Physical modeling of bilabial plosives production

Louis Delebecque, Xavier Pelorson, Denis Beautemps, Xavier Laval

Proceedings of Meetings on Acoustics

Volume 19, 2013 http://acousticalsociety.org/

ICA 2013 Montreal

Montreal, Canada

2 - 7 June 2013

Musical Acoustics

Session 3aMU: Aeroacoustics of Wind Instruments and Human Voice II

3aMU7. Physical modeling of bilabial plosives production

Louis Delebecque*, Xavier Pelorson, Denis Beautemps and Xavier Laval

***Corresponding author's address: GISPA-Lab, Grenoble, 38700, Isère, France, louis.delebecque@gipsa-lab.grenoble-inp.fr**

The context of this study is the physical modeling of speech production. The first step of our approach is to realize in-vivo measurements during the production of the vowel-consonant-vowel sequence /apa/. This measurements concerns intra-oral pressure, acoustic pressure radiated at the lips and labial parameters (aperture and width of the lips) derived from a high-speed video recording of the subject face. In a second time, theoretical models from speech production literature are under investigation to predict the air flow trough the lips. A model is validated by comparing his predictions with results obtained from measurements on a replica of phonatory system. Then, the same experimental set-up is used to introduce an aerodynamic model of supraglottal cavity expansion. Finally, we achieve numerical simulations of a vowel-bilabial plosive-vowel utterance, by using these models. Simulation results highlight the influence of the cheeks expansion during the production of bilabial plosives.

Published by the Acoustical Society of America through the American Institute of Physics

 [DOI: 10.1121/1.4799466]
Received 22 Jan 2013; published 2 Jun 2013

INTRODUCTION

For speech production studies, physical modeling aims at identifying the mechanisms involved and to characterize their importance. The goals are to understand the physical phenomena involved and to achieve modeling in order to predict their consequences. Speech sound signals can be synthesized using numerical simulations based on theoretical models. One of our motivation is to discriminate between the phenomena that are directly and deliberately controlled by the speaker (articulatory motion for example) and those induced by a physical effect.

In this study, we focus on bilabial plosives, especially on the air flow and fluid-walls interaction in the part of the vocal tract downstream of the glottis. During the production of vowel-bilabial plosives-vowel sequences, air flow supplied by the lungs is stopped by the closure of the lips, which causes an increase of the intra-oral pressure. At the release of the occlusion made by lips, this over-pressure is responsible for the formation of jet which causes an audible acoustic disturbance ("burst"). As vocal folds oscillations are driven by pressure difference across the glottis, a volume variation in the upper part of the vocal tract as an influence on voice onset time. According to Westbury (1983), enlargements of the supraglottal cavity helps to sustain oscillations during a voiced stop consonant.

In the first part of this paper, we present in-vivo measurements of an /apa/ utterance. Then two theoretical air flow models are exposed and validated against experimental results obtained from measurements on a replica of phonatory system. This experimental setup allows us to introduce an empirical model to link intra-oral pressure and supraglottal cavity volume. Finally, numerical simulations of a vowel-bilabial plosive-vowel sequence are performed using the air flow model validated before, and compared to in-vivo measurements.

MEASUREMENTS ON SPEAKER FOR AN /APA/ UTTERANCE

Experimental Set-up

In-vivo measurements for an /apa/ utterance are realized using a high speed camera (Mikrotron, EoSens Cube7), a cardioid microphone (AKG C1000S) and a differential pressure sensor (Endevco, 8507c-2). These measurements concern :

- Lips parameter derived from the video recording of the subject face : lips with A and lips height B,
- Acoustic pressure P_a radiated at the lips,
- Intra-oral pressure IOP, measured by a probe tube of length 1.2 m and of diameter 5 mm, inserted inside the mouth cavity.

The sampling rate of the camera is set to 500 Hz in order to measure accurately times at the closure and the opening of the lips. The video signal is synchronized to pressure signals by the means of a trigger signal, send from the A/D converter to camera. The pressure sensor used to measure IOP is calibrated using a water column manometer.

Data Analysis

Lips parameters are extracted by a semi-manually method : edge of intero-labial lips area, computed using the Canny edge detector, is plotted over the picture, then A and B are obtained by clicking manually on each corner of the lips contour. The calibration from pixels to centimeters is performed using a 4 mm long benchmark grid, hold between the lips. An example of results is presented in figure 1.

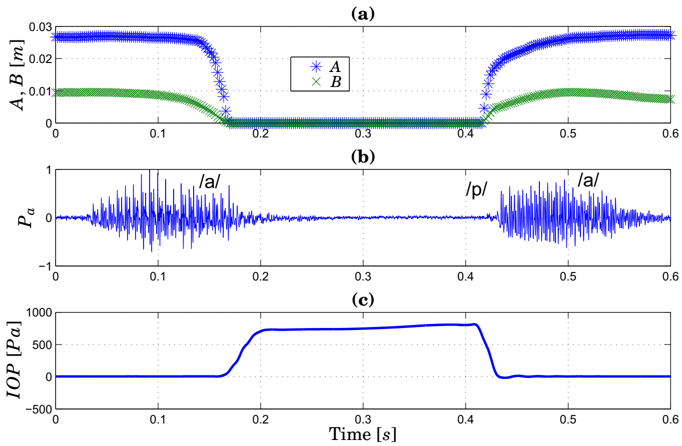

FIGURE 1: In-vivo measurements for an /apa/ utterance. **(a)** : lips parameters (width A and height B) extracted from the video recording. **(b)** : acoustic pressure P_a measured at 50 cm of distance from lips. **(c)** : intra-oral pressure IOP low-pass filtered with butterworth filter of order 6 and cut frequency of 200 Hz.

EXPERIMENTAL VALIDATION OF A FLOW MODEL

Theoretical Models

We present here two quasi-steady theoretical models to describe the air flow through the lips during a plosive. These models are used by Deverge *et al.* (2003) to predict air flow in the glottis. We assumed that the lips shape is rounded, figure 2 illustrates this geometrical model.

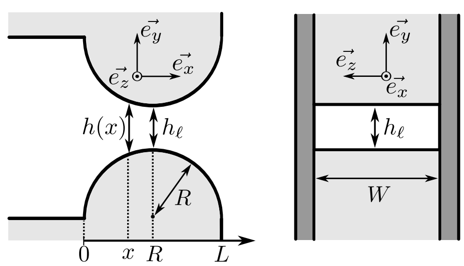

FIGURE 2: Geometrical model for the lips constriction

The height $h(x)$ of the lips constriction along the flow direction is given by

$$h(x) = h_\ell + 2\left(R - \sqrt{R^2 - (x-R)^2}\right) \tag{1}$$

where h_ℓ is the minimum height of the channel, for $x = R$.

Bernoulli Equation Corrected for Friction

The corrected Bernoulli model is composed of the inviscid Bernoulli equation and an additional term $\Delta P_v(x)$ coming from the lubrification theory of Reynolds and accounting for viscous losses. As the height of the mouth constriction is large compared to lips height h_ℓ, the flow velocity upstream to the lips constriction is neglected, the corrected Bernoulli equation becomes :

$$p(x) = p(x=0) - \frac{\rho}{2}\left(\frac{\Phi_V}{W h_\ell}\right)^2 - \Delta P_\nu(x) \quad (2)$$

$$\text{with } \Delta P_\nu(x) = \frac{12\rho\nu\Phi_V}{W}\int_0^x \frac{\mathrm{d}x}{h^3(x)} \quad (3)$$

This model takes into account the formation of a jet in the divergent part of lips constriction. We assume here that the pressure downstream to the separation point is equal to the atmospheric pressure, $P(x > x_s) = 0$, where x_s is the separation point abscissa. We make the assumption that the position of the separation point depends only on the channel geometry and we choose a semi-empirical criterion to predict x_s : $h_s = h(x_s) = \alpha h_L$ with $\alpha = 1.1$. The abscissa of the separation point can be computed analytically using this criterion and equation 1 :

$$x_s = R + \sqrt{R^2 - \left(R - h_\ell \frac{(\alpha-1)^2}{4}\right)} \quad (4)$$

The quadratic equation given by (2) is solved by applying the Bernoulli equation between the beginning of the constriction ($x = 0$) and the separation point ($x = x_s$). Giving volume flow velocity Φ_V, we can computed the pressure $P(x)$ using equation 2 between x and x_s, with $x < x_s$. The term $\Delta P_\nu(x)$, accounting for viscous losses is computed by numerical integration.

Boundary Layer Model

The boundary layer theory assumes that the flow is separated in two different parts : a first one, far enough from the wall, within flow is considered to be frictionless and unidimensional and a second one, called the boundary layer, within which the flow is two-dimensional and viscous. Under the assumption of a thin boundary layer and the equality between viscous and convectional terms, the flow may be predict using the Von Kàrmàn integral formulation, which cannot be solved analytically in case of a complex channel geometry like a rounded one. More details can be founded in Schlichting and Gersten (1999). The Von Kàrmàn equations are solved numerically using the semi-empirical Thwaites method (Vilain (2002)). This method allows to predict the position of the separation point x_s, pressure and the thickness δ of the boundary layer, upstream to this specific point.

Experimental Set-up

The relevance and accuracy of these theoretical models for predicting the air flow trough the lips, are evaluated by the comparison with experimental data obtained from measurements on a mechanical replica of human phonatory system (Figure 3).

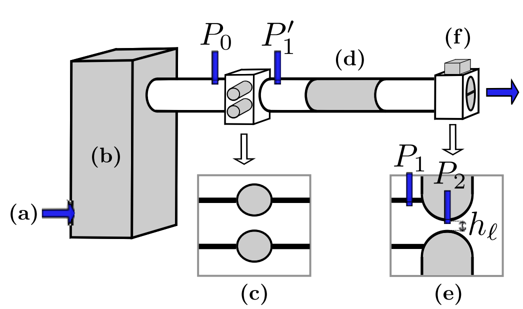

FIGURE 3: Replica of human phonatory system. **(a)** : air compressor. **(b)** : pressure tank. **(c)** : replica of vocal folds. **(d)** : rigid plexiglass tube. **(e)** : motor and mechanical lips displacement sensor. **(f)** : metal replica of lips.

The replica is composed of the following elements :

- a pressure tank (about 0.6 m^3), supplied by a compressor,
- a replica of the vocal folds made of latex tubes filled with water, able to self-oscillate,
- a replica of lips in metal, whose motion of the upper part is controlled by a step motor,
- a metal uniform tube of length 12 cm with an internal diameter of 2.5 cm for modeling the trachea,
- a uniform tube of length 18 cm in metal, including a plexiglass tube of 12 cm length, with the same diameter as the trachea, to represent the vocal tract.

Pressure along replica, is measured in four points using differential pressure sensors. Pressure P_0, P_1', P_1 and $P_2 = P(x = R)$ correspond respectively to subglottal, supraglottal, intra-oral and intra-labial pressure. A latex layer of thickness 0.2 mm, is stuck to the mobile part of the mechanical lips to improve air tightness during the closure. The height h_ℓ of mechanical lips is measured using an optic sensor. As the optic signal is not a linear function of h_ℓ, the sensor is calibrated using a camera in order to extract by image processing, h_ℓ from pictures. The calibration consist in a polynomial interpolation of the sensor characteristic. The motion of mechanical lips, imposed by the motor, is sinusoidal-like and h_ℓ varies from 0 to 1 mm. Other dimensions of the lips channel are kept constant : $L = 2\ cm$ and $W = 3\ cm$.

Results

The relevance of presented theoretical models for modeling flow in the lips channel is evaluated by a comparison between the pressure P_2 measured and the theoretically predicted from P_1. This measurements are performed in case of an unsteady flow to mimic in-vivo conditions for a vowel-plosive-vowel utterance.

Figure 4 presents an example of comparison between prediction of theoretical models and experiment data for a mechanical lips oscillation period $T = 0.564\ s$. We are particularly interested in predicting the pressure in the lips channel during the closure and the opening period.

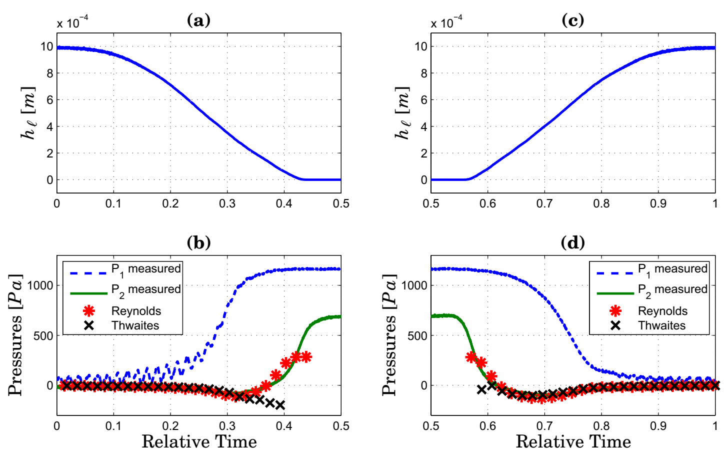

FIGURE 4: Experimental and theoretical results for unsteady flow, with oscillation of vocal folds replica, time axis is normalized by the mechanical lips oscillation period $T = 0.564\ s$. **(a)** : height h_ℓ of the mechanical lips during the closure. **(b)** : experimental and theoretical pressure P_2 computed from experimental pressure P_1, using two theoretical models : Bernoulli equation corrected for friction (Reynold) and boundary layer equations (Thwaites), during the closure, value for which $\delta > h_\ell/2$ are not plotted. **(c)** : height h_ℓ of the mechanical lips during the opening. **(d)** : experimental and theoretical pressures during the opening.

Although Thwaites method is more accurate for $h_\ell > 3\ mm$, Reynolds theory provides good agreement with experimental data for lower apertures. Moreover the boundary layer model fails to predict P_2 for small mechanical lips aperture because under these geometrical conditions, the predicted boundary thickness becomes larger than the height of the channel ($\delta > h_\ell/2$). We can notice that predictions by boundary layer equations are less accurate during the closure than during the opening. This phenomenon might be an unsteady effect due to motion of the lips channel wall.

Comparison between $P'1$ and $P1$ shows a similar evolution between those pressures, the relative difference is less than 2 %. This result allows us to ignore the viscous losses in the vocal tract and consider further that, in the case of bilabial plosives, supraglottal pressure is equivalent to intra-oral pressure.

Empirical Model of Supraglottal Cavity Volume Expansion

In this experiment, we study the effects of change in supraglottal cavity volume, similar to an expansion of the cheeks during a plosive, on the intra-oral pressure. Measurements on replica allow us to observe this effect, by using a rigid or deformable tube to represent the vocal tract.

Experimental Set-up

The experimental set-up is composed of a high speed video camera (Mikrotron EoSens Cube7), with a sampling rate of 500 Hz and the replica of phonatory system. In this experiment, the two configurations are :

- with a rigid tube made of plexiglass, similar to the one presented in figure 3,
- with a deformable latex tube with a thickness $e = 0.2\ mm$, a length $L_t = 7\ cm$ and a diameter equal to 3 cm (latex).

The air pressure supply and vocal folds replica configuration are kept constant between each configuration. The shape of the latex tube is assimilated to a barrel whose revolution axis is parallel to the x-axis. Pictures from video camera allows to extract, using image processing, the diameter $\Delta y(x)$ all along the tube. Under the assumption of a revolution symmetry, the volume V_t inside the latex tube is given by :

$$V_t = \int_0^{L_t} S_{int}(x)\, \mathrm{d}x = \frac{\pi}{4}\int_0^{L_t} (\Delta y(x) - e)^2\, \mathrm{d}x \qquad (5)$$

where $S_{int}(x)$ is the internal area section of the latex tube, at the abscissa x.

For each picture, the volume V_t is computed by numerical integration. In this measurement, the vocal folds replica was not allowed to oscillate. Others parameters are similar to those presented in the previous section.

Data Analysis

A comparison between the pressure P_1, measured with rigid and with the latex tube, is presented in figure 5.

As expected the deformation of the walls yields a perturbation of the pressure P_1. For a similar mechanical lips closing motion, increase slop in P_1 is higher with a rigid tube than with a latex tube. A similar effect, less pronounced, can be observed for the decrease of P_1. The

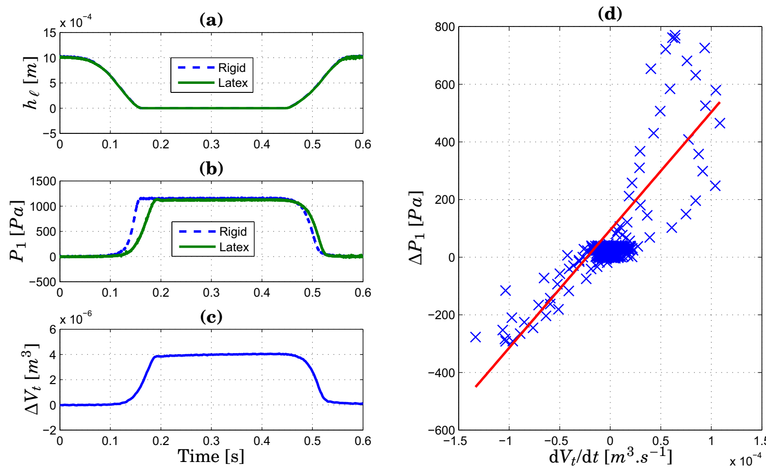

FIGURE 5: Experimental results for the expansion of the latex tube. **(a)** : height h_ℓ of the mechanical lips in rigid and latex tube configurations. **(b)** : experimental pressure P_1 measured in both configurations. **(c)** : volume variation ΔV_t in the latex tube, derived from video recording. **(d)** : experimental relationship between $\Delta P_1 = P_{1\,\mathrm{rigid}} - P_{1\,\mathrm{latex}}$ and time derivative of latex tube volume $\mathrm{d}V_t/\mathrm{d}t$, ΔP_1 is modeled by a linear function of $\mathrm{d}V_t/\mathrm{d}t$, represented by the red straight line.

evolution of the variation in the tube volume seems to be strongly correlated with the pressure P_1 inside the latex tube.

As a first approximation, we consider that the fluctuation ΔP_1, due to the walls deformation, is a linear function of $\mathrm{d}V_t/\mathrm{d}t$. Figure 5.d shows experimental relationship between those parameters and the model resulting from a linear regression of experimental data. For this experiment, the coefficient representing the slope of the straight line is equal to 4.09×10^6 $Pa.s.m^{-3}$.

NUMERICAL SIMULATIONS

In this section, we present numerical simulation of a vowel-bilabial plosive-vowel utterance. We use the physical modeling synthesizer developed by Vilain (2002) and Ruty (2007). This temporal simulation is based on a two-dimensional description of the flow, taking into account for the formation of a jet at the glottis and the dissipation by turbulence. The mechanical behavior of the vocal folds is modeled by a two-mass model, in which geometry of the glottal constriction, represented by a plate model, is driven by the position of the masses. In this study, the effect of acoustic coupling with the vocal tract is neglected.

The flow is modeled by the corrected Bernoulli equation, using Reynolds theory of viscous flows. The input parameters of the simulation are subglottal pressure and labial parameters. Subglottal pressure P_{sub} is chosen as a constant equal to the maximum value of IOP measured in-vivo (810 Pa). Labial parameters A_s and B_s of the simulation are taken from in-vivo measurements using a polynomial interpolation between each data point, since the sampling of the simulation is 30000 Hz.

Simulations of vowel-bilabial plosive-vowel utterance are performed in case of a constant supraglottal cavity volume and in a case where a volume variation is imposed.

Case of a Constant Supraglottal Cavity Volume

Simulation results are presented in figure 6. Although simulation shows that the behavior of the simulated intra-oral pressure P_{io} is qualitatively similar to the measured one, the simulated increase in P_{io} is significantly higher than the experimental one. At lips opening, a similar effect can be observed, decrease in P_{io} is more abrupt for the simulation.

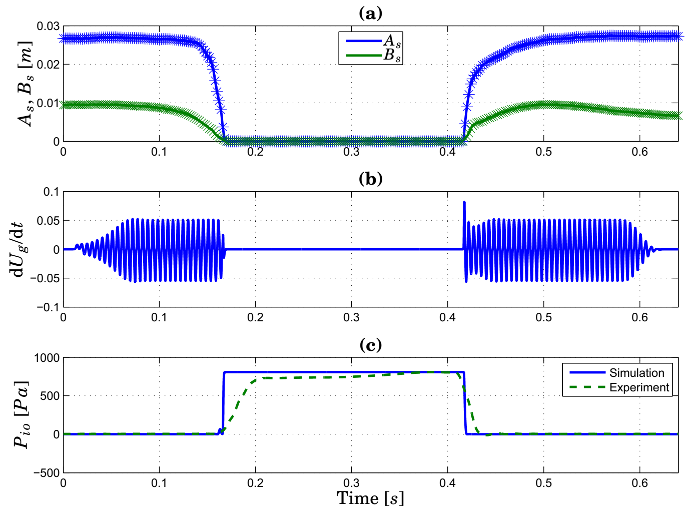

FIGURE 6: Numerical simulations for a vowel-bilabial plosive-vowel utterance in case of a constant supraglottal cavity volume. **(a)** : polynomial interpolation A_s and B_s of lips parameters A and B measured in-vivo, represented by the markers. **(b)** : time derivative of glottis volume velocity simulated, dU_g/dt. **(c)** : comparison between the simulated intra-oral pressure P_{io} and IOP measured in-vivo.

Variations in supraglottal cavity volume appear to be a plausible hypothesis to explain differences in intra-oral pressure evolution.

Case of an Imposed Variation in Supraglottal Cavity Volume

In order to fit the experimental data, McGowan *et al.* (1995) add a volume velocity in the mouth cavity to model the effects of walls motion, for the simulations of a vowel-plosive-vowel sequences. In this paper, we choose to impose a variation ΔV_i in the supraglottal cavity volume. Given a temporal variation in volume, the resulting pressure difference ΔP_i is obtained by using the empirical model presented in figure 5.

Different time domain evolutions for the volume of the mouth cavity were tested. The best fit model is presented in figure 7. The predicted P_{io} fits then the experimental data whereas the imposed volume variation seems consistent with the existing literature (Westbury (1983)).

However, during the opening of the lips, no acceptable fit could be found to explain the measured data. Further research for understanding the aeroacoustic events that could be involved during this opening must be carried out.

CONCLUSION

Lips parameters, acoustical and intra-oral pressure have been measured on a speaker during the production of an /apa/ utterance. Experiments performed on a replica of phonatory system, allow to validate the Bernoulli equation corrected for friction, to predict air flow trough a rounded lips channel. The effects of variation in supraglottal cavity were identified through measurements on the replica using a deformable tube to represent the vocal tract. An empirical model, based on a linear relationship between the temporal variation in volume measured and the resulting pressure difference, is proposed.

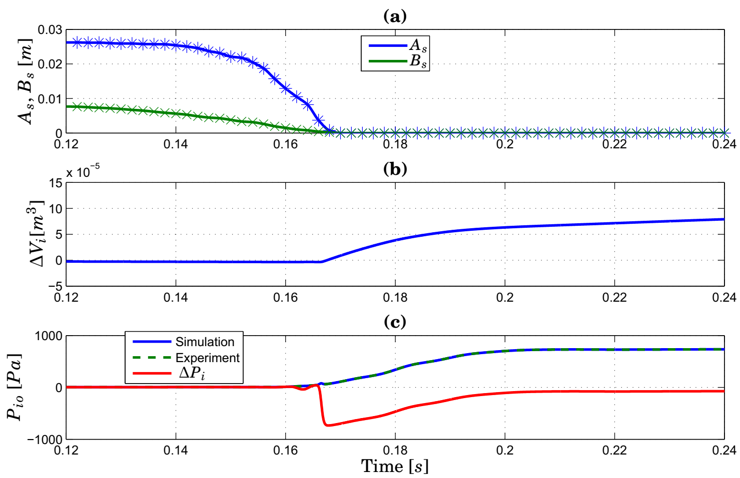

FIGURE 7: Numerical simulations for a vowel-bilabial plosive-vowel utterance in a case of an imposed variation in supraglottal cavity volume. **(a)** : polynomial interpolation A_s and B_s of lips parameters A and B measured in-vivo, represented by the markers. **(b)** : imposed variation in supraglottal cavity volume, ΔV_i. **(c)** : intra-oral pressure simulated P_{io}, measured IOP and pressure difference ΔP_i due to the temporal variation in ΔV_i.

Numerical simulations for a vowel-bilabial plosive-vowel utterance are achieved by using the corrected Bernoulli equation. For an imposed supraglottal cavity expansion during the lips closure, the empirical model proposed explains the difference between the simulated intra-oral pressure and the measured one. However this model does not explain the evolution of the measured intra-oral pressure during the opening of the lips.

Further work will focus on the modeling of supraglottal cavity expansion by using a mass-spring system to represent the elasticity of the walls.

ACKNOWLEDGMENTS

This work is supported by the Rhône-Alpes region by means of CIBLE PhD grant and by ANR Plasmody project.

REFERENCES

Deverge, M., Pelorson, X., Vilain, C., Lagrée, P.-Y., Chentouf, F., Willems, J., and Hirschberg, A. (**2003**). “Influence of collision on the flow through in-vitro rigid models of the vocal folds”, Journal of the Acoustical Society of America **114**, 3354–3362.

McGowan, R., Koenig, L., and Löfqvist, A. (**1995**). “Vocal tract aerodynamics in /aCa/ utterances: Simulations”, Speech Communication **16**, 67–68.

Ruty, N. (**2007**). “Modèles d’interactions fluide/parois dans le conduit vocal. Applications aux voix et aux pathologies”, Ph.D. thesis, Institut National Polytechnique de Grenoble.

Schlichting, H. and Gersten, K. (**1999**). *Boundary Layer theory*, 6 edition (Springer Verlag).

Vilain, C. (**2002**). “Contribution à la synthèse de parole par modèle physique. Application à l’étude des voix pathologiques”, Ph.D. thesis, Institut National Polytechnique de Grenoble.

Westbury, J. R. (**1983**). “Enlargement of the supraglottal cavity and its relation to stop consonant voicing”, Journal of the Acoustical Society of America **73**, 1322–1336.