This is the post print version of the article, which has been published in Journal of Sound and Vibration. 2016, 366, 556–570. http://dx.doi.org/10.1016/j.jsv.2015.12.018

MEASUREMENT AND MATHEMATICAL SIMULATION OF ACOUSTIC CHARACTERISTICS OF AN ARTIFICIALLY LENGTHENED VOCAL TRACT

Vojtěch Radolf [a], Jaromír Horáček [b], Pavel Dlask [c], Zdeněk Otčenášek [d], Ahmed Geneid [e], Anne-Maria Laukkanen [f]

[a] Vojtěch Radolf, Institute of Thermomechanics AS CR, v. v. i., Dolejškova 1402/5, 182 00 Prague, Czech Republic, radolf@it.cas.cz, Tel. +420 26605 3121 **corresponding author**

[b] Jaromír Horáček, Institute of Thermomechanics AS CR, v. v. i., Dolejškova 1402/5, 182 00 Prague, Czech Republic, jaromirh@it.cas.cz

[c] Pavel Dlask, Musical Acoustic Research Center, HAMU, Malostranské nám. 13, 118 01 Prague, Czech Republic, pavel.dlask@hamu.cz

[d] Zdeněk Otčenášek, Musical Acoustic Research Center, HAMU, Malostranské nám. 13, 118 01 Prague, Czech Republic, zdenek.otcenasek@hamu.cz

[e] Ahmed Geneid, Department of Otolaryngology and Phoniatrics - Head and neck surgery, Helsinki University Hospital and University of Helsinki, PL 220, 00029, HUS, Helsinki, Finland, Ahmed.Geneid@hus.fi

[f] Anne-Maria Laukkanen, Speech and Voice Research Laboratory, School of Education, Åkerlundink. 5, University of Tampere, 33100 Tampere, Finland, Anne-Maria.Laukkanen@uta.fi

Abstract: *Phonation into tubes is used for voice training and therapy. In the present study, the formant frequencies were estimated from measurements of the acoustic pressure and the acoustic input impedance for a plexiglass model of the vocal tract (VT) prolonged by a glass tube. Similar transfer function measurements were performed with a human VT in vivo. The experimental results matched the mathematical modelling and confirmed the legitimacy of assuming rigid walls in mathematical simulations of the acoustic characteristics of an artificial VT model prolonged by a tube. However, this study also proved a considerable influence from soft tissues in the yielding walls of human VT cavities on the first formant frequency,* F_1*. The measured* F_1 *for the VT model corresponded to the computed value of 78 Hz. The experiments in a human instead resulted in a much higher value of* F_1*: about 200 Hz. The results confirm that a VT model with yielding walls must be considered for mathematical modelling of the occluded or semi-occluded human vocal tract, e.g. prolonged by tubes or straws. This is explained by an acoustic-structural interaction of the vocal tract cavities with a mechanical low-frequency resonance of the soft tissue in the larynx.*

Voice training and therapy often involve semi-occlusions, i.e. sounds during which the vocal tract is either narrowed, as in the production of closed vowels [i, y, u]; constricted, as in voiced fricatives, e.g. [β, z], or nasals, e.g. [m, n]; completely closed in a rapid succession, as in a lip or tongue trill; or artificially lengthened, as when one phonates into a tube or a straw; e.g. see [4]. Phonation into a small diameter hard-walled “resonance tube” has been used especially in Scandinavia; see [5]. For voice therapy purposes, the outer end of this tube is placed in water (“water resistance therapy”), although, in voice training of normally voiced subjects, the outer end of the tube is kept in the air. During the semi-occlusion exercises, the trainee/patient is instructed to aim at ease of phonation and maximal vibratory sensations in the front part of the vocal tract and the facial structures. The goal of the semi-occlusion exercises is an effortlessly produced, well-resonating voice; see [4, 6], for example.

One reason for the use of semi-occlusions in voice training and therapy lies in the fact that they slow down the airflow in the vocal tract and increase the supraglottal airpressure. Consequently, the intraglottal air pressure also increases. This reduces the vocal fold collision pressure and, hence, the mechanical load related to voice production. Aiming at maximal vibratory sensations in the vocal tract is supposed to help in optimizing the glottal and vocal tract setting for maximal sound energy transfer from the glottis to the outer space. Increased supraglottal air pressure aids the trainee in this optimization as it intensifies the sensations of resonatory vibrations; see [7, 8]. Furthermore, semi-occlusions lower the formant frequencies, especially F_1 [9]. If F_1 comes sufficiently close to f_0 of phonation, the inertive reactance of the vocal tract may mechanically assist vocal fold vibration by reducing the phonation threshold pressure, i.e. the lowest subglottic pressure needed to start and maintain vocal fold vibration; e.g. see [10].

The effect of a particular semi-occlusion is, thus, related to two things: the amount of flow resistance it offers and the amount of inertive reactance within the fundamental frequency range of phonation. Obviously, the narrower a constriction in the vocal tract is, the longer and narrower the tube one phonates into, or the deeper in water the outer end of the tube is inserted, the higher the flow resistance will be. The amount of inertive reactance, in turn, is related to the resonances that will be created in the vocal tract of a subject phonating into a particular tube. To the best of our knowledge, no information on formant frequencies measured directly during phonation into tubes can be found in the literature. Such information would be needed in order to shed further light on the basis for the exercise and to be able to choose the best suited tube for the trainee/patient.

where p and U are the acoustic pressure and volume velocity respectively, $\mathbf{T}_{\mathrm{VT}}$ is a transfer matrix of the vocal tract, $\mathbf{T}_{\mathrm{TB}}$ is a transfer matrix of the tube, and the indices G, L, T respectively mean the position of glottis, lips and open tube end.

The volume velocity is zero ($U_G = 0$) at the closed VT end at the glottis position, and thus we can eliminate U_L from Eq. (1):

$$U_L = C_{VT} / A_{VT} \cdot p_L, \tag{3}$$

and putting Eq. (3) into Eq. (2) yields the transfer function for the pressure between the input and the end of the tube:

$$p_L(\omega) / p_T(\omega) = A_{VT} / \left(A_{TB} A_{VT} + B_{TB} C_{VT} \right). \tag{4}$$

Generally, the transfer function U_T / U_G of the complete system (VT+tube) between the acoustic volume velocity at the glottis position and the tube end can be expressed as (see Eq. (A.13) in Appendix A):

$$U_T(\omega) / U_G(\omega) = 1 / \left(A_{TB} A_{VT} + B_{TB} C_{VT} \right). \tag{5}$$

This equation is valid regardless of the boundary conditions at the glottis when the radiation impedance $Z_{T,RAD}$ at the tube output, defined as the ratio of output pressure to output volume velocity, is omitted. Because the denominator in Eq. (4) is the same as in Eq. (5), resonance frequencies (peaks) of the measured transfer function $p_L(\omega) / p_T(\omega)$ should be approximately equal to the peaks of the transfer function $U_T(\omega) / U_G(\omega)$ of the complete system.

The numerator $A_{VT}(\omega)$ in Eq. (4) is the same as a denominator in the transfer function U_L / U_G of the vocal tract alone (see Eq. (A.16) in Appendix A):

$$U_L(\omega) / U_G(\omega) = 1 / A_{VT} \, . \tag{6}$$

Again, when omitting the radiation impedance $Z_{L,RAD}$ at the lips, Eq. (6) is generally valid regardless of the boundary conditions at the glottis. Antiresonance frequencies (spectral dips) of the transfer function (4) then should be very close to the peaks of the transfer function (6) of the vocal tract without a tube.

2.3 Measurement of the acoustic input impedance

2.3.1 Experimental setup

A measurement of the acoustic input impedance was only carried out on the plexiglass model of the vocal tract using the Brass Instrument Analysis System (BIAS 6) controlled by a PC. A layout of the measuring head of the system is shown in Fig. 4a below. The input of the VT

2.3.2 Processing and interpretation of experimental data from input impedance measurements

The VT input impedance Z_{IN}, calculated as the ratio of the input pressure to the input acoustic volume velocity, is given by the expression (see Eq. (A.17) in Appendix A):

$$Z_{IN} = \frac{p_G}{U_G} = \frac{D \cdot Z_{T,RAD} - B}{A - C \cdot Z_{T,RAD}}, \tag{7}$$

where A, B, C, D are the components of the transfer matrix of the complete system (VT+tube) and $Z_{T,RAD}$ is the radiation impedance at the open end of the tube. The transfer matrix of the complete system relates the input variables with the output variables by the equation (see Eqs. (A.9 and A.10) in Appendix A):

$$\begin{bmatrix} p_T \\ U_T \end{bmatrix} = \mathbf{T_{TB}} \cdot \mathbf{T_{VT}} \cdot \begin{bmatrix} p_G \\ U_G \end{bmatrix} = \begin{bmatrix} A & B \\ C & D \end{bmatrix} \cdot \begin{bmatrix} p_G \\ U_G \end{bmatrix}. \tag{8}$$

Let us consider the situation that the input is closed by a rigid wall ($U_G = 0$) and the output loaded by the impedance $Z_{T,RAD}$. After substituting these boundary conditions, denoted as C-OZ_{RAD} (closed input, open loaded output) in Eq. (8), it can easily be derived that the denominator of the input impedance in Expression (7) is equal to the frequency equation of the complete system (VT+tube) with the boundary conditions C-OZ_{RAD}. Thus the peaks (resonance frequencies) of $|Z_{IN}(\omega)|$ represent resonances or natural acoustic frequencies of the system with boundary conditions C-OZ_{RAD}. These resonances are identical to resonance frequencies obtainable from a commonly used transfer function (A.13) of the complete system.

Considering the input is open ($p_G = 0$), and the output is loaded by the impedance $Z_{T,RAD}$, the substitution of these boundary conditions denoted as O-OZ_{RAD} (open input, open loaded output) in Eq. (8) indicates that the numerator of Eq. (7) is the same as the frequency equation of the system with the O-OZ_{RAD} boundary conditions. It means that the spectral dips (antiresonances) of $|Z_{IN}(\omega)|$ represent natural acoustic frequencies of the complete system with the boundary conditions O-OZ_{RAD}.

2.4 Input parameters of mathematical models

A vocal tract channel with yielding walls was modelled according to [9, 12] as a system of 43 cylindrical elements of 4 mm in length except for the first and last element, whose length was 2 mm, giving the total length of the present subject's vocal tract of 16.8 cm measured during previous experiments [17]. The following formant frequencies estimated from the acoustic recording of the vowel [u:] for the female subject (without a tube) were prescribed to the VT mathematical model: F_1=367 Hz, F_2=730 Hz, F_3=2518 Hz, F_4=3735 Hz (see [17]). Through a tuning procedure [14] carried out by changing the vocal tract shape, i.e. the area of the cylindrical cross-sections, the best-fitting vocal tract configuration was obtained. When connecting the tube to the vocal tract, the diameter of the last element of the VT model was increased to 10 mm (the outer diameter of the tube) and the penultimate element's diameter was recalculated as an average of the two of its neighbouring elements. The resulting model is shown in Fig. 5.

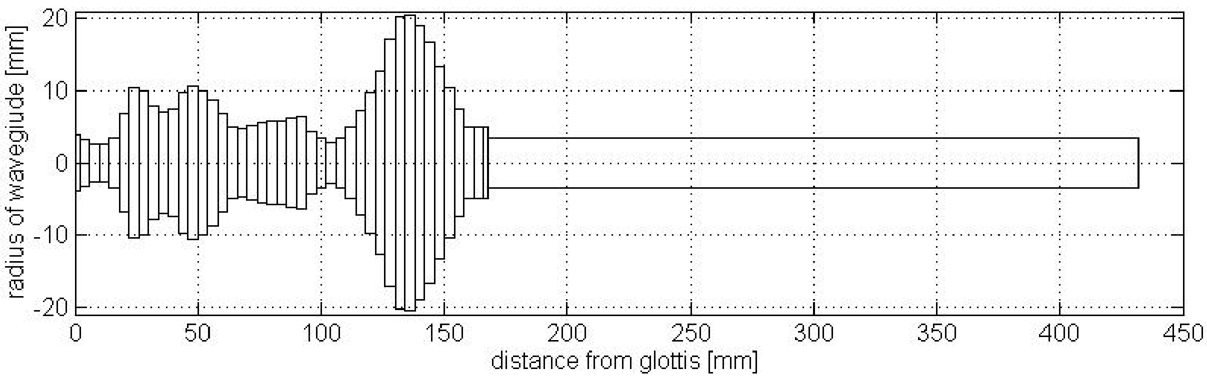
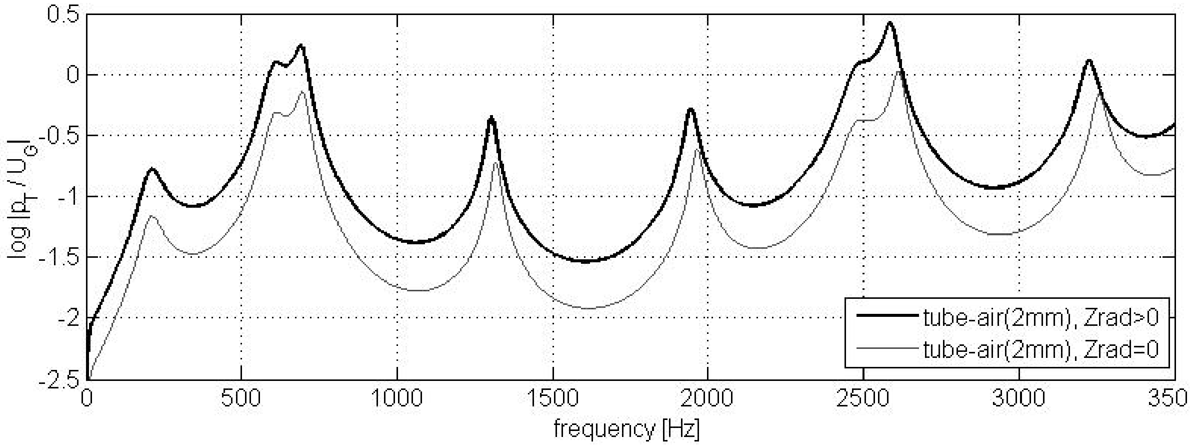


Fig. 5. Geometry of the VT model prolonged by the tube.

A mathematical model of the vocal tract acoustics with rigid walls is described in Appendix A. Its length as well as the length of the plexiglass VT model was 18.9 cm. The tube was modelled as one element with rigid walls.

The following values of fluid density, dynamic air viscosity, and speed of sound were used as input for the models:
ρ_0 = 1.2 kgm^{-3}, μ = 1.8·10^{-5} kgm^{-1}s^{-1}, c_0 = 343 ms^{-1} (considering the temperature in the laboratory) for the artificial VT and c_0 = 353 ms^{-1} for the human VT.

a)



b)

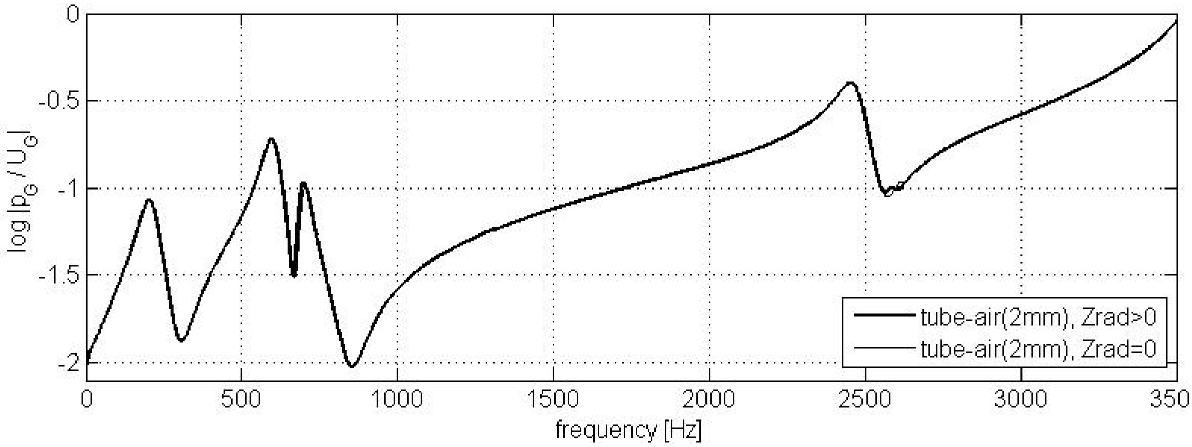


Fig.7. Results of mathematical modelling of the human vocal tract: (a) the magnitude of the transfer function p_T/U_G between the glottis position and the cross-section of the tube, located 2 mm from its open outlet, both considering the radiation impedance $Z_{T,RAD}$ (thick line) and omitting it (thin line) and (b) the magnitude of the acoustic input impedance p_G/U_G considering the radiation impedance $Z_{T,RAD}$ (thick line) and omitting it (thin line).

The resonance frequencies obtained by mathematical modelling were F_1=208 Hz, F_2=602 Hz, F_3=694 Hz, F_4=1304 Hz, F_5=1945 Hz, F_6=2473 Hz, F_7=2588 Hz and F_8=3223 Hz. These values correspond well to the resonances measured (see thick lines in Figs. 6 and 7a), except for the first resonance frequency, F_1, as mentioned above. If F_1 was really 208 Hz, it would suggest that phonation into the resonance tube is assisted by reactance of the vocal tract at fundamental frequencies generally used in speech (i.e. f_0 below 200 Hz).

Figure 7b shows the computed results for the input impedance of the human vocal tract. The three lowest resonances are clearly visible: F_1=208 Hz, F_2=602 Hz and F_3=694 Hz; however, the resonances F_4, F_5, F_7 and F_8 are not identifiable in Fig. 7b. The reason consists

the transfer function near the zero frequency can be caused by a more prominent acoustic structural interaction of the vocal tract with vibrating yielding walls.

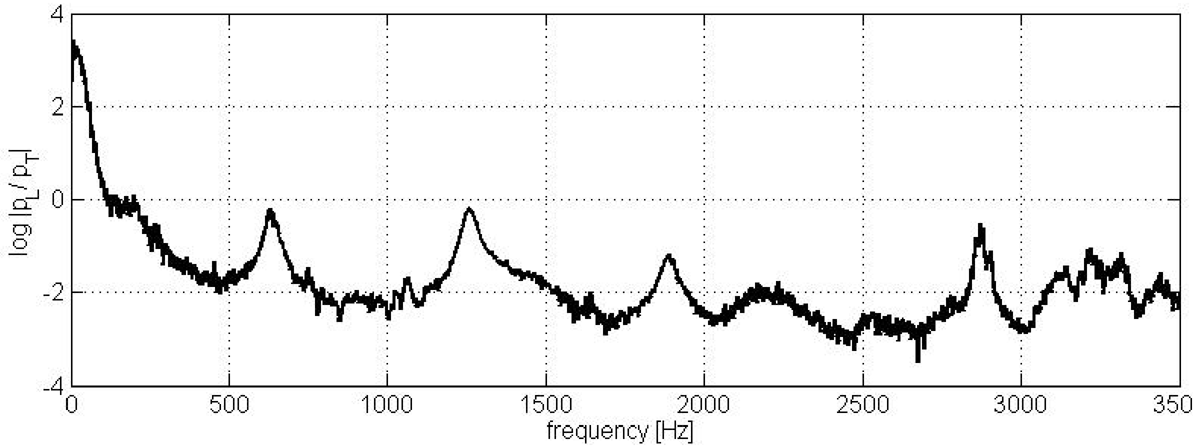


Fig. 9. A measured transfer function between the lips and the open tube end for the human vocal tract prolonged by the tube and excited by random noise from the loudspeaker.

The measured and the computed magnitudes of the transfer function $|p_L(\omega)/p_T(\omega)|$ for the artificial VT model prolonged by the tube can be seen in Fig. 10. The computation was performed according to Eq. (4) assuming hard walls of the vocal tract. First resonance peaks occur at frequencies 73 Hz and 78 Hz for the measured and computed results, respectively. The first two antiresonances, which according to Eq. (6) are the resonance frequencies of the vocal tract alone (without a tube), can be well recognized in the spectra at about 400-500 Hz and 900-1000 Hz.

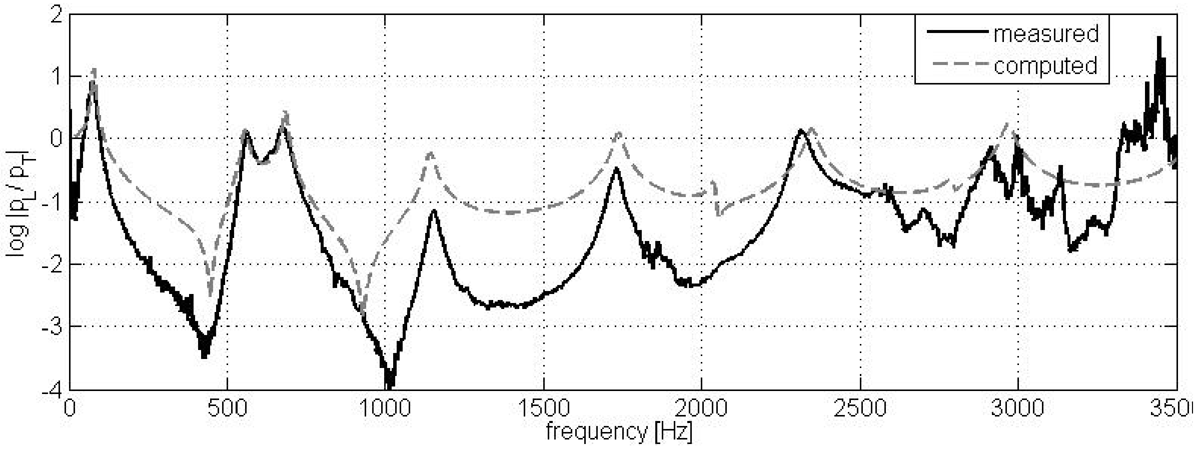


Fig. 10. A comparison of measured (solid line) and computed (dashed line) transfer functions p_L/p_T between the lips and the open tube end for the plexiglass vocal tract

prolonged by the tube. Resonances were excited by random noise using the external loudspeaker, see Fig. 1.

Fig. 11 shows the measured and the computed magnitudes of the acoustic input impedance for the artificial VT model prolonged by the tube. The computation was performed according to Eq. (7), assuming hard walls of the vocal tract. The radiation impedance, $Z_{T,RAD}$, at the tube output was omitted in this mathematical model. The first resonance peak obtained from the measurement occurs at the frequency of F_1 = 99 Hz, whereas the computed value F_1 = 78 Hz is equal to the first resonance frequency of the computed transfer function (see Fig. 10), as expected by comparing Eqs. (4) and (7).

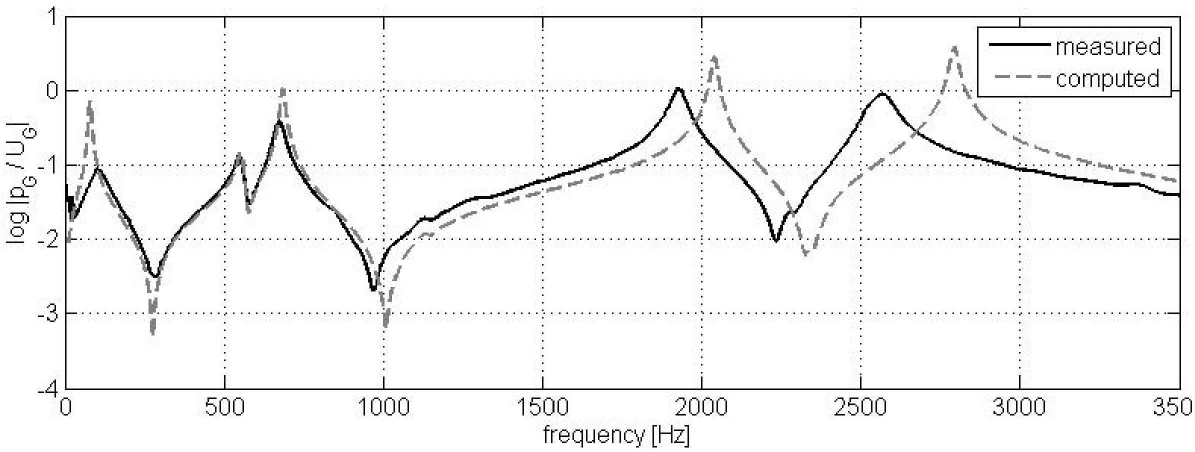


Fig. 11. A comparison of measured (solid line) and computed (dashed line) magnitudes of the input impedance p_G/U_G for the artificial plexiglass vocal tract model prolonged by the tube.

3.3 Summary of the results

Tables 1 and 2 summarize the measured and computed resonance frequencies and their relative differences both for the transfer function and the input impedance of the artificial VT. Similarly, Table 3 compares the measured and computed acoustic resonance frequencies obtained for the human vocal tract. The values marked by “x” could not be identified in the measurements of the human VT. Table 4 compares formant bandwidths obtained from the measured transfer functions (TF) $|p_L(\omega)/p_T(\omega)|$ of the artificial vocal tract and of the human vocal tract, both prolonged by the tube. The formant bandwidth was measured as the frequency difference between the two points 3 dB below the peak level of the solid lines in Figs. 9 and 10. The formant bandwidth values were not identifiable in the case of F_1 (190 Hz) for the human vocal tract (Fig. 9) and in the case of the two close formants (resonances) measured in the model (560 and 672 Hz; see Fig. 10).

	F_1	F_2	F_3	F_4	F_5	F_6
VT model: measured TF	73	560	672	1153	1729	2313
VT model: computed TF	78	555	684	1142	1735	2347
Difference (%)	6.8	-0.9	1.8	-1.0	0.3	1.5

Table 1. Measured and computed resonance frequencies (in Hz) evaluated from the transfer functions (TF) of the artificial vocal tract prolonged by the resonance tube.

	F_1	F_2	F_3	F_4	F_5
VT model: measured Z_{IN}	99	546	673	1926	2564
VT model: computed Z_{IN}	78	555	684	2039	2792
Difference (%)	-21.2	1.6	1.6	5.9	8.9

Table 2. Measured and computed resonance frequencies (in Hz) evaluated from the input impedance Z_{IN} of the artificial vocal tract prolonged by the resonance tube.

	F_1	F_2	F_3	F_4	F_5	F_6
Human VT: phonation	x	680	x	1280	1930	2530
Human VT: measured TF	190	630	x	1262	1889	x
Difference (%)	x	-7.4	x	-1.4	-2.1	x
Human VT: computed TF	208	602	694	1304	1945	2473

Table 3. Formant frequencies (in Hz) obtained from the acoustic spectra of phonation into the resonance tube and from the measured and computed transfer functions (TF) of the human vocal tract prolonged by the tube (x: unidentified values).

	B_1	B_2	B_3	B_4	B_5
VT model: measured TF	12	17	26	21	21
Human VT: measured TF	x	16	x	22	23
Difference (%)	x	-5.9	x	4.8	9.5

Table 4. Formant bandwidths (in Hz) obtained from the measured transfer functions of artificial and human vocal tracts prolonged by the resonance tube (x: unidentified values).

the human vocal tract (see Figs. 6 and 9). This discrepancy between the results from modelling and those obtained from the human may stem in the fact that the frequency resolution of the mathematical models was quite high, given by the frequency step of 1 Hz. Another reason could be a higher damping of the acoustic modes in the human vocal tract due to the interaction of acoustic waves with soft tissues of the VT walls.

A potential error in the input impedance introduced by omitting the radiation impedance $Z_{T,RAD}$ at the tube output is insignificant. This is demonstrated in Fig. 7b, where there is practically no difference between the curves for $Z_{T,RAD} > 0$ and $Z_{T,RAD} = 0$. Even though the transfer function is affected by the omission of $Z_{T,RAD}$, the effect mainly concerns the level of the curve, and just minor shifts in the resonance frequencies can be detected (see Fig. 7a).

The comparison between Figs. 10 and 11 shows that the first three clearly visible resonance frequencies of the artificial vocal tract that resulted from the measured and computed transfer functions fit well with the corresponding resonances that were detectable in the input impedances. However, the higher peaks in the computed Z_{IN} spectrum above 1900 Hz in Fig. 11 were suppressed in the computed transfer function curve by adjacent antiresonances around 2000 Hz and 2800 Hz; see the dashed line in Fig. 10.

Similarly, resonances around 1100 Hz and 1700 Hz that can be seen in Fig. 10 practically vanished in the Z_{IN} spectrum (see Fig. 11) as a result of very tightly adjacent antiresonances. These antiresonances that are around 275 Hz, 574 Hz, 1008 Hz (cf. Fig. 11), 1141 Hz, 1730 Hz, 2326 Hz and 2348 Hz could be computed exactly from the frequency equation given by the numerator of Eq. (7), which is equal to zero as explained in section 2.3.2. These values represent natural frequencies of the system (VT+tube) when the glottis is not closed ($p_G = 0$).

We should note that the frequency range studied here (0-3.5 kHz) fits the domain of validity of 1D mathematical model of the VT for the vowel [u:]. As the maximum width of the mouth cavity was 62 mm for the artificial VT, and the maximum diameter of the mathematical model of the human VT was 41 mm, it can be calculated that the first transversal acoustic mode could exist above 2.77-2.85 kHz.

Since the estimated formant bandwidths B_2, B_4, and B_5 for the formant frequencies F_2 to F_5 are practically the same for the VT model and for the human VT (the difference is less than 10%: see Table 4), it can be concluded that the compliance of the human walls for higher formant frequencies does not influence the acoustic damping of the human VT cavity prolonged by the tube. The values of the measured bandwidths appear to be narrower than the values found in the literature for ordinary phonation without a tube: e.g. see [18]. An explanation can be found in the different methods of measurement. Here we used the transfer function measured between the lips and the open tube end, whereas the spectrum envelope is usually used as it is described, e.g. in [18]. We should also note that the glass tube used in this study has less damping than the human vocal tract itself and, depending on the acoustic mode shape, it can influence the resulting damping of the coupled system (VT+tube).

The computed results shown in Fig. 12 suggest that, for phonation with a resonance tube in the air, F_1 does not go lower than roughly 200 Hz. Thus, it would offer the beneficial effect of increased reactance below this level, i.e. at the fundamental frequencies typically found in speech. To optimize the effect of vocal tract reactance on phonation, males would probably need to use longer or narrower tubes, as suggested by the Story's results [9]. When the tube is long or narrow enough, the effective mechanism eventually changes from acoustic reactance to the magnitude of flow resistance.

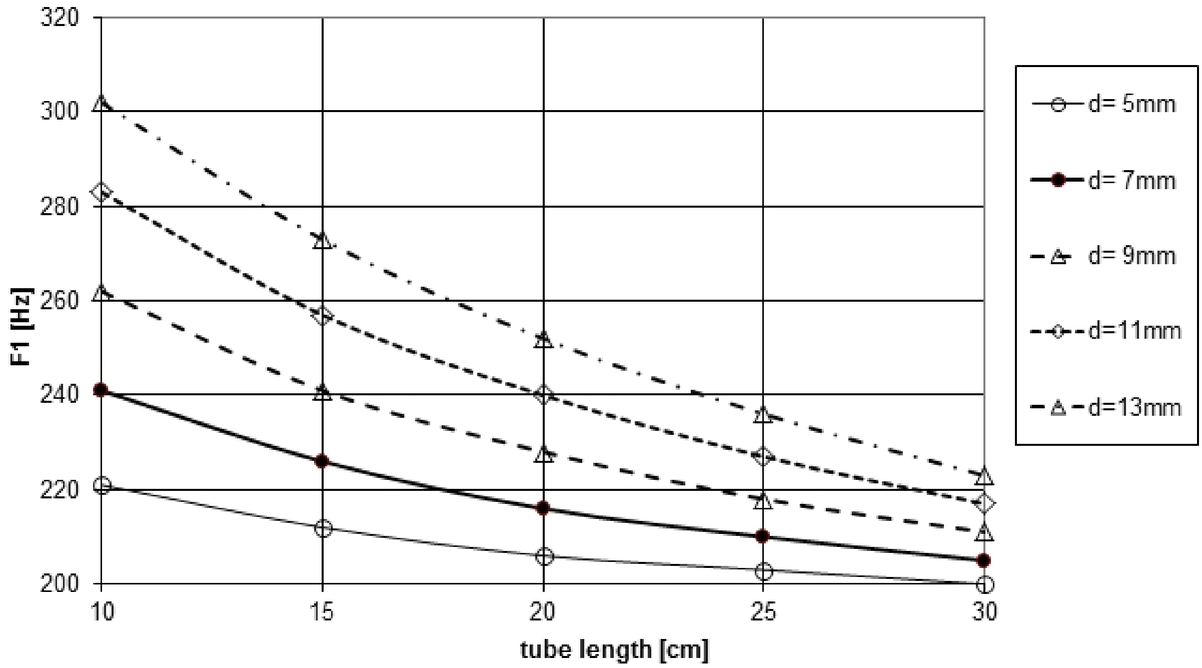


Fig. 12. First resonance frequency of the human VT prolonged by tubes of different lengths and diameters (*d*), calculated according to Story [9, 12].

4.2 Physical background for the influence of yielding walls on the acoustic resonances of the human vocal tract

Following our previous studies of acoustic structural interaction (see, for example, [19 and 20]), we can explain the influence of yielding walls in the vocal tract cavity on the acoustic frequencies by using a simplified model of a coupled mechanical-acoustical system. Similarly, Rabiner and Schaffer [21] considered the effect of yielding walls on acoustical resonances in a simple cylindrical cavity (compliant shell) simulating the vocal tract. However, they did not include the mechanical resonance of the structure in their model.

Let us consider the coupled system shown in Fig. 13, consisting of a simplified vocal tract cavity (1) and a tube (2) with cross-section areas S_1 and S_2 and lengths L_1, L_2. The glottis is closed by a yielding wall having a mass M and vibrating with a displacement $w(t)$ on a spring of stiffness K. The resonance frequencies of such a coupled system are given in the solution of the following transcendent frequency equation derived in Appendix B:

$$\omega^2 + \frac{\omega\rho_0 c_0 S_1}{M} \cdot \frac{S_2 \sin(kL_1)\cos(kL_2) + S_1 \cos(kL_1)\sin(kL_2)}{S_2 \cos(kL_1)\cos(kL_2) - S_1 \sin(kL_1)\sin(kL_2)} - \omega_0^2 = 0\,, \qquad (9)$$

where $k = \omega/c_0$ is the wave number, ω is the angular frequency and $\omega_0 = \sqrt{K/M}$ is the eigenfrequency of the mechanical system.

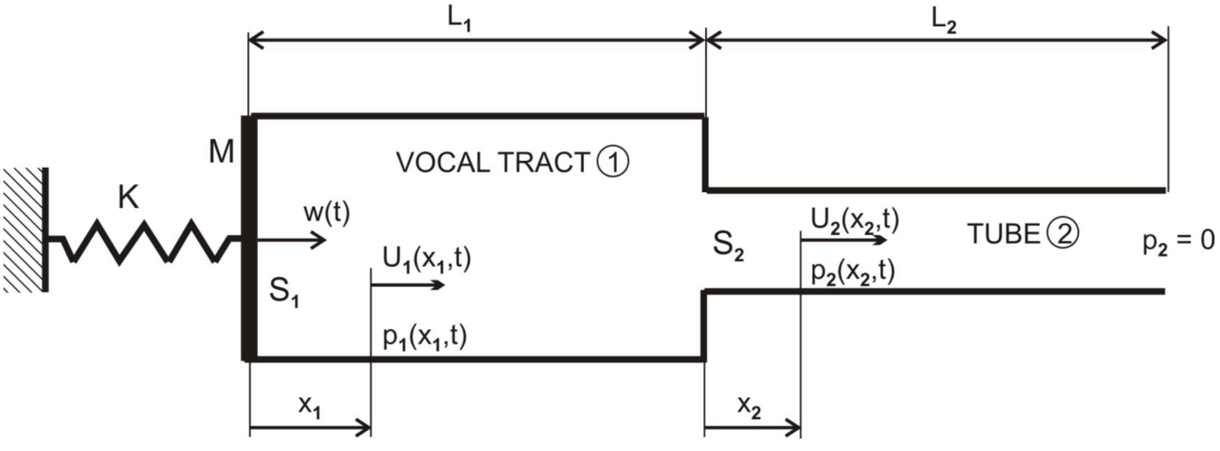


Fig. 13. Simplified model of the vocal tract prolonged by a tube and closed with a yielding wall.

The numerical solution of the frequency equation (9) was found for the parameters ρ_0 = 1.2 kgm^{-3}, c_0 = 353 ms^{-1}, L_1 = 16.2 cm, L_2 = 26.4 cm, S_1 = 3.52 cm^2, S_2 = 0.36 cm^2 in correspondence with the parameters used above for the acoustical system. The mechanical resonance frequency was kept constant at 15 Hz as in [12], $\omega_0 = 2\pi \cdot 15$. The resulting first three natural frequencies are shown in Fig. 14 for the mass M varying from 0 to 3 grams.

The results are in agreement with a general analysis of the solution of Eq. (9) presented in Appendix B for two extreme cases in which the mass M goes to zero or to infinity. The lowest natural frequency F_{struct} corresponds to the mechanical resonance ω_0 and varies from $F_{struct} \rightarrow 0$ Hz for $M \rightarrow 0$ to $F_{struct} \rightarrow 15$ Hz for $M \cong 3$ grams. The second natural frequency of the coupled system corresponds to the first acoustic resonance F_1, which is strongly influenced by coupling with the vibrating wall when the mass M decreases below about two grams. The higher acoustic resonance F_2 is influenced by the vibrating wall in a much smaller range of the mass M = 0-0.2 grams. Similarly, a small influence of the yielding wall on the second acoustic resonance, F_2, was found by Hanna et al. [22].

The acoustic resonances of the system for $M \rightarrow 0$ are equal to frequencies F_1 = 507 Hz and F_2 = 710 Hz, corresponding to the acoustic system with both ends opened: see Eq. (B.16) in Appendix B. The acoustic resonances for M > 3 grams are equal to the acoustical resonances F_1 = 84 Hz and F_2 = 661 Hz, corresponding to the case when the vocal tract is closed by a rigid wall at the glottis: see eq. (B.17). The frequency F_1 in this case corresponds to the artificial model of the vocal tract used in the experiments, in which the first resonance frequency was found to be in the range of 73-99 Hz; see Tables 1 and 2.

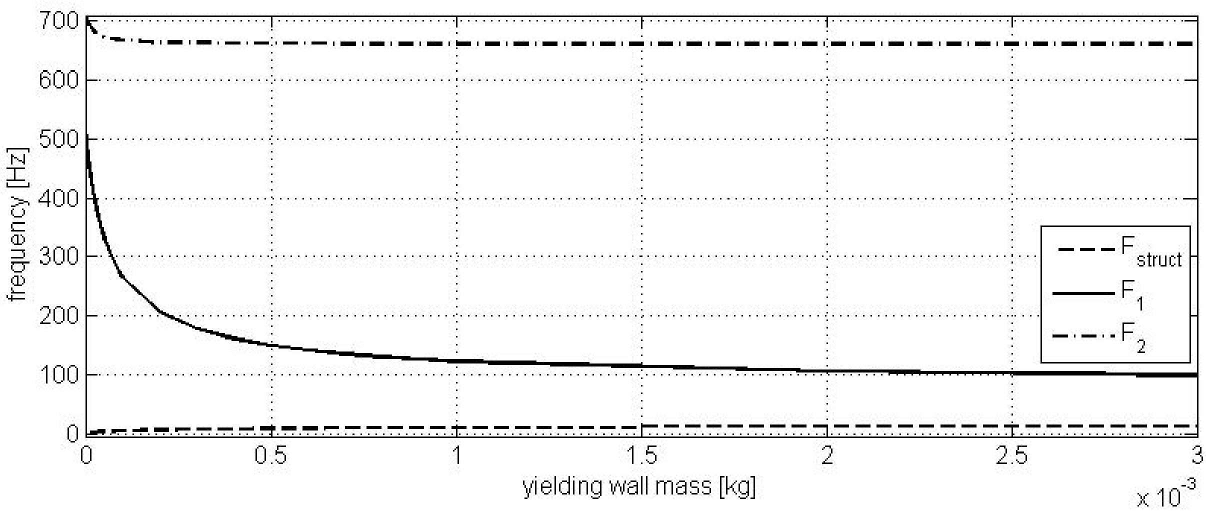


Fig. 14. Natural frequencies of the simplified model of the vocal tract prolonged by the tube and closed with the yielding wall.

Titze et al. [23] presented an empirical relation for the vibrational thickness of a female vocal fold as

Acknowledgements

This study was supported by grants from the Czech Science Foundation (P101/12/P579) and the Academy of Finland (1128095 and 134868).

References

[1] G. Fant, Acoustic theory of speech production, Mouton, S'Gravenage, second edition, 1960.

[2] J. Sundberg, The science of the singing voice, DeKalb, Illinois: N. Illinois Univ. Press, 1987.

[3] I.R. Titze, Principles of voice production, Iowa City, IA: National Center for Voice and Speech, second printing, 2000.

[4] A.M. Laukkanen, On speaking voice exercises. A study on the acoustic and physiological effects of speaking voice exercises applying manipulation of the acousticaerodynamical state of the supraglottic space and artificially modified auditory feedback [doctoral dissertation], Acta Universitatis Tamperensis ser A 445 (1995).

[5] S. Simberg, A. Laine, The resonance tube method in voice therapy: description and practical implementations. Logopedics Phoniatrics Vocology 32 (2007) 165-170.

[6] I.R. Titze, A.K. Verdolini, Vocology. The science and practice of voice habilitation. The National center for voice and speech, Salt Lake City, 2012.

[7] I.R. Titze, Voice training and therapy with a semi-occluded vocal tract: Rationale and scientific underpinnings. Journal of Speech, Language, and Hearing Research 49 (2006) 448–459.

[8] I.R. Titze, A.M. Laukkanen, Can vocal economy in phonation be increased with an artificially lengthened vocal tract? A computer modeling study. Logopedics Phoniatrics Vocology 32 (2007) 147–156.

[9] B.H. Story, A.M. Laukkanen, I.R. Titze, Acoustic impedance of an artificially lengthened and constricted vocal tract. Journal of Voice 14 (2000) 455-469.

[10] I.R. Titze, The physics of small-amplitude oscillation of the vocal folds. Journal of the Acoustical Society of America 83 (1988) 1536-1552.

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

[12] M.M. Sondhi, J. Schroeter, A hybrid time-frequency domain articulatory speech synthesizer. IEEE Transactions on Acoustics, Speech, and Signal Processing ASSP-35 (1987) 955-967.

[13] M.M. Sondhi, Model for wave propagation in a lossy vocal tract. Journal of the Acoustical Society of America 55 (1974) 1070-1075.

[14] T. Leino, A.M. Laukkanen, V. Radolf, Formation of the actor's/speaker's formant: A study applying spectrum analysis and computer modeling. Journal of Voice 25 (2011) 150-158.

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

[16] G. Widholm, BIAS 6 MANUAL Version 1, ART – Acoustics Research Team, 2008, http://www.bias.at/userfiles/files/000809artim.pdf (Accessed 6 October 2015)

[17] A.M. Laukkanen, J. Horáček, R. Havlík, Case-study magnetic resonance imaging and acoustic investigation of the effects of vocal warm-up on two voice professionals. Logopedics Phoniatrics Vocology 37 (2012) 75-82.

[18] G.M. Fant, On the predictability of formant levels and spectrum envelopes from formant frequencies, in: M. Halle, H. Lunt, M. MacLean (Eds.), For Roman Jakobson, Mouton and Company, Gravenhage, The Netherlands, 1956, pp. 109-120.

[19] J. Horáček, I. Zolotarev, Influence of the acoustic-structural couplings upon free vibration of mechanical systems. Archives of Acoustics 18 (1993) 297-309.

[20] D.G. Gorman, J. Horáček, Analysis of the free vibration of a coupled plate/fluid interacting system and interpretation using sub-system modal energy. Engineeering Structures 29 (2007) 754-762.

[21] L.R. Rabiner, R.W. Schafer, Theory and applications of digital speech processing, Prentice Hall, 2011.

[22] N. Hanna, J. Smith, J. Wolfe, Low frequency response of the vocal tract: acoustic and mechanical resonances and their losses, in: T. McGinn (Ed.), Acoustics 2012 Fremantle: Proc. Ann. Conf. Aust. Acoustical Soc., 2012, 7 pages, CD ROM, ISBN: 978-0-646-59039-4.

[23] I.R. Titze, J.G. Švec, P.S. Popolo, Vocal dose measures: Quantifying accumulated vibration exposure in vocal fold tissues. Journal of Speech, Language, and Hearing Research 46 (2003) 919-932.

[24] H. Larsson, S. Hertegård, Vocal fold dimensions in professional opera singers as measured by means of laser triangulation. Journal of Voice 22 (2008) 734-739.

[25] I.R. Titze, Physiologic and acoustic differences between male and female voices. Journal of the Acoustical Society of America 85 (1989) 1699-1707.

[26] V. Radolf, Direct and inverse task in acoustics of the human vocal tract, PhD. thesis, Czech Technical University in Prague, 2010.

Appendix A

A mathematical model of the vocal tract acoustics

A mathematical model of the vocal tract acoustics is based on an analytical solution of a 1D wave equation for acoustic wave propagation in the vocal tract cavity [9, 12, 26]:

$$\frac{\partial^2\phi}{\partial x^2}+\frac{1}{S}\cdot\frac{\partial S}{\partial x}\cdot\frac{\partial\phi}{\partial x}-\frac{1}{c_0^2}\cdot\left(\frac{\partial^2\phi}{\partial t^2}+c_0\cdot r_N\cdot\frac{\partial\phi}{\partial t}\right)=0 \qquad (A.1)$$

where ϕ is the flow velocity potential related to the acoustic pressure p and the acoustic volume velocity U by equations

$$p=-\rho_0\,\partial\phi/\partial t-c_0\rho_0 r_N\phi,\quad U=S\partial\phi/\partial x: \qquad (A.2)$$

x is the longitudinal coordinate along the vocal tract measured from the vocal folds to lips; t is time; r_N *is* the specific acoustic resistance per a unit length; $S(x)$ is the cross-sectional area of the cavity; c_0 is the speed of sound; and ρ_0 is the fluid density.

The relationship between the acoustic pressure p and the volume velocity U at the input and output of each conical acoustic element can be described by the transfer matrix as (see [26])

$$\begin{bmatrix}p_{OUT}\\U_{OUT}\end{bmatrix}=\begin{bmatrix}a&b\\c&d\end{bmatrix}\cdot\begin{bmatrix}p_{IN}\\U_{IN}\end{bmatrix}, \qquad (A.3)$$

where the elements of the transfer matrix are

$$a=\frac{\xi_0}{\xi_0+L_e}\cdot\left(\cosh(\gamma\,L_e)+\frac{1}{\gamma\,\xi_0}\cdot\sinh(\gamma\,L_e)\right),$$

$$b=-\frac{c_0\rho_0(r_N+\mathrm{j}k)\cdot\xi_0}{S_{IN}\cdot\gamma(\xi_0+L_e)}\cdot\sinh(\gamma\,L_e),$$

$$c=S_{OUT}\cdot\frac{\left(1-\gamma^2\xi_0(\xi_0+L_e)\right)\cdot\sinh(\gamma\,L_e)-\gamma\,L\cdot\cosh(\gamma\,L_e)}{\gamma(\xi_0+L_e)^2\cdot c_0\rho_0(r_N+\mathrm{j}k)},$$

$$d=\frac{S_{OUT}}{S_{IN}}\frac{\xi_0}{\xi_0+L_e}\cdot\left(\cosh(\gamma\,L_e)-\frac{1}{\gamma(\xi_0+L_e)}\cdot\sinh(\gamma\,L_e)\right),$$

L_e is the length of the conical acoustic element; S_{IN} and S_{OUT} are the respective cross-sectional areas of the element input and output; γ is a complex exponent given by the formulas

$$\gamma=\overline{\alpha}+\mathrm{j}\overline{\beta}, \qquad (A.4)$$

$$\overline{\alpha}=\frac{r_N}{\sqrt{2+2\cdot\sqrt{1+(r_N/k)^2}}},\quad\overline{\beta}=\frac{k}{2}\cdot\sqrt{2+2\cdot\sqrt{1+(r_N/k)^2}},$$

where $k = \omega/c_0$ is the wave number; ω is the angular frequency of the harmonic signal and $\mathrm{j} = \sqrt{-1}$ is the imaginary unit.

The coefficient ξ_0 is defined by the input (R_{IN}) and output (R_{OUT}) radii of the element

$$\xi_0 = \frac{R_{IN}}{R_{OUT} - R_{IN}} \cdot L, \qquad \text{(A.5)}$$

and the frequency-dependent viscous losses were considered as

$$r_N = \frac{2}{R_{OUT} - R_{IN}} \cdot \sqrt{2k\mu / c_0 \rho_0}\,, \qquad \text{(A.6)}$$

where μ is the dynamic air viscosity.

The acoustic properties of the vocal tract can be described in matrix form as

$$\begin{bmatrix} p_L \\ U_L \end{bmatrix} = \mathbf{T_{VT}} \cdot \begin{bmatrix} p_G \\ U_G \end{bmatrix}, \qquad \text{(A.7)}$$

where $\mathbf{T_{VT}}$ is a transfer matrix obtained by multiplying transfer matrices of all elements from the vocal folds to lips

$$\mathbf{T_{VT}} = \begin{bmatrix} A_{VT} & B_{VT} \\ C_{VT} & D_{VT} \end{bmatrix} = \mathbf{T_N} \cdot \mathbf{T_{N-1}} \cdot \ldots \cdot \mathbf{T_2} \cdot \mathbf{T_1}\,, \qquad \text{(A.8)}$$

and N is the number of conical elements. The transfer matrix of the tube assumed to be one cylindrical element with rigid walls is given by Eqs. (A.3) when $\xi_0 \to \infty$. The complete system (VT+tube) is described by the transfer matrix $\mathbf{T}$ given by the multiplication of transfer matrices of the VT and the tube:

$$\mathbf{T} = \mathbf{T_{TB}} \cdot \mathbf{T_{VT}} = \begin{bmatrix} A_{TB}A_{VT} + B_{TB}C_{VT} & A_{TB}B_{VT} + B_{TB}D_{VT} \\ C_{TB}A_{VT} + D_{TB}C_{VT} & C_{TB}B_{VT} + D_{TB}B_{VT} \end{bmatrix} = \begin{bmatrix} A & B \\ C & D \end{bmatrix}. \qquad \text{(A.9)}$$

The matrix $\mathbf{T}$ describes the relationship between the variables at the glottis and the open tube end:

$$\begin{bmatrix} p_T \\ U_T \end{bmatrix} = \mathbf{T} \cdot \begin{bmatrix} p_G \\ U_G \end{bmatrix}. \qquad \text{(A.10)}$$

The output pressure at the open tube end can be written as the output volume velocity by introducing the radiation impedance (see [7], for example):

$$p_T = Z_{T,RAD} \cdot U_T. \qquad \text{(A.11)}$$

Putting Eq. (A.11) into Eq. (A.10), eliminating p_G and taking into account the known property of the transfer matrices (see e.g. [26]), i.e.

$$\det(\mathbf{T}) = AD - BC = 1, \tag{A.12}$$

yields the expression for the transfer function of the complete system,

$$U_T(\omega)/U_G(\omega) = 1/\left(A - C \cdot Z_{T,RAD}\right), \tag{A.13}$$

or a more frequently used transfer function:

$$p_T(\omega)/U_G(\omega) = Z_{T,RAD}/\left(A - C \cdot Z_{T,RAD}\right). \tag{A.14}$$

Similarly, considering the vocal tract without a tube and loaded by the radiation impedance at the lips,

$$p_L = Z_{L,RAD} \cdot U_L, \tag{A.15}$$

yields the expression for the transfer function of the vocal tract

$$U_L(\omega)/U_G(\omega) = 1/\left(A_{VT} - C_{VT} \cdot Z_{L,RAD}\right). \tag{A.16}$$

To derive the acoustic input impedance, let us insert again Eq. (A.11) into Eq. (A.10). After eliminating U_T, we can express the ratio of the input pressure to input acoustic volume velocity as

$$Z_{IN} = \frac{p_G}{U_G} = \frac{D \cdot Z_{T,RAD} - B}{A - C \cdot Z_{T,RAD}}. \tag{A.17}$$

If the vocal folds were open during the measurement of the transfer function *in vivo*, the pressure at the glottis would be approximately zero ($p_G = 0$). Then we can eliminate U_L from Eq. (1),

$$U_L = D_{VT}/B_{VT} \cdot p_L, \tag{A.18}$$

and putting Eq. (A.18) into Eq. (2) yields the measured transfer function

$$p_L(\omega)/p_T(\omega) = B_{VT}/\left(A_{TB}B_{VT} + B_{TB}D_{VT}\right). \tag{A.19}$$

The denominator in Eq. (A.19) is identical to the frequency equation of the system with O-OZ_{RAD} boundary conditions as mentioned in section 2.3.2 when omitting $Z_{T,RAD}$. Thus, the resonances of the measured transfer function $p_L(\omega)/p_T(\omega)$ would be approximately equal to the resonances of the complete system with the vocal folds open.

We assumed the output loaded by the acoustic radiation impedance of a vibrating circular plate with a radius, R, placed in an infinite wall (see e.g. [14]) to be

$$Z_{RAD} = \frac{c_0 \rho_0}{\pi R^2} \cdot \left[1 - \frac{\mathrm{J}_1(2kR)}{kR} + \mathrm{j}\frac{\mathrm{H}_1(2kR)}{kR}\right], \tag{A.20}$$

where J_1 is the Bessel function of the first kind of order 1 and H_1 is the Struve function of order 1.

Appendix B

A mathematical model of acoustic-structural interaction

Considering the coupled mechanical-acoustical system shown in Fig. 13, the equation of motion for the mass *M* vibrating on a spring of stiffness *K* reads

$$M\ddot{w}(t) + K w(t) = F(t), \qquad \text{(B.1)}$$

where *w* is the translation of the mass, $\ddot{w}$ denotes the second derivative of *w* with respect to time *t* and *F* is the force loading the mass by the pressure in the vocal tract:

$$F(t) = -S_1 \cdot p_1(x_1 = 0, t). \qquad \text{(B.2)}$$

In the simplest case, when omitting the losses, wave equation (A.1) can be written as

$$\frac{\partial^2 \phi}{\partial x^2} - \frac{1}{c_0^2} \cdot \frac{\partial^2 \phi}{\partial t^2} = 0. \qquad \text{(B.3)}$$

Considering a harmonic signal of an angular frequency ω, the solution for the velocity potential can be found in the form

$$\phi_1(x_1, t) = \phi_{01}(x_1) \cdot e^{\mathrm{j}\omega t}, \quad \phi_{01}(x_1) = \alpha_1 e^{\mathrm{j}kx_1} + \beta_1 e^{-\mathrm{j}kx_1}, \qquad \text{(B.4)}$$

$$\phi_2(x_2, t) = \phi_{02}(x_2) \cdot e^{\mathrm{j}\omega t}, \quad \phi_{02}(x_2) = \alpha_2 e^{\mathrm{j}kx_2} + \beta_2 e^{-\mathrm{j}kx_2}, \qquad \text{(B.5)}$$

where indices 1 and 2 correspond respectively to the vocal tract and the tube.

The displacement, *w*, of the wall is given by

$$w(t) = w_0 \cdot e^{\mathrm{j}\omega t}. \qquad \text{(B.6)}$$

The unknown constants α_1, β_1, α_2, β_2 can be obtained from the following boundary and continuity conditions for the pressure, *p*, and the acoustic volume velocity *U*:

$$\dot{w} = U_1(x_1 = 0, t)/S_1, \qquad \text{(B.7)}$$

$$p_2(x_2 = L_2, t) = 0, \qquad \text{(B.8)}$$

$$p_1(x_1 = L_1, t) = p_2(x_2 = 0, t), \qquad \text{(B.9)}$$

$$U_1(x_1 = L_1, t) = U_2(x_2 = 0, t). \qquad \text{(B.10)}$$

Using Eqs. (A.2), i.e. $p = -\rho_0 \, \partial\phi/\partial t$ and $U = S\partial\phi/\partial x$, conditions (B.7-B.10) yield a system of equations:

$$\alpha_1 - \beta_1 = c_0 w_0 \tag{B.11}$$

$$\alpha_2 e^{\mathrm{j}kL_2} + \beta_2 e^{-\mathrm{j}kL_2} = 0 \tag{B.12}$$

$$\alpha_1 e^{\mathrm{j}kL_1} + \beta_1 e^{-\mathrm{j}kL_1} = \alpha_2 + \beta_2 \tag{B.13}$$

$$S_1\left(\alpha_1 e^{\mathrm{j}kL_1} - \beta_1 e^{-\mathrm{j}kL_1}\right) = S_2\left(\alpha_2 - \beta_2\right) \tag{B.14}$$

The solution of Eqs. (B.11-B.14) results in complicated relations for constants α_1, β_1, α_2, β_2. After putting them into Eqs. (B.4-B.5), deriving the acoustic pressure $p_1(x_1, t)$ and inserting it into (B.2) and (B.1), we finally get the frequency equation

$$\omega^2 + \frac{\omega\rho_0 c_0 S_1}{M} \cdot \frac{S_2 \sin(kL_1)\cos(kL_2) + S_1 \cos(kL_1)\sin(kL_2)}{S_2 \cos(kL_1)\cos(kL_2) - S_1 \sin(kL_1)\sin(kL_2)} - \omega_0^2 = 0\,, \tag{B.15}$$

where $\omega_0^2 = K/M$ is a squared angular mechanical resonance frequency.

A brief analysis of frequency equation (B.15) can be done for two extreme cases. When the mass, M, goes to zero (and so does the stiffness, K, because we assume ω_0 to be constant), then the numerator of the second element must be zero and thus, we obtain the solution $\omega = 0$ and the frequency equation for the open-open system:

$$S_2 \sin(kL_1)\cos(kL_2) + S_1 \cos(kL_1)\sin(kL_2) = 0\,. \tag{B.16}$$

If the mass, M, goes to infinity (and so does K), then $\omega = \omega_0$ or the denominator must be zero, which yields the frequency equation for the closed-open system:

$$S_2 \cos(kL_1)\cos(kL_2) - S_1 \sin(kL_1)\sin(kL_2) = 0\,. \tag{B.17}$$