



MÉCANISMES DE PRODUCTION DES CONSONNES FRICATIVES : COORDINATION ENTRE GLOTTE ET CONSTRICTION DU CONDUIT VOCAL

E. Castelli, C. Scully

► To cite this version:

E. Castelli, C. Scully. MÉCANISMES DE PRODUCTION DES CONSONNES FRICATIVES : COORDINATION ENTRE GLOTTE ET CONSTRICTION DU CONDUIT VOCAL. Journal de Physique IV Proceedings, 1992, 02 (C1), pp.C1-303-C1-306. 10.1051/jp4:1992165 . jpa-00251236

HAL Id: jpa-00251236

<https://hal.science/jpa-00251236>

Submitted on 4 Feb 2008

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

MÉCANISMES DE PRODUCTION DES CONSONNES FRICATIVES : COORDINATION ENTRE GLOTTE ET CONSTRICTION DU CONDUIT VOCAL**E. CASTELLI et C. SCULLY****Institut de la Communication Parlée, U.R.A. - CNRS N° 368, I.N.P.G., Université Stendhal, 46 avenue Félix Viallet, F-38031 Grenoble cedex, France***Department of Psychology, The University of Leeds, GB-Leeds LS2 9JT, Great-Britain***Abstract :**

To produce speech simulations of high quality, we need a good knowledge of temporal coordination between all parts of the respiratory tract. The production of fricative consonants depends primarily on a double articulatory gesture : in the vocal tract a constriction is formed; simultaneously the vocal folds are opened. This double gesture is the cause of (1) an increase in the intra-oral pressure and (2) the production of a frication noise due to the turbulent flow downstream from the constriction. Fricative consonants are characterised by their quasi-static spectra and also by rapidly changing acoustic patterns which seem to be important for the characterisation of /f/ versus /../ or /v/ versus /../. Our work is based on multiple analyses for two subjects. With an analog model of the vocal tract in which an automatic noise source was implemented, analysis-by-synthesis will be used to obtain good aerodynamic and acoustic matches between the natural and the simulated speech. The aim of our study is to try to evaluate the influence of the temporal coordination between the formation of the constriction and the opening of the vocal folds.

Introduction :

Pour réaliser des simulations de production de parole de bonne qualité, il est nécessaire de bien connaître les coordinations temporelles entre les différentes parties du conduit respiratoire. La production des consonnes fricatives dépend d'un double geste articulatoire : une constriction dans le conduit vocal est réalisée, simultanément le sujet ouvre ses cordes vocales. Ce double geste articulatoire a pour principales conséquences une augmentation de la pression intra-orale et la production d'un bruit de frication dû à l'écoulement turbulent de l'air en sortie de la constriction. Nous ne cherchons pas à caractériser les influences de la constriction (et donc de la charge acoustique) du conduit vocal sur le comportement de la source vocale [1] mais, sans nier ces effets indissociables du couplage, nous pensons que l'influence de la coordination entre constriction et ouverture des cordes vocales est aussi essentielle dans la production des fricatives. En effet, les fricatives ne sont pas seulement caractérisées par leur spectre quasi-statique, mais aussi par les changements spectraux rapides et de sources acoustiques pendant les transitions. Ces changements semblent être importants pour différencier les consonnes /f/ et /θ/ ou les consonnes /v/ et /ð/ par exemple [2]. Nous allons tenter par des mesures en parole naturelle, puis par des simulations temporelles dynamiques de caractériser ces influences.

Parole naturelle :

Dans le cadre d'un projet européen entre deux laboratoires anglais et un laboratoire français (Leeds, Southampton et Grenoble), un corpus de sons fricatifs en parole naturelle a été réalisé avec deux sujets. Des enregistrements de pression et débit dans le conduit vocal (masque de Rothenberg) et du signal de parole font partie du corpus. Les sujets sont une femme américaine C.S. et un homme d'origine française P.B. Chaque sujet a enregistré des séquences de 10 répétitions environ d'items /pVCV/.

Le bruit de frication est fonction de la nature de la constriction mais aussi de la chute de pression à la constriction et du débit qui la traverse [2]. Nous pensons que les évolutions temporelles de la pression intra-orale et du débit moyen en sortie des lèvres sont dépendantes de la coordination temporelle entre la glotte et la constriction du conduit vocal. Nous nous sommes alors plus particulièrement intéressés à ces évolutions pendant la production des séquences du corpus comportant les voyelles /a/ et /i/ et les consonnes fricatives /s/ et /f/.

Pendant la production des voyelles, le débit n'est modulé que par les cordes vocales, le conduit vocal n'étant pas fortement constrikt. La chute de pression entre la pression subglottique et la pression atmosphérique est localisée à la glotte et la pression intra-orale est pratiquement égale à la pression

atmosphérique. On peut pourtant constater que le débit moyen pendant la production du /i/ est plus faible que pendant la production du /a/ : le débit est un peu diminué par la constriction dans le conduit vocal, plus forte pour le /i/ que pour le /a/, et par les pertes par viscosité qui en résultent.

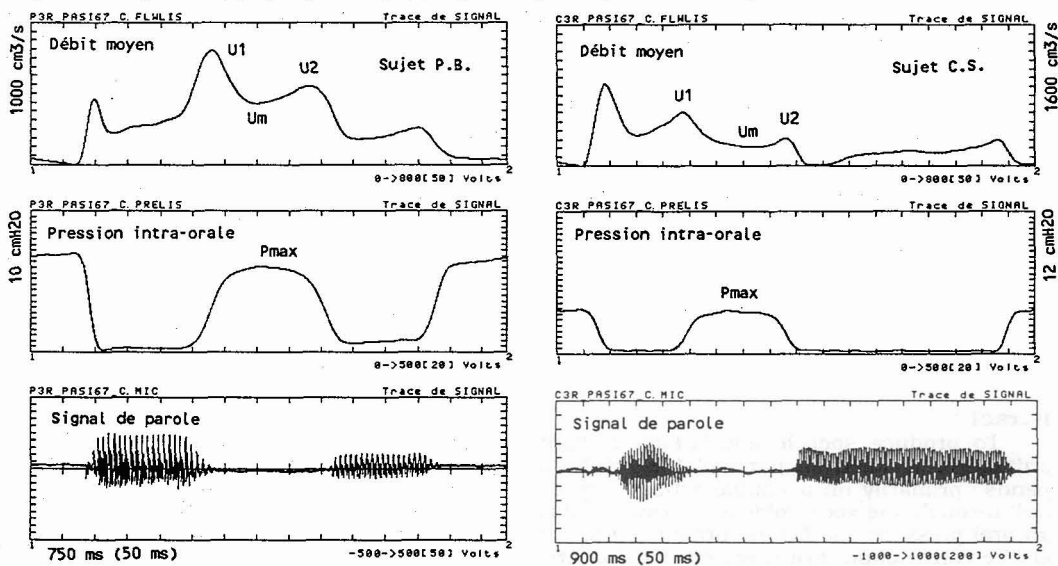
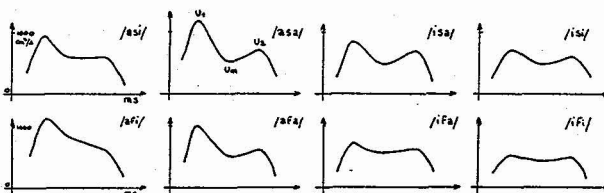


Figure n° 1. Signal, débit moyen et pression intra-orale pour la séquence naturelle /asi/.

Pendant la production des fricatives, les allures de pression et de débit se compliquent (fig. n° 1). Nous avons affaire à une évolution rapide de la géométrie du conduit vocal : fermeture par une constriction localisée. Simultanément les cordes vocales s'ouvrent beaucoup pour les consonnes fricatives non voisées (moins pour les voisées). Ces deux gestes ont une action conjuguée sur la pression comprise entre la glotte/poumons et la constriction de la consonne, elle augmente et son allure prend celle d'une parabole inversée.

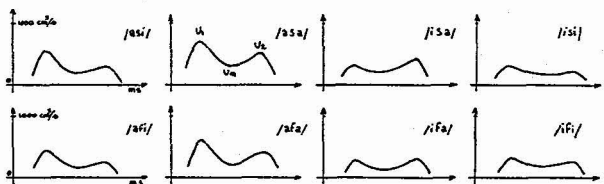
Sujet P.B. :

Item	U1	Um	U2	Pmax
/asi/	870/11	580/17	584/7	5.72/8
/asa/	1162/11	540/11	687/11	4.72/7
/isa/	834/11	485/15	718/8	5.79/7
/isi/	721/11	526/9	619/7	5.33/9
/afi/	1087/10	771/15	625/20	4.62/7
/afa/	1001/16	475/14	589/10	4.52/4
/ifa/	710/15	533/17	611/13	3.87/8
/ifi/	503/10	411/12	451/13	4.02/6



Sujet C.S. :

Item	U1	Um	U2	Pmax
/asi/	572/11	213/12	295/12	4.37/9
/asa/	703/15	300/19	480/14	4.21/16
/isa/	318/18	238/18	415/10	4.13/12
/isi/	305/15	145/18	227/15	5.20/8
/afi/	454/17	164/18	259/18	5.37/9
/afa/	589/16	179/25	372/16	5.52/6
/ifa/	230/19	116/11	268/15	4.86/6
/ifi/	291/12	150/22	285/16	5.24/7



Moyennes de pression et de débit obtenues après filtrage passe-bas sur 10 répétitions des séquences sélectionnées (U en cm³/s, P en cmH2O et écarts-type en %). Figure n° 2. Evolutions temporelles correspondantes du débit.

Les deux gestes articulatoires ont par contre des effets opposés sur le débit. L'ouverture des cordes vocales tend à augmenter celui-ci mais ce même débit est fortement diminué par les turbulences au niveau du pincement du conduit. Ces deux effets antagonistes expliquent les pics de débit que l'on observe sur les évolutions mesurées figures n° 1 et 2 (diminution de la "résistance" aérodynamique de Ag

combinée avec l'augmentation de celle de Ac). Les amplitudes respectives des pics U1 et U2 et de la vallée Um semblent dépendre de la nature de la fricative bien que les tendances soient inversées pour nos deux sujets (Um/s/ > Um/f/ pour C.S. et Um/s/ < Um/f/ pour P.B.) Ces moyennes sont cohérentes avec les mesures faites sur un corpus de consonnes fricatives soutenues réalisées avec les deux mêmes sujets. Les amplitudes dépendent aussi des voyelles coarticulées avec la consonne : le pic de débit placé après ou avant un /a/ est plus important que celui placé après ou avant un /i/. Enfin, elles semblent être très dépendantes du timing entre l'évolution temporelle des cordes vocales et celle de la constriction : le premier pic de débit est dans tous les cas plus important que le deuxième, même dans les séquences où la voyelle est identique /aCa/ ou /iCi/ (sauf /ifa/ et /isa/ C.S. où la 2ème voyelle est un /a/).

Modélisation :

Nous avons implanté dans un modèle de simulation de parole, analogue du conduit vocal [3], une source de bruit. La détection du lieu et de l'aire de la constriction se fait automatiquement et la source génère, si besoin, un bruit de frication à pente spectrale variable dont l'amplitude Sp dépend de Ac et P suivant la formule :

$Sp = k \cdot \Delta P^p \cdot Ac^q$ (k const. empirique, $p=1.43$ et $q=0.29$ pour /s/, $p=0.77$ et $q=0.07$ pour /f/ sujet P.B. [4]).

Afin de piloter ce modèle temporel de manière proche de la parole naturelle, nous l'avons couplé au modèle articulatoire à 7 paramètres de MAEDA [5]. Notre modèle complet, dont les commandes sont alors les paramètres de MAEDA augmentés des commandes du modèle à deux masses des cordes vocales, Ps , Q et $Ag0$, produit, outre le signal de parole synthétique, le débit glottique, l'aire d'ouverture de la glotte $Ag0$, l'aire de la constriction Ac , le débit moyen aux lèvres et la pression intra-orale. Les allures des 10 paramètres ont été déterminées afin de reproduire les chemins dynamiques mesurés sur le corpus de parole naturelle [6]. Pour évaluer l'influence de la coordination source/constriction, nous déplaçons l'instant de réalisation de la constriction par rapport à l'instant d'ouverture des cordes vocales. La figure n° 3a nous montre les évolutions dynamiques de l'aire d'ouverture des cordes vocales $Ag0$ et de l'aire de la constriction Ac , ainsi que des exemples de timing.

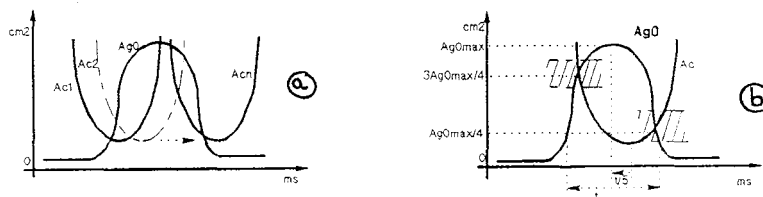


Figure n° 3. Stratégie de variation du timing entre glotte et constriction.

Résultats et discussion :

Sur la figure n° 4 sont reportés les résultats de la simulation d'une séquence /isi/ de 600 ms environ. La coordination entre glotte et constriction choisie ici permet la production d'un signal synthétique de bonne qualité. La figure n° 5 regroupe les allures de débit et de pression produites avec les différentes coordinations pour cette même séquence /isi/. La consonne /s/ est coarticulée avec la même voyelle /i/, cependant l'amplitude des pics U1 et U2 de débit dépend bien de la coordination. La vallée Um, quasiment stable en amplitude, est dépendante des pertes dues aux turbulences : elle pourra prendre une valeur plus faible si la durée de la tenue de la fricative est plus importante.

Les allures de pression sont essentiellement stables en amplitude dans nos cas considérés où constriction et ouverture des cordes vocales se produisent presque simultanément. Pour des cas extrêmes, non montrés ici, où la constriction se produirait bien avant ou bien après l'ouverture de la glotte, le maximum Pm serait bien plus petit. Par contre, nous pouvons constater que la position temporelle de ce maximum Pm et la largeur du pic de pression sont bien fonction de la coordination.

Pour produire des pics de débit proches des mesures, les simulations montrent, sur un graphe où sont reportées simultanément $Ag0$ et Ac (fig. n° 3b), que les zones optimales de croisement des courbes doivent se situer aux alentours de $Ag0 = 3/4 \cdot Ag0max$ pendant l'ouverture de la glotte et $Ag0 = 1/4 \cdot Ag0max$ pendant la fermeture. Ceci correspond à des durées d'ouverture de la glotte et de fermeture de la constriction sensiblement égales mais avec un retard de $Acmini$ par rapport à $Ag0max$ d'environ un quart à un cinquième de la durée de la tenue de la fricative.

Acoustiquement, une fermeture de la constriction *trop en avance* sur l'ouverture de la glotte produit un bruit de frication pendant la fin du voisement de la première voyelle coarticulée : le début de la fricative est alors perçu voisé. En fin de fricative, le bruit de frication cesse (Ac augmente) alors que les cordes vocales n'ont pas encore recommencé à osciller ($Ag0$ toujours grand) : un silence est produit entre la fricative et la deuxième voyelle coarticulée. De même, une fermeture *trop en retard* aurait des conséquences opposées : des oscillations fortement diminuées, voire nulles, avant que le bruit de frication ne soit audible d'où production d'un silence entre la voyelle et le début de la fricative ; une fin

de fricative voisée car le bruit de constriction serait encore produit malgré la fermeture des cordes vocales au début de la deuxième voyelle.

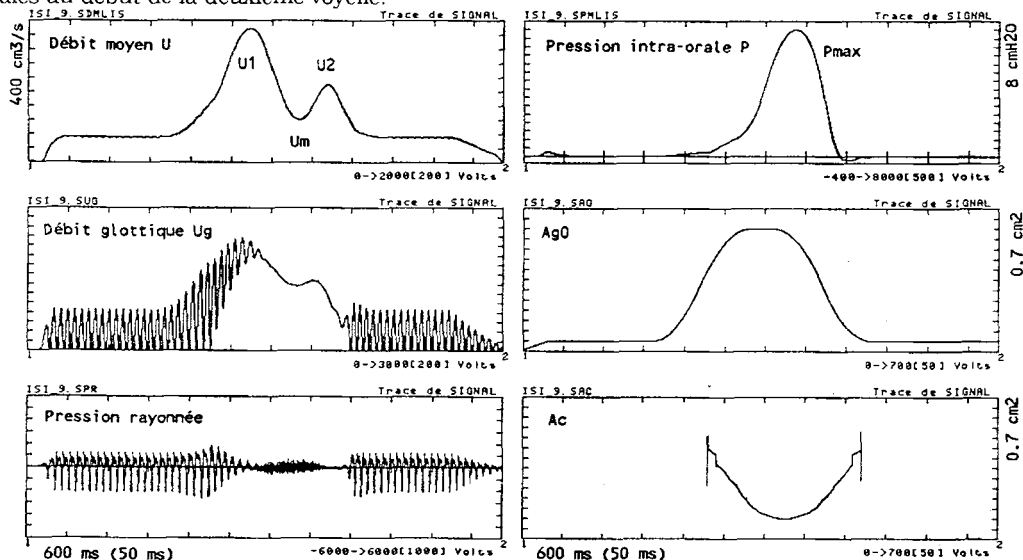


Figure n° 4. Résultats de la simulation de /isi/.

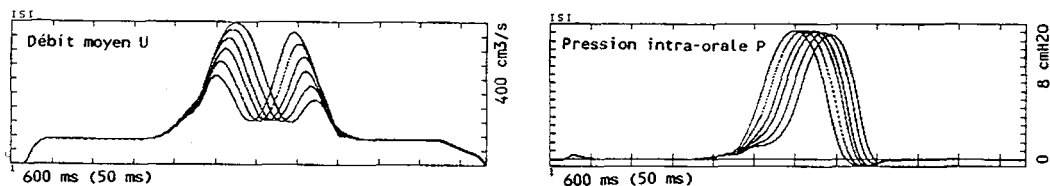


Figure n° 5. Evolutions du débit et de la pression intra-orale pour différentes coordinations

Conclusions :

Les évolutions du débit moyen en sortie des lèvres et de la pression intra-orale pendant la production des fricatives nous ont semblés des paramètres pertinents de notre corpus pour tenter de caractériser la coordination entre l'ouverture des cordes vocales et la réalisation de la constriction. Nous avons montré par simulation que leurs amplitudes dépendent fortement de ce timing et nous proposons une stratégie de commande de Ac et Ag0.

Cette première étude doit bien entendu être complétée par des vérifications sur d'autres sujets, par l'étude des autres consonnes fricatives (en particulier / $\int$ / et / θ /) et en vérifiant que notre stratégie s'applique aussi aux fricatives voisées. Une étude des influences spectrales de cette coordination, sur des spectrogrammes et fonctions de transfert, est envisagée.

Références :

- [1] BICKLEY C.A. & STEVENS K.N. (1986) "Effects of a vocal-tract constriction on the glottal source : experimental and modelling studies." *Journal of Phonetics* n°14, 373-382.
- [2] STEVENS K.N. (1971) "Airflow and Turbulence Noise for fricative and stop consonants : static considerations" *J. Acous. Soc. Amer.* 50, 1180-1192.
- [3] KELLY J.R. & LOCHBAUM C. (1962) "Speech-synthesis". *Speech Communication Seminar*, Stockholm, 127-130.
- [4] BADIN P. (1989) "Acoustics of voiceless fricatives : production theory and data" *STL-QPSR* vol.3, 33-55.
- [5] MAEDA S. (1990) "Compensatory Articulation during speech : evidence from the analysis and synthesis of vocal tract shapes using an articulatory model" in *Speech Production and Speech modeling*, W.J. Hardcastle & A. Marchal, eds., 131-149.
- [6] SCULLY C. , GEORGES E. & CASTELLI E. (1991) "Fricative consonants and their articulatory trajectories" *Proceed. of the XIIIth International Congress of Phonetic Sciences*, Aix en Provence, vol.3, 58-61.