

# CIBLES RESONANTIQUES DES VOYELLES ORALES

**GUERTI Mhania**

Ecole Nationale Polytechnique BP. 182 16200 El Harrach Alger Algérie  
& CRSTDLA - Université de Bouzaréah Alger

**E- mail : mhaniag@hotmail.com**

**Résumé :** L'étude analytique de la propagation des vibrations sonores dans le conduit vocal est si complexe qu'elle nécessite le recours à certaines simplifications a priori et à l'utilisation de modèles. Depuis les travaux de G. Fant, une théorie mathématique s'est développée dans ce sens, sur la base d'une analogie avec des circuits électriques. Dans cette théorie, un mode de résonance dépend de tous les éléments du circuit. Il en est de même dans la production de la parole où tous les éléments du conduit vocal contribuent à la détermination de la fréquence de chacun des formants. Une variation de la dimension d'un élément peut avoir des conséquences sur l'ensemble des paramètres acoustiques (les indices acoustiques).

Dans son modèle à 4 tubes du conduit vocal, Fant a étudié l'évolution de la structure formantique lorsque la coordonnée du centre de la constriction se déplace de la glotte aux lèvres. Ceci est fait pour un ensemble de valeurs formantiques données qui constitue des *nomogrammes*. Conformément à ces derniers, nous proposons une représentation paramétrique du signal vocal qui réduit le nombre de règles de réécriture pour la synthèse, et relie des mouvements spectraux aux gestes articulatoires dans un mode plus naturel.

Nous examinons les caractéristiques essentielles des résonances du conduit vocal et leurs applications aux séquences Voyelles-Voyelles.

**Mots clés :** Modèle de G. Fant ; Paramètres acoustico-articulatoires ; Règles de réécriture ; Résonances ; Voyelles orales.

## 1 Introduction

Un système acoustique se décompose en deux parties appelées résonateur et exciteur :

- le premier est une cavité possédant des modes propres de résonance. C'est celui-ci qui détermine principalement le timbre des sons de la parole et notamment les fréquences formantiques ;

- le second délivre un signal de source, modifié par la fonction de transfert du résonateur. Cette décomposition est difficile lorsque le résonateur réagit sur l'exciteur, et le cas est fréquent dans la phonation.

On ne peut pas relever précisément la forme de l'onde glottique, car il est difficile d'isoler acoustiquement le larynx du conduit vocal. Même si l'on dispose d'un microphone minuscule à quelques millimètres des cordes vocales, on obtient déjà un signal modifié par les résonances du conduit vocal par couplage aérodynamique et acoustique. La fonction donnant l'aire sectionnelle du tube vocal en fonction de la distance par rapport à la glotte est le paramètre d'entrée.

A une telle fonction on peut associer une fonction de transfert exprimant la façon dont l'excitation glottale est transformée en rayonnement d'énergie acoustique à l'ouverture des lèvres. Cette fonction présente des pôles qui définissent les fréquences et les largeurs des résonances.

La théorie de G. Fant a permis des progrès considérables à l'analyse de la corrélation entre les niveaux articulatoires et acoustiques. Elle a montré que la fréquence formantique constitue une information essentielle dans la production de la parole (G. Fant, 1960).

Dans le cadre de l'étude acoustique de la production des voyelles, les nomogrammes permettent d'établir :

- Des relations entre des données articulatoires (physiologiques) concernant la production de la parole et la description des signaux acoustiques ;

- Des affiliations entre formants et cavités, en fournissant des prédictions efficaces des valeurs formantiques qui correspondent à une production vocalique donnée.

Conformément aux nomogrammes de Fant, nous proposons un nouvel espace de représentation spectrale  $R_1$ - $R_2$  ( $R_1$  et  $R_2$  correspondent respectivement aux résonances des deux premières basses fréquences de la cavité postérieure et de la cavité antérieure).

Il existe des Points Focaux  $PF$  où les convergences formantiques se rencontrent et où des affiliations entre formants et cavités sont échangées.

## 2. Rappel sur les résonateurs acoustiques

Un résonateur de Helmholtz - du nom du Physicien Allemand Helmholtz - est une cavité sphérique à parois rigides. Il est muni d'un col fin ayant une longueur " $L$ " et une section " $A$ " de volume " $AL$ " très inférieur au volume " $V$ " de la cavité. Le résonateur est excité directement par la pression acoustique régnant à l'entrée d'une embouchure d'aire quasi nulle. En présence d'un son comportant de nombreuses composantes de fréquences différentes, le résonateur permet de sélectionner celles qui correspondent à sa fréquence de résonance et de les amplifier.

La forme interne de la cavité est sans importance, à une condition que sa plus grande dimension reste très inférieure à la longueur d'onde pour la fréquence d'excitation, de façon que la pression instantanée soit la même en tout point du résonateur. Si cette condition n'est pas respectée, le système ne fonctionne plus en résonateur de Helmholtz, et la détermination de son régime vibratoire devient beaucoup plus complexe.

La résonance associée à un système composé d'un résonateur de volume " $V$ " comportant une embouchure de longueur " $L$ " et de faible section " $A$ ", est appelée *résonance de Helmholtz* (elle correspond à la résonance la plus faible de la fonction de transfert du résonateur).

### 2.1. Propriétés des résonateurs acoustiques

Nous savons que la structure formantique d'une voyelle dépend des volumes et ouvertures de sortie des résonateurs de l'appareil phonatoire, caractéristiques qui dépendent elles-mêmes de l'articulation. Les propriétés principales de ces résonateurs sont donc liées à deux facteurs essentiels :

- La fréquence de résonance d'un résonateur acoustique dépend tout d'abord de son volume : *plus ce résonateur est important, plus la fréquence de résonance est basse, et inversement* ;
- *L'effet de l'ouverture de sortie d'un résonateur acoustique est à l'inverse de l'effet du volume.*

Comme la langue se meut au cours de l'élocution, le volume de ces résonateurs et la grandeur de leurs ouvertures de sortie varient constamment, faisant donc varier leurs fréquences de résonance. Par ailleurs, rappelons quelques principes acoustiques de base des résonateurs. L'air contenu dans un tube uniforme de longueur " $L$ " peut vibrer conformément aux conditions

des extrémités du tube, selon deux modes différents (figure 1) :

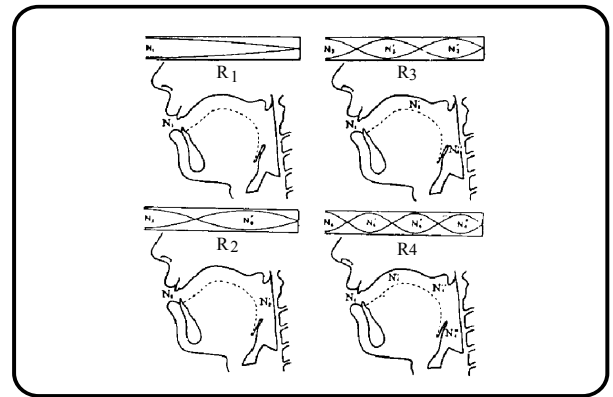
- Si les extrémités sont toutes les deux ouvertes ou fermées, les fréquences de résonance sont des multiples entiers de résonances  $\frac{\lambda}{2}$  (appelés résonances demi-onde).

$$F_n = \frac{nc}{2L} \quad [1]$$

$n$  : entier représentant le nombre de tubes.

- Si une extrémité est ouverte et l'autre fermée, ou inversement, les fréquences de résonance sont des multiples entiers impairs de résonances  $\frac{\lambda}{4}$  (appelés quart d'onde) :

$$F_n = \frac{(2n-1)c}{4L} \quad [2]$$



**Figure 1:** Analogie du conduit vocal avec un tuyau neutre semi-ouvert.

## 3. Modélisation acoustique des voyelles

Pour simuler de plus près les paramètres articulatoires, Fant a tenté de caractériser à l'aide de trois paramètres, la fonction d'aire du conduit vocal (qui sans cela est définie par autant de valeurs que de sections). Ils modélisent les principales caractéristiques de la configuration vocale du conduit vocal par ces paramètres qui sont les équivalents quantitatifs suivants :

- Degré d'ouverture des lèvres ;
- Position de la constriction de la langue (point d'articulation) ;
- Labialisation (avancée des lèvres).

A partir de ces paramètres, il a obtenu les résultats suivants :

- Une fréquence  $F_1$  élevée correspond à une constriction étroite près de la glotte et à une large ouverture des lèvres sans arrondissement. Elle est essentiellement liée à la hauteur de la langue, c'est-à-

dire à l'ouverture du conduit vocal dont elle est fonction croissante ;

- Une fréquence  $F_1$  basse correspond au contraire soit à une constriction étroite près des lèvres soit à une ouverture faible et arrondie des lèvres ;
- La fréquence  $F_2$  est une fonction croissante de l'ouverture des lèvres et du déplacement du point de constriction de la glotte vers les lèvres.

Comme le retrait de la langue ainsi que l'arrondissement des lèvres ont tous deux pour effet d'allonger le résonateur oral (après la constriction), nous pouvons dire que  $F_2$  est fonction décroissante de cette longueur.

Le rôle de  $F_3$  est moins évident. Sa présence est indispensable à la reconnaissance de [i] et inutile à celle de [u]. Pourtant dans une voyelle comme [i] où  $F_2$  et  $F_3$  sont voisins, on peut les remplacer par un formant intermédiaire : un [i] naturel a :  $F_1 \approx 250$  Hz,  $F_2 \approx 2500$  Hz et  $F_3 \approx 3$  kHz. Il est parfaitement simulable par un [i] synthétique qui a le même  $F_1$ , mais un  $F_2 \approx 2750$  Hz.

#### 4. Convergences formantiques

La forme du conduit vocal présente une extrême variabilité articulatoire qui se manifeste au niveau du jeu des articulateurs, mais des possibilités compensatoires sont utilisées pour permettre de réaliser un même phénomène acoustique avec des commandes articulatoires différentes (*système à plusieurs degrés de libertés*). On a montré que des locuteurs arrivent à produire la même voyelle (avec un lieu de rétrécissement et une ouverture quasi identiques), malgré les perturbations apportées sur un articulateur, il y a une convergence entre le niveau des commandes et le niveau acoustique.

Ces expériences de perturbation confirment bien la description du conduit vocal en termes géométriques liés aux trois paramètres ( $X_C$ ,  $A_C$ ,  $A_1$  respectivement : coordonnée du centre de la constriction, aire de la constriction et aire aux lèvres) et confirment bien l'importance de ces variables du conduit vocal déjà proposées par G. Fant dans la caractérisation de son modèle à 4 tubes.

##### 4.1. Modèle à quatre tubes de Fant

Pour modéliser d'une façon simple le comportement acoustique du conduit vocal, Fant a conçu un modèle à quatre tubes (cavité arrière, zone de rétrécissement maximal, cavité avant et lèvres).

Même si la forme du conduit vocal est complexe, il existe toujours un rétrécissement entre la glotte et les lèvres. Celui-ci permet de distinguer de manière générale, deux cavités couplées : la cavité avant, *des lèvres à la constriction* et la cavité arrière *de la constriction au larynx*.

En effet, le modèle utilisé permet des interprétations claires des conséquences acoustiques (en termes de valeurs formantiques) de la variation de la forme du conduit vocal en ses quatre zones (fig. 2). Son modèle est basé sur le principe de la modélisation par ondes planes stationnaires. Ce modèle, bien que simple dans le principe de la modélisation articulatoire, permet une bonne prédiction des formants à partir de la géométrie du conduit vocal.

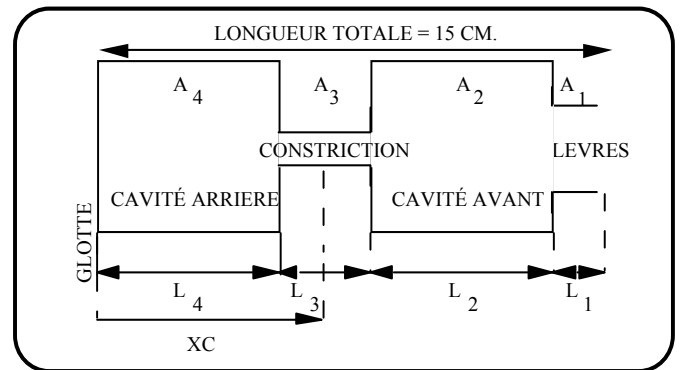


Figure 2 : Modèle à quatre tubes de G. Fant.

$$\begin{aligned} A_1 &= 0.16 \text{ cm}^2 \text{ à } 4 \text{ cm}^2, & L_1 &= 1 \text{ cm}; \\ A_2 &= 8 \text{ cm}^2, & L_2 &\text{ est variable}; \\ A_3 &= 0.65 \text{ cm}^2, & L_3 &= 5 \text{ cm}; \\ A_4 &= 8 \text{ cm}^2, & L_4 &\text{ est variable}. \\ L_2 + L_3 + L_4 &= 15 \text{ cm}. \end{aligned}$$

$$X_C = \frac{(2L_4 + L_3)}{2}$$

$$\text{avec } -2.5 \text{ cm} < X_C < 17.5 \text{ cm}$$

Une étude sur les sensibilités formantiques à des variations des différentes cavités montre que, sauf dans le cas d'une franche constriction du conduit vocal (séparation complète des cavités), un formant sera plus ou moins affilié à telle ou telle cavité. Généralement, on attribue le  $F_1$  à la résonance de la cavité arrière et le  $F_2$  à celle de la cavité avant. En principe, toutes les parties du conduit vocal contribuent à la détermination de tous les formants, avec des degrés variables dépendant de la configuration actuelle. Par conséquent, il n'existe pas de règles simples permettant de lier la fréquence d'un formant à la taille d'une cavité particulière, étant donnée l'importance de leur couplage. Cependant, on peut prévoir le degré d'association de chaque formant avec les parties du conduit vocal à partir de l'estimation sous forme de fonctions de sensibilité (M. Guerti & G. Bailly, 1990).

Compte tenu du couplage acoustique entre les cavités (d'autant plus important que  $A_C$  est importante), on ne peut pas relier strictement une résonance donnée du conduit vocal à une cavité précise ; on parle plutôt *d'affiliation majoritaire* : ces affiliations sont fonction de l'aire à la constriction et sont d'autant plus difficiles à mettre en évidence que l'aire à la constriction est

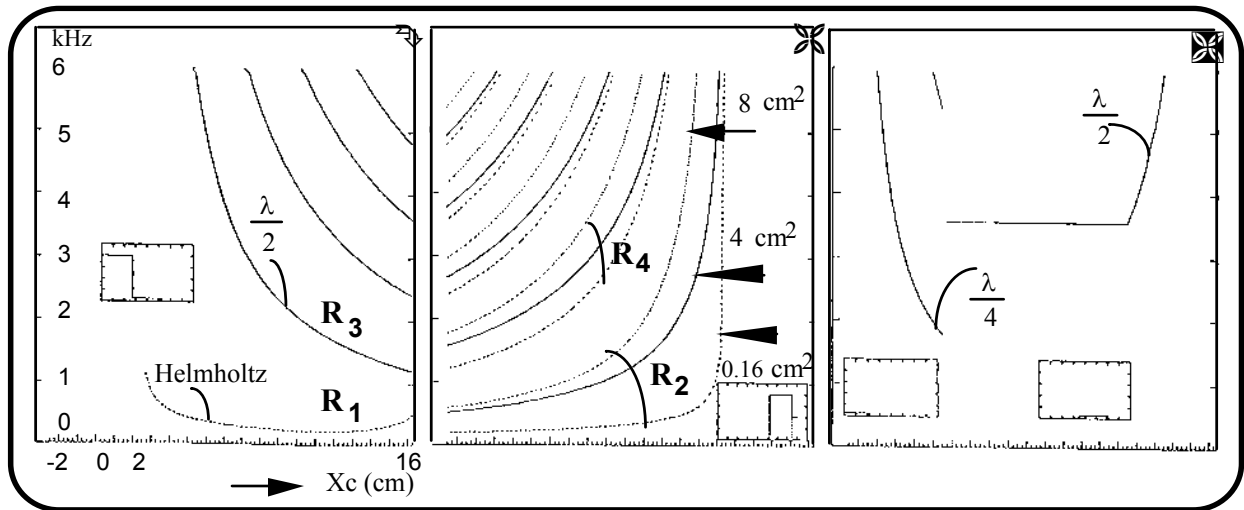
élevée. C'est notamment le cas de [u] qui possède une aire à la constriction importante.

#### 4.2. Nomogrammes générés pour chaque partie du conduit vocal

La figure 3 récapitule le concept de la résonance expliqué grâce aux nomogrammes de G. Fant, à partir de la cavité arrière, de la cavité avant et de la

constriction de la langue. Les formants ont été détectés par un algorithme décrit par Badin et al. , 1989.

Conformément à la stratégie de Fant, nous traçons les trois nomogrammes. L'étude de l'ensemble des cas se fait entre les deux valeurs extrêmes de  $L_2$  et de  $L_4$ . Le  $X_c$  est mesuré avec un pas de 0.25 cm.



**Figure 3 :** Nomogrammes pour les différentes cavités seules. De la gauche vers la droite nous avons : la cavité arrière, la cavité avant et la constriction de la langue.  $R_1, \dots, R_4$  sont les résonances de ces deux cavités

Cette figure montre les comportements des résonances de la :

- ❶ cavité arrière seule qui produit des résonances  $\frac{\lambda}{2}$

c'est-à-dire  $\left(\frac{nc}{2L_4}\right)$  plus une résonance de Helmholtz entre la cavité arrière et la constriction ;

- ❷ cavité avant seule, dans le cas de  $A_1 = 8 \text{ cm}^2$  (maximum d'aperture), cette cavité produit des résonances  $\frac{\lambda}{4}$  correspondant à  $\frac{(2n-1)c}{4L_2}$  ; dans le cas d'une faible aperture aux lèvres  $A_1 = 0.16 \text{ cm}^2$ , nous obtenons des résonances  $\frac{\lambda}{2}$  soit  $\left(\frac{nc}{2L_2}\right)$  puisque la cavité est presque fermée aux deux extrémités. Pour le cas intermédiaire  $A_1 = 4 \text{ cm}^2$ , les résonances ont un comportement intermédiaire ;

- ❸ Constriction du tube quand il est ouvert aux deux extrémités et fermée à la glotte. Elle produit des résonances :

- $\frac{\lambda}{2}$  ou  $\left(\frac{nc}{2L_3}\right)$  pour  $-2.5 \text{ cm} < X_c < 17.5 \text{ cm}$  ;
- $\frac{\lambda}{4}$  ou  $\left(\frac{(2n-1)c}{4L_3}\right)$  pour  $X_c < 2.5 \text{ cm}$ .

Nous adoptons les notations suivantes pour la désignation des résonances :  $R_1$ - $R_3$  et  $R_2$ - $R_4$  qui sont respectivement les deux premières basses fréquences de la cavité postérieure et de la cavité antérieure.

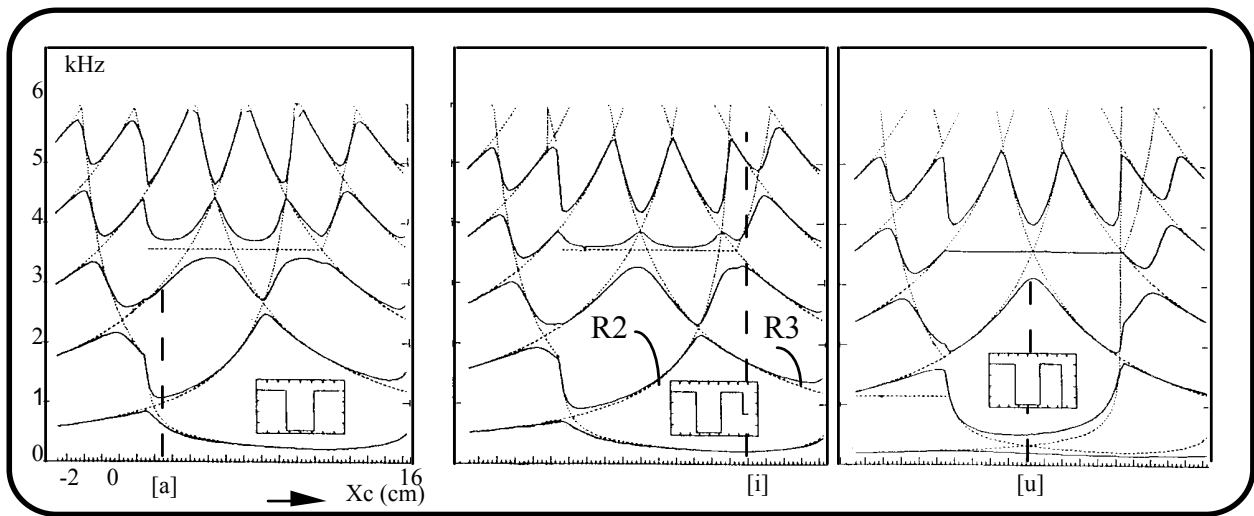
#### 4.3. Nomogrammes d'un système complet : affiliation

Nous essayons de mettre en évidence la notion d'affiliation d'un formant à telle ou telle cavité. Pour cela, nous évaluons les comportements des formants Associés aux différentes cavités du modèle à quatre tubes, puis nous les superposons sur les nomogrammes du modèle afin de mettre en évidence les modifications dues à l'effet du couplage entre cavités.

Les résonances respectives de la cavité avant, de la cavité arrière et de la constriction de la langue sont superposées. Elles représentent un suivi de résonances du conduit vocal entier (figure 4). La proximité d'un formant du système complet avec une résonance d'une certaine cavité démontre clairement l'affiliation. Des croisements de résonances ayant lieu quand une résonance de la cavité arrière se rapproche de celle de la cavité avant, produisant une affiliation qui change de formants correspondants (supposé étiqueté dans un ordre ascendant). Un grand couplage entre la cavité avant et la cavité arrière peut masquer une affiliation qui change (voir le cas d'une constriction au milieu [u] et la constriction arrière [a]). Plus intéressante est la convergence entre  $R_2$ - $R_3$  dans le cas d'une constriction

avant de [i]. Le demi onde de la cavité arrière traverse la résonance basse de la cavité avant

(intermédiaire entre une résonance de Helmholtz et un quart d'onde).



**Figure 4 :** Nomogrammes pour le système complet à 4 tubes. De la droite vers la gauche nous avons l'aperture : maximale ( $A_1 = 8 \text{ cm}^2$ ) ; moyenne ( $A_1 = 4 \text{ cm}^2$ ) et faible ( $A_1 = 0.16 \text{ cm}^2$ ). Celle-ci correspond respectivement aux configurations des voyelles cardinales.

#### 4.4. Quelques exemples de Points Focaux

Nous remarquons deux zones de convergence formantique,  $F_1$ - $F_2$  dans la région arrière et  $F_2$ - $F_3$  dans la région avant, pour les séquences [a] → [i] et [y] → [u]. Nous appelons ces régions de convergence les *Points Focaux* : PF. Ceux-ci étant obtenus quand les affiliations des résonances basses et hautes changent d'une cavité à une autre lors du déplacement de la constriction de la langue (figure 5).

Le PF[1] (tube ouvert) situé aux environs de  $X_c = 3.5 \text{ cm}$ , correspond au croisement d'une courbe de résonance  $\frac{\lambda}{4}$  de la cavité avant avec la courbe de la résonance de Helmholtz cavité arrière + constriction, qui montre une stabilité intermédiaire pour ce PF correspondant à la configuration de la voyelle [a].

Le PF[2] (tube intermédiaire) situé aux environs de  $X_c = 11 \text{ cm}$ , est le résultat d'un croisement d'une courbe de résonance  $\frac{\lambda}{4}$  de la cavité avant avec une résonance  $\frac{\lambda}{2}$  de la cavité arrière. Nous avons localisé la voyelle [i] au voisinage de ce point. Il est à souligner que le positionnement du [i] selon Stevens devrait se situer au point PF[2] alors que l'on verra par

la suite que le phonotype du [i] est toujours post-focal (K. Stevens, 1972).

Le PF[3] situé aux environs de  $X_c = 7 \text{ cm}$ , résulte du croisement des courbes des résonateurs de Helmholtz, constitué par l'ensemble cavité avant + lèvres et celui de cavité arrière + constriction, et la fréquence  $F_3$  affiliée à la cavité arrière et résonnant selon le mode  $\frac{\lambda}{2}$

Connaissant la position de la constriction pour ces voyelles nous pouvons déduire les valeurs approximatives des trois premiers formants.

Le PF[4], situé aux environs de  $X_c = 12 \text{ cm}$ , est le résultat du croisement de la résonance de Helmholtz, constitué par l'ensemble cavité avant + lèvres avec la courbe de résonance  $\frac{\lambda}{2}$  de la cavité arrière. Il correspond à la configuration de la voyelle [y].

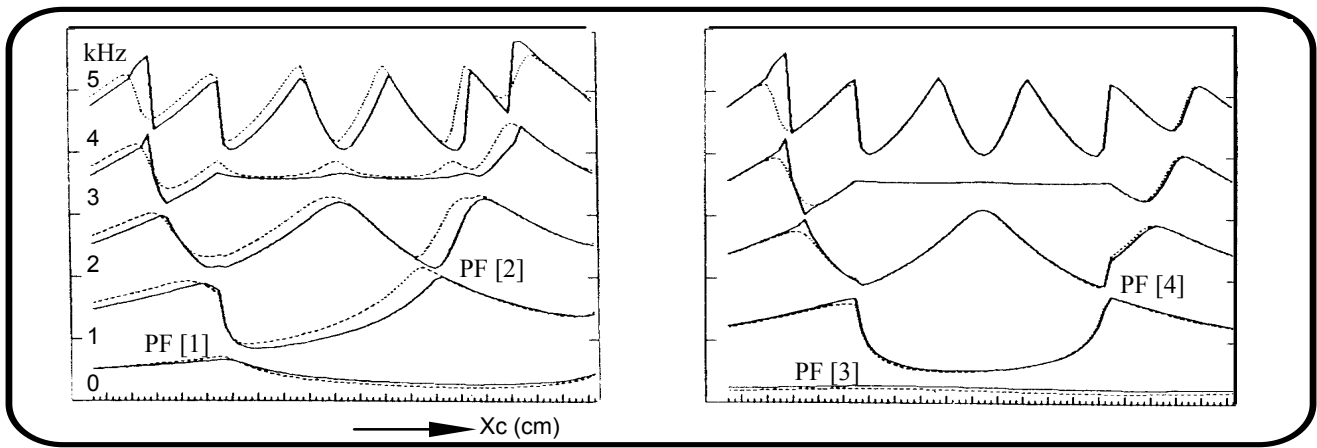


Figure 5 : Nomogrammes des séquences [a] → [i] et [y] → [u] et représentation de leurs Points Focaux.

#### 4.4.1. Remarques sur les stabilités formantiques

L'intérêt des PF réside dans l'explication qu'elle apporte à la notion de stabilité telle qu'elle a été introduite par Stevens. En effet, ces PF présentent dans la transformation articulatoire-acoustique des stabilités très diverses, essentiellement liées aux types de résonances acoustiques mises en jeu et au degré de couplage entre cavités.

#### 4.4.2. Points Focaux et stabilité articulatoire-acoustique

Il existe des causes possibles pour la stabilité articulatoire-acoustique des Points Focaux :

- la première correspond à un grand couplage entre les cavités avant et arrière, s'étend à la convergence formantique, et ainsi augmente la stabilité ;
- la seconde est liée aux résonances de Helmholtz qui se croisent dans une région où ces résonances ne sont pas sensibles aux variations de  $X_c$ , ce qui peut conduire aussi à la stabilité des Points Focaux.

L'exemple d'une petite ouverture aux lèvres montre un PF stable dû aux résonances de Helmholtz. Ces résonances associées au résonateur de la cavité arrière plus la constriction de la langue et avec le résonateur cavité avant plus constriction des lèvres, s'étendent dans  $F_1$  et  $F_2$ , qui sont déjà peu sensibles à  $X_c$  dans cette région. Ceci correspond à la voyelle [u]. Sur le niveau acoustique, ce résultat confirme la forte stabilité de [u] dans le plan  $F_1$ - $F_2$ , observée avec un modèle articulatoire pour de grandes variations de la position de la constriction.

Un modèle articulatoire plus réaliste montre que [i] est une voyelle instable dans la transformation articulatoire-acoustique. Cependant cette instabilité est limitée à la zone du côté de la glotte du Point Focal.

Le seul moyen pour obtenir un PF instable est un petit couplage, c'est-à-dire une petite aire de constriction entre deux cavités dans une région où les résonances de ces deux cavités seules sont déjà très sensibles aux variations de  $X_c$ . Ceci exclut les résonances de Helmholtz.

### 5. Exploitation des continuités pour les cibles de résonances

Les résonances sont un compromis entre une synthèse par règles et celle par contrôle moteur, car les objets manipulés tiennent compte du mouvement de production sous-jacent (affiliation). Nous considérons l'introduction des résonances dans l'élaboration des règles comme une nouvelle représentation du signal de parole qui facilite l'écriture des règles de synthèse (M. Guerti, 1998). Cette représentation est fondée sur le contrôle explicite des configurations acoustiques des résonances du conduit vocal.

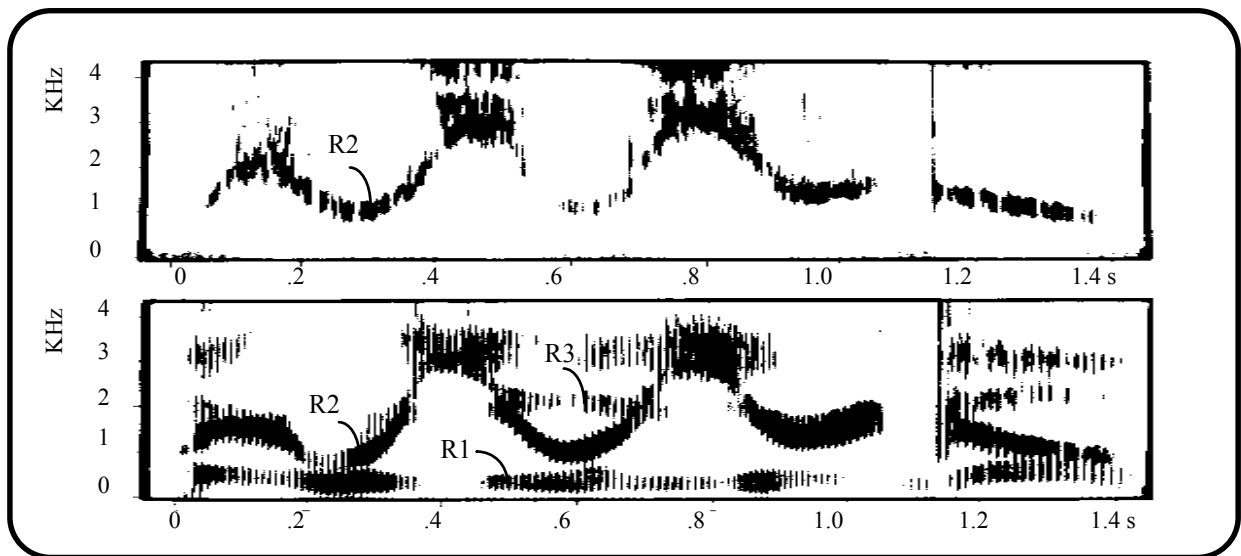
La particularité de notre travail réside dans l'espace de représentation spectrale : les cinq fréquences formantiques sont définies en termes de résonances, mais leur ordre d'étiquetage est modifié en respectant des continuités spectrales : une affiliation de ces résonances à des cavités du conduit vocal peut produire des croisements de formants pour certains couples de sons [Guerti & Bailly, 1991 ; Bailly & Guerti, 1991].

Des données spectrographiques montrent qu'il est possible d'estimer la fréquence de résonance de la cavité avant, à partir d'une information dans le signal de parole : Kuhn a montré que *l'information du lieu d'articulation se trouve dans les transitions et des bursts au niveau de  $F_2$  et parfois de  $F_3$* . (G. M. Kuhn, 1979). Ces transitions peuvent être réinterprétées comme indicateurs de la déformation de la cavité avant. Pour ceci, Kuhn compare l'analyse spectrographique de parole naturelle et chuchotée (figure 8). Deux propriétés intéressantes se dégagent :

- La première est que la parole chuchotée est hautement intelligible et qu'elle est essentiellement à formant unique. Ce dernier est proche de la fréquence d'un formant proéminent ou d'un groupe de formants pour des articulations correspondantes dans une parole naturelle ;
- La seconde montre que ces spectrogrammes illustrent la similarité remarquable du burst au niveau des deux types de parole.

La figure 7 montre des analyses spectrographiques de la phrase : *Where were you a year ago ?* enregistrée dans deux conditions d'excitation : une parole chuchotée (en haut) et une naturelle (en bas). Le spectrogramme d'en haut indique la présence de deux composantes dans cette parole. La composante la plus évidente varie dans les fréquences allant de 700 Hz à 3 KHz. Une autre composante est fixée au-dessus de 3500 Hz est moins visible quand l'arrondissement des lèvres augmente. La composante variable peut être interprétée comme la résonance R<sub>2</sub> de la cavité avant. Les variations de la fréquence de la cavité avant semble

montrer les changements de la position de la constriction (de la vélaire vers la palatale) et des changements dans l'ouverture des lèvres (de la forme arrondie vers la rétractée). Par exemple, une résonance quart d'onde à 700 Hz peut indiquer que la cavité avant a une longueur fonctionnelle d'environ 12.6 cm et à 3 kHz, une longueur d'environ 2.9 cm. Il n'est pas surprenant que la résonance de la cavité avant puisse varier *continuellement* dans une parole chuchotée, donc croiser d'autres fréquences puisque la constriction de la langue est extrême. Cette résonance intéressante peut être tracée facilement dans une parole naturelle enregistrée.



**Figure 7 :** Comparaison spectrographique de la phrase : *Where were you a year ago ?* Pour deux conditions d'excitation : parole chuchotée (en haut) et parole normale (en bas), d'après Kuhn [G. M. Khun, 1975].

Puisque les affiliations des formants à une cavité peuvent changer, il est important de montrer que les différents formants peuvent contribuer hautement à la perception du lieu d'articulation, dépendamment de leur degré d'affiliation avec la cavité avant.

Kuhn met le point sur la résonance fondamentale de la cavité avant qui ne peut pas toujours être associée avec le même formant étiqueté. Elle peut être associée à F<sub>2</sub> dans [u], mais elle est fortement associée à F<sub>3</sub> voire à F<sub>4</sub> dans [i] (si la constriction de la langue est extrêmement avancée). Ces observations permettent de dégager une règle générale sur la cavité avant : *elle est déterminante pour la forme entière du spectre.*

Hermansky & al. vont plus loin encore en émettant l'hypothèse suivante (Hermansky et Broad, 1989) :

*The size and shape of vocal tract's front cavity is the primary carrier of linguistic information. The back cavity geometry is its causal consequence and contributes mainly speaker-dependent information. Both the linguistic and the speaker-dependent information is coded in formants (resonances of the whole vocal tract) in a non-trivial manner. The human auditory system simplifies the formant representation*

*by extracting the effective second formant F<sub>2</sub>'. F<sub>2</sub>' carries in its frequency and bandwidth information about the front cavity and consequently the bulk of the linguistic message.*

La notion de F<sub>2</sub>' provient d'un désir de représenter les voyelles naturelles dans un espace perceptuel équivalent à deux formants. A partir de données expérimentales dans lesquelles des techniques à deux formants ont été utilisées, Carlson et al. donnent des valeurs de F<sub>2</sub>' pour plusieurs voyelles suédoises. Les valeurs expérimentales de F<sub>2</sub>' se trouvent aux environs de 700 Hz pour le [u] et de 3 kHz pour le [i]. Ces valeurs limites et celles qui sont intermédiaires semblent s'étendre d'une manière proche à celle de la fréquence de résonance de la cavité avant (Carlson & al., 1973).

Pour les voyelles [i, e], des expériences ont montré que F<sub>3</sub> et F<sub>4</sub> sont perceptivement très importants. Leur représentation sur le plan F<sub>1</sub>-F<sub>2</sub> n'est donc pas optimale. Carlson & al. ont proposé de remplacer F<sub>2</sub> par une somme pondérée des fréquences de F<sub>2</sub>, F<sub>3</sub> et F<sub>4</sub>, appelée F<sub>2</sub>' (Carlson & Granström, 1973) :

$$F_2' = \frac{F_2 + C\sqrt{F_3 F_4}}{(1+C)}$$

$$\text{avec } C = \left(\frac{F_1}{500}\right)^2 \left(\frac{F_2-F_1}{F_4-F_3}\right)^4 \left(\frac{F_3-F_2}{F_3-F_1}\right)^2 \quad [3]$$

$F_2'$  est largement supérieur à  $F_2$  pour les voyelles antérieures non arrondies, notamment le [i], alors qu'il est égal à  $F_2$  pour les voyelles postérieures et antérieures labialisées.

De la même façon pour certaines études sur la perception, il est souhaitable d'attribuer à  $F_2$  les formants les plus élevés d'une voyelle antérieure comme un groupe formantique unique. L'utilisation de  $F_2'$  aurait sans doute conduit à la réduction ou même à l'élimination des zones communes à [i - y].

L'opinion s'accorde sur le fait que la fréquence de résonance de la cavité avant peut être ce que prédisait  $F_2'$  (c'est-à-dire  $F_2' = R_2$ ). Si c'est ainsi, elle peut être appropriée à la fréquence de résonance de cette cavité estimée et utilise la formule proposée (Equation 3).

Les méthodes pour prédire  $F_2'$  peuvent être prévues pour être plus cohérentes pour les sons les plus constrictifs, où les fréquences peuvent bouger plus continûment. La résonance de la cavité avant peut fournir une articulation correspondante à  $F_2'$ , qui jusqu'ici a été motivée principalement par des considérations psychoacoustiques.

Dans une synthèse bi-formantique, les valeurs moyennes de  $F_2$  et  $F_3$  ont été utilisées pour les voyelles antérieures, où la résonance de la cavité avant dans le cas naturel peut être plus fortement associée avec  $F_3$ . Par exemple le [i] bi-formantique a un  $F_2$  à 2880 Hz, alors que le  $F_2$  naturel semble être localisé aux environs de 2300 Hz et le  $F_3 = 3$  kHz.

## 6. Résonances vocaliques

Nous avons utilisé le modèle de Fant, avec une longueur du conduit vocal égale à 15 cm (plus 2 cm pour les lèvres) et  $2.1 \text{ cm} < X_c < 11.9 \text{ cm}$ .

Un corpus Voyelles-Voyelles (10 voyelles orales) a été enregistré par deux locuteurs masculins et analysé selon la technique : LPC Pitch Synchrone à glotte fermée (T. Barbe, 1990).

Les résonances  $R_1$  et  $R_2$  jouent un rôle important pour suivre des cavités par affiliation :  $R_1$  est affiliée à la cavité arrière et  $R_2$  à la cavité avant. Conformément à cette représentation, des cibles peuvent être données afin de décrire la synthèse formantique, en incluant des règles pour tous les sons.

Une observation minutieuse des transitions Voyelles – Voyelles, en utilisant une interprétation des nomogrammes vocaliques en termes de résonances, nous montre une conception spécifique de l'espace de représentation pour les voyelles (figure 8).

Nous pouvons décrire les transitions concernant ce corpus avec une cible par phonème (table 1) : seulement les voyelles [i] et [e] ont des valeurs différentes par rapport à l'étiquetage formantique habituel. Ces voyelles ont respectivement :  $R_2 \approx 3$  kHz et 2700 Hz. Avec une telle représentation, la troisième résonance (exception faite pour [i] et [e]) est relativement constante. Elle est fortement corrélée avec l'aperture. Les voyelles ouvertes ont  $R_3 \approx 3215$  Hz, par contre les voyelles fermées ont un  $R_3 \approx 2095$  Hz. La représentation est compatible avec l'étude du concept de la théorie perceptuelle motrice :  $F_2'$

A titre d'exemple expliquons le cas [i-y] c'est-à-dire l'effet de la labialisation ou l'arrondissement des lèvres. Quand on passe du [i] vers le [y], le volume et l'ouverture de sortie de la cavité arrière restent inchangés (résonance demi-onde) donc  $R_1$  est le même pour les deux voyelles ( $R_1 \approx 250$  Hz). Par contre l'ouverture de sortie de la cavité avant diminue du fait de la labialisation :  $R_2$  baisse de 3000 Hz à 1700 Hz. Nous remarquons que le  $R_2$  de [i], c'est-à-dire le  $F_3$  continue vers le  $F_2$  de [y] qui devient égal à  $R_3$ .



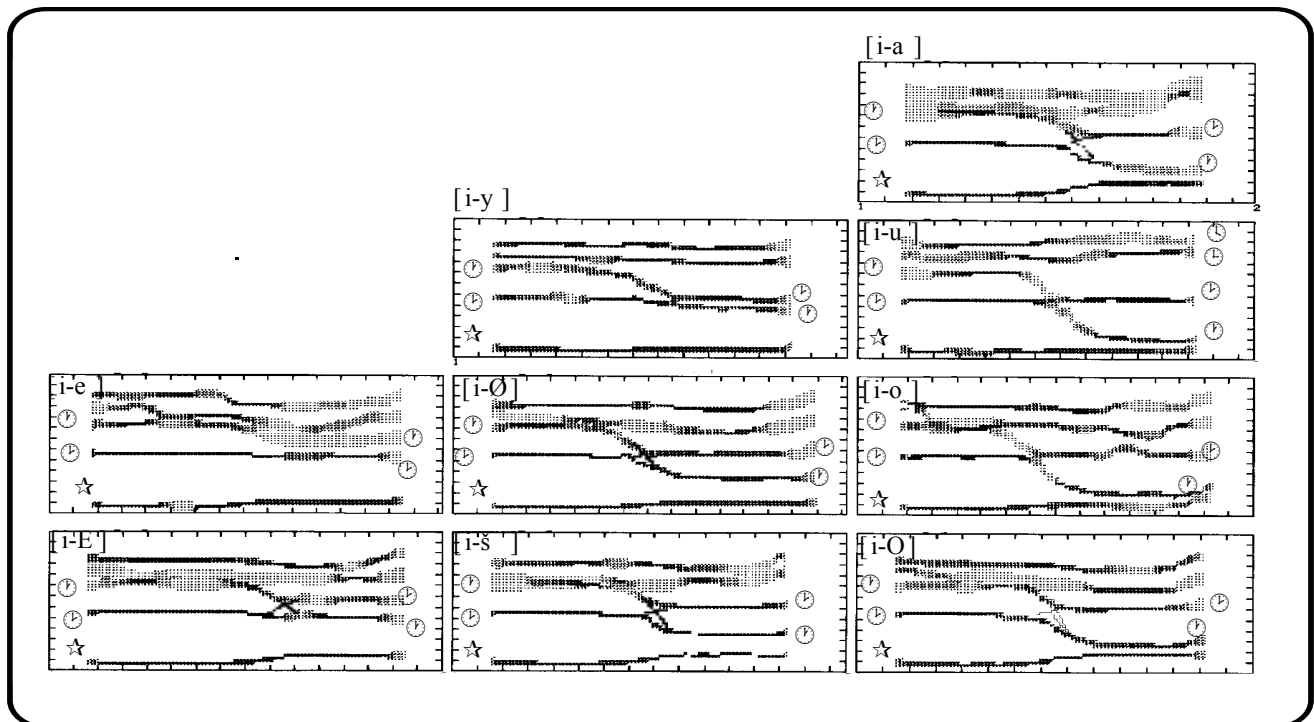


Figure 8 : Suivi de formants étiquetés en résonances dans les séquences [i-voyelles]. Les  $R_i$  ( $i = 1, 2, 3, 4, 5$ ) sont représentées respectivement par : **12345**

Table 1 : Valeurs des résonances vocaliques déterminées à l'aide du modèle à quatre tubes de G. Fant

Phonèmes	[i]	[e]	[ɛ]	[a]	[y]	[ø]	[œ]	[u]	[o]	[ɔ]
R <sub>1</sub>	0250	0320	0590	0620	0250	0360	0480	0250	0370	0460
R <sub>2</sub>	3000	2700	1900	1240	1720	1600	1420	0750	0800	1000
R <sub>3</sub>	2000	2000	3300	3330	2060	2150	3200	2100	2260	3000
R <sub>4</sub>	3400	3300	2300	2500	3000	3000	2300	3000	3000	2600

## 7. Conclusions

Suite aux recherches de Fant sur les affiliations entre les formants et les cavités du conduit vocal, nous proposons de modéliser des trajectoires formantiques en termes de trajectoires de résonances sous-jacentes. Cette représentation permet de suivre l'évolution de la forme du conduit vocal dans la production de la parole et permet une meilleure conversion acoustico-articulaire. Elle facilite l'obtention de règles simples pour les synthétiseurs à formants.

Suivant cette approche - qui est souvent utilisée pour la lecture de sonagrammes - nous cherchons en premier à obtenir les points d'ancrage de notre recherche en étudiant les cibles résonantiques des voyelles orales en vue d'étendre cette description aux voyelles nasales et aux consonnes intervocaliques.

Les nomogrammes obtenus par le modèle à quatre tubes de Fant peuvent éclaircir d'une manière intéressante les systèmes vocaliques acoustiques. Nous avons essayé d'étudier des régions particulières de ces nomogrammes qui correspondent aux échanges cavité

affiliation formantique et qui sont liés aux Points Focaux. Une étude sur des mots artificiels a confirmé l'existence de ces points dans la parole naturelle.

En effet, entre la première description articulaire classique de l'espace vocalique et les problèmes soulevés par l'inversion du conduit vocal (la reconstitution de la forme articulaire à partir du signal acoustique) bon nombre d'explications ont été apportées, mais d'autres problèmes apparaissent, peut-être sans que les relations ne soient faites entre les différents champs de connaissances de la parole. Avec les progrès de la connaissance des phénomènes de production, la tentation est grande de les confronter à des modélisations plus réalistes ou à des mesures directes sur le conduit vocal.

Il est intéressant d'évaluer, dans tous les cas possibles les affiliations cavité-formants, en particulier dans la perspective de vérifier la notion de locus, c'est-à-dire les fréquences correspondant aux cas extrêmes où la constriction est complètement fermée.

## Références

- P. Badin, P. Perrier, L.J. Boë & C. Abry, 1989 : Vocalic nomograms : acoustic and articulatory considerations upon Formant Convergences. *J. Acoust. Soc. Am.* 87 (3), 1290-1300.
- G. Bailly & M. Guerti, 1991, Actes du XIIème ICPhS, 19-24 Août, Aix-En-Provence, France. Vol. 2, 506-509.
- T. Barbe, 1990 : Méthodologie et outils pour la mise en œuvre automatique d'une synthèse de parole de haute qualité. Thèse de Dr INP-Grenoble France.
- R. Carlson., G. Fant & B. Granström, 1973 : Two formant models, pitch, and vowel perception. Paper discussed at the Symp. Aud. Anal. Percept. Speech, Leningrad, 21-24 August.
- G. Fant, 1960 : Acoustic theory of the speech production. Edition The Mouton & Co.
- M. Guerti & G. Bailly, 1990 : Anticipation et rétention dans les mouvements vocaliques du Français. XVIIIème JEP. 28-31 Mai, Montréal-Canada, 292-295.
- M. Guerti & G. Bailly, 1991 : Synthesis by rule using COMPOST : modelling resonance trajectories. European Conference on Speech Technology (EUROSPEECH). 24-25 September, Genova (Italy), 43-46.
- M. Guerti, 1998 : Speech synthesis by rule, 8th International Conference on Computer Theory and Applications ICCTA'98, 15-17 September, Alexandria - EGYPT, III.12-III.15.
- M. Guerti, 1999 : Synthèse par règles à partir de diphtongues formantiques, Signaux Systèmes et Automatique SSA2'99, 10-12 Mai, Blida-Algérie, 110-116.
- G.M. Kuhn, 1975 : On the front cavity resonance and its possible role in speech perception. *J. Acoust. Soc. Am.* 58 (2), 428-433.
- H. Hermansky & D. Broad, 1989 : The effective second formant F2' and the vocal tract front cavity. ICASSP, 480-483.
- K.N. Stevens, 1972 : Bases for universals in the properties of the speech production and perception systems. Proc. of the IX th ICPhS. Vol 2, Copenhagen, 53-59.