

Contribution à l'analyse acoustique du conduit vocal.

X. Pelorson⁽¹⁾, K. Motoki⁽²⁾, R. Laboissière⁽¹⁾.

⁽¹⁾ : Institut de la Communication Parlée,
CNRS, INPG, Université Stendhal,
46 avenue F. Viallet
F-38031, Grenoble Cedex 01.

⁽²⁾ : Faculty of Engineering,
Hokkai-Gakuen University
S-26, W-11, Sapporo, Japan.

Abstract

In this paper an analysis of the vocal tract acoustics at high frequencies, including higher acoustical propagating modes, is presented.

Several examples using vocal tract approximations of increasing complexity are presented and discussed. The effects of the source position as well as of the vocal tract geometry are shown to be of particular importance.

1. INTRODUCTION.

La plupart des modèles physiques appliqués au conduit vocal reposent sur une description de la propagation des ondes acoustiques en ondes planes, c'est-à-dire unidimensionnelle (e.g. [Fan60]). La limite de validité de l'hypothèse unidimensionnelle, bien que formellement difficile à déterminer (elle dépend en grande partie de la géométrie précise du conduit vocal), peut être estimée aux alentours de 4-5 kHz. Cette limitation, longtemps justifiée par le manque de données anatomiques en trois dimensions, devrait cependant pouvoir être levée grâce aux techniques modernes d'imagerie (IRM etc...). Dans cet objectif, nous présentons donc une analyse de la propagation acoustique en trois dimensions qui ne se limite donc plus aux basses fréquences. Il est montré que les effets principaux aux hautes fréquences sont liés à l'apparition de modes d'ordre supérieur qui peuvent avoir une influence considérable sur le champ de pression acoustique. Ces effets sont illustrés et discutés sur la base de quelques exemples synthétiques.

2. ELEMENTS THEORIQUES, PROPAGATION DANS UN CONDUIT UNIFORME.

A titre d'illustration, nous considérons par la suite la représentation la plus schématique possible du conduit vocal : un guide de section rectangulaire et uniforme, de longueur $L_x = 16.8$ cm, rayonnant dans un écran infini. Les parois du conduit vocal sont supposées parfaitement réfléchissantes et l'excitation est

assimilée à une source unique localisée en un point du conduit vocal: $Q = Q_0 \delta(z-z_0) \delta(y-y_0) \delta(x-x_0)$. L'extension de cette analyse aux cas impliquant la présence de plusieurs sources de son dans le conduit vocal ou de sources distribuées dans l'espace est aisée du fait de la linéarité du problème.

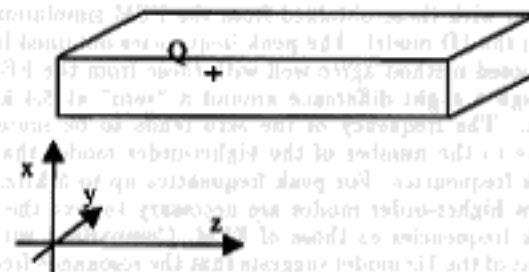


Figure 1: Représentation schématique du conduit vocal comme un guide uniforme.

La solution générale de l'équation de propagation acoustique peut alors être formulée de la manière suivante [Bru98]:

$$P(f, x, y, z) = \sum_{mn} \Lambda_{mn} Q \frac{\Psi_{mn}(x, y) \Psi_{mn}^*(x_0, y_0)}{k_{mn}^2 - k^2} \times (A_{mn} e^{-j k_{mn} z} + B_{mn} e^{j k_{mn} z}) \quad (1)$$

où f est la fréquence imposée par la source Q , A_{mn} et B_{mn} deux constantes déterminées par les conditions aux limites à chaque extrémité du guide, Λ_{mn} une constante de normalisation. Les fonctions propres Ψ_{mn} sont définies par :

$$\Psi_{mn}(x, y) = \cos\left(\frac{m\pi}{L_x} x\right) \cos\left(\frac{n\pi}{L_y} y\right) \quad (2)$$

La constante de propagation k_{mn} est déterminée par l'équation de dispersion:

$$k_{mn}^2 = k^2 - \left(\frac{m\pi}{L_x}\right)^2 - \left(\frac{n\pi}{L_y}\right)^2 \quad (3)$$

Les équations (1) et (3) montrent qu'un mode (m,n) ne sera propagatif (k_{mn} réel) que si la fréquence d'excitation, $f = kc/2\pi$, est supérieure à la fréquence de coupure du mode, f_{mn} :

$$f_{mn} = \frac{c}{2\pi} \sqrt{\left(\frac{m\pi}{L_x}\right)^2 + \left(\frac{n\pi}{L_y}\right)^2}$$

où c la vitesse du son.

En dessous de cette fréquence de coupure, k_{mn} devient imaginaire pur traduisant une onde évanescence. En particulier, lorsque la fréquence de l'excitation est inférieure à la première fréquence de coupure (f_0 ou f_{10}) seul le mode plan (0,0) se propage.

Bien que l'équation (1) implique une double sommation infinie sur m et n , dans la pratique il n'est pas nécessaire de prendre un compte un grand nombre de modes d'ordre supérieur compte tenu de la rapidité de l'atténuation des modes évanescents. A titre d'exemple, il est présenté sur la figure 2 les contours de pression calculés à l'intérieur du guide en prenant en compte deux modes seulement. Pour comparaison, sur la même figure, sont représentés les résultats obtenus par simulation numérique (méthode TLM, [El M98]) pour la même configuration. Compte tenu de la résolution spatiale utilisée, on peut estimer que le nombre de modes pris en compte par la simulation numérique TLM est de 400.

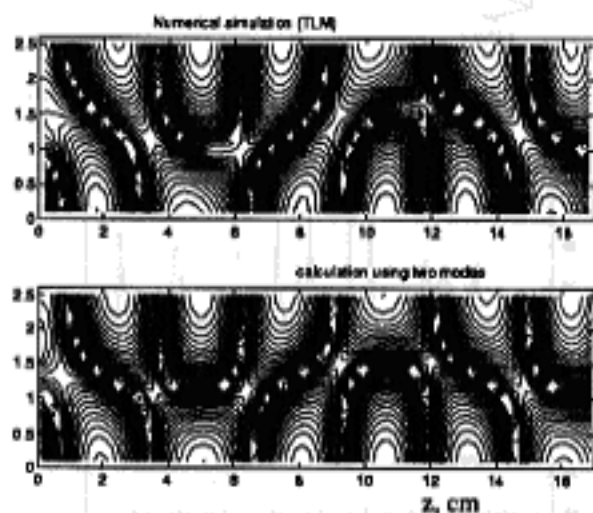


Figure 2 : Contour de pression (par pas d'1 dB) à 9 kHz obtenu par simulation numérique (en haut) et par le calcul en tenant compte de deux modes (bas).

Il est important de noter l'influence de la position de la source sur la génération et l'efficacité des modes d'ordre supérieurs. Ceci s'illustre dans l'équation (1) par le terme $\Psi_{mn}^*(x_0, y_0)$. Ainsi si la source est

placée au centre d'une section du conduit vocal ($x_0 = L_x/2$ ou $y_0 = L_y/2$), comme pour la glotte en première approximation, les modes impairs ne seront pas générés ($\Psi_{mn}^*(x_0, y_0) = 0$). Réciproquement une source placée près d'une paroi excitera les modes impairs avec une amplitude maximum ($|\Psi_{mn}^*(x_0, y_0)| = 1$).

3. CHANGEMENT DE SECTION.

Afin d'illustrer la génération et la transmission des modes d'ordre supérieur, nous considérons maintenant l'approximation décrite dans la figure 2 (approximation à deux sections du conduit vocal). Deux cas sont considérés, l'un parfaitement symétrique et l'autre asymétrique.

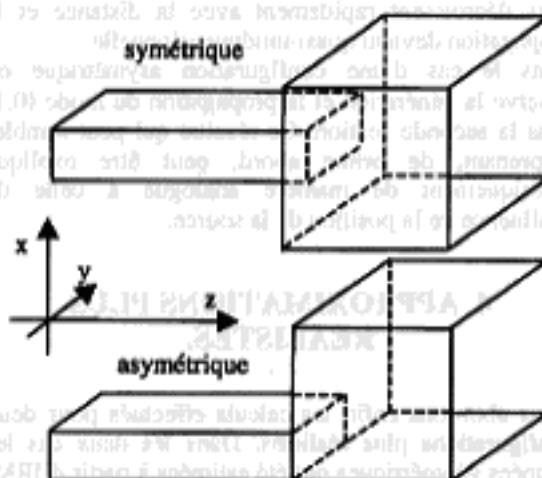


Figure 3: Approximations à deux tubes du conduit vocal. Configurations symétrique (en haut) et asymétrique (en bas).

La résolution de l'équation de propagation acoustique dans de telles structures n'est, en général, pas analytique. La méthode de raccordement modal [Ker91] permet cependant d'obtenir numériquement une solution. Le principe consiste à développer pression et vitesse acoustique sur la base des fonctions propres de chacune des deux sections, l'application du principe de la conservation de la pression et de la vitesse normale à la jonction des deux guides permet alors de déterminer l'amplitude de la pression en tout point de la structure. Dans la figure 4 sont présentés deux exemples de résultats pour les distribution de pression dans le plan $z-y$ à 6 kHz.

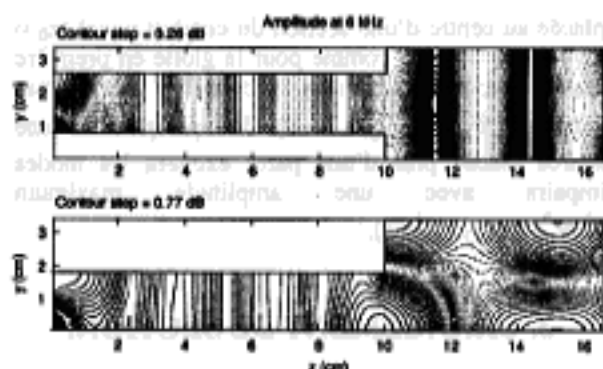


Figure 4 : Comparaison entre les contours de pression obtenus à 6 kHz dans les configuration symétriques et asymétriques.

Dans le cas d'une configuration symétrique, on peut observer que si des perturbations liées à la génération de modes d'ordre supérieurs interviennent au niveau de la source ainsi qu'aux discontinuités de sections, elles décroissent rapidement avec la distance et la propagation devient quasi-unidimensionnelle.

Dans le cas d'une configuration asymétrique on observe la génération et la propagation du mode (0,1) dans la seconde section. Ce résultat qui peut sembler surprenant, de prime abord, peut être expliqué théoriquement de manière analogue à celle de l'influence de la position de la source.

4. APPROXIMATIONS PLUS REALISTES.

Nous abordons enfin les calculs effectués pour deux configurations plus réalistes. Dans les deux cas les données géométriques ont été estimées à partir d'IRM, la source est placée arbitrairement en (0,0,0). La quantité étudiée est la fonction de transfert définie par :

$$H \propto \sqrt{\frac{W_{rad}}{Q}}$$

où W_{rad} est la puissance acoustique totale rayonnée aux lèvres.

Le premier cas concerne la voyelle /a/ simulée au moyen de 15 guides élémentaires comme montré sur la figure 5, dans ce cas il a été supposé que la structure est symétrique par rapport à l'axe z.

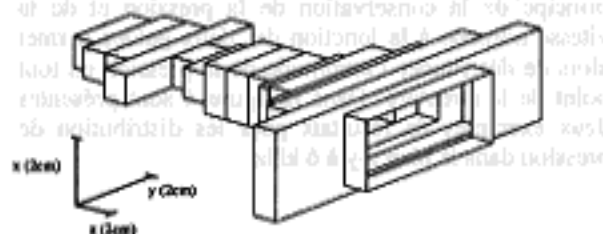


Figure 5: Approximation à 15 sections de la voyelle /a/.

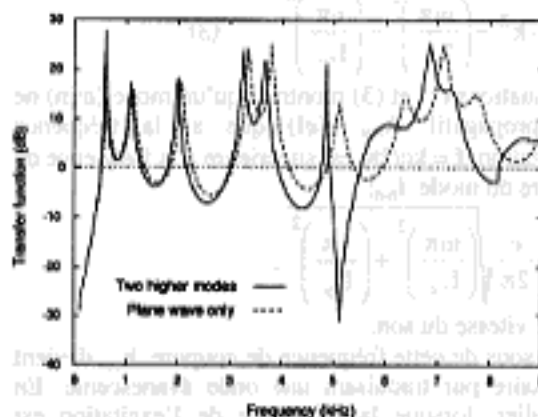


Figure 6 : Fonction de transfert de la voyelle /a/. Comparaison entre une solution unidimensionnelle (onde plane uniquement) et une solution prenant en compte deux modes.

Le second exemple concerne la fricative /f/ approximée par une succession de 38 guides comme décrit sur la figure 7. La fonction de transfert correspondante est présentée sur la figure 8.

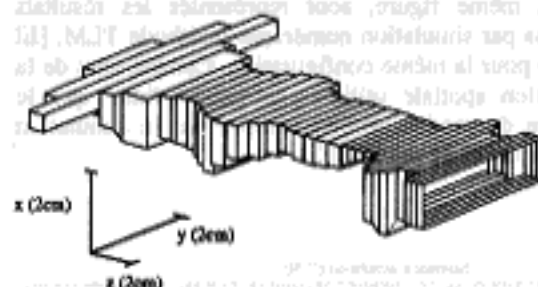


Figure 7: Approximation à 38 tubes de la fricative /f/.

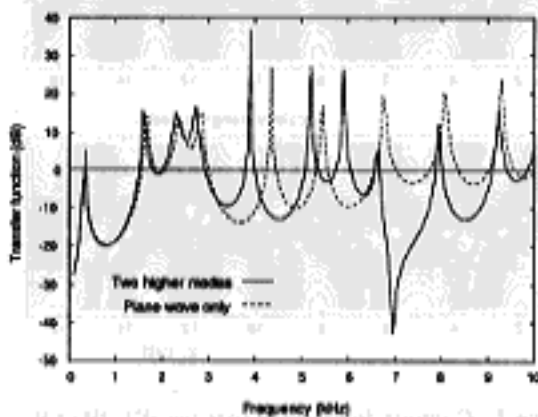


Figure 8 : Fonction de transfert de la fricative /f/.

Les résultats présentés conduisent aux conclusions suivantes :

En dessous de la première fréquence de coupure, la perturbation induite par les modes d'ordre supérieurs est faible : un abaissement des fréquences de résonances de l'ordre de quelques pourcents. A partir

de la fréquence de résonance on peut observer un « zéro » puis une forte perturbation des résonances liée à la présence de modes d'ordre supérieurs propagatifs. La position de ces « zéros » étant fonction de la dimension transverse du conduit vocal, elle est donc différente selon le son considéré (5kHz pour le /a/, 7 kHz pour le /i/). Comme on pouvait s'y attendre l'influence des modes d'ordre supérieurs est plus importante dans une configuration asymétrique du conduit vocal. Cependant, on peut constater au vu de la figure 6 que même pour une approximation symétrique l'influence des modes de propagation est clairement visible.

Enfin, il faut noter que les résultats présentés se réfèrent à une source unique et ponctuelle. Une telle approximation, si elle est acceptable au premier ordre pour les sons voisés, ne l'est certainement plus dans le cas des fricatives.

5. CONCLUSION

Les exemples développés montrent que le champ de pression acoustique aux fréquences élevées est fortement déterminé par la présence de modes d'ordre supérieurs. Les effets observés sont à rapprocher des observations faites sur la parole naturelle : présence de zéros dans la fonction de transfert des voyelles orales ou des fricatives, amplification du contenu spectral du champ rayonné aux fréquences élevées pour les fricatives.

References

[Bru98] Bruneau M. (1998). *Manuel d'Acoustique Fondamentale*, Hermes, Paris.

[Elm98] Elmasri S., Pelorson X., Saguet P., Badin P. (1998). "The use of the Transmission Line Matrix in acoustics and in Speech", *International Journal of Numerical Modelling*, 11, 133-151.

[Fan60] Fant G. (1970) "Acoustic theory of speech production", 2nd Ed., Mouton and co, Den Haag.

[Ker91] Kergomard, J. (1991) "Calculation of discontinuities in waveguides using mode-matching method : an alternative to the scattering matrix approach", *J. Acoustique*, 4, 111-138.

Abstract

The presence of the vocal tract at high frequencies including higher modes of propagation is investigated using three dimensional solutions of the wave equation. Different geometries of articulators and cavities are considered. It is shown in particular that the effect of the position of the source and of the vocal tract geometry appears to play a crucial role in high frequencies. The application of these findings will be reported in vowel and consonant synthesis in a subsequent paper.

Introduction

During the last decades, great efforts in the field of speech modelling have been focusing on the description of the sound source with the assumption that the vocal tract will act as a wave propagation within the vocal tract (see for example [1, 2]). This clearly established the limit of validity of such an approach. It is already well known that in the case of a source which is a discrete radiation in particular for plosives and fricatives involving the use of complex dimensional sources could have been limited by the lack of information about the vocal tract. In particular, however, towards models to new speech synthesis techniques such as MRI, several three-dimensional representations of the vocal tract geometry are available. In this paper the 3-D theory of wave propagation is used in a fairly general way. The application of such a theory to the case of a cascade of tubes -the partition of the vocal tract- is made. The theory is applied as such presented. The authors are aware that this wave propagation theory is a simplification and that the results are not

The one-dimensional approximation of the vocal tract is considered in this section. The simplest representation of the vocal tract is illustrated in figure 1. It is a cascade of tubes with rigid walls. Each tube has one end open and closed in an infinite bottle. In the other end, the tube is closed. The shape of the vocal tract is considered as a cascade of tubes with a cross-section of $0.5 \times 2.5 \text{ cm}^2$ and a length of 18.8 cm. For the sake of simplicity, we only consider that the excitation of the vocal tract is