18th International Conference
ENGINEERING MECHANICS 2012
Svratka, Czech Republic, May 14 – 17, 2012
pp. 1475–1483
Paper #89

# INFLUENCE OF THE GEOMETRIC CONFIGURATIONS OF THE HUMAN VOCAL TRACT ON THE VOICE PRODUCTION

**T. Vampola*, J. Horáček****

**Abstract:** *The three-dimensional (3D) finite element (FE) model of the human vocal tract was constructed, based on CT measurements of a subject phonating on [a:]. A special attention is given to the higher frequency range (above 3.5 Hz) where transversal modes exist between piriform sinuses (PS) and valleculae (VA) and where the higher formants can create a formant cluster known as the speaker's or singer's formant. Since the human ear is most sensitive to frequencies between 2 and 4 kHz concentration of sound energy in this frequency region (F4-F5) is effective for communication.*

**Keywords:** *Boimechanics of voice, 3D FE model, acoustics characteristics, singer's formant*

## 1. Introduction

At present, a considerable attention is given to the computer simulation of voice production in relation to the sound pressure field inside the human vocal tract during phonation. Main effort is focused on understanding the generation of the articulated audio signal and all the factors influencing this process and the voice quality. The mathematical models of the human vocal tract allow much easier and detailed computer analysis of the acoustic pressure in the vocal tract than it can be obtained by acoustic measurements *in vivo*. A number of simplified 1D models of the vocal tract cavities can be found in literature - see e.g. Titze at al. (1997) or Laukkanen at al. (2009). The advantage of these approaches lie in the relatively simple mathematical description of the behaviour of such derived models. However, it is necessary to accept, that these models were derived under the assumption of planar acoustic waves travelling in the vocal tract and don't accept the 3D geometric configuration of the real human vocal tract (Vampola at al., 2008). For lower values of frequencies of the vocal cavities, for which are not excited the transversal shapes of vibrations, are computed pressure fields of these models in a good relationship with measured data. For a precise evaluation of the voice quality due to geometric modification of the vocal tract for which must be taken into account the hire frequency spectrum are these models inadequate. However, for the evaluation of the impact of individual factors that need to be taken into consideration in the preparation of the real 3D model of the vocal tract are these simplified models very useful computation tool.

## 2. Simplified model of the human vocal tract

The influence of geometric configuration of the human vocal tract on the generated acoustic pressure was in the first step simulated by the simplified model shown in *Fig 1*. The method of direct physical discretization was used for derivation of the physical parameters of this model. This method is sufficient for a quantitative assessment of the influence of geometric modification of the vocal tract on the acoustic characteristics. The mass and damping matrixes of the simplified model were derived by the direct physical discretization according to *Fig.1*. The stiffness matrix was derived by the inverse procedure. In the first step was accepted the simplified model without parallel chains (the masses $m_8$ and $m_9$ were neglected).

* Doc. Dr. Ing. Tomáš Vampola: Department of Mechanics, Biomechanics and Mechatronics, Faculty of Mechanical Engineering,CTU, Technická 4; 166 07, Prague 6; CZ, e-mail: tomas.vampola@fs.cvut.cz

** Ing. Jaromír Horáček, DrSc.: Institute of Thermomechanics, Academy of Science of the Czech Republic; Dolejškova 5, 182 00 Praha 8; CZ, e-mail:jaromirh@it.cas.cz

In the position defined by the coordinate $x_0$ was assumed the inhomogeneous pressure boundary condition $p(x_0,t)=p_0(t)$. In the position of mouth was used the homogeneous pressure boundary condition $p(x_{10},t)=0$.

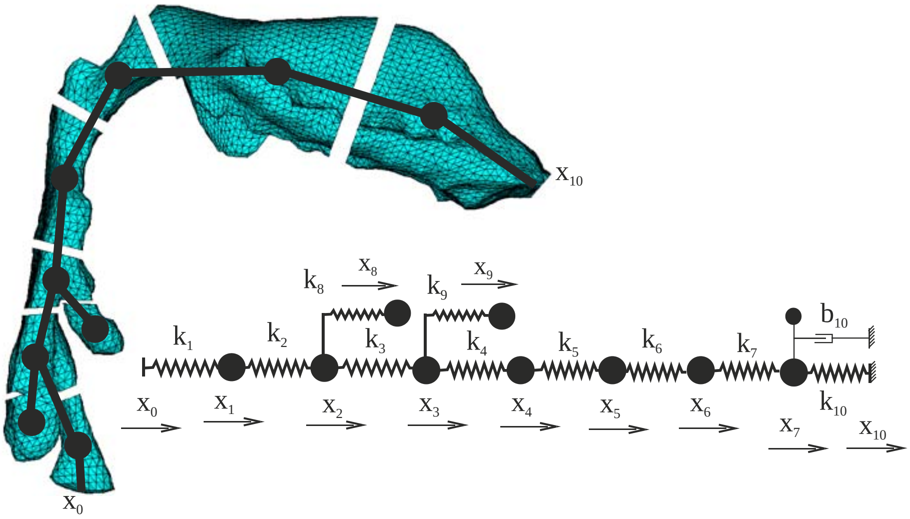



*Fig.1 Simplified model of the human vocal tract*

The condition of the static pressure equilibrium can be reformulated as

$$\begin{bmatrix} k_{12} & -k_2 & 0 & 0 & 0 & 0 & 0 \\ & k_{23} & -k_3 & 0 & 0 & 0 & 0 \\ & & k_{34} & -k_4 & 0 & 0 & 0 \\ & & & k_{45} & -k_5 & 0 & 0 \\ & sym. & & & k_{56} & -k_6 & 0 \\ & & & & & k_{67} & -k_7 \\ & & & & & & k_{710} \end{bmatrix} \begin{bmatrix} p_1 \\ p_2 \\ p_3 \\ p_4 \\ p_5 \\ p_6 \\ p_7 \end{bmatrix} = \begin{bmatrix} k_0 p_0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} \tag{1}$$

where, the symbols $k_{ij}=k_i+k_j$ were used for simplification of notation. By using the global stiffness matrixes of the separate flexible members of the model according to *Fig.1* can be the equation of the pressure static equilibrium (1) rewritten to:

$$\left(k_1\mathrm{K}_{1g} + k_2\mathrm{K}_{2g} + \cdots + k_{10}\mathrm{K}_{10g}\right)\mathrm{p} = \mathrm{f} \tag{2}$$

where $k_i$ $i=1,\ldots 10$ are the unknown stiffness parameters of the simplified model and vector **p** was assembled from the known values of the pressure field computed by the FE analysis of the 3D volume model in positions defined by coordinates $x_i$ $i=1,\ldots 10$. Vector **f** contains the kinematic excitation of the simplified model. In the following analyses was used the harmonic excitation in position of vocal folds - position $x_0$. The form of the matrixes $\mathbf{K}_{ig}$ can be derived from (1). For example the matrix $\mathbf{K}_{2g}$ can be defined in the form:

$$\mathrm{K}_{2g} = \begin{bmatrix} 1 & -1 & 0 & 0 & 0 & 0 & 0 \\ & 1 & 0 & 0 & 0 & 0 & 0 \\ & & 0 & 0 & 0 & 0 & 0 \\ & & & 0 & 0 & 0 & 0 \\ & sym. & & & 0 & 0 & 0 \\ & & & & & 0 & 0 \\ & & & & & & 0 \end{bmatrix} \tag{3}$$

Equation (2) can be reformulated to the final form:

$$\left(\mathrm{K}_{1g}\mathrm{p}, \mathrm{K}_{2g}\mathrm{p}, \cdots, \mathrm{K}_{10g}\mathrm{p}\right)\mathrm{k} = \mathrm{f} \tag{4}$$

The order of the stiffness matrix in the equation (4) is $n*(n+1)$, where $n$ is the number of discrete masses in the linear chain. Choosing the value of stiffness $k_1=\xi$ can be the equation (4) rewritten to the final form:

$$\mathrm{K}(\mathrm{p})\,\mathrm{k} = \mathrm{f} \tag{5}$$

where **K(p)** is the stiffness matrix whose elements are functions of the pressure obtained from the analysis of 3D FE model. The vector **k** contains unknown stiffness parameters of the linear chain (*Fig.1)* and the vector **f** results from the inhomogeneous boundary condition $p(x_0,t)=p_0(t)$ at the position of the vocal folds. The stiffness parameters of the simplified model are computed by using the additional condition on the first eigen-frequency that is identical with the 3D model. The similar procedure can be used for determination of the stiffness parameters of the PS and the VA, where the shortened chain between the vocal folds and the PA or VA was used. The volumes of the PS and the VA are connected with the basic frame by the very soft stiffness. The matrixes derived in this way are only approximate but maintain the proportionality of the stiffness parameters of the 3D volume model. The dominant mechanism of the energy dissipation from the human vocal tract is due to emitting the acoustic energy from the mouth to the open space. In the simplified model was this mechanism of energy dissipation modeled by the acoustic impedance $Z_a$. The model of the circular plate vibrating in the infinite wall was used (Vampola at al., 2008)

$$Z_a = \frac{c_0\rho_0}{S}\left(A + iB\right), \quad A = 1 - \frac{J_1(2kR)}{kR},\ B = \frac{H_1(2kR)}{1+kR}, \tag{6}$$

where functions $J_1$ and $H_1$ can be expressed by means of the infinite series

$$J_1(2kR) = \left(\frac{2kR}{2}\right)\sum_{m=0}^{\infty}\frac{(-1)^m}{m!(m+1)!}\left(\frac{2kR}{2}\right)^{2m}, \tag{7}$$

$$H_1(2kR) = \sum_{m=0}^{\infty}\frac{(-1)^m}{\Gamma\left(m+\frac{3}{2}\right)\Gamma\left(m+\frac{3}{2}+1\right)}\left(\frac{2kR}{2}\right)^{2m+2}.$$

In the equation (6) is $k = \omega/c_0$ wave number. $R,S$ are the radius and area of the vibrating circular plate simulating the mouth area, $c_0$ is speed of the sound in the air cavity of the vocal tract and $\rho_0$ is density of the air for the defined conditions. The expression (6) can be normalized due to standardized wave resistance $\rho_0 c_0$ and area of the vibrating circular plate:

$$z_{nv} = A + iB, \tag{8}$$

where the real part describes the emitting resistance and the imaginary part is the fictional mass air-column vibrating with the human vocal tract. Using the validity of the relations between the acoustic and mechanical impedance can be parameters used in the model of the vocal tract in *Fig. 1*, defined by the expressions:

$$b_{10}(\omega) = S^2 r_a^*(\omega), \quad m_{10}(\omega) = S^2 m_a(\omega) \quad , \tag{9}$$

where $r_a^*(\omega)$, $m_a(\omega)$ are the frequency dependent values of the acoustic resistance and the acoustic mass following from (8). The character of the acoustic radiation impedance for the area of the mouse approximately 7e-3 m$^2$ is presented in the *Fig.2*.

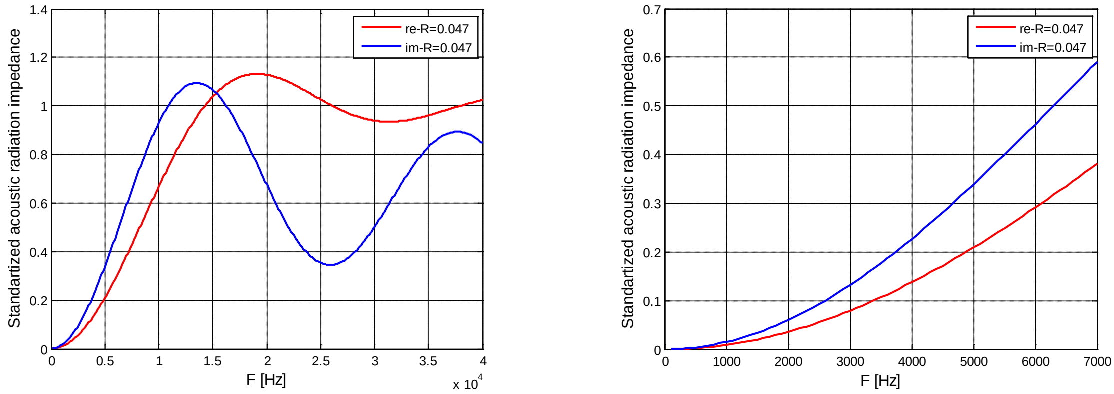



*Fig.2 Standartized acoustic radiation impedance*

It is necessary to keep in mind that the using of the simplified model of the human vocal tract is limited to the lower frequency span. The model acts as a frequency filter, which removes from the frequency spectrum the frequencies of the transversal vibrating modes, which the simplified model is not able properly modeled. With regard to this fact, it is evident that for the lower values of the frequency range, can be the standard radiation impedance approximated with sufficient precision by the polygon line.

## 3. Influence of geometric configuration of the model to the position of the resonant peaks

The influence of the shape modification of PS and VA to the generated pressure characteristics of the human vocal tract was in the first step used the simplified model according to *Fig.1*. This model is sufficient for the quantitative assessment influence of the geometric modification to the predicted pressure fields of the human vocal tract. The character of the pressure field in the position defined by the coordinate $x_7$ was simulated. In the first step can be this value accepted as the output pressure from the human vocal tract. The pressure fields were calculated using the method of harmonic analysis. The vector of unknown pressures amplitudes is considered in the form

$$\mathrm{p}(t) = \mathrm{r}e^{j\omega t} . \tag{10}$$

The amplitudes of excited vibration ***r*** can be derived from the equation of motion:

$$\mathrm{r} = (\mathrm{K} - \omega^2\mathrm{M} + i\omega\mathrm{B})^{-1}\mathrm{h}(\omega) = \mathrm{A}^{-1}\mathrm{h}(\omega) . \tag{11}$$

where h(ω) is vector of the kinematic excitation of the system. The complex amplitudes of the excited vibration can be derived from the expression (11):

$$\mathrm{r}_i = \frac{\mathrm{q}_i}{\det(\mathrm{A})}\mathrm{h}(\omega) . \tag{12}$$

From equation (12) results

$$q_7 = k_2 k_3 k_4 k_5 k_6 k_7 \,. \tag{13}$$

for the model, when were not used the masses $m_8$ and $m_9$ in the linear chain. In case, that the mass $m_9$ is not used, we got:

$$q_7 = k_2 k_3 k_4 k_5 k_6 k_7 \left(k_8 - m_8 \omega^2\right). \tag{14}$$

For the full model according to *Fig.1* was

$$q_7 = k_2 k_3 k_4 k_5 k_6 k_7 \left(k_8 k_9 - \omega^2 \left(k_8 m_9 + k_9 m_8\right) + \omega^4 m_8 m_9\right). \tag{15}$$

From the equations (14 and 15) it is apparent that by adding a "parallel" branch to a serial chain the state of antirezonance can occurs, when the output pressure amplitude of the forced vibrations on the output from the model are zero. The equations (14 and 15) from the equation (11) differ by the modifier

$$mo_1 = \left(k_8 - m_8 \omega^2\right) \tag{16}$$

$$mo_2 = \left(k_8 k_9 - \omega^2 \left(k_8 m_9 + k_9 m_8\right) + \omega^4 m_8 m_9\right)$$

In the *Fig.3* is presented the dependence of the amplitude of the output pressure at the point defined by the coordinate $x_7$ due to the excitation frequency for the nominal size of the volume of the PA and VA

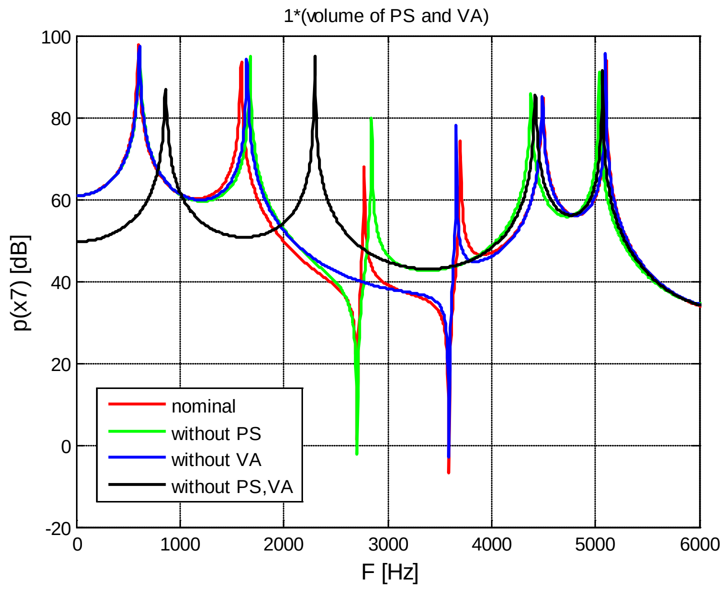



*Fig.3 The dependence of the amplitude of the output pressure in position $x_7$ on the excitation frequency*

It is evident that, in accordance with the (13) that for the model without the "parallel" branches is not in the output signal presented the antirezonance frequency. On the contrary, by adding one or two parallel branches arise in the output pressure signal one or two antirezonance frequency. The dependence of the relationship (16) on the excitation frequency is presented in *Fig.4*.

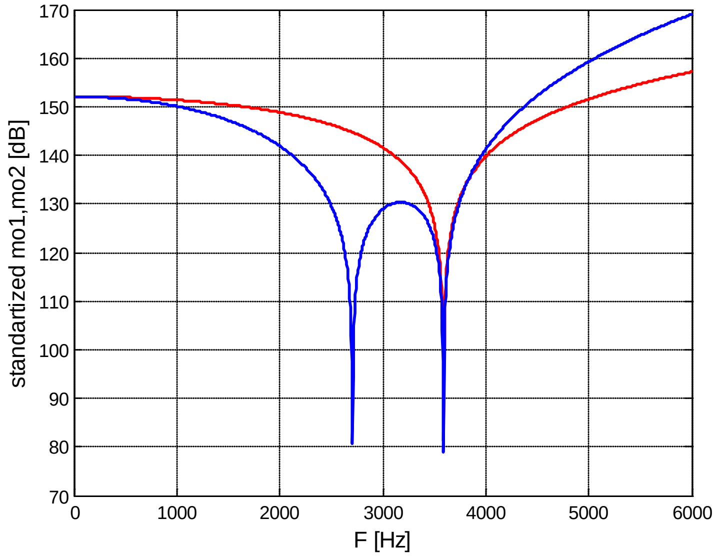



*Fig.4 The dependence of "modifiers" on the output pressure in position $x_7$ on the excitation frequency*

In the *Fig. 5* is introduced the dependence of the resonant and antirezonant peaks for the simplified model, where the volumes of PA and VA were changed.

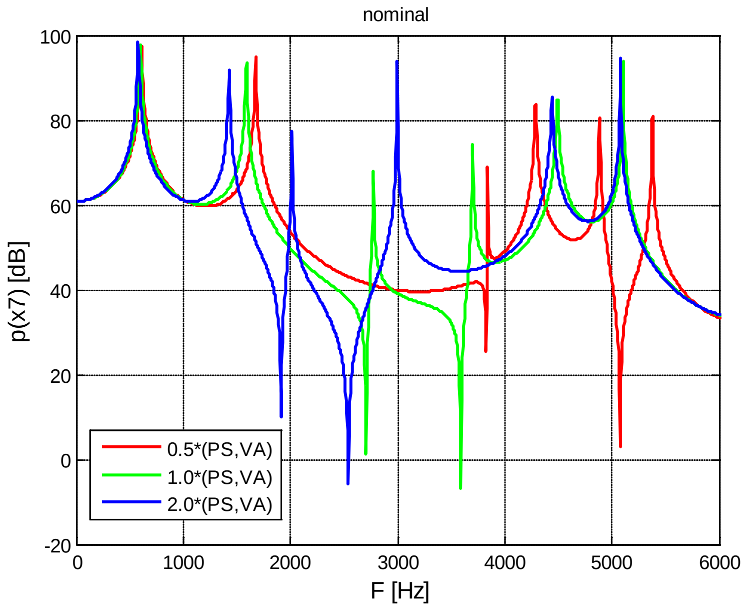



*Fig.5 The dependence of the position of the resonant peaks for the modified volume of PS and VA*

It is apparent that for increasing volume of the "parallel" branches the frequency range between antirezonance frequencies decreases and shift to the lower frequency values. The position of the antirezonance peaks can therefore be used to evaluating the size of the volume of the "parallel" branches of the human vocal tract. In the *Fig. 6* can be seen that if in the output signal is presented the antiresonance frequency then the energy coming into the system is "consumed" by the relatively heavy vibration of the parallel branches. It is to be noted that in these considerations do not take into account the influence of transversal shapes of vibrations.

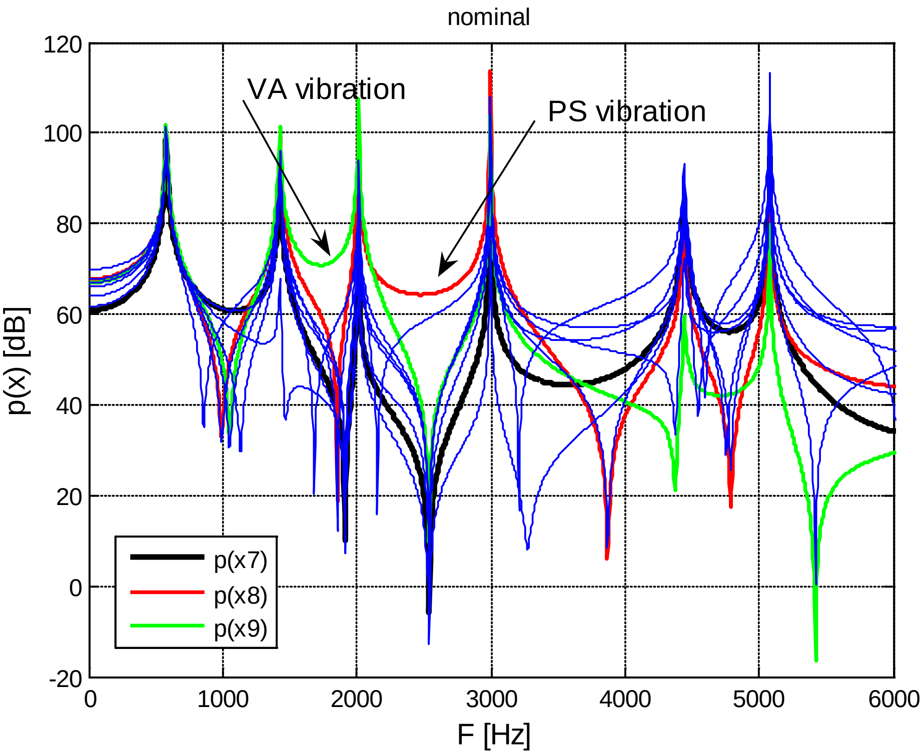



*Fig.6 Pressure field of the simplified model of the human focal tract*

Therefore these findings were proved on the 3D volume models.

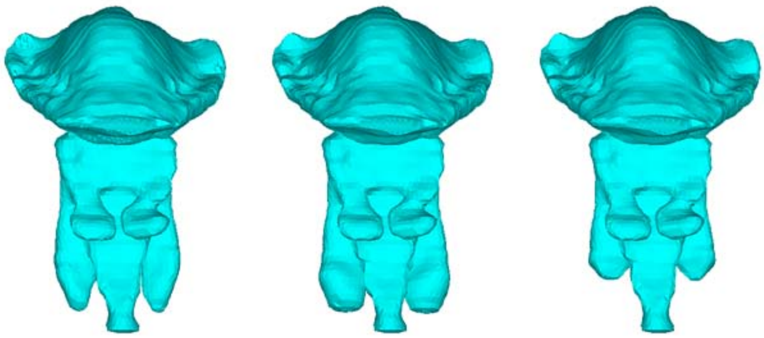
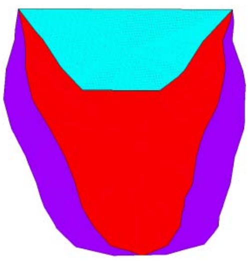


*Fig.7 Volume modification of PS*

The geometric modification of the PS is presented on the *Fig.8* in more detail



| | Volume [$mm^3$] |
|---|---|
| smaller | 304 |
| nominal | 1317 |
| bigger | 2437 |

*Fig.8 Volume modification of the PS*

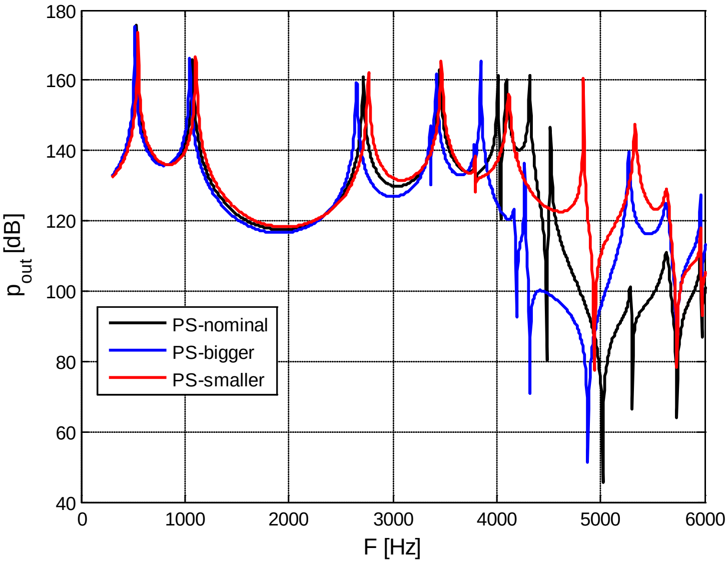



*Fig.9 The output pressure field of the 3D model*

In the *Figs. 10* and *11* are introduced the characters of the vibration of the 3D volume model of the human vocal tract during the antiresonace frequencies.

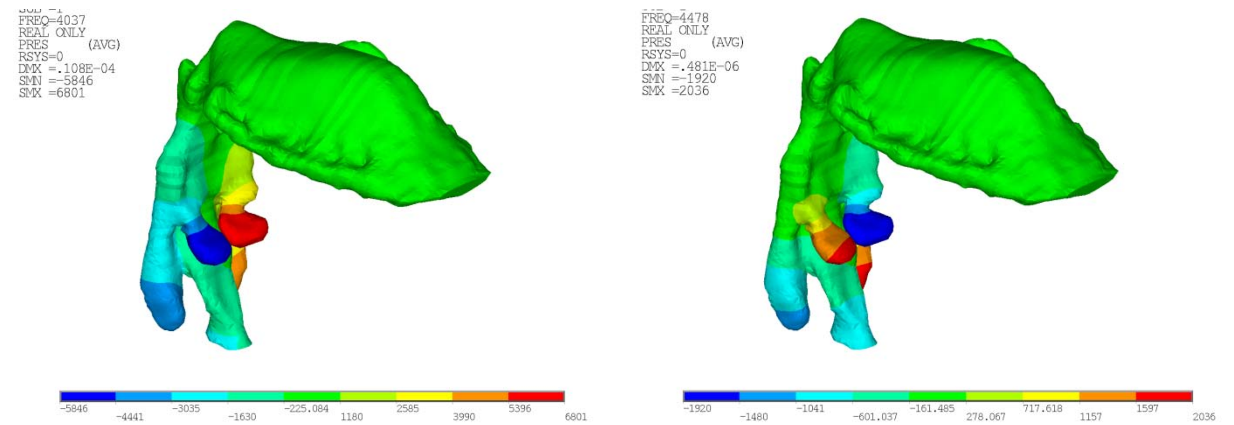



*Fig.10 The first and the second significant antiresonance frequencies for the nominal volume PS*

In accordance with the *Fig.9* it can be seen the significant (rarely damped) vibration of the „parallel branches“ of the vocal tract.

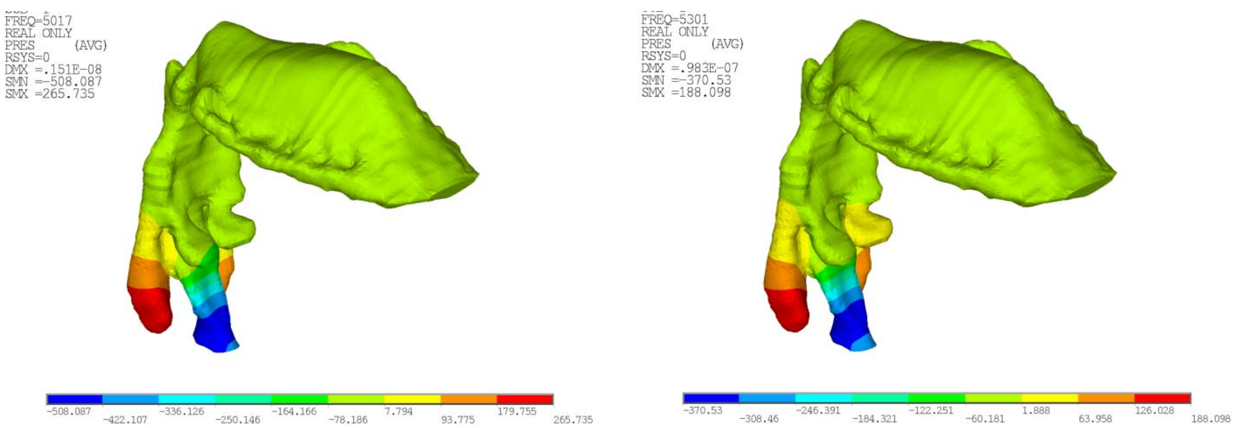



*Fig.11 The third and fourth significant antiresonance frequency for the nominal volume PS*

From the above it is apparent that in the frequency region about 4kHz the accumulation of the energy occurs by increasing of the volume of parallel branches. So called the singer's formant is formed. For frequencies values about 4kHz are the human ears the most sensitive. The accumulation of energy in this frequency span contributes to the clarity of communication. On the contrary, when is disproportionately enlarged PS the shift in the frequency spectrum occurs and the level of the emitted energy is lower. It can be seen that for optimal phonation it is necessary to reach the appropriate "optimal" volume configuration of the parallel branches of the vocal tract. For too small PA and VA is voice energy concentrated in other frequency ranges than would be appropriate for the sensitivity of the ear. For very large volumes of PS and VA occurs in the required frequency range decreasing of the energy level. In the frequency spectrum (*Fig.9*) is interesting the frequency peek about 3.8kHz. It is the first excited transversal mode of vibration, when the energy emitted from the vocal tract is decreased.

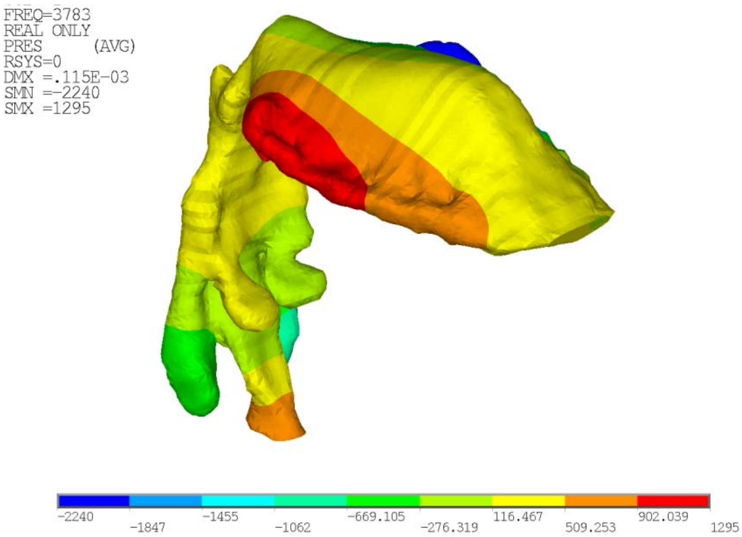



*Fig.12 The first transversal shape of vibration*

## 4. Conclusions

The simplified, computationally efficient 1D model of the vocal tract was assembled and used for prediction of the pressure fields for a more clear explanation of effects of geometrically changed configurations of the human vocal tract resulted from the size of piriform sinuses and valleculae. The findings were compared with the results acquired from the 3D FE models with the good correlation.

## Acknowledgement

The research was supported by the grant GACR No. P101/12/1306 *Biomechanical modeling of human voice production – way to artificial vocal folds.*

## References

Titze, I.R. and Story, B.H. (1997). "Acoustic interactions of the voice source with the lower vocal tract", *J. Acoust. Soc. Am.,* 101, 2234-2243.

Vampola, T., Horáček, J. and Švec, J.G. (2008). "FE modeling of human vocal tract acoustic. Part I: Production of Czech vowels", *Acta Acust. United Acust.*, 94, 433-447.

Laukkanen, A.M., Radolf, V., Horáček, J., Leino, T. (2009). "Estimation of the origin of a speaker's and singer's formant cluster using an optimization of 1D vocal tract model", *Proceedings of 3rd Advanced Voice Function Assessment*. Madrid: EUIT Telecomunicación, 2009, pp. 1-4. ISBN 84-95227-64-9.