

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/320516553>

# Modélisation du conduit vocal

Book · January 2007

CITATION

1

READS

81

1 author:



[Alain Ghio](#)

French National Centre for Scientific Research

149 PUBLICATIONS 730 CITATIONS

[SEE PROFILE](#)

Some of the authors of this publication are also working on these related projects:



Speech disorders assessment [View project](#)



Voice Disorders Assessment [View project](#)

A. Ghio - Ingénieur de recherche

Laboratoire Parole et Langage, CNRS UMR 6057  
Université de Provence  
29, Avenue Robert Schuman  
13621 Aix-en-Provence, Cedex 1.

Courriel : alain.ghio@lpl.univ-aix.fr

## I. De l'intérêt et de la diversité des modèles du conduit vocal

### 1. De l'intérêt des modèles

Un modèle est une représentation abstraite, simplifiée, d'un objet observable, qui peut se ramener à des paramètres et des grandeurs. Par exemple, dans le cas d'une voiture, le modèle correspondant à cet objet peut être résumé par un ensemble de grandeurs définies : longueur, largeur, hauteur, masse, puissance, coefficient de traînée aérodynamique ( $c_x$ ), couleur... Il s'agit là d'un modèle descriptif de l'objet. Si à présent le modèle incorpore non seulement des grandeurs descriptives mais aussi des lois de la mécanique, il pourra de plus prédire la vitesse de pointe du véhicule à partir de la puissance, de la masse, du  $c_x$ ... Une fois la modélisation établie, l'objet et son modèle peuvent être étudiés ou comparés, et ce de façon plus simple qu'en manipulant les données réelles souvent variables de façon opaque.

Dans l'élaboration d'un modèle, on oppose généralement deux approches : prédictive ou descriptive. Une approche prédictive a pour principe d'obtenir un résultat inconnu à partir de données explicatives. Ainsi, les modèles prédictifs de météorologie sont prévus pour fournir le temps à venir à partir de données barométriques, de température, de courants... Dans ce cas-là, ces données explicatives, qui sont les paramètres d'entrée du modèle, sont considérées comme primordiales et incontestables sachant qu'une erreur de prédiction sera alors vue comme une faiblesse de la modélisation. Dans les approches descriptives, les modèles servent à représenter, de façon organisée et synthétique, des données réelles et l'objectif est de rendre compte,

si possible de manière interprétable, de cette masse d'informations. Dans ce cas-là, ce sont les données observées qui sont considérées comme des données d'entrée, primordiales et incontestables. Les paramètres du modèle sont alors ajustés pour rendre compte au mieux de ces données, quitte à perdre une partie de la capacité d'interprétation du modèle. Les modèles les mieux aboutis sont ceux qui mêlent les deux approches : une bonne modélisation prédictive suppose au moins la prédiction de situations observées, c'est-à-dire une bonne description. Inversement, une bonne description serait parfaitement vaine si elle ne servait pas au moins de diagnostic, d'explication des mécanismes ou de généralisation du modèle. Dans la plupart des cas, les modèles peuvent être abordés sous la forme de schémas fonctionnels composés de blocs connectés par des lignes d'action. On y retrouve des variables d'entrée qui peuvent être les commandes du modèle, et des variables de résultats en sortie. Les blocs sont le moteur du modèle, intégrant la connaissance ou la description du phénomène modélisé. Selon l'approche choisie, ces boîtes peuvent s'avérer transparentes, si les connaissances sont explicitées ou alors totalement noires, sans capacité d'interprétation des mécanismes.

### 2. La diversité des modèles de production de la parole

La modélisation de la production de parole est un champ d'investigation qui a donné lieu à de nombreux travaux et il apparaît difficile de présenter ces contributions de façon exhaustive. Un bilan historique est toutefois disponible (Boë *et al.*, 1994). Les premiers modèles sont fondés sur des approches purement acoustique et géométrique avec un degré plus ou moins important d'anthropomorphisme. Plus tard, sont apparus des modèles qui visaient

la simulation physique du comportement des articulatoires en tenant compte de connaissances sur le fonctionnement des organes et en introduisant des contraintes articulatoires explicites. Plus récemment, des modélisations prennent leur source dans la biomécanique et les neurosciences avec une approche de type « contrôle moteur ». Dans tous les cas, selon le type de modélisation considéré, la finalité peut être très différente : simuler le conduit vocal pour comprendre les mécanismes mis en jeu dans la production de parole ou avoir pour ligne de mire des techniques de synthèse de la parole. Le premier objectif concernera plus la modélisation prédictive alors que le deuxième sera plus descriptif.

Dans le cadre de l'étude de la production de parole, l'intérêt d'une modélisation du conduit vocal se porte plus sur une approche prédictive de type articulatoire la plus proche possible de la réalité, ceci pour mieux appréhender les mécanismes de production de parole avec sa projection acoustique. Si les modèles de fonctionnement des cordes vocales (cf. chapitre 7), ou plus généralement du larynx, sont relativement pris en considération dans l'étude des dysphonies, les modèles articulatoires sont pour le moment quasi absents dans l'exploration des dysarthries. Cela provient essentiellement du fait qu'ils n'atteignent pas un niveau suffisant de similarité avec la réalité anatomique, fossé qui pourrait se combler avec l'apparition récente de modèles neurophysiologiques tridimensionnels (3D).

De façon plus générale, plusieurs niveaux de modélisation du conduit vocal peuvent être considérés (Fig. 1). Un modèle acoustique se pilote, en entrée, par des indices phonétiques ou acoustiques afin de générer le signal sonore en sortie. Un modèle géométrique considérera des formes de cavités ou fonctions d'aire du conduit vocal comme variables prédictives du résultat acoustique. Un modèle plus articulatoire utilisera des paramètres d'entrée de position/vitesse des organes. Un modèle prédictif neurophysiologique aura des commandes d'entrée d'activation/inhibition de modèles de muscles actionnant les organes phonatoires ou articulatoires. Les modèles peuvent être mis en cascade et on peut imaginer combiner plusieurs modules, les sorties de l'un étant les entrées d'un autre. Ainsi, il est envisageable d'initier des commandes d'activation/inhibition qui, via un modèle de muscles, permettent d'obtenir des positions et vitesse d'organes qui, via un modèle articulatoire, génèrent une forme géométrique de conduit vocal ou une fonction d'aire qui, via un modèle géométrique, fait émerger des indices acoustiques pour finalement générer un signal sonore (Fig.1).

Nous passerons en revue les différentes modélisations en commençant par les modèles acoustiques, pour finir avec des modèles neurophysiologiques. En revanche, nous passerons sous silence les approches faisant une analogie entre le conduit vocal et l'équivalent électrique (Stevens *et al.*, 1953) car, bien qu'intéressantes dans la caractérisation des phénomènes physiques, ces méthodes s'éloignent du cadre de cet ouvrage.

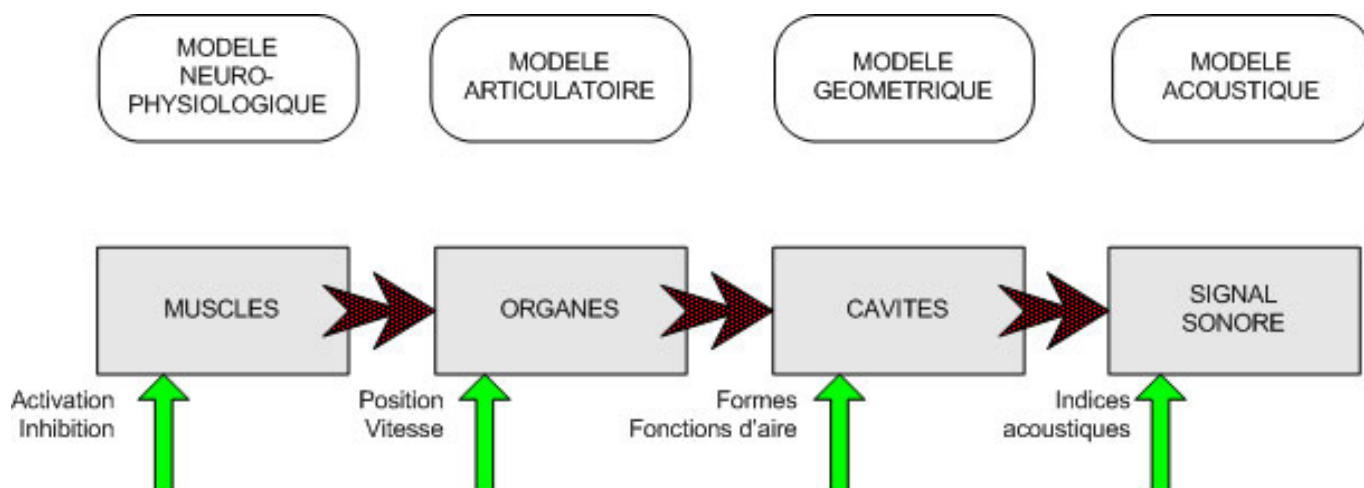


Fig.1. Les différents niveaux de modélisation du conduit vocal.

La mise en cascade des modèles est possible. Les paramètres d'entrées d'un modèle sont soit des variables explicatives externes (flèches verticales), soit des sorties d'un modèle de plus haut niveau (flèches horizontales).

## II. La modélisation acoustique du conduit vocal

### 1. Le synthétiseur de Klatt

La modélisation acoustique de la parole consiste à produire une séquence de sons de parole en fonction de paramètres acoustiques (Fig. 1). Ce type de modélisation a été abordé essentiellement à des fins de synthèse vocale. Le synthétiseur de Klatt (1980) reste le système le plus représentatif et comporte dans sa version de 1980 une quarantaine de paramètres de contrôle dont la valeur doit être modifiée dans le temps à intervalles très courts lors de la production d'une séquence sonore dynamique (Fig. 2). Hormis la source périodique et le générateur de bruit, pilotés par une dizaine de paramètres, l'expérimentateur doit définir explicitement la fréquence, l'amplitude et la bande passante des résonateurs qui modélisent le comportement des cavités du conduit vocal (cf. chapitre 8). Pour obtenir un comportement dynamique, il est alors nécessaire de renouveler périodiquement (p. ex. : toutes les millisecondes) cette définition des paramètres. Le modèle général intègre un petit ensemble de règles formulées dans le domaine acoustique uniquement et permet une grande liberté quant au résultat sonore, autorisant même la génération de configurations phonétiques aberrantes s'éloignant d'une production réelle de parole.

Au vu de la figure 2, nous pouvons formuler quelques réserves. La possibilité d'utiliser une fonction de transfert du conduit oral par la connexion de résonateurs placés en parallèle démontre une finalité de simulation plus que de « reproduction » de la réalité. En effet, l'anatomie du conduit repose sur une mise en série des résonateurs, excepté pour la nasalité. Inversement, la possibilité de connexion des résonateurs en cascade se rapproche plus de l'anatomie mais le résonateur nasal, placé en tête de la série, est une entorse à la morphologie du conduit. L'approche de Klatt est clairement descriptive, avec une finalité de synthèse plus que de modélisation des mécanismes sous-jacents. Il le mentionne d'ailleurs clairement dans son introduction de 1980 : « Il existe le besoin en psychologie et en sciences du langage d'un outil de recherche maniable permettant d'étudier la perception de la parole par le biais de la synthèse de parole ou de sons ayant des caractéristiques proches de la parole. » Son synthétiseur a été développé pour répondre à ces besoins de génération de stimuli, avec un objectif de perception plus que de modélisation de la production de parole.

### 2. Le synthétiseur HLSyn

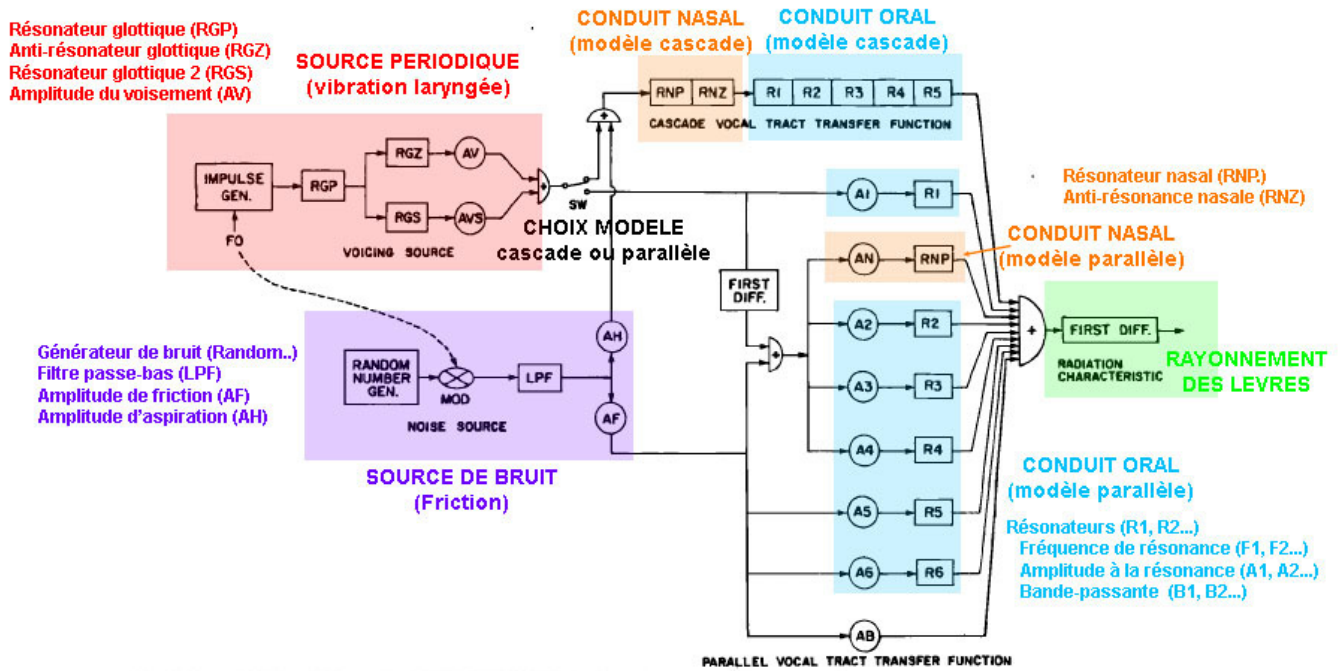
Le cas du dispositif HLSyn (Hanson et Stevens, 2002) nous apparaît ambigu dans la mesure où il annonce, par son nom, une vocation de High-Level Speech Synthesis qui laisse présager une synthèse pilotée par des paramètres de haut niveau au sens phonétique du terme. Ce point de vue est confirmé d'une part par le nombre réduit de paramètres de contrôle (de 10 à 13 selon la version) qui milite pour un niveau de représentation supérieur, d'autre part par certains paramètres indéniablement représentatifs du niveau articulatoire : section de l'ouverture vélaire, surface moyenne de l'ouverture glottique au niveau membranaire et au niveau cartilagineux, section de la constriction labiale, section de la constriction apicale... Mais parallèlement à ces paramètres articulatoires apparaissent les 4 premiers formants comme données d'entrée du modèle, de façon similaire à l'approche de Klatt, ce qui enlève une certaine crédibilité non pas au résultat, mais à l'approche dite de « haut niveau ». Le modèle HLSyn avec ses paramètres mixtes acoustique et articulatoire doit être considéré, à notre sens, comme un modèle *ad hoc* permettant un pilotage simplifié d'un synthétiseur à formant comme celui de Klatt. Ce point de vue est d'ailleurs confirmé par le schéma fonctionnel du dispositif (voir <http://sens.com/hlsyn/>). Il est bien évident que ces réserves ne remettent aucunement en cause l'intérêt scientifique d'un tel dispositif, notamment pour la génération de stimuli synthétiques de parole.

## III. La modélisation à tubes du conduit vocal

La modélisation géométrique se fonde sur le principe « source-filtre » de production de parole qui postule que les sons de parole sont le résultat de la modification (filtrage), par le conduit vocal, de la source laryngée et/ou d'une source de bruit de friction. Les caractéristiques acoustiques des phonèmes sont liées d'une part aux sources, mais surtout à la configuration géométrique du conduit vocal, cette conformation étant dynamique au cours du temps (cf. chapitre 8). La connaissance des lois qui régissent la relation entre géométrie et conséquences acoustiques permet ainsi de proposer un modèle géométrique du conduit vocal capable de générer, à partir de paramètres géométriques, un signal acoustique en sortie.

Nous avons détaillé au chapitre 8 le modèle de Fant (1960) que l'on peut considérer comme un précurseur





Symbole	C/V	Min.	Typ.	Max.	Name
DU	C	30	200	5000	Durée du signal (ms)
NWS	C	1	5	20	Intervalle de mise à jour des paramètres (ms)
SR	C	5000	10000	20000	Fréquence d'échantillonnage du signal (Hz)
NF	C	1	5	6	Nombre de formants dans la configuration cascade
SW	C	0	0	1	0=Cascade, 1=Parallèle excité par AV
G0	C	0	47	80	Facteur de gain général du système (dB)
FO	V	0	100	500	Fréquence de la fondamentale (Hz)
AV	V	0	60	80	Amplitude du voisement (dB)
AVS	V	0	0	80	Amplitude du voisement quasi-sinusoidal (dB)
FGP	V	0	0	600	Fréquence du résonateur glottique "RGP" (Hz)
BGP	V	50	200	2000	Bande passante du résonateur glottique "RGP" (Hz)
FGZ	V	0	2500	5000	Fréquence de l'antirésonateur glottique "RGZ" (Hz)
BGZ	V	100	1500	9000	Bande passante de l'antirésonateur glottique (Hz)
BGS	V	100	200	1000	Fréquence du résonateur glottique "RGS" (Hz)
AH	V	0	0	80	Amplitude de l'aspiration (dB)
AF	V	0	0	80	Amplitude de la friction (dB)
F1	V	180	500	1300	Fréquence du 1er formant (Hz)
B1	V	30	60	1000	Bande passante du 1er formant (Hz)
F2	V	550	1500	3000	Fréquence du 2e formant (Hz)
B2	V	40	90	1000	Bande passante du 2e formant (Hz)
F3	V	1200	2500	4800	Fréquence du 3e formant (Hz)
B3	V	60	150	1000	Bande passante du 3e formant (Hz)
F4	V	2400	3800	4990	Fréquence du 4e formant (Hz)
B4	V	100	300	1000	Bande passante du 4e formant (Hz)
F5	V	3000	4500	6000	Fréquence du 5e formant (Hz)
B5	V	100	300	1500	Bande passante du 5e formant (Hz)
F6	V	4000	4990	6500	Fréquence du 6e formant (Hz)
B6	V	100	500	4000	Bande passante du 6e formant (Hz)
FNP	V	180	280	700	Fréquence du pôle nasal (Hz)
BNP	V	40	90	1000	Bande passante du pôle nasal (Hz)
FNZ	V	180	280	800	Fréquence du zéro nasal (Hz)
BNZ	V	40	90	1000	Bande passante du zéro nasal (Hz)
AN	V	0	0	80	Amplitude du formant nasal (dB)
A1	V	0	60	80	Amplitude du 1er formant (dB)
A2	V	0	0	80	Amplitude du 2e formant (dB)
A3	V	0	0	80	Amplitude du 3e formant (dB)
A4	V	0	0	80	Amplitude du 4e formant (dB)
A5	V	0	0	80	Amplitude du 5e formant (dB)
A6	V	0	0	80	Amplitude du 6e formant (dB)

Fig. 2. Le synthétiseur de Klatt (1980), un modèle acoustique de production de parole. Schéma fonctionnel (en haut) et paramètres de contrôle (en bas).

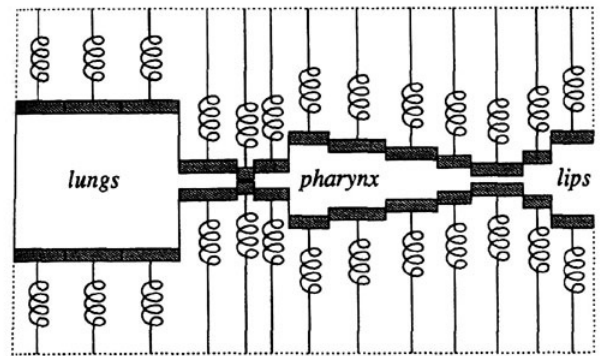
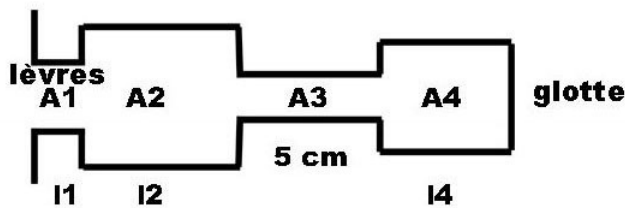


Fig. 3. Modèles à tubes du conduit vocal. À gauche, modèle de Fant (1960). À droite, modèle de Boersma (1993). Ce dernier intègre aussi la soufflerie pulmonaire et le larynx.

dans le domaine. Ainsi, en décrivant le conduit vocal comme 4 tubes (Fig. 3, gauche) de dimensions variables connectés entre eux, et en faisant varier les dimensions de ces tubes, le modèle de Fant est capable de prédire et générer un résultat acoustique en accord avec la réalité. Cette modélisation à l'aide de tubes a été reprise dans divers travaux, en particulier sous l'impulsion d'acousticiens retrouvant dans l'appareil vocal des analogies avec certains instruments de musique. Les innovations ont consisté à augmenter le nombre de tubes et à tenir compte de plus en plus de phénomènes aérodynamiques comme la viscosité, les pertes aux parois... Un modèle « spectaculaire » a été proposé par Boersma (1993) où c'est l'ensemble de l'appareil de production de la parole, du diaphragme aux lèvres, qui est vu comme une cascade de nombreux tubes (Fig. 3, droite). Chaque tube possède une section à géométrie variable dont la surface est réglée par un ensemble masse-ressort amorti, à l'image des modèles de cordes vocales (cf. chapitre 7). La sortie acoustique est calculée à partir de la dérivée temporelle du flux d'air aux lèvres et aux narines en fonction du mouvement des parois mobiles

Par généralisation, le tractus vocal étant vu comme la concaténation de tubes, on appelle fonction d'aire la représentation mettant en relation les caractéristiques géométriques de ces tubes (surface de la section) en fonction de la distance par rapport à la glotte (Fig. 4c). Il s'agit d'une figure à 2 dimensions (Fig. 4c :  $x$  = distance en cm,  $y$  = surface en  $\text{cm}^2$ ) mais qui explore l'espace tridimensionnel du tractus en tenant compte, de façon plus ou moins précise, de la dimension transverse du conduit, propriété non visible sur les coupes sagittales. À partir de telles données d'entrée, il est possible, par le biais d'équations mathématiques relatives à la propagation des sons et en opérant des simplifications, d'obtenir la fonction de transfert du tube, c'est-à-dire la fonction qui définit les propriétés de filtrage du conduit transformant un signal source (p. ex. : vibration laryngée) en signal de sortie, c'est-à-dire en données sonores articulées (Fant, 1960 ; Mrayati et Carré, 1976 ; Maeda, 1982 ; Badin et Fant, 1984). À ce niveau-là, il faut rappeler que la relation articulation/acoustique n'est pas bijective (cf. chapitre 8) car on peut montrer que, si à une forme précise du conduit correspond un résultat acoustique déterminé, en revanche, il existe une multitude de fonctions d'aire qui produisent le même résultat acoustique (Atal *et al.*, 1978).

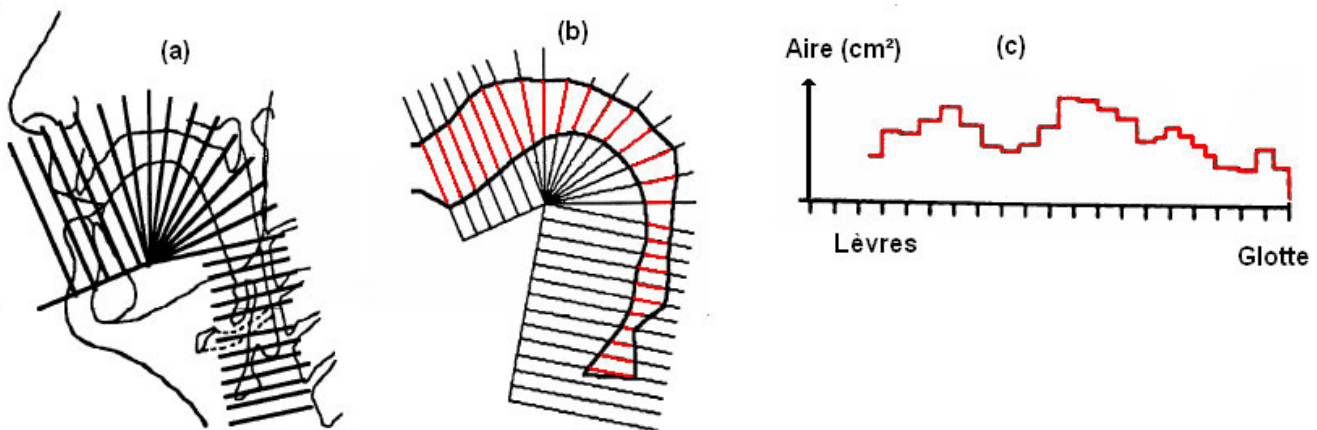


Fig. 4. (a) Coupe sagittale du conduit vocal. (b) Contour simplifié du conduit vocal. (c) Fonction d'aire (aire des coupes du conduit vocal en fonction de la distance par rapport à la glotte).

Dans ce type d'approche, l'obtention de données géométriques réelles du conduit est essentielle. Les premiers travaux associant fonction d'aire et fonction de transfert du conduit ont été proposés par Chiba et Kajiyama (1941) à partir de moulages. Les travaux de Sanchez et Boë (1984) ont eux aussi porté sur le moulage du conduit vocal d'un cadavre, procédé qui a autorisé la mesure d'aire et de dimensions transversales (Fig. 5b). Les travaux de l'Institut de Phonétique de Strasbourg restent, au niveau national, la référence en la matière (Bothorel *et al.*, 1986), notamment dans la mise à

disposition d'images du conduit vocal obtenues par cinéradiographie (Fig. 5a). Plus récemment (Baer *et al.*, 1987 ; Demolin *et al.*, 1996 ; Badin *et al.*, 1998), l'apparition des techniques d'Imagerie par Résonance Magnétique (IRM) permet d'obtenir des images anatomiques du conduit vocal sans exposer le sujet à des rayonnements ionisants, comme c'est le cas en cinéradiographie X interdite à présent en France sur des sujets sains. Un autre atout de l'IRM est qu'elle permet d'obtenir des coupes non seulement dans le plan sagittal (Fig. 5c) mais aussi dans des plans

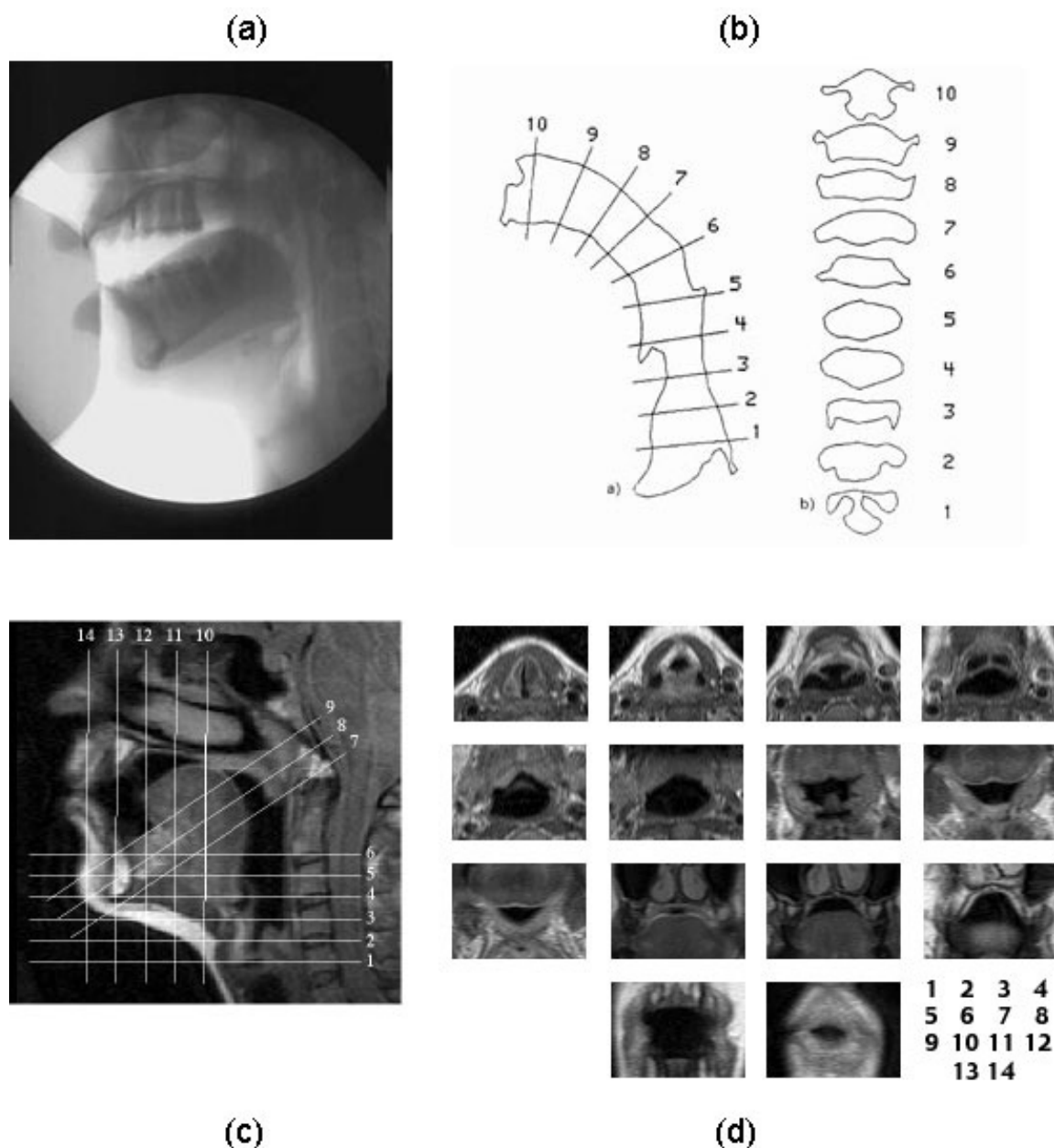


Fig. 5. Imagerie du conduit vocal. (a) Cinéradiographie X (Source : Institut de Phonétique, Strasbourg). (b) Coupes transverses du conduit vocal réalisées dans un moulage de cadavre (Sanchez et Boë, 1984, dans Tubach, 1989). (c) Coupe sagittale par IRM anatomique et (d) coupes axiales, obliques et frontales du conduit (Source : Lab. Phono. ULB, Belgique).

On peut remarquer l'analogie des coupes transverses issues du moulage (b) et celles de l'IRM (c). Attention, la numérotation et l'orientation des coupes n'est pas la même.



perpendiculaires à l'axe antéro-postérieur du conduit vocal et d'observer ainsi la forme de la section du tractus (Fig. 5d). Ainsi peut être résolu le problème du passage de la coupe sagittale du conduit (Fig. 4a et b) à la fonction d'aire (Fig. 4c), autrement dit le passage de deux à trois dimensions, en ajoutant une mesure transverse du conduit que les coupes sagittales ne fournissent pas. Les dimensions X (axe antéro-postérieur), Y (ouverture du conduit mesurable sur la coupe sagittale) et Z (dimension transversale du conduit) permettent d'obtenir la surface  $S(X) = Y(X) \cdot Z(X)$ , c'est-à-dire la fonction d'aire, donnée d'entrée des modèles géométriques.

Comme le notent Boë *et al.* (1994), l'intérêt acoustique de ces méthodes est évident mais « il faut souligner la nécessaire relecture des fonctions d'aires qu'il ne faudrait manipuler qu'à la lumière des contraintes de production de la parole : elles ne peuvent être atteintes que par le déplacement des organes de l'articulation (et non l'inverse) ». Cette mise en garde vise essentiellement l'approche géométrique purement descriptive, où la forme du conduit est vue déconnectée des organes, ce qui

peut donner lieu à des configurations aberrantes d'un point de vue anatomique ou à des spéculations fonctionnelles uniquement justifiées par une synthèse acoustique correcte. Ce sont ces réserves qui ont poussé la recherche vers une modélisation plus « anthropomorphique ».

## IV. La modélisation de la position des organes articulatoires

La modélisation articulaire du conduit vocal consiste à utiliser des paramètres de position des articulateurs comme paramètres de contrôle du modèle et à générer en sortie de façon indirecte le signal de parole. Nous utiliserons la distinction faite par Scully (1990, p.154) et n'aborderons que quelques modèles articulaires représentatifs.

### 1. Les modèles statistiques

Les modèles statistiques de l'articulation utilisent une approche montante, fondée sur un

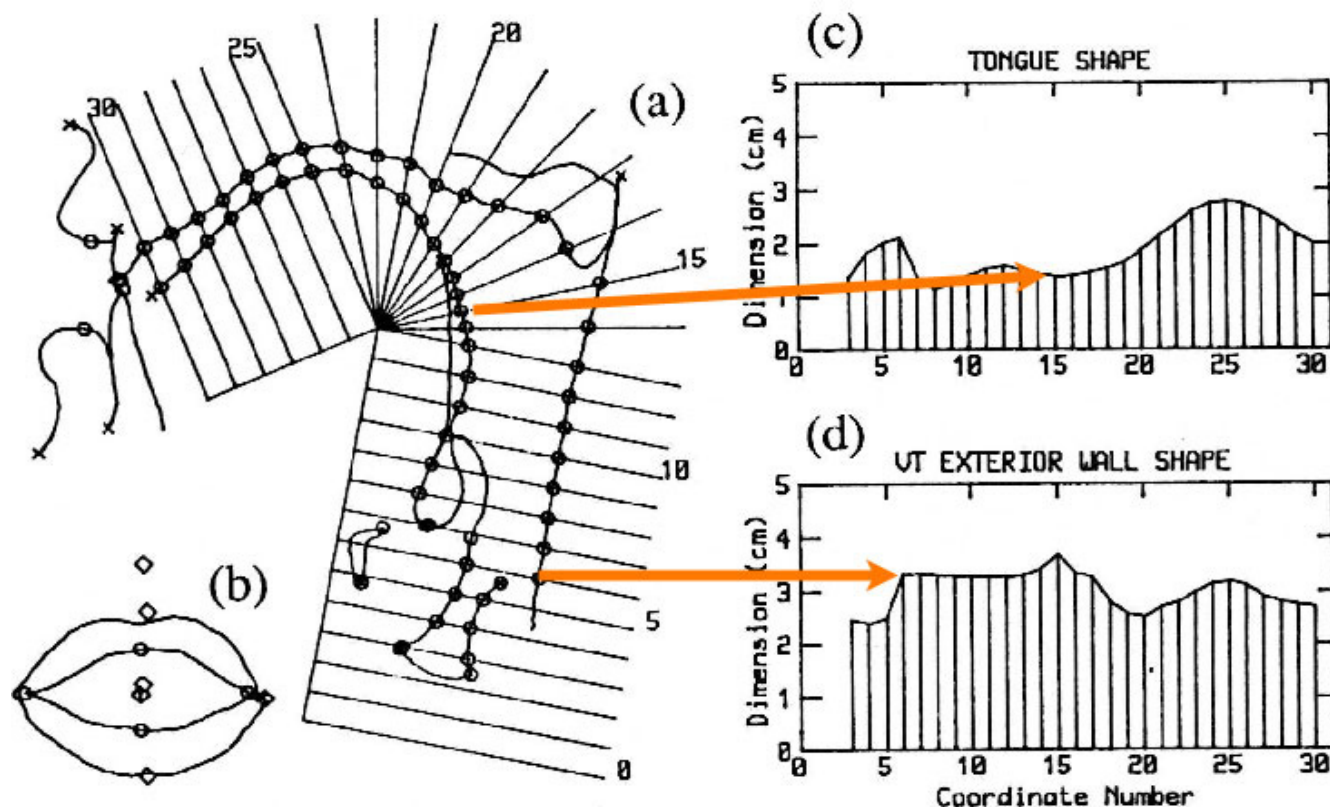


Fig. 6. (a) Contour de conduit vocal avec son système de coordonnées semi-polaires dans lequel les cercles noirs indiquent les points de mesure. (b) Vue frontale des lèvres. (c) Représentation vectorielle (dans le repère semi-polaire) du contour mesuré de la langue, incluant la zone laryngée supérieure. (d) Représentation vectorielle de la paroi supérieure du conduit (Source : Maeda, 1990). Les flèches illustrent le passage des mesures relevées sur la coupe sagittale aux coordonnées résultantes dans le repère semi-polaire, la représentation dans le repère semi-polaire étant redressée dans (c) et (d).



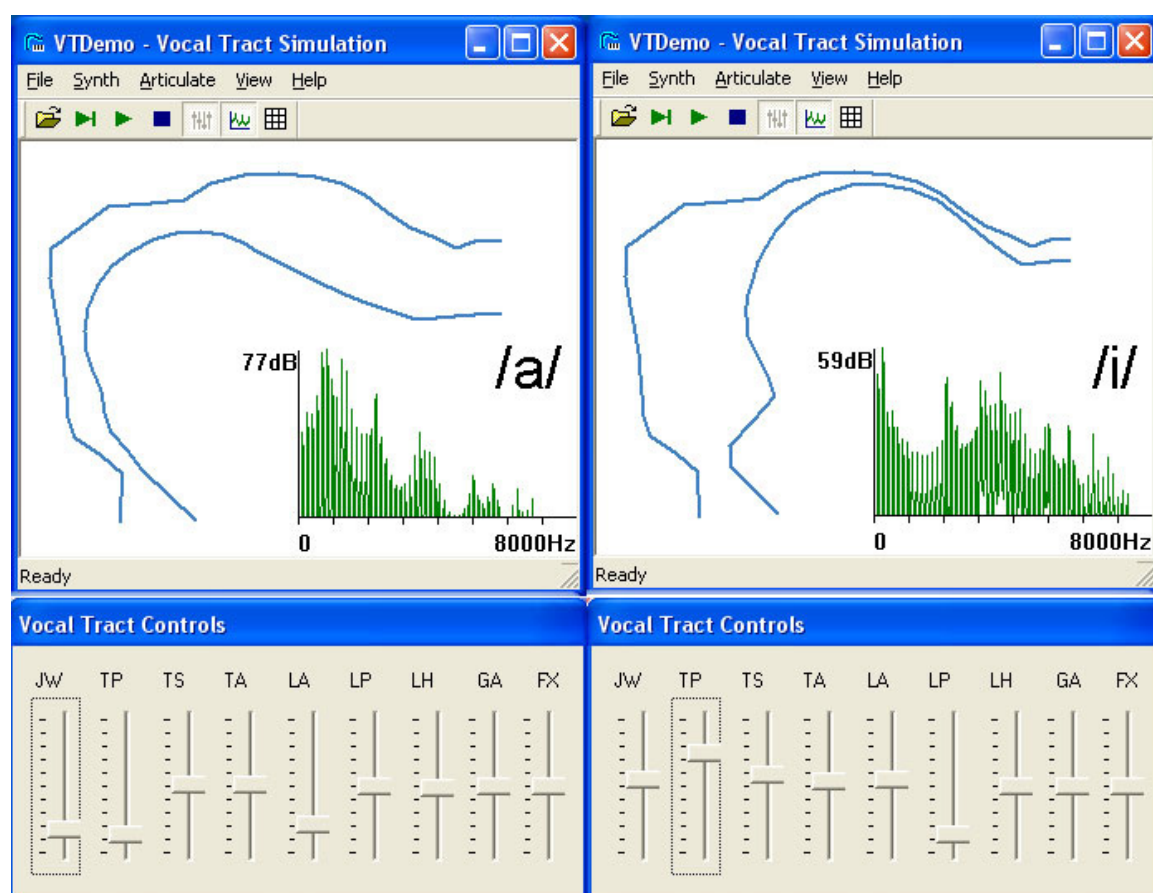


Fig. 7. Logiciel de synthèse articulatoire à partir du modèle de Maeda (Source : <http://www.phon.ucl.ac.uk/resource/vtdemo>). A gauche, /a/ ; à droite, /i/. En haut, la configuration du tractus et le spectre du signal synthétisé. En bas, les paramètres de contrôle du modèle : JW (jaw = mandibule, ouverte/fermée), TP (Tongue-Position = position la langue, avant/arrière), TS (Tongue-Shape = forme de la langue, voûtée/plate), TA (Tongue-Apex = apex de la langue, relevé/abaissé), LA (Lips Area = aire des lèvres), LP (Lips Protrusion = protrusion des lèvres), LH (Larynx Height = hauteur du larynx), GA (Glottal Area = aire glottique), FX (fréquence fondamentale).

corpus de coupes sagittales dont l'analyse statistique permet d'extraire des caractéristiques articulatoires pertinentes et interprétables. Comme ils sont fondés sur des données observées, ils ont, en général, une bonne capacité descriptive. L'exemple le plus représentatif est le modèle de Maeda.

A l'origine, on peut considérer que les travaux de Maeda se situent dans une approche géométrique descriptive du conduit vocal. Les analyses de Maeda (1978) reposent sur l'observation de 400 contours de conduits vocaux extraits d'images cinéradiographiques (Fig. 5a) enregistrés dans la production de 12 voyelles. Les positions des parois du conduit vocal sont alors mesurées dans un système de coordonnées semi-polaires (Fig. 6a), ce repère épousant la courbure longitudinale du tractus. Dans ce travail initial, l'auteur s'est intéressé essentiellement

à l'étude statistique des contours de langue (Fig. 6c) dans les différentes productions de voyelles en utilisant une analyse en composantes principales (ACP). L'ACP est une méthode statistique qui a pour but de comprendre et de visualiser comment les effets de phénomènes a priori indépendants se combinent en un résultat observé opaque. Le résultat d'une ACP est de déterminer les axes qui expliquent au mieux la dispersion des observations disponibles. Autrement dit, cette analyse permet de décomposer un phénomène complexe observé en un ensemble réduit de facteurs « explicatifs ». Les résultats de Maeda (1978) montrent que 2 composantes suffisent à rendre compte de 90 % de la variance des données de contours de langue. Avec 3 composantes retenues, ce taux s'élève à 98 %.

Là où ce type d'approche géométrique descriptive s'est transformé en un véritable modèle articulatoire, c'est dans sa capacité d'interprétation des composantes de l'ACP en mouvements élémentaires des articulateurs : position de la mandibule (ouverte/fermée), position du corps de la langue (avant/arrière), forme du dos de la langue (voûtée/plate), position de l'apex (relevé/abaissé).

L'ajout d'un modèle de lèvres issu de données de l'Institut de Communication Parlée (ICP) de Grenoble (d'après Tubach, 1989), ainsi que des compléments au niveau de la zone laryngée, ont permis d'aboutir à un véritable modèle articulatoire prédictif, fondé sur un positionnement « réaliste » des organes articulatoires (Maeda, 1990). L'obtention de la fonction d'aire puis du résultat acoustique (Maeda, 1982) a fini de transformer le modèle en authentique synthétiseur articulatoire dont l'implémentation informatique a été diffusée sous le nom de VTcalcs (Vocal-Tract Calculations). Une version de VTcalcs pour Matlab est disponible à l'adresse <http://www.cns.bu.edu/~speech/VTCalcs.php>. De même, nous invitons le lecteur à essayer le logiciel VTdemo développé par Mark Huckvale, University College of London. Cette application, très pédagogique, est téléchargeable librement à l'adresse <http://www.phon.ucl.ac.uk/resource/vtdemo/>. La synthèse sonore en temps réel permet d'entendre les modifications acoustiques engendrées par la modification des paramètres de contrôle des articulateurs (Fig. 7).

## 2. Les modèles à structures solides

Les modèles à structures solides utilisent une approche descendante, fondée sur la représentation solide de l'organe. Cette approche a le mérite de reproduire, au mieux, un fonctionnement anatomique. Les simplifications nécessaires à une telle modélisation peuvent entraîner un pouvoir de description appauvri, dans lequel le résultat acoustique produit s'éloigne des données observées.

Coker et Fujimura (1966) ont été les premiers à proposer un modèle articulatoire à 5 paramètres de contrôle (Fig. 8). La langue est modélisée, de façon un peu simpliste, par un cercle qui se déplace selon deux directions du plan sagittal et dont l'incursion est limitée par un autre cercle plus important représentant le palais et la paroi pharyngale. Une troisième variable est prévue pour spécifier la déformation du cercle dans le but de simuler un contrôle indépendant de la position de l'apex de la langue. Enfin, deux paramètres d'ouverture et de protrusion pilotent les lèvres. À partir de cet ensemble à 5 variables de

contrôle est calculée la fonction d'aire (Fig. 4) qui permet d'obtenir les propriétés acoustiques dérivées de la géométrie, autorisant la synthèse de la séquence sonore correspondante. Une notion intéressante introduite par les auteurs est la modélisation de la dynamique des organes. La réponse en position de l'organe est déduite d'un filtrage passe-bas de la commande et ce avec des constantes de temps différentes selon les organes (50 ms pour la pointe de langue et la fermeture labiale, 70 ms pour le corps de la langue, 110 ms pour la mandibule, 200 ms pour la protrusion). Cette propriété du modèle permet de rendre compte des phénomènes d'inertie ou de grande mobilité selon les organes.

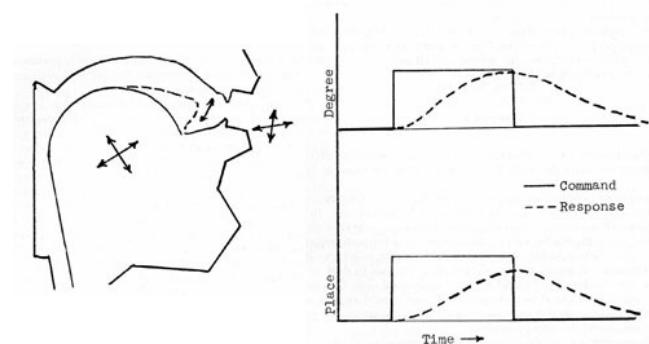


Fig. 8. Modèle de Coker (Source : Coker, 1968).

À gauche, le modèle articulatoire indiquant les degrés de liberté des 5 paramètres de contrôle. À droite, les coordonnées de la langue pour la production d'une séquence /VgV/ et la mise en évidence de la relation commande-réponse intégrant une modélisation contrainte de la dynamique des gestes.

Initié par Mermelstein (1973) dans les laboratoires Bell et plus tard affiné par Rubin *et al.* (1981), le modèle du tractus vocal de Haskins, implémenté sous la forme du logiciel ASY (<http://www.haskins.yale.edu/facilities/asy.html>) fournit une description cinématique des articulateurs de la parole à partir de la position dynamique de six structures principales : la mâchoire, le voile du palais, le corps de la langue, l'apex de la langue, les lèvres et l'os hyoïde (Fig. 9). La position de chaque articulateur n'est pas totalement libre dans l'espace et suit un ensemble de contraintes géométriques « réalistes ». Ainsi, le point F est vu comme l'axe de rotation de la mâchoire, cet axe étant considéré comme fixe. Le modèle de la mâchoire ne permet alors qu'une rotation simple autour de ce point, le déplacement étant contrôlé par la variable angulaire  $\theta_j$  (Fig. 9a). La position C du corps de la langue est repérée par des coordonnées polaires ( $Sc$ ,  $\theta_c$ ) par rapport à une ligne de base FJ liée à la mâchoire. Puisque la mâchoire tourne autour du point F, la coordonnée angulaire du centre de la langue est alors  $\theta_j + \theta_c$  par rapport à un repère fixe. De même,

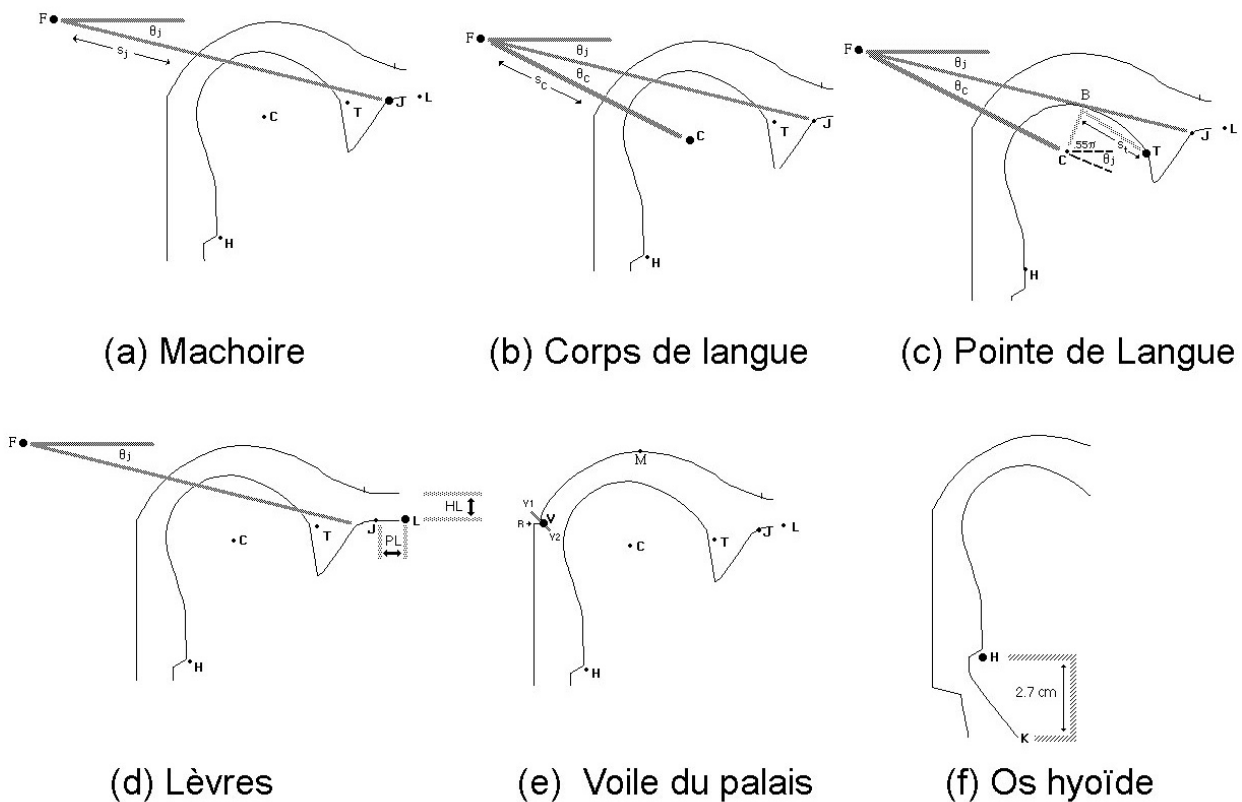


Fig. 9. Le contrôle des articulateurs dans le modèle articulatoire d'Haskins.

Les lignes représentent des contraintes géométriques et non l'action de muscles. (Source : <http://www.haskins.yale.edu/facilities/ee.html>).

la position de la pointe de la langue est repérée par le point T (pour « tip » en anglais) dont les coordonnées dérivent de l'angle effectué entre la pointe et la ligne de référence BC qui joint le centre C et le corps B de la langue (Fig. 9c).

Dans ce modèle articulatoire, les positions des articulateurs peuvent être commandées soit manuellement pour des productions statiques, soit au moyen d'une

table de caractéristiques évoluant au cours du temps et autorisant une séquence dynamique. La définition de la position des différents organes permet d'obtenir la coupe sagittale correspondante, puis la fonction d'aire, suivie de la fonction de transfert du conduit et enfin la génération de la séquence sonore (Fig.10).

Ce qui est intéressant dans cette approche, c'est la prise en compte, par principe, d'effets collatéraux

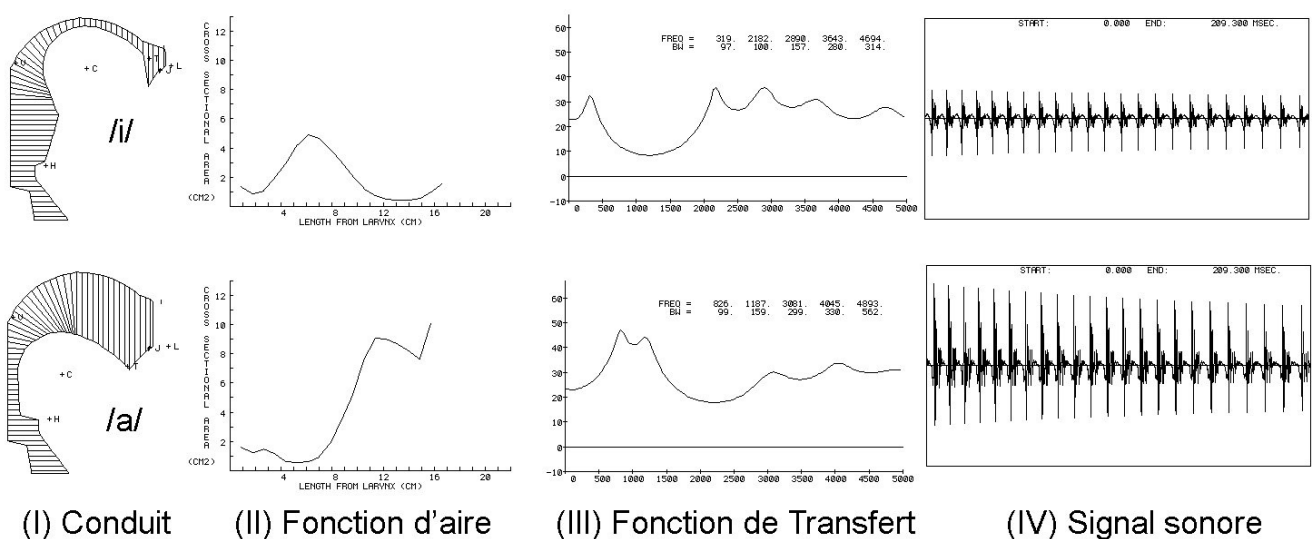


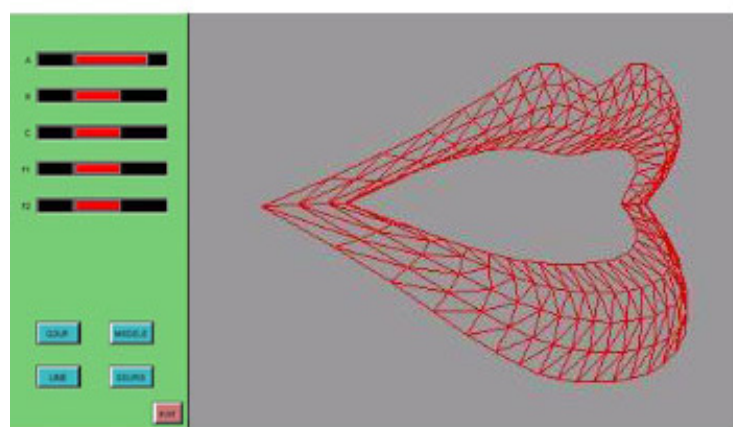
Fig. 10. Les étapes de la synthèse articulatoire (Source : <http://www.haskins.yale.edu/facilities/ee.html>) pour la production d'un /i/, en haut, et d'un /a/, en bas. (I) Configuration du conduit obtenue par réglage des positions des articulateurs. (II) Obtention de la fonction d'aire. (III) Calcul de la fonction de transfert du conduit. (IV) Synthèse de la séquence sonore.

d'un articulateur sur un autre (p. ex. : mâchoire sur langue, corps de langue sur apex), ce qui permet de distinguer les mouvements réellement contrôlés par rapport aux mouvements « involontaires », conséquences d'autres mouvements. Ainsi, le point T relatif à la pointe de la langue (Fig. 9) bougera de façon « involontaire » lors de mouvements de la mâchoire et/ou du corps de la langue, mais il pourra aussi se mouvoir de façon active en actionnant ses propres paramètres de contrôle. On ne retrouve pas ce type de contraintes dans des approches purement géométriques, uniquement fondées sur des fonctions d'aire. Il faut toutefois noter que les lignes visibles sur la figure 9 sont des repères géométriques et non des lignes d'action de muscles. Ce ne sont que des attributs fonctionnels. On peut aussi reprocher à ce modèle une simplification certaine comme le fait de considérer circulaire le corps de la langue, excluant ainsi sa capacité d'aplatissement, propriété physiologiquement bien connue.

Ce sont ces réserves qui ont poussé la recherche vers une modélisation d'une part plus « biomécanique », et d'autre part en 3D.

## V. La modélisation physiologique des articulateurs

La modélisation biomécanique des organes du conduit vocal consiste à déterminer un certain nombre de propriétés et mouvements de base qui permettent de représenter simplement, mais avec une très bonne précision, les mouvements complexes des articulateurs tels que la mâchoire, la langue, le voile du palais, les lèvres... Un premier constat s'impose.



Dans un tel contexte, il semble inévitable de se placer dans un espace de modélisation 3D, notamment :

- pour tenir compte de la géométrie transverse du conduit (Fig. 5), qui est sous-évaluée dans les approches 2D ;
- pour tenir compte des propriétés myo-élastiques de certains tissus modélisés et dont le comportement ne peut être correctement rendu que dans un espace en 3D, notamment pour la langue, muscle quasi incompressible à volume constant (Boë *et al.*, 1994) ;
- parce que l'anatomie du conduit est clairement tridimensionnelle.

De plus, dans une finalité de modélisation la plus proche possible de la réalité physiologique, il apparaît intéressant de prendre en compte non seulement la forme et la position statique des organes, mais aussi la dynamique du système en termes de vitesse de déplacement et d'inertie. Enfin, la prise en considération des effecteurs neuromusculaires à la source de ces mouvements permet d'atteindre un stade de modélisation très avancé.

### 1. Modèles de lèvres

La modélisation des lèvres est facilitée par l'obtention aisée d'images de locuteurs filmés de face et de profil, et dont, généralement, les lèvres sont maquillées en bleu pour faciliter les procédés de traitement d'image. Un état des lieux portant sur la question a été proposé (Boë *et al.*, 1994). Plus récemment, les actions tournées vers les robots parlants (Réveret *et al.*, 2000) ont grandement contribué au développement de modèles de lèvres. Les travaux de Réveret et Le Chevalier (1998) ayant porté sur l'analyse en composante principale (cf. la section IV. 1.

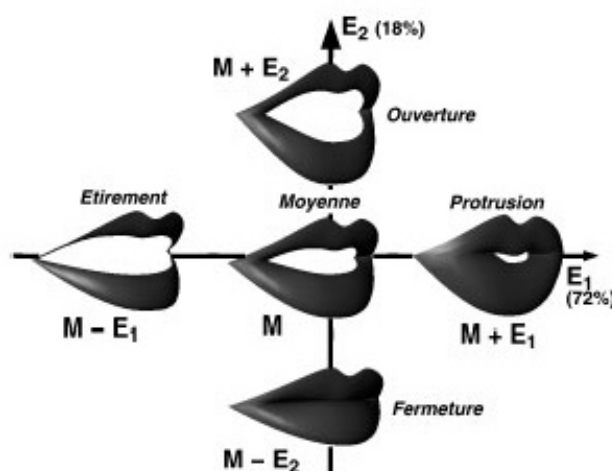


Fig. 11. A gauche, modèle de lèvres (Adjoudani, 1994). (Source : [http://www.mediaport.net/CP/CyberScience/BDD/fich\\_013.en.html](http://www.mediaport.net/CP/CyberScience/BDD/fich_013.en.html)). A droite, analyse en composantes principales du mouvement des lèvres (Réveret et Le Chevalier, 1998).



précédente) de mouvements filmés de lèvres montrent que le geste d'étirement/protrusion compte pour 72 % de la variance, le geste de fermeture/ouverture pour 18 %, 2 % restant inexpliqués (Fig. 11). Il faut toutefois noter que, du fait de la finalité de ces travaux (l'application à des visages synthétiques parlants), la plupart des modèles sont géométriques sous forme de maillage (Fig. 11). Ils présentent malgré tout l'intérêt de mettre en évidence l'importance du mouvement transversal d'étirement, propriété souvent négligée dans les modèles 2D du conduit.

## 2. Modèles de mâchoire

Par rapport aux autres organes articulatoires et notamment la langue, on peut considérer en première approximation que la mâchoire, de par sa structure osseuse, possède un fonctionnement assez simple en terme de positionnement : elle peut être correctement représentée par un corps solide qui peut effectuer principalement une rotation autour de l'articulation temporo-mandibulaire et éventuellement une translation parallèle à la surface de l'os temporal (Laboissière *et al.*, 1996). Ce dernier mouvement s'avère utile dans la production de consonnes labio-dentales (/f/, /v/) pour lesquelles la lèvre inférieure vient en contact avec les incisives supérieures.

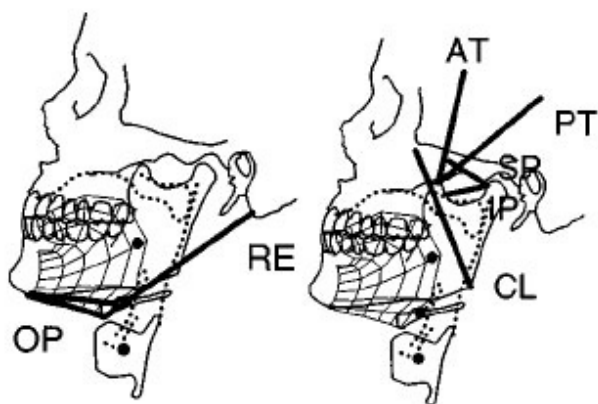


Fig. 12. Modélisation des muscles contrôlant le mouvement de la mâchoire (d'après Sanguinetti *et al.*, 1998).

A gauche : opener (OP) et retractor (RE); à droite : closer (CL), anterior temporalis (AT), posterior temporalis (PT), superior pterygoid (SP) et inferior pterygoid (IP). Les lignes épaisses représentent les macrofibres qui sont utilisées pour modéliser chaque muscle.

Bien plus qu'une modélisation géométrique de l'organe déjà présente dans Mermelstein (1973), l'approche de Sanguinetti *et al.* (1998) est intéressante car elle intègre de façon explicite les propriétés et la disposition réaliste des muscles de la mandibule (Fig. 12), ainsi que les commandes réflexes nerveuses et la dynamique des structures osseuses et des tissus mous

adjacents. Son travail est clairement orienté vers l'organisation des signaux de contrôle à l'origine des mouvements visibles du tractus, et le comportement dynamique des articulateurs.

## 3. Modèles de langue

Parmi tous les articulateurs, la langue apparaît comme l'organe le plus important pour l'articulation de la parole. Son fonctionnement reste aussi le plus complexe et permet une multitude de mouvements rapides et précis. D'ailleurs, ces deux constats sont probablement liés par une relation de dépendances réciproque.

La modélisation biomécanique 3D de la langue a fait l'objet de divers travaux (Wilhelms-Tricarico, 1995; Sanguinetti *et al.*, 1998 ; Perrier *et al.*, 2003). L'intérêt des modèles 3D de la langue par rapport à une approche 2D réside dans les points suivants :

- la prise en compte des interactions mécaniques entre la langue et les parties anatomiques qui l'entourent (dents, palais dur, voile du palais, paroi pharyngale...), ces interférences pouvant clairement influencer les mouvements linguaux ;
- l'évidente nécessité d'un espace à 3 dimensions pour assurer certaines propriétés physiologiques de la langue comme la conservation de son volume, cet organe étant considéré comme quasi incompressible de par la structure de ses tissus (Boë *et al.* 1994). On imagine très bien alors qu'un aplatissement ou plus généralement, une réduction de la surface de la langue dans le plan sagittal médian s'accompagne d'un élargissement dans le plan transverse ;
- la prise en compte de la dimension transverse du tractus vocal en ne limitant plus la description de la forme de la langue uniquement sur le plan sagittal médian, permettant ainsi d'obtenir une image plus fine de la géométrie du conduit, notamment dans la production de certaines consonnes où le tractus s'éloigne de caractéristiques tubulaires (p. ex. : consonnes latérales).

Les modèles de Wilhelms-Tricarico (1995) et Perrier *et al.* (2003) sont fondés sur la méthode des éléments finis qui est utilisée depuis longtemps en résistance des matériaux. Ce procédé permet de résoudre numériquement des équations mathématiques représentant analytiquement le comportement dynamique de l'objet modélisé. C'est ainsi que peut être testée la réaction de structures architecturales, de pièces mécaniques, de véhicules automobiles, à des contraintes. Cette méthode nécessite la décomposition de l'objet observé sous forme de maillage en petites structures. Dans le modèle de Perrier *et al.* (2003),

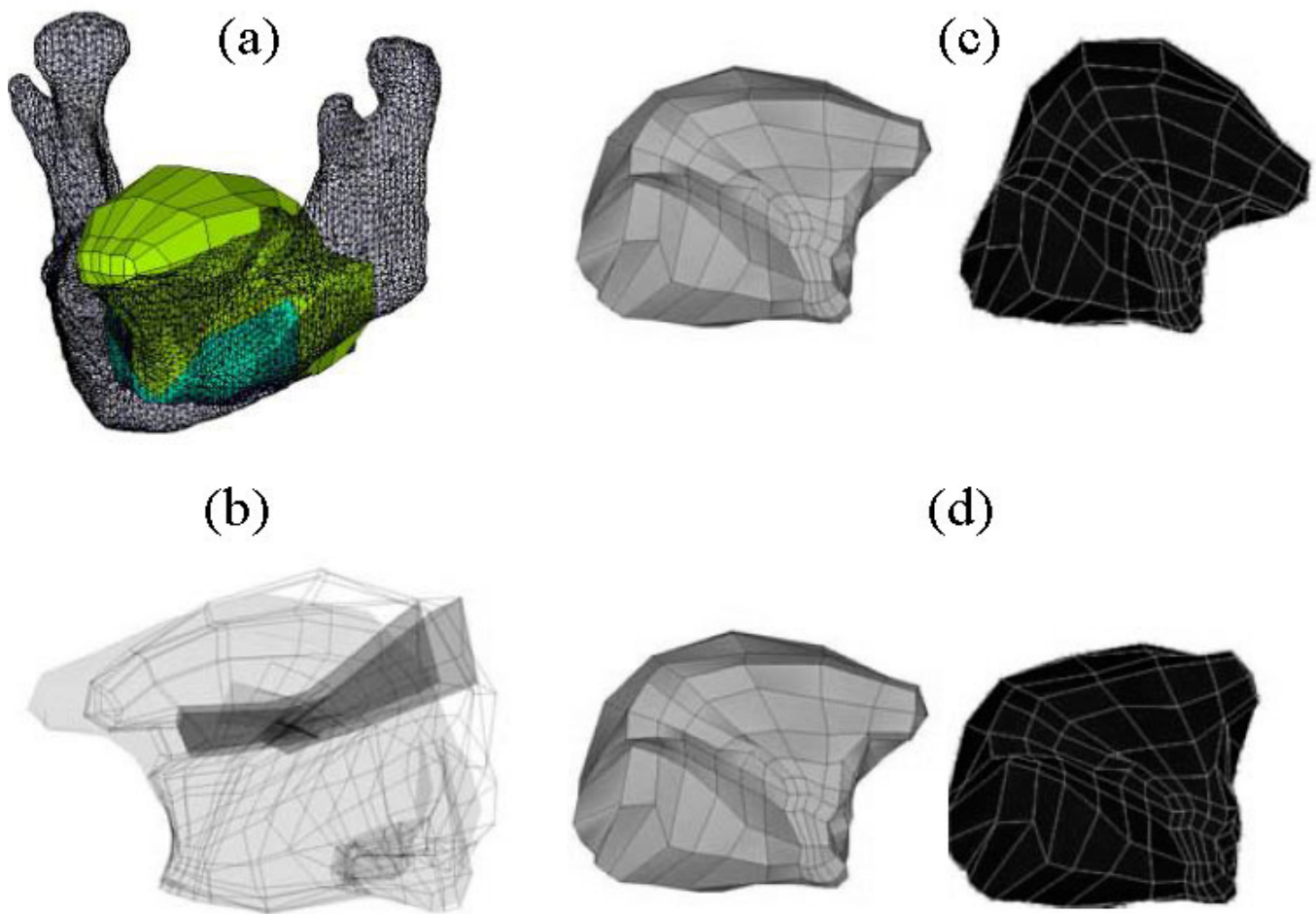


Fig. 13. (a) Modèle biomécanique 3D de la langue (Perrier *et al.*, 2003). (b) Elévation du dos de la langue sous l'effet d'une simulation de l'action du muscle stylo-glosse. (c) Déplacement vers l'avant et élévation du corps de la langue sous l'effet du génio-glosse. (d) Mouvement de l'apex sous l'effet du longitudinal supérieur.

Source (a) : [www-timc.imag.fr/gmcao/fr-fiches-projets/modele-langue.htm](http://www-timc.imag.fr/gmcao/fr-fiches-projets/modele-langue.htm) ; Source (b, c et d) : Gérard *et al.*, 2004.

le maillage est structuré pour respecter au mieux la distribution anatomique des muscles de la langue. L'action d'un muscle est simulée par l'application d'une force exercée sur un ou plusieurs éléments (Gérard *et al.*, 2004), ce qui entraîne un déplacement et une déformation de l'ensemble ou d'une partie de la structure linguale (Fig. 13).

#### 4. Modèles de velum

Il existe peu de travaux de modélisation du conduit nasal, essentiellement par manque de données anatomiques et du fait de la complexité géométrique des cavités nasales. Ce constat déjà effectué par Boë *et al.* (1994) se vérifie toujours actuellement. On peut toutefois noter l'initiative de Serrurier et Badin (2005) qui proposent un modèle articulatoire 3D du velum (Fig. 14) fondé sur l'analyse en composantes

principales de contours du voile obtenus à partir d'images IRM. Leurs résultats montrent qu'une description 3D de la forme du voile peut être réduite, au moyen d'un modèle linéaire, à un unique paramètre de commande, interprété comme la hauteur d'un point situé dans le plan sagittal médian sur la partie descendante du voile (Fig. 14g). Ce paramètre a pour effet de déplacer le velum selon un axe postérieur haut versus antérieur bas qui pourrait être interprété comme l'action du muscle élévateur vélo-palatin fermant ou ouvrant le port vélo-pharyngé, cette hypothèse devant être malgré tout vérifiée. De plus, cette étude mentionne que l'effet, probablement non linéaire, d'une action d'appui de la langue sur le voile a été écarté.

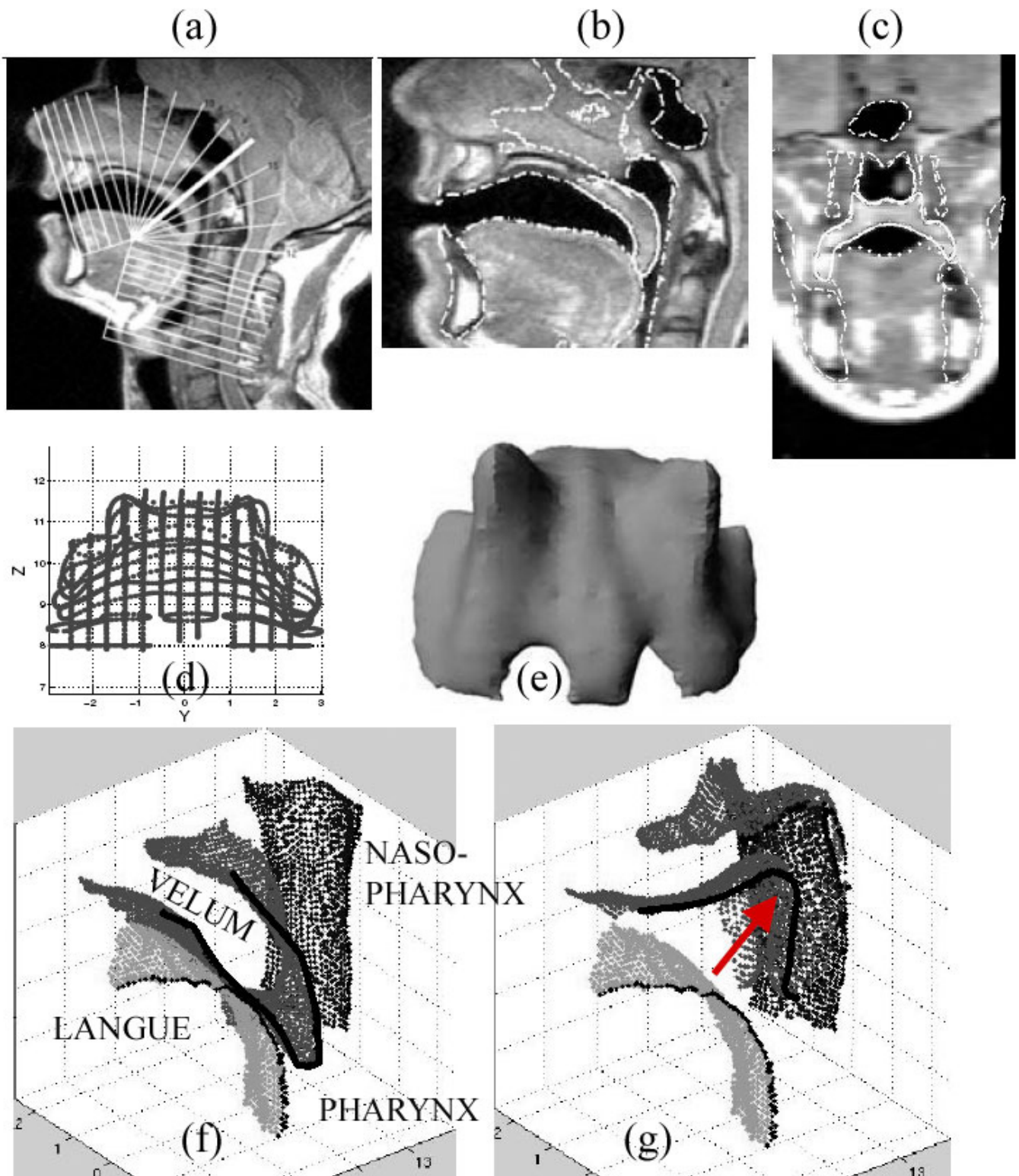


Fig. 14. Modélisation 3D du velum (Serrurier et Badin, 2005).

A partir des images IRM, (a) repérage des contours du voile (en trait blanc plein) ;(b) en coupe sagittale ; (c) en orientation perpendiculaire au plan sagittal et distribuées dans un repère semi-polaire comme visible sur (a). Cette opération est renouvelée pour les différentes coupes non médianes ; (d) juxtaposition des contours extraits de chaque coupe ; (e) vue postérieure du voile après application d'un maillage et un rendu de surface ; (f et g) représentation des parties droites de la langue, du velum et du nasopharynx (arrière-plan). En (g), on observe nettement l'effet du paramètre de contrôle du velum actionnant un mouvement arrière montant.



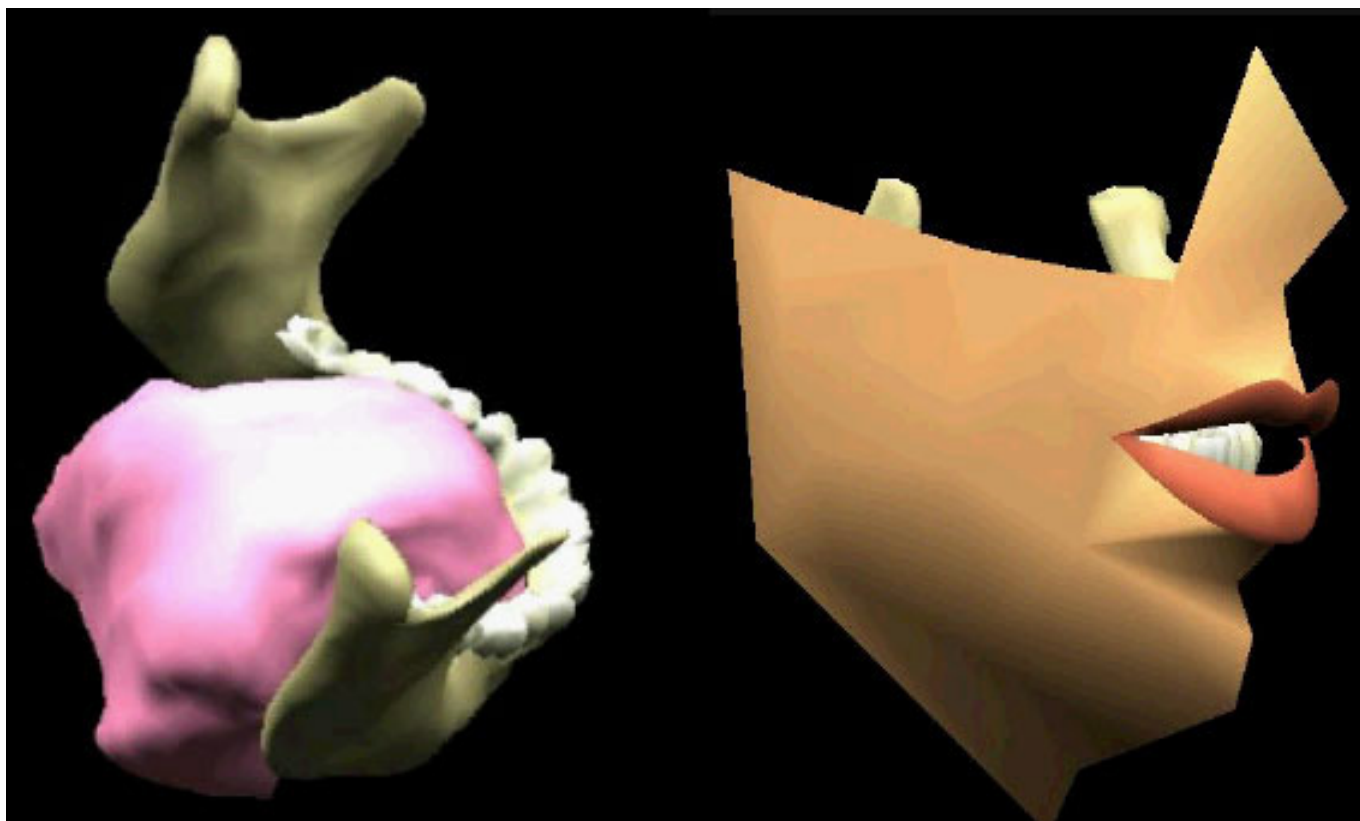


Fig.15. Modélisation 3D des organes articulatoires (Badin *et al.*, 2002).

A gauche, la langue, la mâchoire et les dents. A droite, vue externe (Source : <http://www.icp.inpg.fr/~badin/VTH-SPS5.html>).

## 5. Modèles complets du conduit

Les modèles 3D intégrant l'ensemble des articulatoires sont rares. On peut citer celui de Badin *et al.*, (2002) qui est fondé sur l'analyse d'images anatomiques IRM pour les organes internes et de films vidéo pour les parties externes (mâchoire, lèvres). A partir de ces données descriptives, les auteurs ont extrait par une analyse multilinéaire guidée des facteurs explicatifs élémentaires ayant une réalité articulatoire (hauteur de la mandibule, protrusion des lèvres...).

## VI. Conclusion

Ce tour d'horizon des modèles du tractus vocal ne se veut pas exhaustif. Il permet néanmoins de faire un constat général : plus le modèle est de haut niveau, moins il y a de paramètres d'entrée mais plus complexe est la modélisation. Inversement, plus le modèle est de bas niveau (p. ex. : acoustique), plus la modélisation est simple mais plus le nombre de paramètres de contrôle est important. Dans une approche anthropomorphique, nous aurions

tendance à pencher vers des modèles complexes de type biomécanique, possédant peu de paramètres. Dans tous les cas, une bonne description des données observées est un gage de bonne modélisation, de même que la capacité d'interprétation reste nécessaire à l'enrichissement des connaissances sur les mécanismes de production mais aussi de perception.

Dans le cadre de l'étude des dysarthries, des modèles comme ceux décrits précédemment, seraient d'un grand secours pour mieux appréhender des observations infiniment variables, souvent opaques mais qui pourraient être mieux comprises à la lumière d'un modèle physiopathologique. De même, la synthèse artificielle de productions pathologiques pourrait permettre d'une part de tester nos outils d'évaluation sur des matériaux parfaitement calibrés, et d'autre part, de fournir un outil pédagogique pour la formation des cliniciens et autres scientifiques préoccupés par les troubles de la voix et de la parole.



# Références

1. Adjoudani A, Guiard-Marigny T, Le Goff B, Benoît C (1994). Un modèle 3D de lèvres parlantes. XXèmes Journées d'Etudes sur la Parole. Trégastel, France. 1994, pp 143-146.
2. Atal BS, Chang JJ, Mathews MV, Tukey JW. Inversion of articulatory-to-acoustic transformation in the vocal tract by a computer-sorting technique. J Acoust Soc Am 1978, 63: 1535-1555.
3. Badin P, Fant G. Notes on vocal tract computations. Speech Transmission Laboratory. Quarterly Progress Status Report. Stockholm, Sweden. 1984, 2-3: 53-107.
4. Badin P, Bailly G, Raybaudi M, Segebarth C. A three-dimensional linear articulatory model based on MRI data. Proceedings of the Workshop on Speech Production. Kuchan, Japan. Hokkaido University. 1998, pp. 1-7.
5. Badin P, Bailly G, Revéret L, Baciú M, Segebarth C, Savariaux C. Three-dimensional articulatory modeling of tongue, lips and face, based on MRI and video images. J Phonetics 2002, 30: 533-553.
6. Baer T, Gore JC, Boyce S, Nye PW. Application of MRI to the Analysis of Speech Production. Magn Reson Imaging 1987, 5: 1-7.
7. Boë LJ, Maeda S, Perrier P. La modélisation articulatoire : un demi-siècle d'évolution entre fonctionnel, physique et biomécanique. Actes des XXèmes Journées d'Etudes sur la Parole. Trégastel, France. 1994, pp 41-54.
8. Boersma P. An articulatory synthesizer for the simulation of consonants. Proceedings of Eurospeech, Berlin, Germany. 1993, pp. 1907-1910.
9. Bothorel A, Simon P, Wioland F, Zerling JP. (1986). Cinéradiographie des voyelles et consonnes du français. Institut de Phonétique (ed). Strasbourg.
10. Chiba T, Kajiyama M. (1941). The Vowel : its Nature and Structure. Tokyo-Kaseidan Pub. Tokyo.
11. Coker C, Fujimura O. Model for specification of the vocal tract area function. J Acoust Soc Am 1966, 40: 1271.
12. Coker C. (1968). Speech Synthesis with a Parametric Articulatory Model. In Flanagan J, Rabiner L (eds). Speech Synthesis. Dowden, Hutchinson, Ross, Inc. Stroudsburg, USA. 1973, pp 135-139.
13. Demolin D, Metens T, Soquet A. Three-dimensional Measurement of the Vocal Tract by MRI. Proceedings of the 4th International Conference on Spoken Language Processing (ICSLP). Philadelphia, USA. 1996, pp 272-275.
14. Dudley H. (1955). Fundamentals of speech synthesis. In Flanagan J, Rabiner L (eds). Speech Synthesis. Dowden, Hutchinson, Ross, Inc. Stroudsburg, USA. 1973, pp 43-58.
15. Fant G. (1960). Acoustic Theory of Speech Production. Mouton & Co. La Haye.
16. Gérard JM, Perrier P, Payan Y. Modélisation biomécanique tridimensionnelle de l'articulateur lingual pour étudier le contrôle de la parole. Actes des XXVèmes Journées d'Etude sur la Parole, Fès, Maroc. 2004, pp 237-240.
17. Hanson HH, Stevens KN. A quasiarticulatory approach to controlling acoustic source parameters in a Klatt-type formant synthesizer using Hlsyn. J Acoust Soc Am 2002, 112 : 1158-1182.
18. Klatt DH. Software for a cascade/parallel formant synthesizer. J Acoust Soc Am 1980, 67: 971-995.
19. Laboissière R, Ostry DJ, Feldman AG. Control of multimuscule systems: Human jaw and hyoid movements. Biol Cybern 1996, 74.3: 373-384.
20. Maeda S. Une analyse statistique sur les positions de la langue : Etude préliminaire sur les voyelles françaises. Actes des IXèmes Journées d'Etude sur la Parole. Lannion, France. 1978, pp 191-199.
21. Maeda S. A digital simulation method of vocal-tract system. Speech Communication 1982, 1: 199-229.
22. Maeda S. Compensatory Articulation During Speech: Evidence From The Analysis And Synthesis Of Vocal Tract Shapes Using An Articulatory Model. In Hardcastle WJ, Marchal A (eds). Speech Production and Speech Modelling. Coll. NATO ASI, Series D, vol 55. Academic Press Inc. 1990, pp 131-149.
23. Mermelstein P. Articulatory model for the study of speech production. J Acoust Soc Am 1973, 53: 1070-1082.
24. Mrayati M, Carré R. Relations entre la forme du conduit vocal et les caractéristiques acoustiques des voyelles françaises. Phonetica 1976, 33 : 285-306.
25. Perrier P, Payan Y, Zandipour M, Perkell J. Influences of tongue biomechanics on speech movements during the production of velar stop consonants: a modeling study. J Acoust Soc Am 2003, 114: 1582-1599.
26. Revéret L, Bailly G, Badin P. MOTHER: a new generation of talking heads providing a flexible articulatory control for video-realistic speech animation. International Conference on Speech and Language Processing. Beijing, China. 2000, pp 755-758.

27. Revéret L, Le Chevalier L. Un modèle géométrique de lèvres 3D adaptable au locuteur. *Actes des XXIIèmes Journées d'Etudes sur la Parole*. Martigny, Suisse. 1998, pp 213-216.
28. Rubin PE, Baer T, Mermelstein P. An articulatory synthesizer for perceptual research. *J Acoust Soc Am* 1981, 70: 321-328.
29. Sanchez H, Boë LJ. (1984). De la coupe sagittale à la fonction d'aire du conduit vocal. *Actes des XIIIèmes Journées d'Etude sur la Parole*. Bruxelles, Belgique. 1984, pp 23-25.
30. Sanguineti V, Laboissière R, Ostry DJ. A Dynamic Biomechanical Model for Neural Control of Speech Production. *J Acoust Soc Am* 1998, 103:1615-1627.
31. Scully C. Articulatory Synthesis. In Hardcastle WJ, Marchal A (eds). *Speech Production and Speech Modelling*. Coll. NATO ASI, Series D, vol 55. Academic Press Inc. 1990, pp 151-186.
32. Serrurier A, Badin P. A Three-Dimensional Linear Articulatory Model of Velum Based on MRI Data. *Proceedings of Interspeech*. Lisbon, Portugal. 2005, pp 2161-2164.
33. Stevens KN, Kasowski S, Fant CGM. (1953). An Electrical Analog Of The Vocal Tract. In Flanagan J, Rabiner L (eds). *Speech Synthesis*. Dowden, Hutchinson, Ross, Inc. Stroudsburg, USA. 1973, pp. 108-116.
34. Tubach JP. (1989). *Calliope. La parole et son traitement automatique*. Masson. Paris.
35. Whilhems-Tricarico R. Physiological modelling of speech production: methods for modelling soft tissues articulators. *J Acoust Soc Am* 1995, 97: 3085-3098.