Abstract: Finite element (FE) models of acoustic spaces corresponding to the human nasal and vocal tract for vowel /a/ are used for numerical simulations. Simplified FE model of the vocal tract for English vowel was created from geometrical data published in literature and for the Czech vowel by transferring data directly from MRI images. The nasal cavities were added to the models manually according to anatomical literature. The acoustic signal for the vowel /a/ is simulated using transient analysis of the FE models in time domain. The vocal tract is excited by time dependent displacement of a small circular plate moving at the position of the vocal folds. The time response and frequency response functions are calculated near the lips, nostrils and at the vocal folds. Effects of velopharyngeal insufficiency are simulated and compared to results from acoustic measurements.

Keywords: Biomechanics of voice, acoustic transient and modal analysis, supraglottal spaces, cleft palate

I. FINITE ELEMENT MODELS OF SUPRAGLOTTAL SPACES

In the previous papers of the authors [1,2], the acoustic frequency-modal characteristics of the human vocal tract were studied by FE modelling including the effects of cleft palate [3]. Here the study is extended to the time domain analysis using a real type of excitation of the supraglottal spaces by pulses generated at the vocal folds. The simplified FE model of a male vocal tract for the English vowel /a/ was developed according to the MRI data published by Story et al. [4]. The FE model approximating the human supraglottal tract including the added nasal cavity spaces is presented in Fig. 1. The total length of the vocal tract from the vocal folds (on the right) to the lips (on the left) is 174.58 mm. The FE model used for simulation of phonation of the Czech vowel /a/ is shown in Fig. 2a [1].

A small connection (size of 20 finite elements) of the nasal and oral cavities was considered in the back area of the soft palate modelling the velopharyngeal insufficiency. The acoustic transient analysis was realised by the system ANSYS 5.7 using the acoustic finite elements FLUID30 considering the speed of sound $c_0 = 353$ m/s and the air density $\rho_0 = 1.2$ kg/m$^3$. Zero acoustic pressure ($p = 0$) was assumed at the lips and nostrils. Other boundary walls of the acoustic spaces were considered to be acoustically absorptive.

The acoustic damping, which is associated with the fluid-structure interface on the boundary between the air and the walls (tissues) of the vocal tract, was modelled by the boundary admittance coefficient $\mu = 0.006$ for supraglottal acoustic spaces and $\mu = 0.008$ for the nasal cavity. This coefficient defined as $\mu = x/\rho_0 c_0$ is a dimensionless quantity between 0 and 1 that is equal to the ratio of the real component of the specific acoustic impedance (resistance term $x$) associated with the sound absorbing material to the fluid characteristic impedance. Another frequently used characteristic of the sound absorption of the material is the dimensionless absorption coefficient $\alpha$, which is related to the boundary admittance coefficient $\mu$ as

$$\alpha = \left[0.5 + 0.25(\mu + 1/\mu)\right]^{-1}.$$ 

The pulse excitation of the supraglottal spaces was realised by a small rigid circular plate (a piston) translating in the axial direction along the axes $z$. The plate was situated in the position of the vocal folds, and its diameter was equal to $1/3$ of the diameter of the cross-section area of the FE model of the acoustic space at this point (see the detail in Fig. 2b). The translation motion of the plate in time was given by integration of the shape of volume velocity that approximately corresponds to the airflow through the vocal folds (see Fig. 3). Five subsequent excitation pulses with the period corresponding to the fundamental (pitch) frequency $F_0=100$ Hz were considered in the transient

**Fig. 1** FE model of male vocal tract for English vowel /a/ including the nasal cavity.
analysis. The interaction between the plate and the acoustic space was realised by the interactive acoustic finite elements.

In the model of the English vowel /a/ the effect of hard palate compliancy was included in the study. The material properties of the bone were assumed as follows: Young modulus $E_1 = 6.50 \times 10^9$ Pa, Poisson ratio $\mu_1 = 0.21$, density $\rho_1 = 1.41 \times 10^3$ kg/m$^3$ and wall thickness $h = 0.6$ mm. The bone of hard palate was modeled using two separated parts. The first (lower) part of the finite elements SHELL63 was directly joined with the acoustic finite elements of the vocal tract using the material properties $E_1$, $\mu_1$ and $\rho_1$. The second (upper) part of the finite elements SHELL63 was joined with the acoustic finite elements of the nasal tract on its lower boundary area. The material properties corresponding to the second part of the FE model of the bone were identical with the first part of the bone model except the Young modulus $E_2 = 0.01 E_1$ respecting a much more compliant material. Each node of the lower part of the bone was connected with the corresponding node at the upper part of the hard palate FE model. This connection of corresponding nodes guarantees identical motion of the nodes in both parts of the FE model.

II. MATHEMATICAL FORMULATION

Wave equation for the acoustic pressure can be written as:

$$\nabla^2 p = \frac{1}{c_o^2} \frac{\partial^2 p}{\partial t^2},$$  

(1)

where $c_o$ is the speed of sound, with the possible boundary conditions as follows

- on acoustically hard area and at the open end

$$\frac{\partial p}{\partial n} = 0, \quad p = 0,$$  

(2)

- between the flexible structure and the fluid elements:

$$\frac{\partial p}{\partial n} = -\rho_s \frac{\partial^2 w_s}{\partial t^2},$$  

(3)

where $n$ is normal to the boundary area and $w_s$ is the displacement of the structure in the normal direction to the vibrating surface.

Equations of motion for the elasto-acoustic system after discretization can be written as

$$\begin{bmatrix} M_s & 0 \\ \rho_s R' & M_f \end{bmatrix} \ddot{u} + \begin{bmatrix} B_s & 0 \\ 0 & B_f \end{bmatrix} \dot{u} + \begin{bmatrix} K_s & -R \end{bmatrix} u = 0$$  

(4)

where $M$, $B$, $K$ are the global mass, damping and stiffness matrices, $P$ is the vector of nodal acoustic pressures, subscripts $s$ or $f$ denote the structure or fluid, $u$ is the structural displacement, $R$ is the coupling matrix and $\rho_0$ is the air density.

For the special case of kinematic excitation by the moving structure the following equations for the pressure describe the air vibration

$$RP = M \ddot{u} + K u$$

$$M \ddot{P} + B \dot{P} + K_P P = \rho_s R' \ddot{u},$$  

(5)

where the structural motion $u(t)$ is prescribed. The Newmark method of solution in time was used.

III. NUMERICAL RESULTS

The results of the transient dynamic analysis of the FE models are the time responses of the acoustic pressure in selected points of supraglottal spaces near the vocal folds, the lips and the nostrils. The spectra of the exciting acoustic pressure pulses and the pressure time responses were calculated by MATLAB using FFT.
Fig. 3 presents excitation pulses of the airflow velocity through the glottis from where the corresponding displacement of the rigid plate was calculated and afterwards used for excitation of the vocal tract in the time domain.

The results of transient analysis of the FE models for English vowel /a/ are presented in frequency domain in Fig. 4 showing the calculated acoustic pressure near the nostrils. The formant frequencies $F_1=823$ Hz, $F_2=1164$ Hz and $F_3=2826$ Hz calculated by modal analysis of the FE model can be detected in the spectrum. A nasopharynx (oro-nasal) resonant frequency $f_{naso}=2143$ Hz is embodied in the frequency response function between the formants $F_2$ and $F_3$.

The results of the transient analysis of the FE model for the Czech vowel /a/ are presented in Figs. 5 and 6 showing the spectra of the acoustic pressure calculated near the vocal folds and lips. The pressure levels near the vocal folds are much higher than the acoustic pressure near the lips. The formant frequencies $F_1=623$ Hz, $F_2=890$ Hz and $F_3=2935$ Hz can be found in the frequency response functions in Fig. 6. These formant frequencies are in good agreement with the data known from the Czech literature [6] as well as with calculations by modal analysis for the same FE models – see, e.g. [1,3]. Another resonant frequency $f_{naso}=1707$ Hz appears in the Fig. 6 due to the velopharyngeal insufficiency.

The differences between the Czech and English formants originate mainly in the fact that two very different types of the FE models were used, however, the results obtained are in a range of variability of the vocal /a/ production.

IV. EXPERIMENTAL VERIFICATION OF THE MODEL

The first experiment took into account the well-known phenomenon of connection vowel - nasal consonant - vowel. The passageway between the oral and nasal cavities of the first vowel is closed or almost closed in the Czech language for a clear sound to be pronounced. The velopharyngeal passageway must be opened when producing nasal consonants.

The vowels following the nasal consonants are nasalized because the passageway is still not closed. The differences in the velopharyngeal opening between the first and the second vowel should result in changes of the formant frequencies. Five normal subjects were asked to pronounce the interconnection /ama/ and the changes of the formants between the two vowels were studied.

The nasal and oral signals were picked up by microphones of the headset part of Nasometer 6200-3 (Kay Elemetrics Corp.) and analysed by Multi-Speech (Kay Elemetrics Corp.) programme.
V. EXPERIMENTAL RESULTS

Examples of results from the practical experiments are shown in Fig. 7. The spectrogram of the interconnection /ama/, where the second vowel is more nasal than the first one, is shown in Fig. 7a. The signal was picked up in front of the nose. The position of the formants F1=800 Hz and F2=1100 Hz and F3= 3700 Hz is stable. The position of the oro-nasal formant changes from \( f_{\text{naso}} \approx 2600 \) Hz to 2950 Hz as approximately predicted by the FE models. Effects of increasing the cleft area of the hard palate were theoretically studied in detail in previous publications [2,3].

The continual changes of the soft palate closing for vowel /a/ are demonstrated in Fig. 7b. The signal was picked up in front of the mouth and the measurement started from the soft palate opening. The formants F1=680 Hz, F2=1100 Hz, F3=3950 Hz remain practically unchanged. The oro-nasal formant changes its position from \( f_{\text{naso}} \approx 2700 \) Hz to 2350 Hz.

The second nasalized vowel /a/ in the interconnection /ama/ corresponded to an opening of the soft palate and simulated a velofaryngeal insufficiency.

VI. CONCLUSION

The transfer functions were obtained as the results of the transient analysis of the FE models of the vocal tract. The models were excited by a transient translation of a small rigid plate situated in the area of vocal folds and driven by a time signal which shape in the time domain approximately corresponds to a volume velocity of the air flowing through the vocal folds during phonation. The formant frequencies F1 – F3 evaluated from the resonances of the calculated frequency response functions for the pressure are in good agreement with the experimental data known for the formants from the literature [4-6] as well as with the results of the modal analysis performed [1-3]. The existence of calculated oro-nasal formants was verified by the measurements when the velofaryngeal insufficiency was simulated by the normal subjects.

ACKNOWLEDGEMENTS

The authors are very grateful to Doc. MUDr. Petr Krupa from the Hospital U svaté Anny in Brno for making possible the special measurements and providing the original MRI data sets for human vocal tract during phonation of the Czech vowels.

The research is supported by the Grant Agency of the Czech Republic by the project No 106/98/K019 “Mathematical – Physical Modeling of Vibroacoustic Systems Important on Biomechanics of Voice and Hearing.”

REFERENCES