Numerical simulation of human voice production using aeroelastic model of self-oscillations of the vocal folds and finite element model of the vocal tract

J. Horáček[a,*], P. Šidlof[a], J.G. Švec[b], F. Griffond-Boitier[c]

[a]*Institute of Thermomechanics, Academy of Sciences of the Czech Republic, Dolejškova 5, 182 00 Prague 8, Czech Republic*
[b]*Medical Healthcom, Ltd, Centre for Communication Disorders, Řešovská 10/491, 181 00 Prague 8, Czech Republic*
[c]*Applied Mechanics Laboratory RC, University of Besancon, 24 rue de l'Epitaphe, 25000 Besancon, France*

Abstract

A complex mathematical model for simulating phonation of the Czech vowel /a/ is presented. The model is based on simulation of the airflow-excited self-sustained vocal fold oscillations, where the non-linear aerodynamic terms and the Hertz model of the contact (collision) forces between the vocal folds are taken into account. The output signals (intraglottal pressure or glottal airflow rate) in the time domain are used for excitation of the finite element (FE) model of the acoustic spaces of the human vocal tract obtained from magnetic resonance imaging (MRI).

Keywords: Flow-induced vibrations; Human voice biomechanics; Non-linear dynamics; Self-sustained vibrations; Vibroacoustics

1. Introduction

A two-degrees-of-freedom aeroelastic model was developed recently [1]. It allows the study of the *post-critical* behaviour of the system after losing the stability and numerical simulations of self-oscillations of the vocal folds during phonation. The Hertz model for impact forces between the vocal folds and non-linear dynamic and aerodynamic forces is implemented into the aeroelastic model. The inviscid incompressible one-dimensional fluid flow theory is used to express the unsteady aerodynamic forces. The parameters of the model, i.e. the mass, stiffness and damping matrices, are related approximately to the geometry, size and material density of the real vocal folds as well as to their known or prescribed fundamental natural frequencies and damping. In this paper, the outputs of the numerical simulations, i.e. the intraglottal pressure or airflow velocity flow rate in the glottis, are used for excitation of the finite element (FE) model developed by Dedouch et al. [2] for the acoustic spaces of the human vocal tract in the Czech vowel /a/.

A more clinically oriented study, based on different models of the vocal folds and vocal tract, was published by Titze [3] focusing on similar time domain simulations of radiated mouth pressure excited by the glottal airflow simulated by a bar-plate model for vibration of the vocal folds.

2. Mathematical model of the vocal folds

A vibrating element of the length L with mass m and moment of inertia I with two degrees of freedom (rotation and translation ${}^{T}\mathbf{V} = [V_1(t), V_2(t)] = [(w_2 - w_1)/2l, (w_1 + w_2)/2]$) supported by an elastic foundation and vibrating in the wall of a channel conveying air is used to approximate the vocal fold oscillations (Fig. 1). Vibrations of one vocal fold are modelled by the equations of motion of an equivalent three-mass system on two springs [1]:

$$\bar{\mathbf{M}}\ddot{\mathbf{V}} + \bar{\mathbf{B}}\dot{\mathbf{V}} + \bar{\mathbf{K}}\mathbf{V} + \mathbf{F} = \mathbf{0}, \tag{1}$$

where $\bar{\mathbf{M}}$, $\bar{\mathbf{B}}$ and $\bar{\mathbf{K}}$ are the mass, damping and stiffness (2×2) matrices, respectively, and the aerodynamic

*Corresponding author. Tel.: +420 266 890383/053125; Fax: +420 286 584695; E-mail: jaromirh@it.cas.cz

Computational Fluid and Solid Mechanics 2005
K.J. Bathe (Editor)

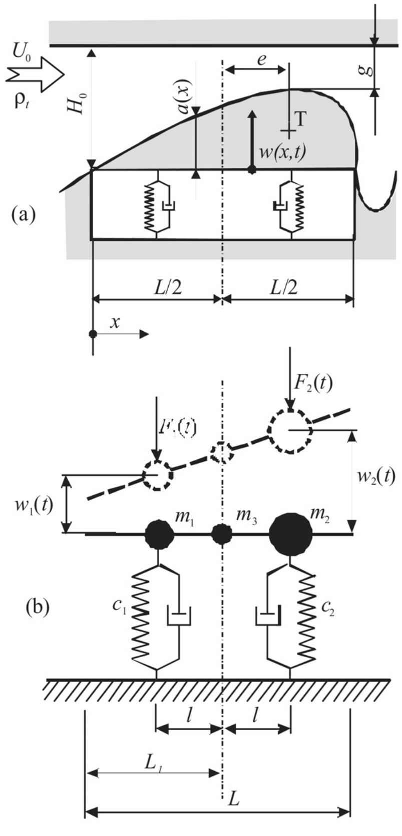

Fig. 1. Model of the vocal fold.

excitation forces **F** are given by the perturbation pressure $\tilde{p}(x,t)$ of the fluid flow in the glottis:

$$F_1(t) = \frac{h}{2}\int_0^L \left(1 - \frac{x}{l} + \frac{L_1}{l}\right)\tilde{p}(x,t)dx,$$

$$F_2(t) = \frac{h}{2}\int_0^L \left(1 + \frac{x}{l} - \frac{L_1}{l}\right)\tilde{p}(x,t)dx; \quad (2)$$

where h is the channel depth, the distances l and L_1 define the two springs positions, and $\bar{\mathbf{B}} = \bar{\varepsilon}_1\bar{\mathbf{M}} + \bar{\varepsilon}_2\bar{\mathbf{K}}$ is assumed.

Aerodynamic forces are calculated from the unsteady Euler and continuity equations:

$$\frac{\partial A}{\partial t} + \frac{\partial (AU)}{\partial x} = 0, \quad \frac{\partial (AU)}{\partial t} + \frac{\partial \left(AU^2\right)}{\partial x} + \frac{A}{\rho_t}\frac{\partial P}{\partial x} = 0, \quad (3)$$

where $A(x,t) = hH(x,t)$ is the channel cross-sectional area, ρ_t and $U(x,t)$ are the fluid density and velocity, respectively, and $P(x,t)$ is the pressure. After separation of steady and unsteady components:

$$U(x,t) = \bar{U}_0(x) + \tilde{u}(x,t), \quad P(x,t) = P_0(x) + \tilde{p}(x,t),$$
$$H(x,t) = H_0 - w(x,t) - a(x). \quad (4)$$

Eq. (3) yields the following equation for the perturbation velocity and pressure:

$$\frac{\partial \tilde{u}}{\partial t} + \frac{\partial\left[\bar{U}_0(x)\tilde{u}\right]}{\partial x} + \tilde{u}\frac{\partial \tilde{u}}{\partial x} = -\frac{1}{\rho_t}\frac{\partial \tilde{p}}{\partial x}. \quad (5)$$

Introducing the velocity potential $\tilde{u} = \partial\Phi/\partial x$, the perturbation pressure can be expressed as

$$\tilde{p}(x,t) = -\rho_t\left[\frac{\partial \Phi}{\partial t} + \bar{U}_0(x)\frac{\partial \Phi}{\partial x} + \frac{1}{2}\left(\frac{\partial \Phi}{\partial x}\right)^2\right]. \quad (6)$$

Considering small vibration amplitudes $|w| \ll H_0$ and the following boundary conditions at the channel inlet and outlet $\tilde{u} = 0\ |_{x=0}$, $\tilde{p} = 0\ |_{x=L}$, the non-linear perturbation pressure can be expressed as

$$\tilde{p}(x,t) = -\rho\sum_{i,j=0}^{2}\sum_{k,l=0}^{2} K_{i,j,k,l}(x)\left[V_1^{(i)}(t)\right]^k\left[V_2^{(j)}(t)\right]^l, \quad (7)$$

where the superscripts above V_1 and V_2 denote the order-of-time derivatives and the coefficients $K_{i,j,k,l}(x)$ are complicated algebraic expressions and functions given by the integrals [1].

The Hertz model of impact force F_H is considered for the vocal fold collisions:

$$F_H = k_H y^{3/2}(1 + b_H\dot{y}), \qquad k_H \cong \frac{4}{3}\frac{E}{1-\mu^2}\sqrt{r}, \quad (8)$$

where E is the Young modulus, μ is the Poisson ratio and r is the radius of the impacting surfaces.

The input parameters for numerical analysis were derived from the data known from the literature for the vocal folds. The geometry of the vocal fold was approximated by the function $a(x) = a_1 x + a_2/2x^2 = 1.858 - 159.86\ x^2$[m]. From here, the coordinates of the contact point can be determined:

$$x_{\max} = \min\{L, \max[0, -(V_1 + a_1)/a_2]\},$$
$$y_{\max} = y(x_{\max}) = a(x_{\max}) + (x_{\max} - L_1)V_1 + V_2.$$

The impact Hertz force can be expressed as $F_H = k_H(y_{\max} - H_0)^{3/2}$, where $k_H \cong 730\,\mathrm{Nm}^{-3/2}$ for $E = 8\,\mathrm{kPa}$ and $\mu = 0.4$ and where the damping coefficient was neglected $(b_H = 0)$; $H_0 = \max_{x\in<0,L>} a(x) + g$ is the height of the channel and g is the glottal half-width. A correction on the static subglottal pressure, which is constant during the vocal folds collision, gives after integration of the pressure p_{sub} in the interval $x \in <0,$

$x_{\max} >$ the resulting forces in Eq. (1) during vocal folds contact:

$$F_1 = F_H \frac{L_1 + l - x_{\max}}{2l} + p_{sub} h x_{\max} \frac{L_1 + l - \frac{x_{\max}}{2}}{2l},$$
$$F_2 = F_H \frac{x_{\max} - L_1 + l}{2l} + p_{sub} h x_{\max} \frac{\frac{x_{\max}}{2} - L_1 + l}{2l}. \quad (9)$$

For numerical simulations, Eq. (1) was transformed into the system of four first-order ordinary differential equations:

$$Z_1' = \hat{f}_1(Z_1,Z_2,V_1,V_2), \quad Z_2' = \hat{f}_2(Z_1,Z_2,V_1,V_2), V_1' = Z_1, \quad V_2' = Z_2 \quad (10)$$

and the fourth-order Runge–Kutta method was used for the calculations. The functions $\hat{f}_1$, $\hat{f}_2$ are determined differently for non-contact and contact regimes by Eqs. (2) and (9).

The density, thickness and length of the vocal folds were taken as follows: $\rho_h = 1020\,\mathrm{kg/m^3}$, $L = 6.8$ mm and $h = 10$ mm. From these data, we calculated the eccentricity e, the total mass m and the moment of inertia I; the air density was considered as $\rho_t = 1.2\,\mathrm{kg/m^3}$, $L_1 = L/2$ and $l = 0.344\,L$. A tuning procedure was used to adjust the stiffness (c_1,c_2) of the elastic foundation and the damping coefficients $\bar{\varepsilon}_1$, $\bar{\varepsilon}_2$ in order to approximate the pitch frequency $F0 \cong 100$ Hz, setting the natural frequencies $f_1 = 100$ Hz, $f_2 = 105$ Hz and 3 dB half-power bandwidths of both resonances in correspondence to the physiological data ($\Delta f_1 = 23$ Hz and $\Delta f_2 = 29$ Hz).

3. Finite element model of the vocal tract

The geometry of vocal tract shape for the vowel /a/ was obtained from a native Czech speaker using magnetic resonance imaging (MRI). An automatic procedure was developed that enabled us to transform the MRI data directly into the FE model (Fig. 2). The acoustic pressure p inside a cavity is described by the equation

$$\nabla^2 p = \frac{1}{c_0^2} \frac{\partial^2 p}{\partial t^2}, \quad (11)$$

where c_o is the speed of sound. For the vocal tract walls, the boundary condition on the acoustically hard area $\frac{\partial p}{\partial \mathbf{n}} = 0$ is considered, where $\mathbf{n}$ is the direction of the normal to the boundary area. At the mouth, the boundary condition is, in this first approximation, assumed as $p = 0$. The equation of motion for the acoustic pressure in the FE formulation can be written in matrix form in the global coordinate system as

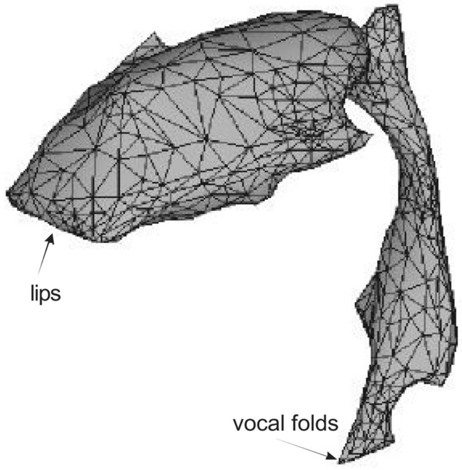

Fig. 2. FE model of the vocal tract for the vowel /a/.

$$\mathbf{M}\ddot{\mathbf{P}} + \mathbf{B}\dot{\mathbf{P}} + \mathbf{K}\mathbf{P} = \mathbf{f}(t) \quad (12)$$

where **M, B** and **K** are the mass, acoustic boundary damping and stiffness matrices, respectively, and **P** and **f** are the vectors of nodal acoustic pressures and excitation forces. The effect of outgoing acoustic energy was modelled by an absorption boundary condition at the lips, where the boundary admittance was prescribed in correspondence to the 3-dB half-power bandwith known for formant (acoustic resonant) frequencies. The transient analysis with the Newmark integration method was used for numerical simulation of the acoustic signals near the lips when the excitation was applied at the position of the vocal folds. The excitation signal was the intraglottal pressure $\tilde{p}(t) = \tilde{p}(x,t)|_{x=L-0.5_{\min}}$ or the airflow volume velocity $Q(t)$ resulting from the aeroelastic model of the vocal folds.

4. Results

An example of the excitation signal $Q(t)$ received from the aeroelastic model and the simulated time response of the vocal tract obtained for the acoustic pressure $p(t)$ near the lips are shown in Fig. 3. It is possible to see that after two to three airflow volume velocity pulses, the response is stabilized and becoming periodical. The input and output signals in the frequency domain are shown in Fig. 4, together with the transfer function $H_{\mathrm{Qp}}(f)$ between them, from which it is possible to detect the first two formant frequencies near 670 Hz and 1070 Hz, which are in a reasonable agreement with known phonetic data [4].

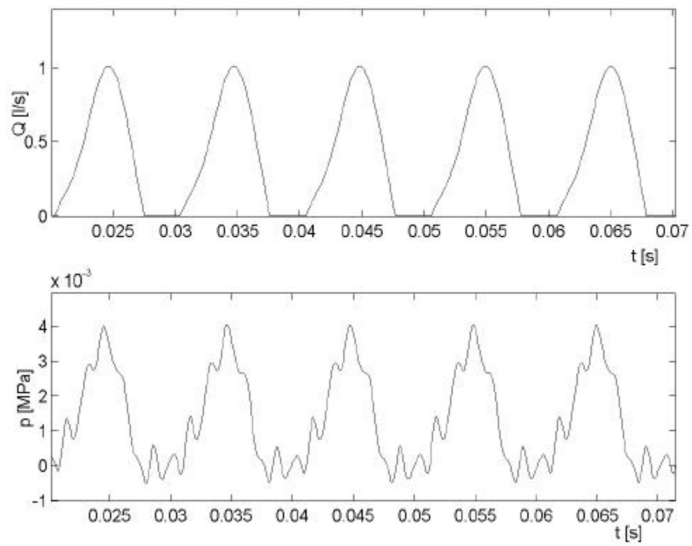

Fig. 3. Glottal airflow rate and pressure near the lips.

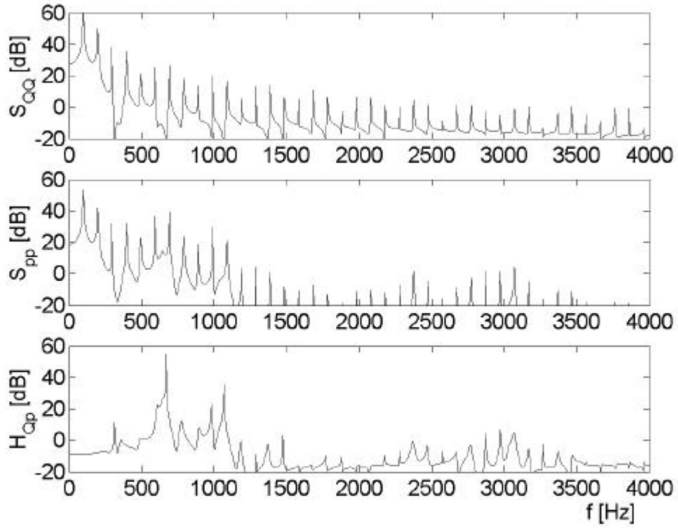

Fig. 4. Spectra of the input flow pulses, output pressure and vocal tract transfer function.

5. Conclusions

A joint aeroelastic model for the vocal fold self-oscillations with the FE model of the vocal tract was presented. The aeroelastic model generates the intraglottal pressure and airflow rate. These signals were used for acoustic excitation of the vocal tract at the position of the vocal folds. Both ways of excitation, i.e. by pressure or airflow volume velocity pulses, resulted in a good approximation of the voice signal for the modelled vowel /a/. The spectrum of the acoustic pressure near the lips contains many harmonics and the first two formants are in good agreement with measurement in humans. This was also checked perceptually by listening to the sound from the wav files generated by the output signals. The generated signal sounded more natural when excited by the airflow volume rate than when excited by the glottal pressure.

Acknowledgement

Research is supported by the Grant Agency of the Czech Republic within the project 106/04/1025

Modelling of vibroacoustic systems focusing on human vocal tract.

References

[1] Horáček J, Šidlof P, Švec JG. Numerical simulation of self-oscillations of human vocal folds with Hertz model of impact forces. In: Flow-Induced Vibration Langre E, Axisa F, editors, Paris: École Polytechnique, 2004; pp. 143–148.

[2] Dedouch K, Horáček J, Vampola T, Švec JG, Kršek P, Havlík R. Acoustic modal analysis of male vocal tract for Czech vowels. In: Proc. Interaction and Feedbacks. Prague, 26–27 Nov., Zolotarev I, editor, Institute of Thermomechanics, ASCR, Prague, 2002, pp. 13–20.

[3] Titze IR. Regulating glottal airflow in phonation: application of the maximum power transfer theorem to a low dimensional phonation model. J Acoust Soc Am 2002;111:367–376.

[4] Titze IR. Principles of Voice Production. London: Prentice Hall, 1994.