

Modélisation physique de la production de séquences voyelle-plosive-voyelle

Louis Delebecque, Xavier Pelorson, Denis Beautemps, Xavier Laval

▶ **To cite this version:**

Louis Delebecque, Xavier Pelorson, Denis Beautemps, Xavier Laval. Modélisation physique de la production de séquences voyelle-plosive-voyelle. CFA 2014 - 12ème Congrès Français d'Acoustique, Apr 2014, Poitiers, France. hal-01158994

HAL Id: hal-01158994
https://hal.science/hal-01158994

Submitted on 2 Jun 2015

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.



Modélisation physique de la production de séquences voyelle-plosive-voyelle

L. Delebecque, X. Pelorson, D. Beautemps et X. Laval
GIPSA-lab UMR 5216, 11 rue des mathématiques, BP 46, 38402 Saint-Martin-D'Hères, France
louis.delebecque@gipsa-lab.fr

Le contexte de cette étude est la modélisation physique de la production de séquences voyelle-plosive-voyelle. Cette étude s'intéresse aux effets induits par la présence d'une constriction dans la cavité supra-glottique, sur le voisement, dans le cas d'une absence de contrôle « actif »de la part du locuteur. L'objectif est de déterminer le rôle des différents phénomènes physiques intervenant dans la mise en oscillation des plis vocaux après le relâchement de la constriction. Des mesures aéroacoustiques de séquences voyelle-plosive-voyelle sont réalisées sur une maquette de l'appareil phonatoire, qui comprend une maquette déformable de plis vocaux, capable d'auto-osciller, un résonateur représentant le conduit vocal, constitué d'un tube de section constante, et d'une constriction rigide placée en aval, dont le mouvement est imposé par un moteur pas à pas. Le Voice Onset Time (temps entre le relâchement de l'occlusion et la mise en oscillation de la maquette de plis vocaux), est mesuré pour différentes longueurs de résonateur. Ces résultats expérimentaux sont analysés au moyen de simulations numériques. Ces simulations reposent sur un modèle à deux masses pour décrire le comportement mécanique des plis vocaux, un modèle d'écoulement visqueux basé sur l'équation de Bernoulli et un modèle acoustique basé sur la propagation d'ondes planes dans un réseau de cylindres approximant la géométrie du conduit vocal (synthèse par guide d'onde). Ces résultats sont également mis en relation avec les études existantes sur l'influence de la longueur du résonateur sur la pression d'alimentation au seuil de l'oscillation.

1 Introduction

Le contexte de cette étude est la modélisation physique de la production de la parole et plus spécifiquement de la production de plosives bilabiales. L'intérêt est plus particulièrement porté sur l'influence de la modification de la géométrie du conduit vocal sur l'oscillation des plis vocaux. Ces modifications induisent des variations de l'écoulement d'air à travers l'ensemble de l'appareil phonatoire ainsi qu'une modification du couplage acoustique avec le conduit vocal.

Le comportement des plis vocaux peut être assimilé à celui d'un système dynamique non-linéaire qui résulte de l'interaction entre la structure bio-mécanique des plis vocaux et l'écoulement d'air à travers la glotte. L'auto-oscillation des plis vocaux apparaît lorsque la différence entre la pression en amont et la pression en aval de la constriction glottique atteint une valeur seuil.

Dans le cas de la clarinette, de récents travaux [1] ont mis en évidence l'existence d'un seuil de pression dynamique, obtenu lors d'une augmentation progressive de la pression d'alimentation, dont la valeur est toujours supérieure à celle du seuil de pression statique. Cet écart entre les seuils de pression statique et dynamique est interprété comme un retard à la bifurcation du système dynamique.

Lors de l'ouverture de la constriction formée au point d'articulation, la différence de pression à travers la glotte augmente rapidement puis l'oscillation des plis vocaux apparaît. Ainsi, la notion de seuil dynamique est adaptée au cas des plosives sourdes.

Le VOT (Voice Onset Time) est un indice utilisé en phonétique pour caractériser les plosives. Il désigne la durée entre l'ouverture de l'occlusion et la mise en auto-oscillation des plis vocaux. Plusieurs études ont montré que la valeur du VOT augmente lorsque le point d'articulation est déplacé des lèvres vers le vélum [16, 2], indépendamment du contexte phonétique.

Dans le cas des voyelles, des mesures sur maquette de l'appareil phonatoire ont montré que le couplage acoustique avec le résonateur situé en aval des plis vocaux influe considérablement sur la valeur de la pression sous-glottique nécessaire à la mise en oscillation [12]. Dans le cas où la première fréquence de résonance du résonateur est supérieure à la fréquence d'oscillation des plis vocaux, augmenter la longueur du résonateur a pour effet d'abaisser la pression au seuil de l'oscillation.

Ces travaux se basent sur l'hypothèse que le geste d'abduction des plis vocaux, observé durant la fermeture du conduit vocal lors de la production d'une plosive sourde, est un phénomène passif, provoqué par l'équilibrage des pressions de part et d'autre de la constriction glottique [15, 4], plutôt que par un geste d'abduction induit musculairement par le locuteur. Cette hypothèse suppose également que la reprise du voisement après relâchement de la constriction n'est pas directement contrôlée mais est due à la bifurcation du système dynamique que sont les plis vocaux, sous l'effet de l'augmentation de la différence de pression à travers la glotte. L'objectif de cette étude est de déterminer l'influence du couplage acoustique sur le VOT dans le cas de la production d'une plosive bilabiale sourde suivie d'une voyelle.

La méthode employée dans cette étude consiste, dans un premier temps, à reproduire une séquence voyelle-plosive bilabiale sourde-voyelle sur une maquette de l'appareil phonatoire qui comprend une réplique de plis vocaux auto-oscillante, et à modifier la longueur du résonateur aval. Dans une deuxième partie, ces mesures sont reproduites au moyen de simulations numériques basées sur des modèles théoriques comprenant un modèle d'écoulement visqueux, un modèle à deux masses pour décrire le mouvement des plis vocaux, et un modèle acoustique basé sur l'hypothèse de la propagation d'ondes planes dans les résonateurs amont et aval.

2 Mesures sur maquette

2.1 Dispositif expérimental

La maquette de l'appareil phonatoire humain est composée des éléments suivants :

- un réservoir de pression alimenté par un compresseur, d'environ $0,6\ m^3$ de volume, dont les parois intérieures sont couvertes de mousse acoustique,
- un résonateur « amont », un tube cylindrique de $6,5\ cm$ de long, pour représenter la trachée,
- une réplique de plis vocaux capable d'auto-osciller, faite de deux boudins de latex remplis d'eau,
- un résonateur « aval », un tube cylindrique de longueur variable pour représenter le conduit vocal,

- une réplique de lèvres rigides, dont le mouvement est contrôlé par un moteur pas-à-pas.

La figure 1 présente l'ensemble du dispositif expérimental. La réplique de plis vocaux utilisée est composée de deux tubes en latex remplis d'eau, ainsi que d'un tube en latex central guidant l'écoulement d'air entre les deux tubes remplis d'eau [11]. La réplique de lèvres rigides est formée de deux demi-cylindres de 1 *cm* rayon, et de largeur 3 *cm*. La hauteur minimale de la constriction labiale, notée h_ℓ, varie de 0 à $1,1$ *mm*. Le mouvement imposé par le moteur pas-à-pas est choisi de manière à reproduire la fermeture puis l'ouverture des lèvres lors de la production d'une plosive bilabiale. La durée d'ouverture et de fermeture des lèvres mécaniques est d'environ 150 *ms*. Elles restent closes pendant 350 *ms*.

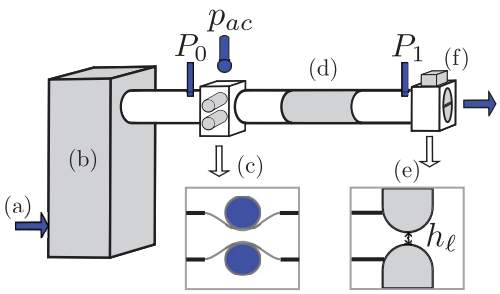


Figure 1 – Schéma de la maquette de l'appareil phonatoire. (a) : Compresseur. (b) : Réservoir de pression. (c) : Réplique de plis vocaux. (d) : Tube en plexiglas. (e) : Moteur pas-à-pas et capteur de déplacement. (f) : Réplique de lèvres.

La pression est mesurée au moyen de capteurs de pression différentielle (Endevco, 8507c-2), à deux endroits différents de la maquette. P_0, mesurée en amont de la réplique des plis vocaux est équivalente à la pression sous-glottique et P_1, mesurée en amont de la réplique de lèvres correspond à la pression intra-orale. La calibration des capteurs de pression est réalisée à l'aide d'un manomètre à eau. Un microphone (Brüel & Kjær, 4192) est placé à 10 *cm* au dessus du centre de la réplique de plis vocaux.
La valeur de h_ℓ est donnée par un capteur optique. La caractéristique non-linéaire de ce capteur est prise en compte en réalisant une calibration au moyen d'une caméra (Baumer, TXG50). La hauteur h_ℓ est estimée à partir de photos de l'ouverture des lèvres mécaniques, en utilisant une méthode de traitement d'images automatique. La précision finale sur l'estimation de h_ℓ est de $\pm 0,05$ *mm*. Les différents signaux mesurés sont numérisés à une fréquence d'échantillonnage de 10 *kHz*. La longueur du résonateur aval est modifiée en utilisant différents tubes en plexiglas de longueur L_p, tels que $L_p = \{6; 12; 16; 22\}$ *cm*. La longueur totale du résonateur aval est donnée en ajoutant 19 *cm* à L_p.

2.2 Analyse des résultats

La figure 2 présente un exemple de mesure dans le cas où $L_p = 12$ *cm*. Ce cas est utilisé comme mesure de référence dans la suite de cette étude. L'évolution de la fréquence fondamentale $F0$ est déterminée à partir du signal de pression acoustique p_{ac}, en utilisant une méthode donnée par [3]. Cette technique est basée sur l'auto-corrélation du signal temporel. Les oscillations contenues dans le signal sont détectées à partir d'un certain seuil.
La mesure présentée en figure 2 montre que l'oscillation de la réplique de plis vocaux est interrompue lors de la fermeture des lèvres mécaniques, en raison de l'augmentation de P_1 jusqu'au niveau de P_0. La pression P_1 retrouve une valeur proche de zéro environ 75 *ms* après l'ouverture des lèvres, alors que l'oscillation est détectée seulement 200 *ms* après l'ouverture des lèvres mécaniques.

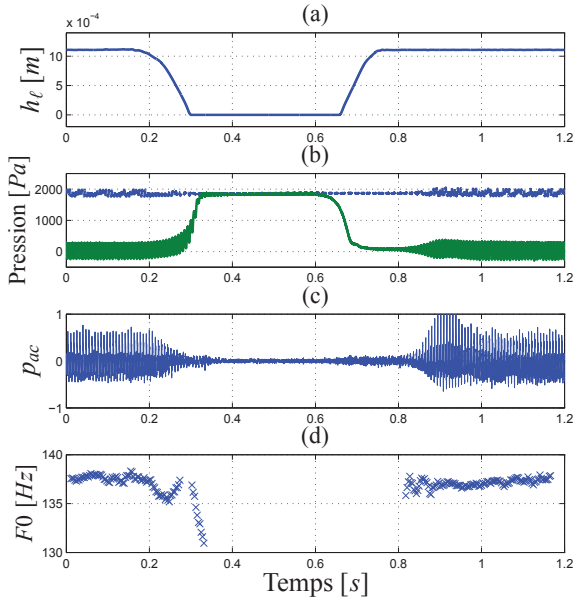


Figure 2 – Exemple de mesure réalisée pour $L_p = 12$ *cm*. (a) : Ouverture h_ℓ des lèvres mécaniques. (b) : Pressions P_0 (en pointillés bleus) et P_1 (en traits pleins verts). (c) : Pression acoustique p_{ac}. (d) : Évolution de la fréquence fondamentale $F0$ du signal p_{ac}.

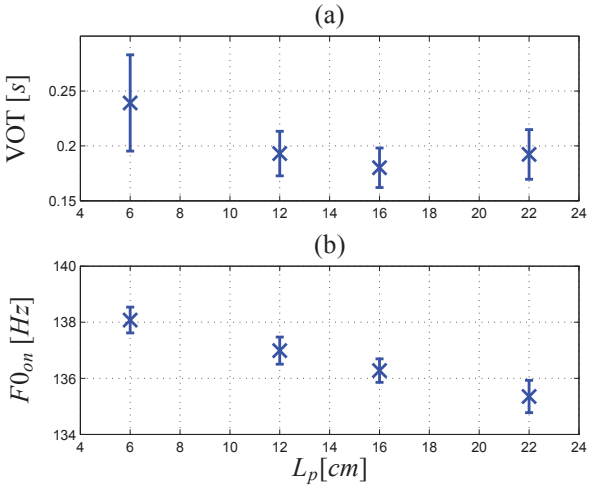


Figure 3 – (a) : VOT mesurés en fonction de L_p. (b) : Fréquence fondamentale $F0_{on}$ au seuil de l'oscillation en fonction de L_p. Les croix représentent la valeur moyenne sur 8 mesures et les barres verticales représentent l'écart-type.

Le VOT et la fréquence fondamentale $F0_{on}$ au seuil de l'oscillation sont déterminés pour chaque mesure. Le VOT est défini ici comme l'écart de temps entre l'instant de l'ouverture des lèvres mécaniques et l'instant pour lequel

la première valeur de $F0$ est détectée par algorithme. La fréquence $F0_{on}$ est définie comme la valeur moyenne de $F0$, calculée sur 50 *ms* à partir de la première valeur détectée. La figure 3 présente les VOT et les fréquences $F0_{on}$ mesurés pour les différentes valeurs de L_p. L'ordre de grandeur des valeurs de VOT obtenues avec ce dispositif expérimental est plus important que celui des valeurs de VOT mesurées sur des locuteurs (entre 40 et 100 *ms* pour les plosives sourdes [9, 8]). La figure 3 montre que l'augmentation de la longueur du résonateur aval a pour effet de diminuer le temps nécessaire à l'émergence des oscillations. Cependant les mesures effectuées dans le cas où $L_p = 22$ *cm* ne confirment pas cette tendance. La fréquence fondamentale $F0_{on}$ mesurée au seuil de l'oscillation décroît avec l'augmentation de la longueur du résonateur aval.

3 Simulations numériques et modèles théoriques

Des simulations numériques sont réalisées à partir des mesures sur maquette dans le but de vérifier le rôle du couplage acoustique dans l'émergence des oscillations lors de la production de plosives sourdes. Les modèles théoriques qui alimentent ces simulations sont basés sur ceux utilisés par Lous [10]. Les modifications apportées ici par rapport au modèle de Lous sont la prise en compte des pertes dues à la viscosité dans l'écoulement d'air [13, 12] et l'ajout d'une constriction an aval des plis vocaux, représentant les lèvres. Cette dernière modification implique que la pression P_{io} dans la cavité supra-glottique n'est pas considérée comme nulle, mais est calculée en fonction de la pression sous-glottique P_{sub} et de la géométrie des constrictions glottique et labiale. Les paramètres d'entrée de la simulation sont P_{sub} et la hauteur minimale de la constriction labiale h_ℓ. La figure 4 représente la géométrie et les paramètres aérodynamiques de la simulation.

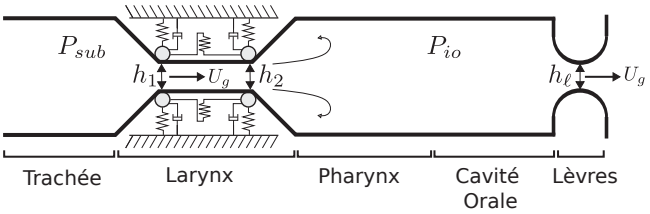


FIGURE 4 – Schéma de la géométrie du modèle aérodynamique de l'appareil phonatoire, incluant le modèle à deux masses et un modèle de lèvres rigides à géométrie ronde.

3.1 Modèle aérodynamique

Le modèle d'écoulement utilisé est l'équation de Bernoulli stationnaire à laquelle est ajouté un terme issu de la théorie de la lubrification de Reynolds et qui représente les pertes de charge dues aux frottements visqueux. Pour une constriction de largeur constante W et de hauteur variable $h(x)$, dans le sens de l'écoulement x, le modèle donne le débit volumique U, constant sur l'ensemble de la constriction tel que

$$\begin{aligned} P(0,t) + \frac{\rho}{2}\left(\frac{U_g(t)}{Wh(0,t)}\right)^2 &= P(x,t) + \frac{\rho}{2}\left(\frac{U_g(t)}{Wh(x,t)}\right)^2 + \Delta P_v(x) \\ \text{avec } \Delta P_v(x,t) &= \frac{12\rho\nu U_g(t)}{W}\int_0^x \frac{\mathrm{d}x}{h^3(x,t)} \end{aligned} \quad (1)$$

où ρ est la masse volumique de l'air et ν le coefficient de viscosité cinématique de l'air. L'équation 1 est appliquée pour chacune des deux constrictions, entre l'origine de la constriction et le point de séparation de l'écoulement. Pour le modèle de plis vocaux, le point de séparation se déplace le long de la plaque centrale (celle comprise entre les deux masses). Sa position est déterminée en utilisant le critère empirique défini par [10] tel que

$$h_s(t) = h(x_s,t) = \alpha \min\{h(x,t)\}, \quad (2)$$

où $\alpha = 1,1$. Dans le cas de la constriction labiale, cette même condition est appliquée en considérant que le point de séparation est situé en aval du point d'ouverture minimale de la constriction.

3.2 Modèle mécanique de plis vocaux

Le modèle mécanique utilisé pour décrire le mouvement des plis vocaux est le modèle à deux masses symétrique. La position y_i de la masse m_i est donnée par l'équation

$$m_i\ddot{y}_i + r_i\dot{y}_i + k_i y_i + k_c(y_i - y_j) = F_i \quad (3)$$

où la notation $\dot{y}_i$ désigne la dérivée temporelle de y_i. k_i et r_i sont respectivement la raideur et le coefficient d'amortissement associés à la masse m_i et k_c représente la raideur du ressort entre les masses m_i et m_j.

La force aérodynamique F_i qui s'applique sur la masse m_i est donnée par

$$\begin{aligned} F_i(t) &= W_g \int_{x_{i-1}}^{x_i} \frac{x - x_{i-1}}{x_i - x_{i-1}} p(x,t)\,\mathrm{d}x \\ &+ W_g \int_{x_i}^{x_{i+1}} \frac{x_{i+1} - x}{x_{i+1} - x_i} p(x,t)\,\mathrm{d}x, \end{aligned} \quad (4)$$

où W_g est la largeur de la constriction glottique, x_0, x_1, x_2 et x_3 sont respectivement les abscisses correspondant aux positions du début de la constriction glottique, des positions des masses 1 et 2 et de fin de la constriction glottique.

3.3 Modèle de propagation acoustique

Le modèle de propagation acoustique est un modèle mono-dimensionnel, basé sur l'hypothèse que la longueur d'onde est grande devant les dimensions du tube considéré. Il consiste en l'association de l'équation d'onde et de l'équation d'Euler linéarisées au premier ordre,

$$\left(\partial_x^2 - \frac{1}{c^2}\partial_t^2\right) p(x,t) = 0 \quad (5)$$

$$\frac{\rho}{S(x)}\partial_t u(x,t) + \partial_x p(x,t) = 0 \quad (6)$$

où $S(x)$ est la section du tube, c est la célérité des ondes acoustiques et ∂_x désigne l'opérateur de dérivée partielle par rapport à la variable x. L'état acoustique en un point x du

tube est décrit en fonction de la pression acoustique p et du débit acoustique u.
La géométrie des deux résonateurs est approximée par une agrégation de tronçons cylindriques. Les variables acoustiques p et u sont décomposées en ondes p^+ et p^- qui correspondent aux ondes planes progressives dont la propagation est découplée dans les tubes cylindriques [5]. Les ondes p^+ et p^- sont définies par

$$\begin{pmatrix} p^+(x,t) \\ p^-(x,t) \end{pmatrix} = \frac{1}{2} \begin{pmatrix} 1 & Z_c(x) \\ 1 & -Z_c(x) \end{pmatrix} \begin{pmatrix} p(x,t) \\ u(x,t) \end{pmatrix} \quad (7)$$

avec $Z_c(x) = \rho c / S(x)$.

À la jonction entre deux cylindres notés n et $n+1$, de longueur L, les conditions de continuité de pression et de débit font apparaître les fonctions de transmission et de réflexion, qui dans le cas des cylindres se présentent comme de simples gains. L'expression de ces fonctions de transfert, en fonction des ondes p^+ et p^-, sous une forme qui respecte le principe de causalité, fait apparaître une factorisation par la fonction de réflexion k_n [7] telle que

$$\begin{pmatrix} p^+_{n+1}(0,t) \\ p^+_n(L,t) \end{pmatrix} = \begin{pmatrix} 1+k_n & -k_n \\ -k_n & 1-k_n \end{pmatrix} \begin{pmatrix} p^-_n(L,t) \\ p^-_{n+1}(0,t) \end{pmatrix} \quad (8)$$

avec $k_n = (S_n - S_{n+1})/(S_n + S_{n+1})$.

La fonction de réflexion aux extrémités de chaque résonateur est calculée à partir de l'impédance de rayonnement d'un piston plan bafflé dans un écran infini. La fonction de réflexion est déterminée grâce à la formule approchée donnée par [14]. Le couplage entre les grandeurs acoustiques et aérodynamiques se fait de part et d'autre de la constriction glottique en considérant qu'à l'entrée des deux résonateurs la pression totale est la somme de la pression aérodynamique et de la pression acoustique (par exemple $P_{io} + p_{supra}$ à l'entrée du résonateur aval) et que le débit acoustique u est égal au débit glottique U_g.

3.4 Implémentation numérique

L'évolution de la section interne le long de chacun des deux résonateurs est déterminée à partir des dimensions des différents éléments utilisés dans le montage. Cette fonction d'aire est discrétisée en tronçons cylindriques de 5 mm de longueur (figure 5). En prenant $c = 247\ m.s^{-1}$, ce choix donne une fréquence d'échantillonnage $Fe = 69378\ Hz$. Cette discrétisation est réalisée en prenant la valeur moyenne de la fonction d'aire sur chaque tronçon.

Le modèle de propagation acoustique (équations 5 et 6), la condition de continuité de p et de u à la jonction entre deux cylindres, et l'expression de l'impédance de rayonnement, permettent d'obtenir l'expression de l'impédance d'entrée $Z_e(f)$ du résonateur aval dans le domaine fréquentiel. La figure 5 présente l'évolution du module de $Z_e(f)$ au niveau de la première fréquence de résonance. Pour des fréquences proches de celle de la fréquence fondamentale des oscillations produites par la réplique de plis vocaux (environ 140 Hz), plus le résonateur est long et plus la valeur du module de $Z_e(f)$ est important.
La pression P_{sub} est obtenue en filtrant passe-haut la pression P_0 mesurée pour la mesure de référence, au moyen d'un filtre de Butterworth d'ordre 2 et de fréquence de coupure

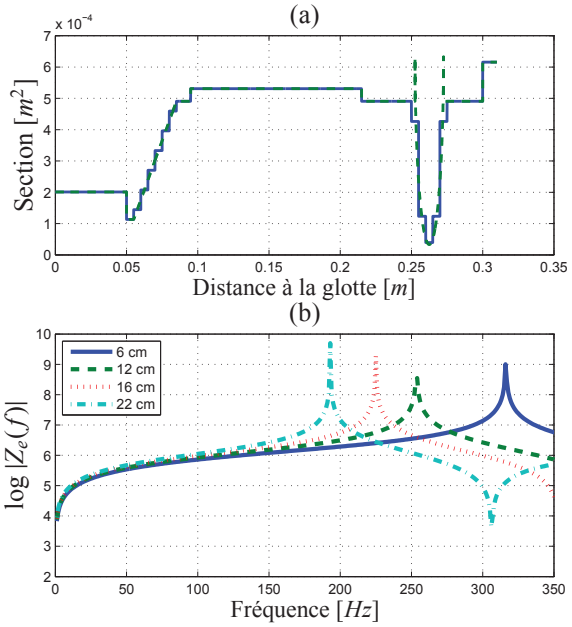


FIGURE 5 – (a) : Évolution de la section du résonateur aval en fonction de la distance à la glotte pour $L_p = 12\ cm$ et $h_\ell = 1,1\ mm$ (version continue en pointillés verts et version discrétisée en traits pleins bleus). (b) : Module de l'impédance d'entrée $Z_e(f)$ calculée à la glotte, pour le résonateur aval et pour différentes valeurs de L_p. $Z_e(f)$ est calculée à partir de l'impédance de rayonnement et de la géométrie discrétisée du résonateur, dans le cas où $h_\ell = 1,1\ mm$.

30 Hz. Le signal de l'ouverture h_ℓ des lèvres mécaniques est directement issu de la mesure de référence. Une interpolation polynomiale est réalisée pour adapter ces deux signaux à la fréquence d'échantillonnage de la simulation.
La fonction d'aire est recalculée puis discrétisée à chaque itération temporelle en prenant en compte la nouvelle valeur de h_ℓ. La fonction de réflexion aux extrémités de chaque résonateur est approximée par un filtre RIF d'ordre 20 en utilisant la méthode de la fenêtre.
Pour chaque pas de temps, le calcul de P_{io} et U_g est réalisé en deux étapes. Dans un premier temps, P_{io} est déterminée en utilisant uniquement le modèle aérodynamique (équation 1). Le débit U_g est ensuite calculé en considérant la différence de pression à travers la glotte $P_{sub} - P_{io}$ ainsi obtenue et la pression acoustique due aux ondes planes incidentes. Le déplacement des deux masses du modèle est donné par l'équation 3 qui est discrétisée par différence finie avec un schéma implicite.

3.5 Résultats

La valeur des paramètres mécaniques reste constante pour l'ensemble des simulations réalisées. Les paramètres du modèle à deux masses sont fixés tel que $m_1 = m_2$, $k_1 = k_2$, $r_1 = r_2 = 0,1 \times 2\sqrt{m_1 k_1}$ et $k_c = k_1/2$. L'ordre de grandeur des paramètres de masse et de raideur de la réplique de plis vocaux sont obtenus par une méthode d'identification expérimentale [6]. Les valeurs de k_1 et k_2 sont ajustées pour obtenir une fréquence fondamentale semblable à celle de la mesure de référence. La hauteur initiale et la largeur de la constriction glottique sont choisies de manière à obtenir

une oscillation des plis vocaux sans collision. La fréquence fondamentale $F0$ est déterminée de la même façon que pour les mesures réalisées sur maquette, à partir de la pression acoustique p_{supra} simulée à l'entrée du résonateur aval.

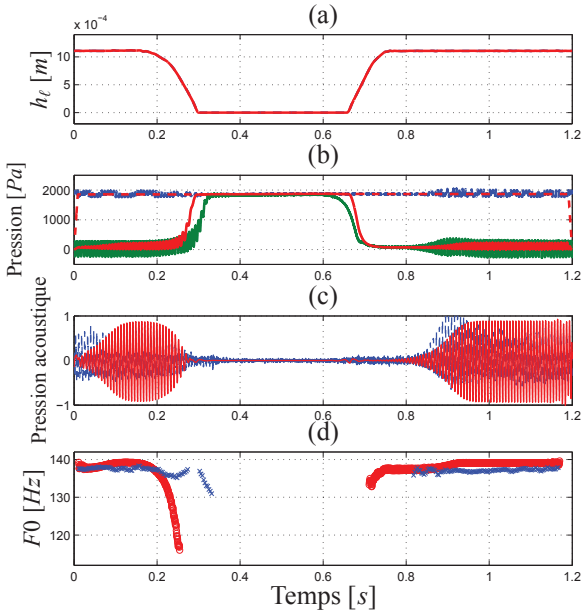


FIGURE 6 – Exemple de simulation réalisée à partir de la mesure présentée en figure 2. (a) : Ouverture h_ℓ des lèvres mécaniques. (b) : Pressions P_{sub} (en pointillés rouges), P_{io} (en traits pleins rouges). (c) : Pression acoustique p_{supra}. (d) : Évolution de la fréquence fondamentale $F0$ du signal p_{supra}. Les signaux simulés sont superposés aux signaux mesurés.

Les simulations sont effectuées pour des longueurs L_p allant de 4 à 24 *cm* par pas de 2 *cm*. La figure 6 présente un exemple de simulation basée sur la mesure de référence (pour $L_p = 12$ *cm*). Lors de la fermeture des lèvres, les oscillations simulées sont maintenues légèrement moins longtemps et la chute de $F0$ simulée est plus importante que celle observée pour la mesure. Même si l'évolution de la forme d'onde des pressions acoustiques mesurées et simulées est similaires, la présence de bruit aérodynamique dans le signal de pression acoustique mesuré induit que les oscillations sont détectées plus tard dans ce cas.
La figure 7 présente l'ensemble des résultats de simulation obtenus. Les valeurs de VOT ainsi obtenues sont plus faibles que celles mesurées sur la maquette. Plus le résonateur aval est long et plus la valeur du VOT simulé diminue. L'évolution de la fréquence $F0_{on}$ suit la même tendance que celle observée par la mesure, $F0_{on}$ décroit avec l'augmentation de L_p. Dans les deux cas, cette décroissance semble quasi-linéaire, cependant les variations de $F0_{on}$ sont plus importantes dans le cas de la simulation. Cette différence peut s'interpréter comme l'effet d'un couplage acoustique plus important pour les simulations.

4 Conclusion

Les mesures sur maquette de l'appareil phonatoire et les simulations réalisées dans cette étude montrent que, dans le cas d'un contrôle laryngé inexistant, le couplage

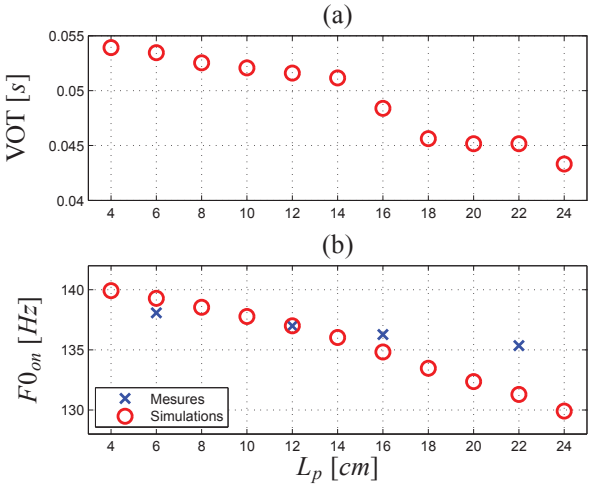


FIGURE 7 – (a) : VOT simulés en fonction de L_p. (b) : Comparaison entre les fréquences fondamentales $F0_{on}$ simulée et mesurée moyennes, au seuil de l'oscillation, en fonction de L_p.

acoustique a une influence non négligeable sur le VOT (Voice Onset Time) d'une plosive sourde bilabiale. En effet, des variations relatives de la valeur du VOT de l'ordre de 20% ont été observées pour la mesure et pour la simulation. L'augmentation de la longueur du résonateur situé en aval des plis vocaux (assimilable au conduit vocal) a pour effet de diminuer la valeur du VOT. L'interprétation proposée est que l'augmentation de la longueur du conduit vocal fait diminuer la valeur de sa première fréquence de résonance, qui se rapproche ainsi de la fréquence fondamentale de l'oscillation des plis vocaux. Le rapprochement entre ces deux fréquences favorise l'émergence des oscillations.
Cependant cet effet n'est pas observé pour les mesures réalisées sur maquette, dans le cas d'un résonateur de 41 *cm* de longueur. Les simulations numériques effectuées ne permettent pas d'expliquer ce cas de figure. Une explication possible de cette différence entre la mesure et la simulation est que le modèle de propagation acoustique n'est pas assez raffiné pour décrire l'influence des parois déformables au niveau de la maquette de plis vocaux. L'analyse des différentes valeurs de fréquence fondamentale au seuil de l'oscillation montre que les modèles théoriques utilisés surestiment l'effet du couplage acoustique avec le conduit vocal.
Cette étude montre que, dans le cas des consonnes plosives sourdes, l'influence du couplage acoustique est similaire à celle observée dans le cas des voyelles : l'augmentation de la longueur du conduit vocal abaisse la valeur de la pression au seuil de l'oscillation dans le cas statique, et réduit le retard à la bifurcation du système dans le cas dynamique.
Ce résultat peut être mis en parallèle avec les observations issues de différentes études phonétiques [16, 2] qui montrent que le VOT moyen augmente lorsque la longueur de la portion du conduit vocal comprise entre la glotte et le point d'articulation diminue. Une perspective intéressante de cette étude serait de déterminer si les variations de la position du larynx peuvent avoir une influence significative sur le VOT. L'effet de la prise en compte des pertes visco-thermiques et de la vibration des parois du conduit vocal dans le modèle de propagation acoustique reste également à analyser.

Remerciements

Ce travail est supporté par la région Rhône-Alples au moyen d'une bourse de thèse CIBLE et du projet ANR Plasmody.

Références

[1] B. Bergeot, A. Almeida, C. Vergez et B. Gazengel : Prediction of the dynamic oscillation threshold in a clarinet model with a linearly increasing blowing pressure : influence of noise. *Nonlinear Dynamics*, 74(3):591–605, octobre 2013. 14 pages.

[2] T. H. Crystal et A. S. House : Segmental durations in connected-speech signals : current results. *Journal of Acoustical Society of America*, 83(4):1553–1573, 1988.

[3] A. de Cheveigné et H. Kawahara : YIN, a fundamental frequency estimator for speech and music. *The Journal of the Acoustical Society of America*, 111(4):1917–1930, 2002.

[4] L. Delebecque, X. Pelorson, D. Beautemps et X. Laval : Physical modeling of bilabial plosives production. *In Proceedings of Meetings on Acoustics - POMA*, volume 19, page 035047 (9pages), Montréal, Canada, juin 2013.

[5] G. Fant : *Acoustic theory of speech production*. Mouton and co. édition, 1960.

[6] N. Hermant, X. Pelorson, N. Ruty et B. Maillou : New measurement method to derive mechanical parameters on a vocal fold replica. *In Pan European Voice Conference (PEVOC9)*, Marseille, France, août 2011.

[7] J. K. Kelly et C. C. Lochbaum : Speech synthesis. pages 1–4. 4th International Congress on Acoustics, 1962.

[8] D. H. Klatt : Voice onset time, frication, and aspiration in word-initial consonant clusters. *Journal of Speech, Language, and Hearing Research*, 18:686–706, 1975.

[9] L. Lisker et A. S. Abramson : A cross-language study of voicing in initial stops : Acoustical measurements. *Word*, 20:384–422, 64.

[10] N.J.C. Lous, G.C.J. Hofmans, R.N.J. Veldhuis et A. Hirschberg : A symmetrical two-mass vocal-fold model coupled to vocal tract and trachea, with application to prosthesis design. *Acta Acustica*, 84: 1135–1150, 1998.

[11] X. Pelorson et X. Laval : An experimental replica of the vocal folds to study normal and pathological voice. *In* Société Française d'Acoustique, éditeur : *Proceedings of the Acoustics 2012 Nantes Conference*, Nantes, France, avril 2012.

[12] N. Ruty : *Modèles d'interactions fluide/parois dans le conduit vocal. Applications aux voix et aux pathologies*. Thèse de doctorat, Institut National Polytechnique de Grenoble, Novembre 2007.

[13] D. Sciamarella et C. D'alessandro : On the acoustic sensitivity of a symmetrical two-mass model of the vocal folds to the variation of control parameters. *Acta Acustica*, 90:746–761, 2004.

[14] F. Silva, P. Guillemain, J. Kergomard, B. Mallaroni et A. N. Norris : Approximation formulae for the acoustic radiation impedance of a cylindrical pipe. *Journal of Sound and Vibration*, 322:255–263, 2009.

[15] K. N. Stevens : Models for the production and acoustics of stop consonants. *Speech communication*, 13(3):367–375, 1993.

[16] V. Zue : *Acoustic characteristics of stop consonants : A controlled study*. Thèse de doctorat, Massachusetts Institute of Technology, 1976.