

LARYNGEAL ADJUSTMENT IN THE PRODUCTION OF VOICELESS CONSONANTS: II. PHYSICAL MODELLING.

Annemie Van Hirtum^{1,2}, Nicolas Ruty², Xavier Pelorson², Susanne Fuchs³, Pascal Perrier²

¹TUE - Technical University Eindhoven, Eindhoven, The Netherlands

²ICP - Institut de la Communication Parlée, UMR CNRS 5009 - INPG, Grenoble, France

³ZAS - Centre for General Linguistics, Berlin, Germany

ABSTRACT

In this paper, we present an attempt to describe flow past the glottis and to predict the pressure inside the vocal tract during phonation. Different theoretical models to describe the pressure distribution inside the whole vocal tract will be presented and compared. Theoretical models will be validated on ‘in-vitro’ measurements performed on a mechanical replica of the human phonatory system. Next, an application of this theoretical approach to the simulation of vowel-plosive-vowel (VCV) sequences is presented. It is shown that even using a very crude mechanical model for the vocal folds (such as a 1 or 2-mass model) one can already replicate some important features with a surprising accuracy. This will be illustrated by examples of predictions for the onset-offset glottal pressure and of the fundamental frequency of oscillation. Lastly, the simulations are compared to ‘in-vivo’ observations as e.g. presented in a companion paper.

1. INTRODUCTION

In this paper we discuss the effect of the vocal tract on the vocal folds oscillations. While such a study has already been carried out extensively in the case of vocal coupling [1, 2, 3, 4], very little is known about the behaviour of flow past the glottis. The possibility of having a pressure recovery downstream of the glottis is particularly crucial during the production of the voiceless consonant such as a plosive. Due to the presence of a closure of the vocal tract the supraglottal pressure increases which can explain, in a part, the offset of the vocal folds oscillations. Previous attempts to simulate such voiceless consonants tend indeed to show what a precise coordination between the glottal source and the constriction is crucial [5, 6]. As the matter of fact, without such a description, acceptable acoustic results can only be simulated using unrealistic glottal gestures.

In this paper, we present an attempt to predict the pressure inside the vocal tract during phonation. Different theoretical models to describe the pressure distribution inside the whole vocal tract will be presented and compared. A particular attention will be devoted to the relative balance between inertia and viscosity inside the vocal tract.

These theoretical models will be compared to ‘in-vitro’ measurements performed on a mechanical replica of the human phonatory system. This set-up consists of a pressure reservoir (‘the lungs’), a self-oscillating mechanical ‘glottis’ [7] and a ‘vocal tract’ modeled by pipes with varying sections and length. Using this set-up, the generation of a voiced-plosive sequence can thus be simulated by closing one part of the ‘vocal tract’. Using pressure sensors

placed along the replica, the pressure distribution can be measured (subglottal, supraglottal and inside the vocal tract). A systematic study of the acoustical coupling and of the flow recovery will be presented and compared to the theoretical predictions. Next, we present an application of this theoretical approach to the simulation of vowel-plosive-vowel (VCV) sequences. It is shown that even using a very crude mechanical model for the vocal folds (such as a 1 or 2-mass model) one can already replicate some important features with a surprising accuracy. This will be illustrated by examples of predictions for the onset-offset glottal pressure and of the fundamental frequency of oscillation.

Lastly, the simulations are compared to ‘in-vivo’ observations as e.g. presented in a companion paper [8].

2. MODELING VOCAL FOLDS DYNAMICS

The interaction of expiratory airflow with the vocal folds tissues is known to be the primary source of human voiced sound production. The airflow through the larynx induces instability of the vocal folds. The resulting vocal fold vibrations modulate the airflow giving rise to a periodic sequence of pressure pulses which propagates through the vocal tract and is radiated as voiced sound. Consequently physical modeling of the 3D fluid-structure interaction between the living vocal folds tissues and the expiratory airflow is essential in the study of phonation. Simplifications of the physical reality are favoured due to a historical interest for speech control and synthesis applications requiring a limited number of physiological meaningful and measurable parameters. Therefore physical models like vocal fold two-mass models strive to represent the main features of phonation while assuming severe simplifications in the biomechanical structure and fluid mechanical flow modeling. The description of the aerodynamics in the glottis assumes a simplified one-dimensional quasi-stationary incompressible flow as described by the stationary Bernoulli’s equation. Usually several corrections are applied accounting for 1) flow separation using Liljencrants ‘ad-hoc’ criterium, 2) viscosity in the glottis (Poiseuille flow), 3) inertance of air [9, 10] and 4) downstream pressure recovery.

The possibility of having a pressure recovery downstream of the glottis is particularly crucial during the production of voiceless consonants concerning the onset and offset of vocal fold oscillation. In [9] the pressure recovery is estimated by evaluating the quasi-steady momentum equation depending on the area ratio at the position of flow separation in the glottis and the vocal tract area past the glottis. The same area ratio is presented in [11, 12] as a geometrical basis for quantifying the pressure recovery in a diffuser.

In the following the pressure distribution, and so the pressure recovery in the downstream resonator, representing the vocal tract geometry or the experimentally assessed pipe, is modelled applying the same simplified description of the aerodynamics as used for the flow through the glottis. This way the flow through the downstream pipe geometry is described by the stationary Bernoulli's equation and correcting for viscosity of inertance. Hereby it is assumed that the downstream pipe geometry is described by uniform sections with changing cross-section.

This approach, requiring the upstream pressure P_{sub} as an input parameter, is applied to the simulation of vowel-plosive-vowel (VCV) sequences for which the temporal variation of the different vocal tract uniform cross-sections is obtained as described in [13].

3. EXPERIMENTAL PROCEDURE

The experimental set-up consists of a pressure reservoir ('lungs'), a self-oscillating mechanical replica ('glottis') and a downstream pipe ('vocal tract') with varying sections and length. The glottal replica and experimental set-up are detailed in [4, 7]. Briefly the mechanical replica (width 0.024m) consists of two connected latex tubes filled with water representing the vocal folds. An internal pressure $P_{in} = 5200Pa$ is imposed and controlled before and during each experiment. The replica is connected to a pressure reservoir ($V_i = 0.68m^3$) representing the lungs and supplying a static pressure P_{sub} in the range of 0 to 700Pa. The vocal tract is represented by a downstream cylindrical pipe of length 50cm and diameter 0.025m. In order to attempt the study of production of a 'vowel-plosive-vowel' sequence, a rectangular constriction of length 2cm with variable aperture height is added at the downstream pipe end. Constriction aperture heights h_c of 0.82mm, 1.2mm, 2.2mm, 2.7mm, 3.5mm and 4.05mm are assessed corresponding to respectively 4%, 6%, 11%, 14%, 18% and 21% of the uniform downstream pipe area. The pressures just upstream P_{sub} and downstream P_d of the replica as well as the pressure at the constriction level P_c are measured. The imposed internal pressure of $P_{in} = 5200Pa$ corresponds to an initial closure of the vocal fold replica for $P_{sub} = 0$, which is the optimal condition to obtain sustained oscillation of the glottal replica, i.e. a minimum effort or threshold P_{sub} [3, 4]. A photograph and schematic overview of the experimental set-up is given in Figure 1.

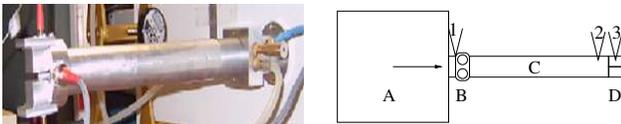


Fig. 1. Picture and schematic overview of experimental setup: A 'lungs', B 'glottal' replica, C 'vocal tract' and D constriction. The position of 3 distinct pressure taps are numbered in the denoted airflow direction: (1) P_{sub} , (2) P_d and (3) P_c .

4. RESULTS

Figure 2 illustrates the influence of the constriction aperture height on the oscillatory behaviour of the mechanical vocal folds replica. The measured P_d just upstream the constriction for an aperture of 1.2mm and 3.5mm are shown as function of P_{sub} in respectively

part (a) and part (c). The absence and presence of sustained oscillations of the replica is clearly illustrated. The oscillation frequency as well as threshold P_{sub} required to get onset or offset of sustained oscillation derived from the corresponding spectra given in part (b) and (c) of figure 2. The P_{sub} -threshold for on- and off-set of sustained oscillation for all assessed apertures h_c is plotted in figure 3 illustrating the hysteresis behaviour. It is easily seen that the minimum pressure required to get sustained oscillations decrease with increasing minimum aperture. Therefore the applied aerodynamic description might describe the off- and onset of vocal folds oscillation by consecutive decreasing and increasing vocal tract cross-section as is e.g. the case in vowel-plosive-vowel sequences. Part (a) of figure 4 illustrates a first attempt to reproduce experimentally a vowel-plosive-vowel sequence to perform controlled validation. This qualitatively corresponds to the 'aka' simulation shown in part (b). To study the influence of P_{sub} on glottal abduction ($h[m]$) during the plosive the simulation is performed for P_{sub} yielding 500Pa and 1000Pa. The retrieved duration of glottal abduction is not changed which corresponds to 'in-vivo' findings as presented in [8].

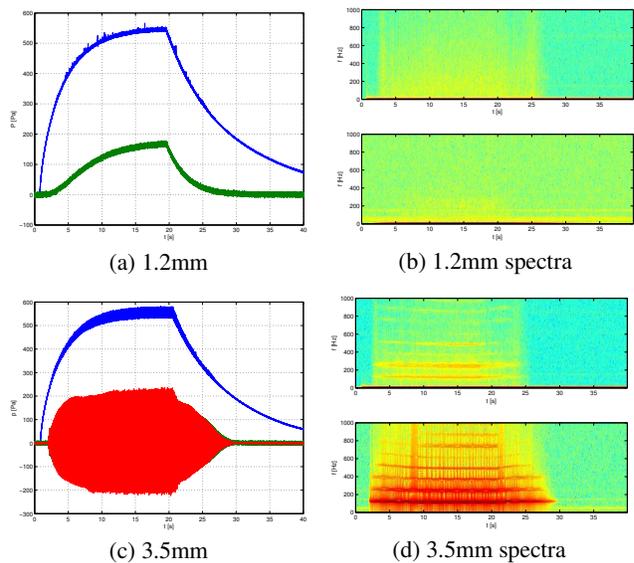


Fig. 2. Exemplar data of P_{sub} [Pa] (top) and P_d [Pa] (bottom) and corresponding spectrograms for an obstruction height of $h_c=1.2mm$ (a,b) and $h_c=3.5mm$ (c,d) illustrating absence ($h_c=1.2mm$) and presence ($h_c=3.5mm$) of sustained oscillations. In the last case the P_{sub} -threshold for on- and offset of sustained oscillation (frequency 130Hz) is illustrated.

5. CONCLUSION

We presented a first attempt to describe flow past the glottis and in particular the pressure recovery based on simple aerodynamic principles. We performed a quantitative validation on 'in-vitro' measurements performed on a mechanical replica of the human phonatory system. Next, this theoretical approach is applied to the simulation of vowel-plosive-vowel (VCV) sequences. It is shown that even using a very crude mechanical model for the vocal folds (such as a 1 or 2-mass model) one can already replicate some important features with a surprising accuracy. This is illustrated by

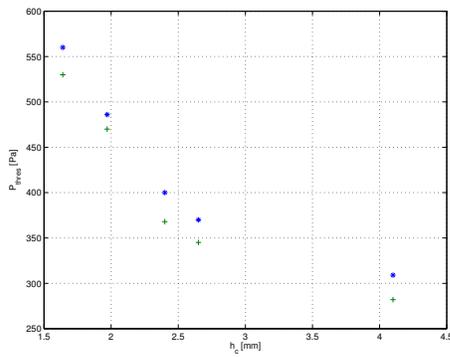


Fig. 3. P_{susb} -threshold for onset (*) and offset (+) of sustained oscillation as a function of h_c .

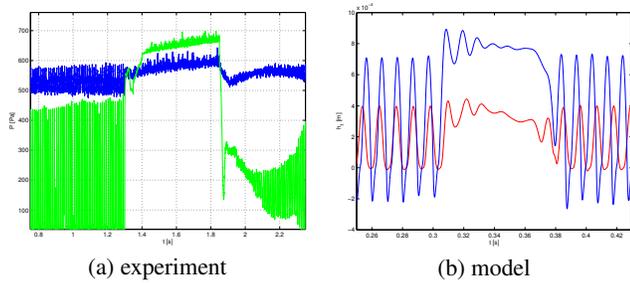


Fig. 4. Illustrative example of (a) preliminary attempt of an ‘in-vitro’ experimentally ‘vowel-plosive-vowel’ sequence (P [Pa]) and (b) modelled ‘vowel-plosive-vowel’ (‘aka’) sequence showing the influence of P_{susb} , i.e. 500Pa(bottom) and 1000Pa(top), on glottal abduction ($h[m]$) during the plosive.

examples of predictions for the onset-offset glottal pressure and of the fundamental frequency of oscillation. Lastly, the simulations are compared to ‘in-vivo’ observations as e.g. presented in a companion paper.

6. REFERENCES

- [1] R. Laje, T. Gardner, and G.B. Mindlin, “Continuous model for vocal fold oscillations to study the effect of feedback,” *Physical Review E*, vol. 64, pp. 1–7, 2001.
- [2] P. Mergell and H. Herzel, “Modelling biphonation - the role of the vocal tract,” *Speech Comm*, vol. 22, pp. 141–154, 1997.
- [3] A. Van Hirtum, I. Lopez, M.H. Schellekens, X. Pelorson, N. Driessen, and A. Hirschberg, “The effect of acoustical feedback on buzzing. from lips to vocal folds ?,” in *Proc. Cfa/daga*, Strasbourg, France, 2004, pp. 1–2.
- [4] I. Lopez, M.H. Schellekens, N.M. Driessen, A. Hirschberg, A. Van Hirtum, and X. Pelorson, “Buzzing lips and vocal folds: the effect of acoustical feedback,” in *Proc. Flow induced vibrations*, Paris, France, 2004, pp. 1–6.
- [5] A. Lofqvist, L. Koenig, and R. McGowan, “Vocal tract aerodynamics in /aca/ utterances: Measurements,” *Speech Comm*, vol. 16, pp. 49–66, 1995.
- [6] R. McGowan, L. Koenig, and A. Lofqvist, “Vocal tract aerodynamics in /aca/ utterances: Simulations,” *Speech Comm*, vol. 16, pp. 67–88, 1995.
- [7] C.E. Vilain, X. Pelorson, A. Hirschberg, L. Le Marrec, W. Op’t Root, and J. Willems, “Contribution to the physical modeling of the lips. influence of the mechanical boundary conditions,” *Acta Acoust.*, vol. 89, pp. 882–887, 2003.
- [8] S. Fuchs, P. Hoole, X. Pelorson, A. Van Hirtum, P. Perrier, K. Dahlmeier, and J. Creutzburg, “Laryngeal adjustment in voiceless consonant production: I. an experimental study of glottal abduction in loud versus normal speech,” in *Proc. ICVPB*, Marseille, France, 2004, pp. 1–4.
- [9] K. Ishizaka and J.L. Flanagan, “Synthesis of voiced sounds from a two-mass model of the vocal cords,” *Bell Syst. Tech. J.*, vol. 51, pp. 1233–1267, 1972.
- [10] N.J.C. Lous, G.C.J. Hofmans, N.J. Veldhuis, and A. Hirschberg, “A symmetrical two-mass vocal-fold model coupled to vocal tract and trachea, with application to prosthesis design,” *Acta Acoustica*, vol. 84, pp. 1135–1150, 1998.
- [11] R. Blevins, *Applied Fluid Dynamics Handbook*, Krieger publishing company, Malabar, 1992.
- [12] S. Candel, *Mécanique des fluides*, Dunod, Paris, 1995.
- [13] Y. Payan and P. Perrier, “Synthesis of v-v sequences with a 2d biomechanical tongue model controlled by the equilibrium point hypothesis,” *Speech Comm*, vol. 22, pp. 185–205, 1997.