Physical modeling of vowel-bilabial plosive sound production

L. Delebecque, X. Pelorson, D. Beautemps, C. Savariaux and X. Laval

Grenoble Images Parole Signal Automatique, Gipsa-lab 961 rue de la Houille Blanche BP 46
F - 38402 Grenoble Cedex
louis.delebecque@gipsa-lab.grenoble-inp.fr

Downstream of the vocal folds, it is often assumed that all the kinetic energy of the glottal jet is dissipated by turbulence. Although this assumption is reasonable for normal voicing, there are many configurations for which a constriction or even an occlusion of the vocal tract occurs and leads to a significant pressure recovery. Such a phenomenon occurs during the production of vowel-bilabial plosive sequences. In this paper, we will first present results from measurements on a human speaker during production of vowel-bilabial plosive-vowel sequences. These measurements concern intra-oral pressure, acoustical pressure, EGG and lips video recording of a female subject. The relative timings between the motion of the lips, the glottal signals and the intra-oral pressure are analyzed and related with Voice Onset/Offset Times. A theoretical model for air flow in the lips is validated by experiments on a mechanical replica of the vocal tract including a self-oscillating glottis and a controlled moving constriction to mimic the closure of the lips. This theoretical model is used to achieve numerical simulations of an /apa/ utterance. The differences between the intra-oral pressure observed in vivo and the simulated one lead to an investigation on the influence of the cheeks expansion during bilabial plosive production.

1 Introduction

The context of this work is physical modeling of speech production. This approach allows to study the mechanical and aeroacoustic mechanisms governing speech production. The goal is to understand the physical phenomena but also to achieve modeling in order to predict the consequences. For speech, numerical simulations based on these theoretical models are used to synthesize sound signals. The development of such a tool has applications in many fields (medical, communication, teaching ...).
Speech sounds are produced by various disturbances of the flow of air through the vocal tract. The present study focuses on the bilabial plosives, especially on the air flow in the upper part of vocal tract, downstream to the vocal folds. Many studies on the fluid mechanical interactions in the vocals folds have been published. They propose theoretical air flow models which are useful to model voiced sound. In this paper, we investigate on the air flow in the constriction shaped by lips, during the production of a bilabial plosive. The physical modeling of the flow in the whole vocal tract provides prediction of the evolution of the intra-oral pressure, which is, in bilabial plosive utterance, responsible for the creation of an audible acoustic disturbance.
Otherwise, we study the effect of the expansion of oral cavity volume under the intra-oral pressure effect, during bilabial plosive production. The physical origin of this phenomenon is related to the elasticity of the cheeks. This deformation of the cheeks causes a variation in the volume of the oral cavity and therefore in the intra-oral pressure and in oral flow.
In the first part, we present in vivo measurements during /apa/ utterances. Then three different theoretical air flow models are exposed and compared to results from in vitro measurements. Experimental data allows to evaluate the relevance of this models for the air flow in the lips during the production of vowel-plosive-vowel. Numerical simulations of an /apa/ sequence using the air flow model validated before, are achieved and compared to in vivo measurements. Finaly, the influence of the cheeks expansion during bilabial plosive production is identified and quantified through in vivo and in vitro measurements.

2 In vivo measurements

2.1 Experimental setup

In vivo measurements for /apa/ sequences presented in this paper are part of a data base initially achieved by [3]. These measurements were performed using EVA™system and a video camera. The EVA device [2] allows to perform aeroacoustic measurements during speech production. These measurements concern 4 synchronized signals :

- acoustic pressure, recorded by a microphone placed in front of the mouth of the subject,
- electroglottographic signals, measured by two electrodes placed on the larynx,
- Intra-Oral Pressure (IOP) measured by means of a tube of 5 *mm* diameter, inserted through the corner of the lips,
- lip parameters derived from the video recording of the subject face.

The pressure sensor of EVA have been calibrated in static and dynamic configurations.

2.2 Data analysis

The figure 1 shows a typical measurement observed during vowel-bilabial plosive-vowel production. IOP rises significantly, when the lips close, until a maximum value of 1000 *Pa*. At the lips opening, IOP drops quickly to atmospheric pressure. At this moment, a burst, which is audible as a bilabial plosive, appears in the acoustic pressure signal. As IOP falls, electroglottographic and acoustic pressure signals shows that the vocal folds start to oscillate again to produce the second vowel.

3 Theoretical models and experimental validation

3.1 Theoretical models

The objective of this section is to validate a theoretical model for the air flow in the upper part of the vocal tract, during a plosive production. The constriction shaped by the lips is modeled a rectangular duct of length L (in the flow direction), of aperture h_l and of width W.
We make the assumption of a one-dimensional flow, $\vec{u} = u(x,y,t)\vec{e}_x$, where $\vec{u}$ is the air velocity field and $\vec{e}_x$ is the unit vector parallel to the vocal tract axis. Thus, pressure are supposed to be uniform in a cross section normal to the flow direction. $p(x)$ represents pressure relative to the atmospheric pressure. According to the Reynolds number, based on the aperture h_l ,the air flow is expected to be laminar. Moreover, we make the assumption of a quasi-steady flow (low Strouhal

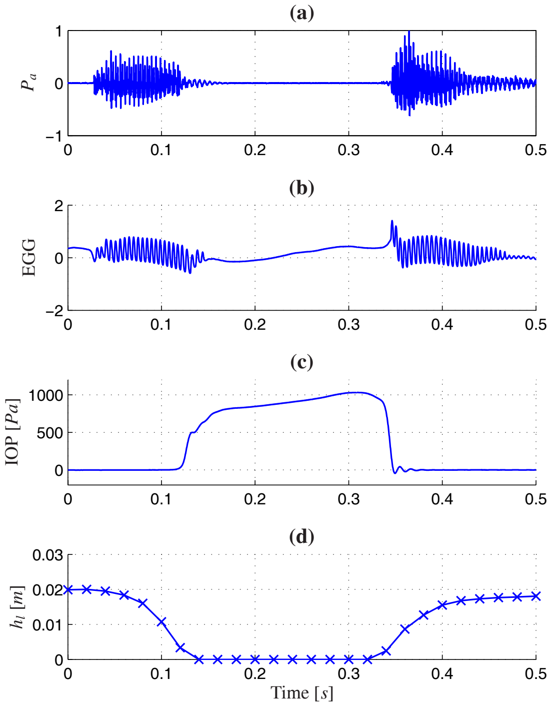

Figure 1: In vivo measurements for /apa/ sequence, acoustic pressure P_a, electroglottographic signal (EGG), intra-oral pressure (IOP) and labial aperture h_l

number approximation).
The pressure $p(x > 0)$ inside the lip constriction is predicted from the pressure $P_1 = p(x = 0)$ in the vocal tract, upstream of the lips.

Three different theoretical models are under investigation for the flow in the lips :

- the steady Bernoulli equation, (Bernoulli),
- boundary layer model for steady flow, using linear velocity profile (Van Zon),
- the steady Bernoulli equation corrected for viscous pressure losses, given by the lubrification theory of Reynolds (Reynolds).

Details of these models, can be founded in [1].

Inviscid Bernoulli equation

The steady and inviscid Bernoulli equation links the pressure difference P_1 across the lips to the volume velocity Φ_l or velocity $u = \Phi_l/(LW)$, inside the lips.

$$p(x) = P_1 - \frac{\rho}{2}\left(\frac{\Phi_l}{Wh_l}\right)^2 \qquad (1)$$

The vocal tract area is considered to be large compared with the labial area, thus, we assume that the volume velocity in the vocal tract (for $x = 0$) is negligible. Since the lips constriction is represented by a uniform duct, the volume velocity conservation implies that $p(x = L) = p(x) = 0$. In this case, the steady inviscid Bernoulli equation will always predict a intra-labial pressure equal to zero.

Boundary layer equations

The boundary layer theory assumes the existence of a frictionless core , far enough from the wall, with a uniform velocity $u_e(x)$. Van Zon proposes to uses a linear velocity profile $u(x, y) = u_e(x)y/\delta(x)$, where $\delta(x)$ is the thickness of the boundary layers. The corresponding volume velocity becomes :

$$\Phi_l = Wu_e(x)(h_l - \delta(x)), \qquad (2)$$

$\delta(x)$ satisfies the non-linear equation :

$$\frac{4\delta(x)}{h_l} + 9\ln\left(1 - \frac{\delta(x)}{h_l}\right) + \frac{5\delta(x)}{h_l - \delta(x)} = \frac{6\nu xW}{h_l\Phi_l}, \qquad (3)$$

where ν denotes the kinematic viscosity. Equations (2) and (3) are applied at the exit of the lip constriction ($x = L$). Then the frictionless core velocity is determined by applying Bernoulli equation (1) across the lips, $u_e(x = L) = \sqrt{2P_1/\rho}$. The set of two equations (2, 3) can be solve to determinate $\delta(x = L)$ and Φ_l. The value of volume velocity Φ_l allows to calculated the viscous boundary layer $\delta(x)$ at any position x using equation 3. Finally $u_e(x)$ is obtained by using (2) and the pressure is calculated by using Bernoulli equation.

$$p(x) = P_1 - \frac{\rho}{2}u_e^2(x) \qquad (4)$$

Steady Bernoulli equation corrected for friction

The principle of this model is to add to inviscid Bernoulli equation, a term $\Delta P_\nu(x)$ accounting off viscous losses. The Bernoulli equation becomes :

$$p(x) = \frac{1}{2}\rho\left(\frac{\Phi_l}{Wh}\right)^2 - P_1 + \Delta P_\nu(x) \qquad (5)$$

For a uniform duct the lubrification theory of Reynolds yields :

$$\Delta P_\nu(x) = \frac{12\rho\nu\Phi_l}{Wh_l^3}x \qquad (6)$$

Combination of equations (5) and (6) gives a quadratic equation for Φ_l which is solved for $x = L$. The pressure profile $p(x)$ inside the lips is obtained by using equation (5).

3.2 Experimental Setup

Validation measurements were performed on a 1:3 scale replica of the human phonatory system. Figure 2 schematises the experimental setup.

The replica is composed of various elements:

- a pressure tank, cuboid about 0.6 m^3 volume, supplied by a compressor,
- a replica of the vocal folds made of latex tube filled with water, able to self-oscillate,
- a replica of lips in metal, whose motion of the upper part is controlled either by a motor or manually,
- a metal uniform tube of length 12 cm with an internal diameter of 2.5 cm for modeling the trachea,

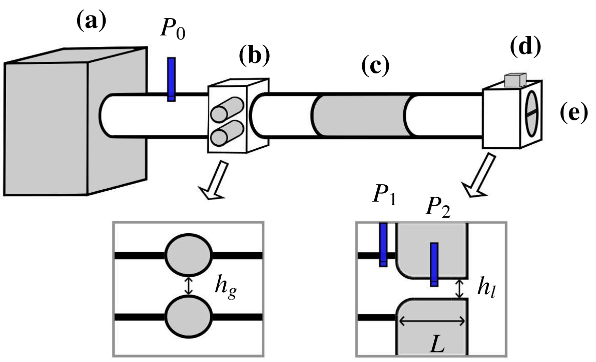

Figure 2: Replica of human phonatory system. **(a)** : pressure tank. **(b)** : replica of vocal folds. **(c)**: plexiglas tube. **(d)**: motor and mechanical lips displacement sensor. **(e)** : metal replica of lips.

- a uniform tube of length 24.5 *cm* in metal, including a plexiglas tube of 17 *cm* length, with the same diameter as the trachea, to represent the vocal tract.

The dimension of the straight channel is $L = 2$ *cm* and $W = 3$ *cm*. Three differential pressure sensors are used to measure P_0, $P_1 = p(x = 0)$ and $P_2 = p(x = L/2)$ which correspond respectively, to subglottal, intra-oral and intra-labial pressure. Without oscillations, the initial aperture of the glottis h_g is 0.7 *mm*. The aperture h_l varies from 0 to 1.2 *mm*. The motion imposed by a motor is sinusoidal for a frequency of 0.5 *Hz*.

Theoretical models are evaluated in two different experimental conditions : in case of steady flow and in case of an unsteady flow (with vocal folds oscillations).

3.3 Results

Steady flow experiments

In figure 3 is presented an example of the comparison between experimental data and the prediction obtained using the theoretical models described in section 3.1.

Pressure P_2 measured inside the replica of lips (for $x = L/2$) stay finite even when the channel is closed. This result shows that a complete closure can not be achieved because of the roughness of the mechanical lips surface.

Although boundary layer model seems to be more accurate, this theory fails to explain the experimental pressure P_2 for $h_l < 0.3$ *mm*. This result is due to the fact that, below this value of aperture, the predicted boundary layer becomes larger than the aperture.

Corrected Bernoulli equation using Reynolds theory reproduces an acceptable behavior of pressure in mechanical lips. However this theory overestimates the value of P_2. The relative difference with experimental data, estimated for $h_l = 0$ is of order of 20 %. When $h_l(t) = 0$, the corrected Bernoulli equation predicts $P_1/P_2 = 1/2$.

The inviscid steady Bernoulli equation is not able to reproduce the behavior since this prediction is always $P_2 = 0$.

Unsteady flow experiments

In this part, the pressure of water within the latex tubes which represent the vocal folds, is increased to allow self-

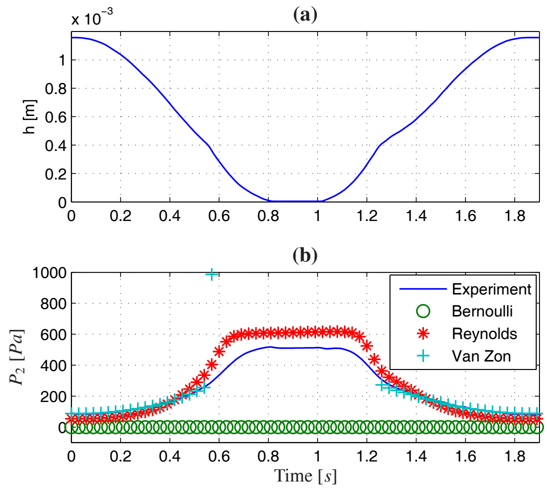

Figure 3: Experimental and theoretical results for steady flow. **(a)** : aperture h_l of the mechanical lips. **(b)** : Experimental pressure P_2 and theoretical pressure P_2 computed from experimental pressure P_1 using three theoretical models : Bernoulli inviscid equation (Bernoulli), Bernoulli equation corrected for friction (Reynold), boundary layer equations (Van Zon)

sustained oscillations of vocal folds replica. An example of results obtained for unsteady flow is shown in figure 4.

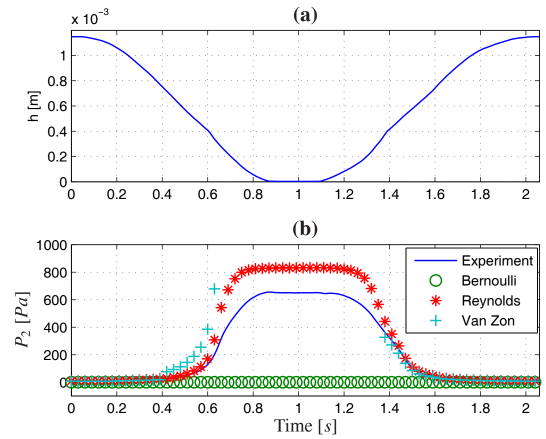

Figure 4: Experimental and theoretical results for unsteady flow, during vocal folds replica oscillation. **(a)** : aperture h_l of the mechanical lips. **(b)** : Experimental pressure P_2 and theoretical pressure P_2 computed from experimental pressure P_1 using three theoretical models : Bernoulli inviscid equation (Bernoulli), Bernoulli equation corrected for friction (Reynold), boundary layer equations (Van Zon)

For small values of apertures ($h_l < 0.3$ *mm*), predictions by Van Zon theory is less accurate than for steady flow experiments, especially during the closure of mechanical lips ($0.4 < t < 0.7$). The prediction from corrected Bernoulli equation, as in the steady flow case, shows good agreement with measured pressure P_2 but overestimates the value. The relative difference with experimental data is of order of 20 % again.

We conclude from this result that the corrected Bernoulli equation for friction, using Reynolds theory seems to be a reasonable approximation for predict air flow in the lips during bilabial plosive production. This theory will be used in numerical simulations for vowel-plosive-vowel utterance.

4 Numerical simulations

In this part, we achieve numerical simulation of an /apa/ utterance. We use the physical modeling synthesizer developed by [4]. Mechanical behavior the vocal folds is modeled by a two-mass model. The aerodynamic part of the 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 flow at the lips 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 (aperture and width). Subglottic pressure is chosen as a constant equal to the maximum value of IOP measured in vivo (1000 Pa).

Labial parameters, h_l and W are directly taken from in vivo measurements. Since the video rate was $50Hz$ only, a polynomial interpolation was performed between each data point. Simulation results are presented in figure 5.

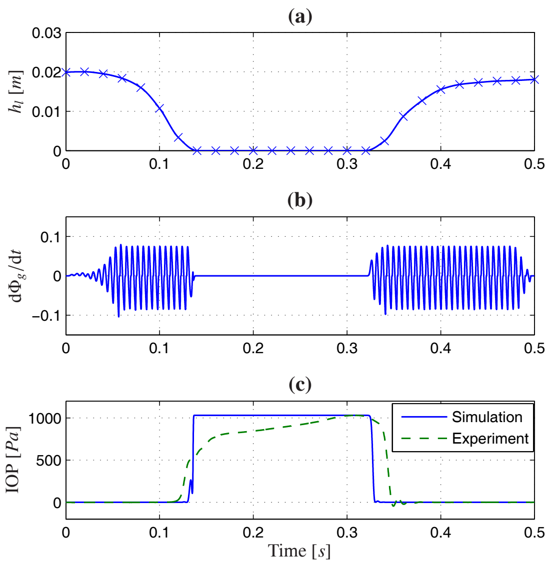

Figure 5: Numerical simulation of an /apa/ utterance. **(a)**: h_l polynomial interpolation of lips aperture measured in vivo, represented by the crosses. **(b)**: $d\Phi_g/dt$ time derivative of glottis volume velocity Φ_g simulated. **(c)**: Comparison between simulated IOP and IOP measured in vivo.

The behavior of simulated IOP is qualitatively similar to the experimental one. However, differences can be observed.

- Simulated onset and offset times of IOP are different for these measured on the subject but these differences might not be relevant since they are within the accuracy given by the video acquisition system (20 ms).
- During the closure of the lips, the simulated increase in IOP is significantly higher than the measured one. A possible explanation for this departure could lie in the inflation and deflation of the subject's cheeks during the vowel-plosive-vowel production. This hypothesis will be test in the next section.

5 Influence of the cheeks expansion

The objective of this section is to highlight the influence of the cheeks expansion on the intra-oral pressure during the production of bilabial plosives. The assumption made here is that the increase of oral cavity volume causes an additional depression which concurrence the increase of intra-oral pressure. Firstly this phenomenon will be illustrated by in vivo measurements and reproduced in the laboratory.

5.1 In vivo measurements

These measurements consist in recording simultaneously acoustic pressure and intra-oral pressure during /apa/ utterances using EVA station. A subject is asked to repeat 10 times the /apa/ sequences in two different ways :

- in a natural way: **N**,
- by placing hands on the face to prevent the cheeks from inflating : **H**.

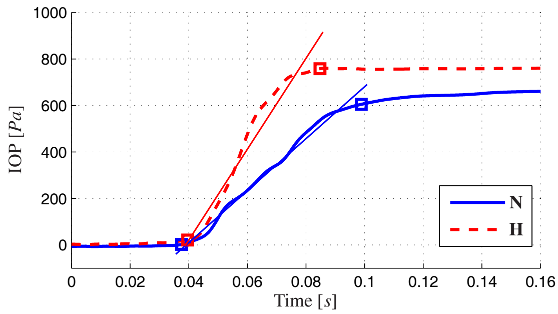

Figure 6: Influence of cheeks expansion for in vivo measurements during /apa/ utterances : evolution of IOP evolution at the closure of the lips for a natural condition (**N**) and with hands on the face to prevent the cheeks from inflating (**H**). The increase in IOP is modeled, between the square markers, as a straight line.

We want to quantify the differences in IOP time evolution under both conditions. The increase in IOP that appears at the lips closure, is measured using a straight line whose slope is determined by a linear regression. Figure 6 illustrates the determination of the increase rate coefficient a_{iop}, presented in table 1. A student test made on the 10 repetitions allows to discriminate, at 1% level, the IOP increase between the two conditions.

Conditions	ΔIOP [Pa]	a_{iop} [$Pa.s^{-1}$]
N	520	11100
H	740	19600

Table 1: IOP differences ΔIOP and increase rate coefficients a_{iop} of IOP for /apa/ utterance during in vivo measurement for both conditions of production.

According to table 1, the increase rate in IOP is stronger for the condition **H**. The expansion of the cheeks has thus a measurable influence in the case of bilabial plosives. In the next part, this experience is reproduced in the laboratory.

5.2 In vitro measurements

These measurements are made with the replica of human phonatory system presented in section 3.2. We focus on the evolution of P_1 at the lips closure. The vocal tract is represented either by a rigid plexiglas tube of 16 cm of length or deformable tube made of latex of 10 cm of length and of thickness 0.2 mm. We compare the increase of P_1 with the rigid or the deformable tubes, for a similar lips movement, $h_l(t)$

The increase rate coefficient a_{P1} is estimated from P_1, by a linear regression. An example of results is shown in figure 7 and the corresponding increase rate coefficients are presented in table 2.

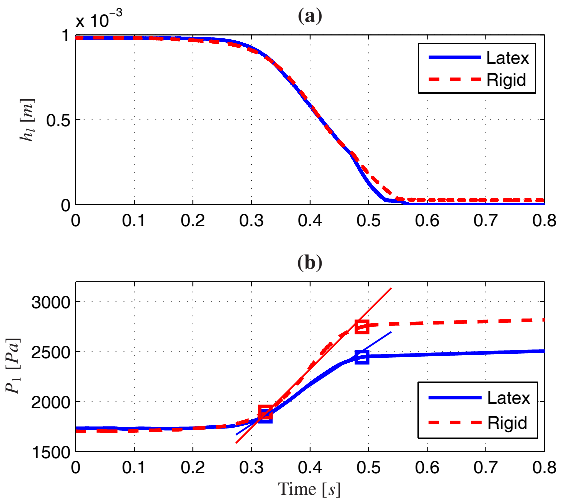

Figure 7: Influence of the expansion of latex tube during in vitro vowel-plosive-vowel utterances for the rigid tube and the deformable tube made of latex, **(a)**: lips aperture h_l, **(b)**: evolution of pressure P_1

Tubes	ΔP_1 $[Pa]$	a_{P1} $[Pa.s^{-1}]$
Latex	590	3900
Rigid	850	5700

Table 2: Pressure differences ΔP_1 and increase rate coefficients a_{P1} of P_1, for the rigid tube and the deformable tube made of latex

In vivo and in vitro increase rate coefficients are not directly comparable because of the difference in duration of the lips closure. This duration is 200 ms for in-vitro measurement (figure 7) and 50 ms in human (figure 6). Variation of IOP slope due to the stiffness of the oral cavity is 51% for in vitro measurements and 77% for in vivo measurements. This difference can be explained by the fact that latex is stiffer than human cheeks.

6 Conclusion

Acoustical pressure, electroglottographic signals and intra-oral pressure have been measured on a female speaker during /apa/ utterances. Experiments on a replica of phonatory system allow to validate the Bernoulli equation corrected for friction, using the Reynolds lubrification theory, for the air flow in the lips. Numerical simulations of aerodynamical phenomenons governing speech production are achieved for a vowel-bilabial plosive-vowel utterance, by using this theoretical model. The comparison between simulated IOP and IOP measured on human highlights the importance of the expansion of supra-glottal cavity volume during the production of a bilabial plosive. The role of the cheeks expansion has been identified and quantified through in vivo and in vitro measurements. This expansion of the walls of oral cavity causes a reduction of increase of the intra-oral pressure at lips closure.

The obtained results are stimulating and lead to the conclusion for the need of the physical modeling of the cheeks expansion. For this purpose, the estimation of elasticity parameters of the latex tube which represents the cheeks is a necessary step to validate a elasticity model of oral cavity walls. In addition, the use of a high sampling rate camera for lip parameters extraction could allow to determinate if the timing difference between the simulated IOP and the measured one is due to a limited time precision of the measure or a lack in physical model used for simulation.

Acknowledgments

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

References

[1] M Deverge, Xavier Pelorson, Coriandre Vilain, P.Y. Lagrée, F. Chentouf, J. Willems, and A. Hirschberg. Influence of collision on the flow througgh in-vitro rigid models of the vocal folds. *J. Acoustic. Soc. Am.*, 114(6):3354–3362, December 2003.

[2] A. Giovanni, P. Yu, J. Révis, M.D. Guarella, B. Teston, and M. Ouaknine. Analyse objective des dysphonies avec l'appareillage eva, état des lieux. *Revue Oto-Rhino-Laryngologie Franćaise*, 90:183–192, 2006.

[3] Balbine Maillou. Approche physique de la voix d'une personne sourde. Master's thesis, Université du Maine, July 2011.

[4] Nicolas Ruty. *Modèles d interactions œuide/parois dans le conduit vocal. Applications aux voix et aux pathologies.* PhD thesis, Institut National Polytechnique de Grenoble, Novembre 2007.