225-262.
- [5] JOSPA, P. (1977), "Théorie acoustique du conduit vocal de forme variable dans le temps", Thèse de Doctorat, Université Libre de Bruxelles.
- [6] MAEDA, S (1979), "Un modèle articuloire de la langue avec des composantes linéaires", 10èmes JEP, GALF, 152-164.
- [7] MRAYATI, M. et CARRE, R. (1976), *Phonetica*, 33, 285-306.
- [9] POWELL M.J.D. (1970), "A fortran subroutine for solving systems of nonlinear algebraic equations", in: "Numerical methods for nonlinear algebraic equations", RABINOWITZ, Ph. Ed., London: Gordon and Breach.
- [10] SONDHI, M.M. (1974), *J. Acoust. Soc. Am.*, Vol.55, 5, 1070-1075.

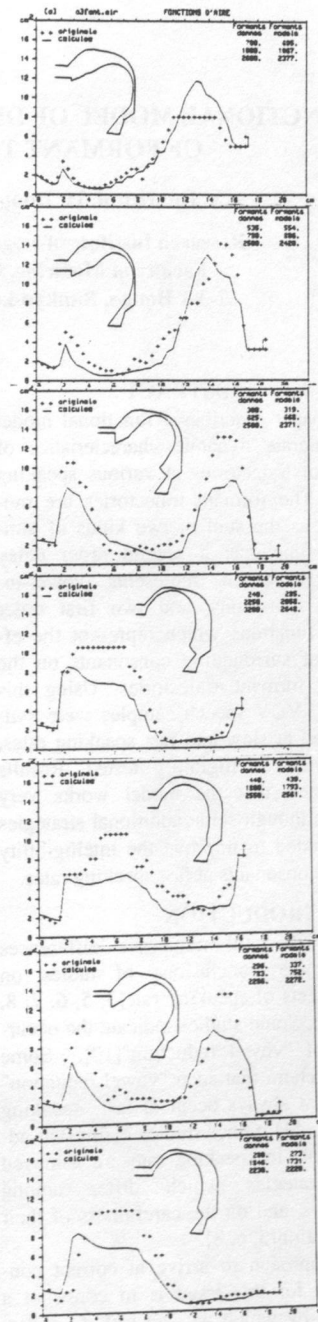


Fig. 1
[a]
Fant

Fig. 2
[o]
Fant

Fig. 3
[u]
Fant

Fig. 4
[i]
Fant

Fig. 5
[e]
Fant

Fig. 6
[u]
Mrayati

Fig. 7
[y]
Mrayati