19th International Conference
ENGINEERING MECHANICS 2013
Svratka, Czech Republic, May 13 – 16, 2013

EFFECT OF THE SIZE OF PIRIFORM SINUSES ON THE VOICE QUALITY

T. Vampola*, J. Horáček**

Abstract: *The influence of piriform sinuses (PS) on the resonance and antiresonance characteristics of the vocal tract is investigated. The change in sizes of PS cavities alters the resulting voice quality. Pilot studies reveal that additional formants caused by PS can occur in the frequency range of 3 – 5 kHz, i.e., in the range which is important for the production of the so called singer's or speaker's formant. This contribution therefore aims at investigating the influence of the side cavities of the vocal tract in more detail using two computational models of the vocal tract. First, is presented analysis of the influence of the acoustic spaces of PS on the existence of resonances and antiresonances in the spectra of the acoustic signal simulated using a reduced finite element (FE) model of the human vocal tract. Then the full FE model is used for the analysis by using direct numerical simulations of phonation.*

Keywords: ***Voice quality, singers's formant, biomechanics of voice, human vocal tract, numerical simulation of phonation***

1. Introduction

The influence of the vocal tract on vocal output has been presented in many articles. The influence of side cavities of human vocal tract, such as the piriform sinuses (PS) has received much less attention. Generally, these cavities have been reported to cause antiresonances in the resulting vocal spectrum, i.e., largely decreasing the radiation of some frequencies out of the mouth, particularly around 4-5 kHz. (Motoki, 2002; Mokhtari, *et al.*, 2008). As such, their role for the resulting vocal intensity may be considered undesirable, since it contrasts with the general goal of achieving maximum vocal output with the smallest vocal effort. However, a more detailed analysis reveals that besides the antiformants there are also new formants which occur due to these side cavities (Honda *et al.*, 2004). In technical terms, the effect is side cavities can be described as creating a zero-pole pair in the overall transfer function of the vocal tract (Titze and Story, 1997). Pilot studies reveal that additional formants caused by PS can occur in the frequency range of 3 – 5 kHz, i.e., in the range which is important for the production of the so called singer's formant (in operatic voices) or speaker's formant (observed in professional speakers) (Vampola *et al.,* 2013). It can thus be speculated that these cavities can help in establishing a "better" voice quality in speaking and singing. Considerable progress has been made lately in medical investigation methods like computer tomography which enables to create high quality three-dimensional (3D) models of the vocal tract taking into account details of the laryngeal cavity, valleculae and piriform sinuses. Use of these precise models for the evaluation of acoustic characteristics due to changed geometric configuration of the vocal tract is rather time consuming process.

2. Computational efficient model of the human vocal tract

For a quick evaluation of geometric configuration of PS on the generated acoustic characteristics the simplified one-dimensional (1D) model was developed (Vampola & Horáček, 2012). It should be realized that this model was 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. The acoustic pressure fields computed inside the vocal cavities by using the 1D models correspond to the measured

* 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

data only for lower frequencies, for which the transversal acoustic mode shapes of vibration are not excited. This model was derived from the 3D FE model by the reduction process (Vampola *et al.*, 2008). Because the relatively dramatic reduction of the original 3D FE model was used, it was necessary to tune the basic acoustic resonance characteristics of the reduced model so that they correspond to the fundamental resonances and antiresonances of the full 3D FE model. The global multicriteria optimization method was used for tuning of the stiffness parameters of the reduced 1D model according to the chosen fundamental eigenfrequencies and anti-resonance frequencies of the 3D model. The objective function was chosen as

$$F_{obj} = \sum_{i=1}^{Q} (f_{i,res\_demand} - f_{i,res})^2 + \sum_{i=1}^{P} (f_{i,antires\_demand} - f_{i,antires})^2 , \tag{1}$$

where Q and P were considered for in the computation for the first Pth and Qth resonant and anti-resonant frequencies, respectively. Because the global optimization method generates many solutions we applied the additional constraint functions in order not to change much the found stiffness parameters of the reduced model of the vocal tract. The following constraint functions were used

$$m_i = const., \; i = 1,\cdots,11 \quad \text{and} \quad \sum_{i=1}^{11} \left(k_{i,new} - k_{i,orig}\right)^2 \Rightarrow minimum , \tag{2}$$

where m_i, i = 1,…,11 superelements were used. The dramatic reduction of stiffness parameters were used. The relatively comprehensive stiffness matrixes of separate super-elements were substituted by the single “springs”.

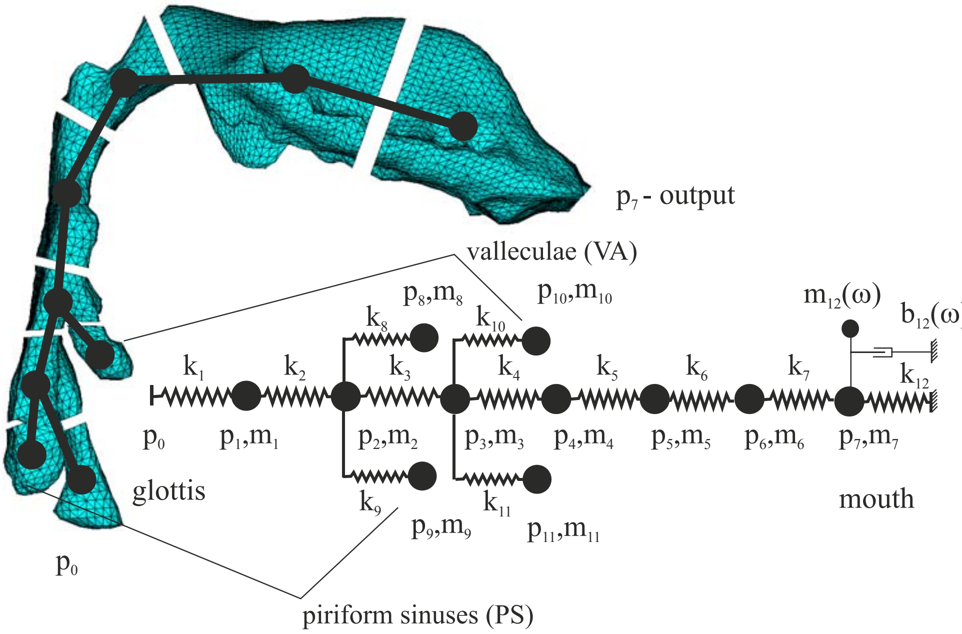

Fig. 1. Schema of simplification of the 3D FE model of the human vocal tract into the reduced model. The p_i and m_i symbols denote the acoustic pressures and masses of the different cavities along the vocal tract.

The dominant mechanism of the energy dissipation from the human vocal tract is given by emitting of the acoustic energy from the mouth to the open space. In the reduced model, the mechanism of the energy dissipation was modeled by the standardized radiation acoustic impedance (see Vampola *et al.*, 2008):

$$z_{nv} = Z_a \frac{S}{c_0 \rho_0} = r_a(kR) + j\, m_a(kR), \tag{3}$$

where ρ_0 is the air density, c_0 is the speed of sound, $j = \sqrt{-1}$ is the imaginary unit and $r_a(kR)$ and $m_a(kR)$ are nonlinear functions of the wave number $k = 2\pi f\, c_0^{-1}$, f is the frequency, R is the radius which corresponds to a circular plate vibrating in the infinite wall, $S=\pi R^2$ is the area between lips. The real part r_a describes the emitting resistance and the imaginary part m_a (reactance) is a fictional mass of an air-column vibrating with the human vocal tract in front of the lips.

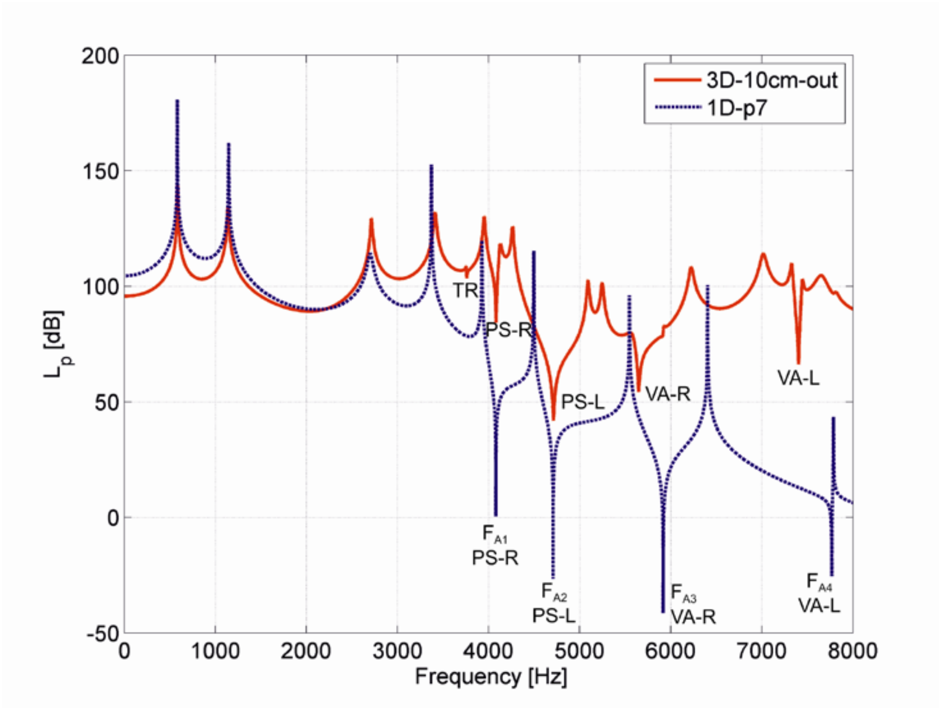

Fig. 2. Acoustic pressure response computed at the distance of 10 cm in front of the mouth using the full 3D FE model of the vocal tract (red full line) and at the output pressure p_7 from the reduced model (blue dashed line).

Comparison of the acoustic pressure response computed at the distance of 10 cm in front of the mouth using the full 3D FE model with the output pressure p_7 radiated from the reduced model of the vocal tract is shown in Fig. 2. The first four resonant frequencies up to 4 kHz and first two antiresonant frequencies of the reduced model are in very good agreement with the resonances obtained from the full 3D model. Four prominent antiresonances in the pressure response are present in the frequency region from 4 to 8 kHz in both models. There is also a small antiresonance-resonance pair at the frequency of 3756 Hz in the full 3D model which is missing in the reduced model. This is because this resonance belongs to the first transversal acoustic mode shape in the oral cavity which cannot be modeled by the reduced model and which was identified using a modal analysis of the full 3D model. This transversal mode is shown in Fig. 3. At the frequencies above 4 kHz the two models show different resonances; there are many more resonances in the full 3D model. This can be attributed to the limitation of the reduced model which is not able to model correctly the resonances higher than 4 kHz due to a small number of the superelements considered.

3. Influence of the size of the piriform sinuses to the quality of the vouce production

The influence of the shape and size modifications of the piriform sinuses (PS) and valleculae (VA) on the generated acoustic pressure characteristics inside and outside of the human vocal tract was, in the first step, analyzed by means of the reduced model of the vocal tract. This model is sufficient for a quantitative assessment of the influence of the geometric modifications of the vocal tract on the predicted pressure field, which were numerically simulated using the method of harmonic analysis. The vocal tract cavities were harmonically excited at the location of the vocal folds by the acoustic pressure p0. The acoustic pressure response is given by the solution of the following non-homogeneous equation for the reduced model:

$$\mathbf{M_r}\ddot{\mathbf{p}} + \mathbf{B_r}\dot{\mathbf{p}} + \mathbf{K_r}\mathbf{p} = \mathbf{p_0}e^{j\omega t}, \tag{4}$$

where the mass $\mathbf{M}_\mathrm{r}$, damping $\mathbf{B}_\mathrm{r}$ and stiffness $\mathbf{K}_\mathrm{r}$ matrices were assembled accordind *Fig. 1*.

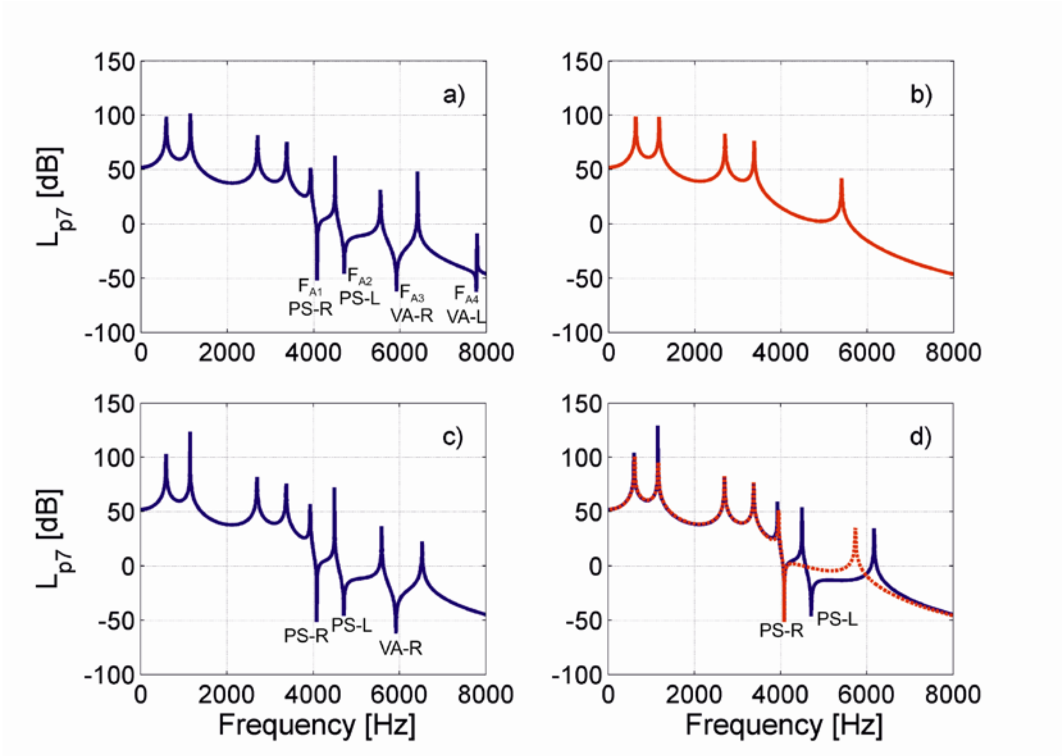

Fig. 3. Output acoustic pressures (transfer functions at m_7) of the reduced vocal tract excited at the vocal folds with a sinusoidal pressure signal (amplitude 1 Pa) computed by harmonic analysis for several variants of the model:
a) Complete model with both PS and both VA, b) model completely without PS and VA (m_8=m_9=m_{10}=m_{11}=0), c) model without the right, smaller VA (m_{11}=0), d) model without both VA (m_{10}=m_{11}=0), (blue solid line) and without both VA and the smaller PS (m_8=m_{10}=m_{11}=0), (red dotted line). The antiresonances appear as minima and the resonances as peaks in the output signal.

The resulting acoustic pressures at the lips of the reduced vocal tract are shown in *Fig. 3* for several variants of the model. For the complete model with two PS and two VA four antiresonances (A_1 – A_4) can be seen in the transfer function (*Fig. 3a*). These antiresonances disappear when neither PS nor VA are included in the model (*Fig. 3b*). Together with the antiresonances also their four paired resonances disappear (*Fig. 3b*). When one of the VA is excluded from the model only one resonance-antiresonance pair disappears (*Fig. 3c*). When both VA are excluded, two resonance-antiresonance pairs disappear (*Fig. 3d*). Comparing *Figs. 3a* and *3d* it can be deduced that the first two antiresonances in the complete model of the vocal tract in *Fig. 3a* belong to the PS and the third and fourth antiresonances correspond to the two VA. Since each of the antiresonance troughs is paired with a resonance peak, it is worth noticing how they are arranged. *Fig. 3* reveals that the antiresonancies A_1 – A_4 are close to the internal resonant frequencies of PS and VA. The three lowest antiresonance-resonance pairs (ARP), corresponding to the both PS cavities and the larger VA cavity, their resonance maxima are at lower frequencies than their corresponding antiresonance minima. For the smaller VA cavity, the situation becomes reversed - the resonance maximum occurs at higher frequency than the resonance minimum.

Summarizing the results of the analytical analysis it can be concluded, in general:

- By adding any "parallel" branch for one VA or PS to the main serial chain of the masses in the vocal tract model shown in *Fig. 1*, one antiresonance appears in the spectrum of the output pressure (i.e., the amplitude of the excited acoustic pressure at the output from the model at the antiresonant frequency is zero).
- For the complete reduced model with two PS and two VA four antiresonances are present in the output pressure signal.
- If the mass and stiffness parameters of the VA or PS are identical then the number of the antiresonances is reduced.

• No antiresonances occur in the output pressure signal radiated from the vocal tract when no parallel branches (i.e., no VA or PS) are included.

The resonance and antiresonance phenomena are different inside of the vocal tract from those detected outside of the vocal tract. This can be seen in *Fig. 4* which compares the transfer function for the acoustic pressure p_7 computed at the outlet of the simplified vocal tract (mass m_7) with the transfer functions for the pressures p_8, p_9, i.e., the acoustic pressures inside the PS. The following phenomena can be noted here:

1) The levels outside the vocal tract are generally lower than inside (notice that the curve for for the outlet signal at the lips is lower than those for the PS cavities). The only frequencies at which this is not true are at the antiresonances occurring inside the vocal tract.

2) The number of resonance peaks in the studied frequency region (0 - 8 kHz) is larger outside versus inside (9 vs. 8). There extra resonance peak is the third one in the outlet transfer function. The frequencies of the remaining eight resonances are identical outside and inside.

3) The number of the antiresonance troughs is smaller outside than inside (4 vs. 6). The frequencies of the antiresonance troughs are not constant and depend on the measurement location.

4) Relatively high levels of the pressure amplitudes in PS are at the antiresonant frequencies A_1 – A_4 in the output pressure. It is caused by the close inner resonances of the PS. Therefore, a relatively high part of the input energy is consumed by the pressure oscillations in the parallel branches that behave similarly as a Hemholtz resonator, which works as a frequency-selective sound absorber and an energy storage device (Fahy, 1989).

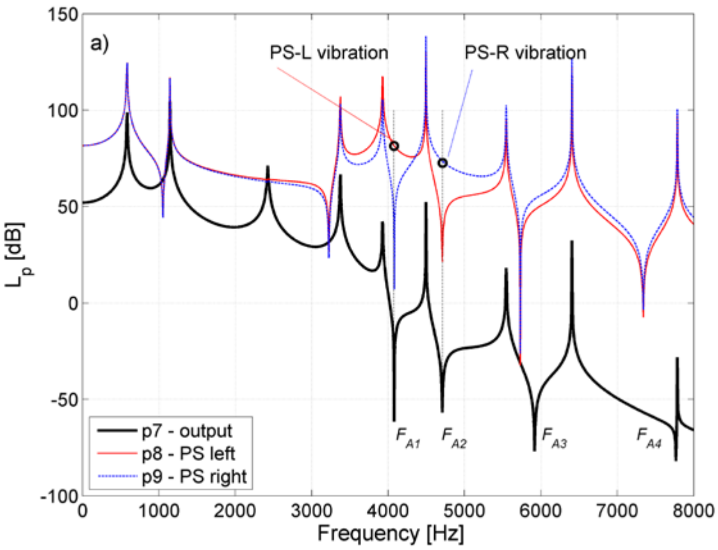

Fig. 4. Comparison of the resonance and antiresonance phenomena inside and outside of the vocal tract, calculated with the reduced model: a) sound pressure level inside the left piriform sinus (PS left) and the right piriform sinus (PS right), The sound pressure level outside (at the lips) is shown in black (thick full line). In all cases, sinusoidal pressure excitation signal with 1 Pa amplitude (i.e. c. 91 dB related to 20 μPa sound pressure level) at the vocal fold level was considered for all the frequencies. The vertical lines denote the antiresonance frequencies belonging to individual cavities inside the vocal tract. The black circles show the level of the signal inside the cavities where the corresponding antiresonances appear.

In order to see how the resonances and antiresonances can change, it is advantageous to study the changes in size of the PS cavities and the corresponding shifts of the resonance peaks and anti-resonance troughs in the outlet signal. For clarity purposes, here we considered may be exaggerated volume changes from the nominal state, i.e., -20, +40 %, for PS. The influence of the PS size on the acoustic pressure p_7 computed at the output from the simplified vocal tract is shown in *Fig. 5*. Here, the arrows show the frequency shifts of the antiresonance troughs of PS. When increasing the volumes of the PS cavities, the corresponding first two antiresonance frequencies A1 and A2 as well as the associated resonant frequencies R5 and R6 move downwards (see Fig. 5b). It is worth also to notice how the order of the peaks and troughs within their antiresonance-resonance pair (ARP) changes. Fig. 5a shows that at the nominal and smaller size of PS, their resonance frequencies are lower than their antiresonance frequencies (1st and 2nd ARP). At about +125 % of the PS volume, however, the first PS resonance frequency becomes higher than the corresponding first antiresonance frequency A1. This change occurs when the first antiresonance frequency A1 becomes lower than the 4th resonance frequency. The 3rd and 4th antiresonance troughs (A3 and A4) remain at the same frequencies, because the volumes of the corresponding VA cavities were kept the same.

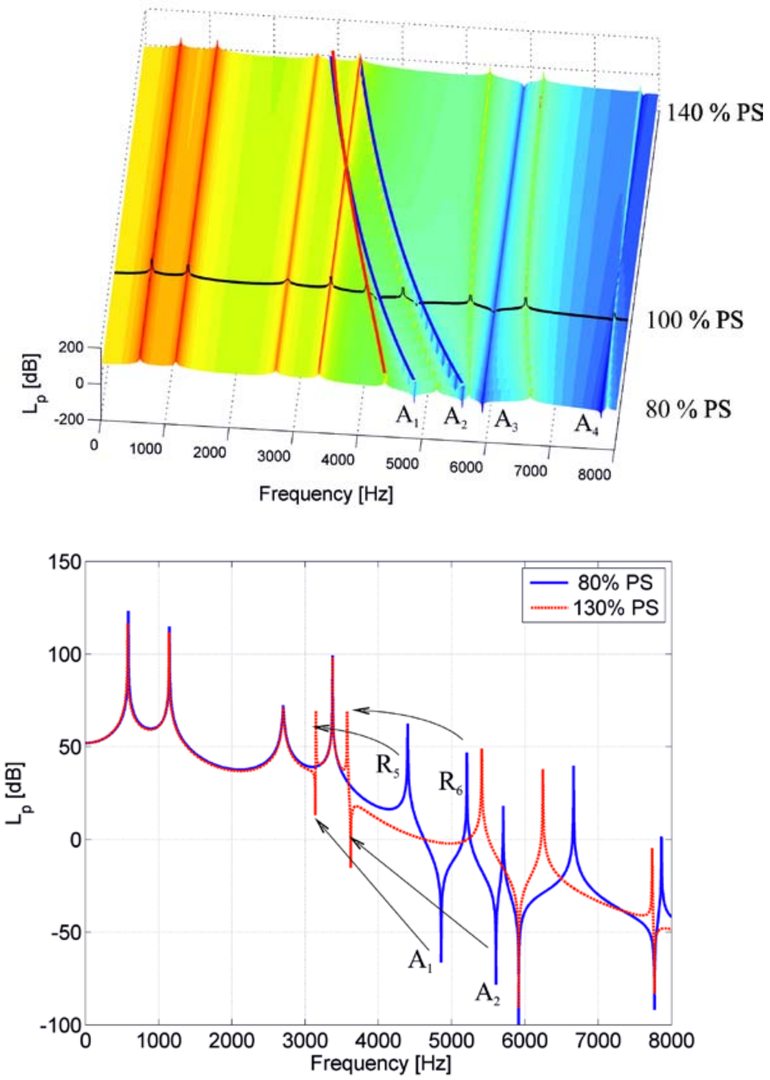

Fig. 5. Spectra of the simulated acoustic pressure in front of the lips using the reduced model of the vocal tract harmonically excited at the vocal folds for increased PS volumes from 80 to 140 % of the nominal volumes (upper panel) and two detail spectra for two chosen sizes of PS (lower panel). The shifts of the antiresonances and resonances belonging to the both PS are marked by arrows. The VA volumes were fixed at the nominal values.

Figure 6 demonstrates that an optimum size of the PS can be found when the first ARP is approaching the 4th resonance frequency and consequently when the acoustic response in front of the lips results in the highest SPL in the frequency range 3-4 kHz, where a cluster of three resonances can be created in this way.

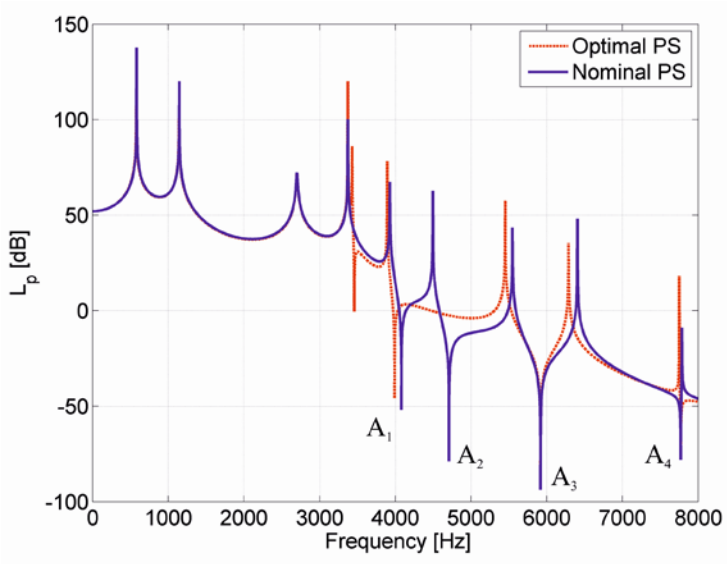

Fig. 6. Spectrum of the simulated acoustic pressure in front of the lips using the reduced model of the vocal tract harmonically excited at the vocal folds for increased PS volumes up to an optimum volumes. The VA volumes were fixed at the nominal values.

4. Conclusions

The results show that human vocal tract is a very complex resonator. Generally, the PS and VA cavities act as "side branches" of the vocal tract, i.e., dead-end acoustic paths branching out of the main acoustic path between the vocal folds and the lips. Side branches are generally known to cause antiresonances, i.e., sharp local minima in the resulting transfer function (Titze & Story, 1997). In speech research the antiresonance phenomenon is well known from the studies of nasalized vowels where the nose acts as the side branch of the vocal tract (Hattori *et al.*, 1958; Pruthi *et al.*, 2007; Vampola *et al.*, 2008b; Rong & Kuehn, 2010). In this study the nasal cavity was not considered, however, and the antiresonance phenomena caused only by the PS and VA cavities were studied.

The PS and VA cavities act as antiresonators which severely decrease the sound level radiating out of the mouth around the antiresonance frequency. Simultaneously, however, they act also as resonators which amplify the acoustic output at different frequencies. The larger the volume of the PS and VA cavities, the lower their antiresonance and resonance frequencies are. There has been some evidence that humans can actually change the size of these cavities. Considerable expansion of the PS cavities has been reported in operatic singers (Sundberg, 1974). Expansion of the VA cavities has also been observed: ca. 60% volume increase of VA was measured in a female subject after semioccluded voice exercises using a resonance tube (Vampola *et al.*, 2011a). These findings suggest that the PS and VA cavities may play a beneficial role in producing the "resonant voice". Many questions remain to be answered, however. For instance, it has not been well known how far apart the paired resonance and antiresonance frequencies of the PS and VA cavities are. A thorough observation of the *Figs*. 4-6 reveals that the frequency interval between the paired resonance peak and antiresonance trough of the PS a VA cavities can shrink as well as expand. The exact mechanism for this phenomenon remains to be clarified but the changes in the PS and VA volumes can be expected to play an important role here. Based on the results presented here we hypothesize that it is advantageous for a speaker or singer to have the resonance of each of the PS and VA cavity at a lower frequency than their corresponding antiresonance frequency, and as far apart as possible. In this case the resonance phenomenon of the PS or VA cavities can be better combined with the resonances of the main path of the vocal tract in the lower part of the spectrum, leaving the antiresonances in the upper part of the spectrum (i.e., above 4-5 kHz). Our pilot results indicate that the resonance frequency is more likely to be lower than the corresponding antiresonance frequency when the volumes of the PS cavities are expanded (recall *Fig.* 5) up to an optimal size when the radiated acoustic pressure level in the frequency range of resonance clustering around 3-4 kHz has a maximum (recall Fig. 6). In this sense, the expansion of the PS volumes would be beneficial for a "more resonant" voice production.

The rearrangement of the resonances may contribute to forming a complex singers'/speakers' formant cluster around 3-5 kHz frequency range. This is the most sensitive frequency region for the human ear, contributing largely to the perceived loudness of voice. The full 3D model simulations reveal that up to six resonance peaks can be concentrated in this frequency region (recall *Fig. 2*): this number is larger than the classic 3 resonance peaks obtained by the 1D vocal tract modeling (Titze, 2000; Story, 2006). However, since the resonance phenomena related to the PS and VA cavities are highly complex, more research is needed to confirm these pilot observations.

The piriform sinuses (PS) and the valleculae (VA) have been previously mostly connected with antiresonances in vocal spectrum which are attenuating the voice signal. Using a modeling approach, this study identified how these antiresonances are related to the PS and VA cavities and their sizes. It also identified the resonances caused by these cavities, which can amplify the resulting voice signal. The resonance-antiresonance frequencies have proven to be possible to change by changing the volumes of the PS and VA cavities. These phenomena may play an important role for producing a resonant voice: when adjusted properly, these extra resonances may join the other resonances of the vocal tract to form a complex speaker's / singer's formant cluster.

Acknowledgement

This work is supported by the Grant Agency of the Czech Republic by the project No P101/12/1306 "Biomechanical modeling of the human voice production – way to artificial vocal folds".

References

Fahy, F. J. (1989). *Sound Intensity*. Elsevier.

Hattori, S., Yamamoto, K., & Fujimura, O. (1958). Nasalization of vowels in relation to nasals. *Journal of the Acoustical Society of America*, 30, 267-274.

Honda, K., Takemoto, H., Kitamura, T., Fujita, S., & Takano, S. (2004). Exploring Human Speech Production Mechanisms by MRI. *IEICE Transactions on Information and Systems*, E87-D, 1050-1058.

Mokhtari, P., Takemoto, H., & Kitamura, T. (2008). Single-matrix formulation of a time domain acoustic model of the vocal tract with side branches. *Speech Communication*, 50, 179-190.

Motoki, K. (2002). Three-dimensional acoustic field in vocal-tract. *Acoustical Science and Technology*, 23, 207-212.

Pruthi, T., Espy-Wilson, C. Y., & Story, B. H. (2007). Simulation and analysis of nasalized vowels based on magnetic resonance imaging data. *Journal of the Acoustical Society of America*, 121, 3858-3873.

Rong, P. & Kuehn, D. P. (2010). The effect of oral articulation on the acoustic characteristics of nasalized vowels. *Journal of the Acoustical Society of America*, 127, 2543-2553.

Story, B. H. (2006). Technique for "tuning" vocal tract area functions based on acoustic sensitivity functions (L). *Journal of the Acoustical Society of America*, 119, 715-718.

Story, B. H. (1995). *Physiologically-based speech simulation using an enhanced wave-reflection model of the vocal tract.* (Ph.D. dissertation). University of Iowa.

Sundberg, J. (1974). Articulatory interpretation of the "singing formant". *Journal of the Acoustical Society of America*, 55, 838-844.

Titze, I. R. (2000). *Principles of voice production* (second printing). Iowa City, IA: National Center for Voice and Speech.

Titze, I. R. & Story, B. H. (1997). Acoustic interactions of the voice source with the lower vocal tract. *Journal of the Acoustical Society of America*, 101, 2234-2243.

Vampola, T., Horáček, J. (2012). Influence of geometric configurations of the human vocal tract on the voice production. In Náprstek, J.; Fischer, C. (ed.). *Engineering Mechanics 2012*. Prague: ITAM AS CR, v.v.i., *2012*, pp. 1475-1483. ISBN 978-80-86246-40-6.

Vampola, T., Horáček, J., Laukkanen, A.-M., & Švec, J. G. (2013). Human vocal tract resonances and the corresponding mode shapes investigated by three-dimensional finite-element modelling based on CT measurement. *Logopedics Phoniatrics Vocology*, pp.1-10, DOI: 10.3109/14015439.2013.775333.

Vampola, T., Horáček, J., & Švec, J. G. (2008). FE modeling of human vocal tract acoustics. Part I: Production of Czech vowels. *Acta Acustica United with Acustica*, 94, 433-447.

Vampola, T., Laukkanen, A.-M., Horáček, J., & Švec, J. G. (2011). Vocal tract changes caused by phonation into a tube: A case study using computer tomography and finite element modeling. *Journal of the Acoustical Society of America*, 129, 310-315.