This is the accepted manuscript of the article, which has been published in *Journal of Speech, Language, and Hearing Research*. 2019, 62(7), 2227-2244. https://doi.org/10.1044/2019_JSLHR-S-17-0490

EXPERIMENTAL AND COMPUTATIONAL MODELING OF THE EFFECTS OF VOICE THERAPY USING TUBES

J. Horáček[1], V. Radolf[1], A.-M. Laukkanen[2]

[1]Institute of Thermomechanics, Czech Academy of Sciences,

Dolejškova 1402/5, 182 00 Prague, Czech Republic,

jaromirh@it.cas.cz, radolf@it.cas.cz

[2]Speech and Voice Research Laboratory, Faculty of Education,

Åkerlundink. 5, University of Tampere, 33100 Tampere, Finland,

Anne-Maria.Laukkanen@uta.fi

Corresponding Author:

Jaromír Horáček, Institute of Thermomechanics, Czech Academy of Sciences, Dolejškova 1402/5, 182 00 Prague, Czech Republic, jaromirh@it.cas.cz,

Conflict of Interest: There are no relevant conflicts of interest.

Funding Statement:

This study was supported by a grant from the Czech Science Foundation: No. 16-01246S "Computational and experimental modelling of self-induced vibrations of vocal folds and influence of their impairments on human voice."

Abstract:

Phonations into a tube with the distal end either in the air or submerged in water are used for voice therapy. This study explores the effective mechanisms of these therapy methods.

The study applied a physical model complemented by calculations from a computational model, and the results were compared to those that have been reported for humans. The effects of tube phonation on vocal tract resonances and oral pressure variation were studied. The relationships of transglottic pressure $P_{trans}(t)$ variation in time vs. glottal area variation $GA(t)$, were constructed.

The physical model revealed that, for the phonation on [u:] vowel through a glass resonance tube ending in the air, the first formant frequency F_1 decreased by 67%, from 315 Hz to 105 Hz, thus slightly above the fundamental frequency F_0 that was set to 90–94 Hz. For phonation through the tube into water, F_1 decreased by 91-92%, reaching 26–28 Hz, and the water bubbling frequency $F_b \cong 19$–24 Hz was just below F_1. The relationships of $P_{trans}(t)$ vs. $GA(t)$ clearly differentiate vowel phonation from both therapy methods, and show a physical background for voice therapy with tubes. It is shown that comparable results have been measured in humans during tube therapy. For the tube in air, F_1 descends closer to F_0, while for the tube in water, the frequency F_b occurs close to the acoustic-mechanical resonance of the human vocal tract.

In both therapy methods, part of the airflow energy required for phonation is substituted by the acoustic energy utilizing the first acoustic resonance. Thus, less flow energy is needed for vocal fold vibration, which results in improved vocal efficiency. The effect can be stronger in water resistance therapy if the frequency F_b approaches the acoustic-mechanical resonance of the vocal tract, while simultaneously F_0 is voluntarily changed close to F_1.

Keywords: Biomechanics of voice; maximum glottal area declination rate (MADR), vocal tract acoustics; formant frequencies; phonation into tubes; water resistance voice therapy; water bubbling frequency, vocal efficiency.

INTRODUCTION

Phonation into different kinds of tubes is widely used for voice training and therapy (Sovijärvi, 1969; Laukkanen, 1992; Titze, Finnegan, Laukkanen, & Jaiswal, 2002; Granqvist et al., 2015; Amarante et al., 2016; Guzman et al., 2017). According to practical experience, phonation into a tube may feel easier than ordinary vowel phonation, and speech immediately after tube phonation often sounds louder and less strained. Voice therapy has exploited phonation through a tube into water – i.e., water resistance therapy – for treating both hypofunctional and hyperfunctional voice disorders. Two variants of tubes have been increasingly widely used in water resistance therapy: a glass tube called a "resonance tube" (Sovijärvi, 1969; Laukkanen, 1992; Simberg & Laine, 2007) with a length of 24–28 cm and an inner diameter of ca. 8–9 mm, and a Lax Vox tube made of flexible silicon, 35 cm in length and 1 cm in inner diameter (Sihvo, 2006; Sihvo & Denizoglu, 2007; Sihvo, 2017).

During the last two decades, many studies have been carried out to explain the basis of tube training and therapy. It is known that artificial prolongation and narrowing of the vocal tract (VT) increases the impedance. In particular, phonation through narrow tubes and straws and phonation through a tube into water increase supraglottic resistance (air pressure/flow) (Titze et al., 2002; Horáček, Radolf, Bula, & Laukkanen, 2014; Amarante et al., 2016). Increased supraglottic resistance offers training for breath support (Carroll & Sataloff, 1991) and decreases transglottic pressure – i.e., the difference between subglottic and supraglottic pressure – which provides the driving force of vocal fold vibration (Titze & Laukkanen, 2007). The phonation threshold and amplitude of the vocal folds' (VFs) vibration decrease; thus, the biomechanical loading related to phonation decreases (Titze, 2006b). Simultaneously, especially if the first acoustic resonance of the VT plus the tube is lowered effectively close to the fundamental frequency (F_0) of the VFs' vibration, the positive inertance of the VT may improve the voice quality by assisting the VFs' vibration and/or strengthening the amplitude of the harmonics (Story, Laukkanen, & Titze, 2000; Fant & Lin, 1987; Rothenberg, 1981).

Earlier modelling studies have shown that positive inertance of the vocal tract promotes vocal fold vibration as the oscillation of the air pressure above the vocal folds is in phase with the velocity of the vocal folds, and the phonation threshold is also reduced (Titze, 2001). There are also findings for a female singer showing that the vocal fold vibration is positively affected when F_0 is close to F_1 (Rothenberg, 1988).

Additionally, the increased supraglottic resistance intensifies sensations of resonatory vibrations in the VT (Titze, 2006b). Thus, the exercise with tubes and other semiocclusions of the VT offers the trainee sensations of more economic and efficient voice production. These sensations may then be a guide to establish similar acoustic-mechanical conditions in the VT also after the exercise, since the trainee may learn to apply epilaryngeal narrowing as a source of impedance-matching between the glottis and the VT (Titze, 2006b). Moreover, water resistance therapy offers the element of water bubbling, which is reflected in the oscillation of supraglottic pressure and results in variation of the amplitude of the VFs' vibration (Granqvist et al., 2015). This bubbling effect has been reported to feel like a massage of the VT and the larynx, and it has been suggested that it could offer similar beneficial effects as massage – e.g., improved blood circulation in the tissue (Mori et al., 2004).

Although water bubbling has been regarded as relaxing for the muscles and potentially healing for the tissues, there are also some results showing that the relative glottal closing speed and the closed time of the glottis may increase (Guzman et al., 2017) when the tube immersion depth is high (e.g., 10–18 cm below the water surface) which suggests higher mechanical loading of the VFs. It is also possible for the water bubbling frequency to approach the acoustic–mechanical resonance of the laryngeal tissues, which may increase the effect of bubbling both beneficially and adversely (Horáček, Radolf, & Laukkanen, 2017a).

This study aims to shed more light on the similarities and differences between tube training with the distal end of the tube in air and tube training with the distal end of the tube immersed in water. This study focuses on the following research questions:

1. Are there any essential differences between *the glottal area variation* in time $GA(t)$ versus transglottic pressure $P_{trans}(t)$ variation measured for phonation on the vowel [u:] and phonation into the resonance tube with the distal end either in air or in water? An important reason for studying this relation is the fact that it gives an estimate of the average work done by airflow during vocal fold vibration thus allowing estimation of changes in vocal efficiency. Such a study is difficult to realize in humans, although investigations have been conducted with high speed filming of the VF using larger tubes (Laukkanen et al., 2007; Guzman et al., 2017) and filming through the nose (Granqvist et al., 2015). Both cases obviously result in unnatural VT conditions compared to ordinary phonation into a resonance tube.

2. Are there any important differences between the evaluated *glottal area time derivatives* for all three cases of phonation considered? Our focus is especially on the potential differences in the maximum area declination rate (MADR), which is considered to be a measure of the stress loading imposed on the VFs during collisions (Titze & Laukkanen, 2007).

3. What are the effects of *the acoustic-mechanical resonance and the first acoustic resonance* in phonation through the tube in air and in water? Are there any fundamental differences in the principles of these two therapy methods? This information would be useful to guide the proper choice and use of the tubes.

In this study, a physical model will be applied that includes a lung model, artificial VFs, and a plexiglass model of the VT (Horáček, Bula, Košina, & Radolf, 2016a). The results obtained with this hard-walled VT model are further compared with (1) those obtained previously in humans and with (2) calculations obtained from a computer model, where yielding walls were implemented to better resemble the VT of humans.: Yielding walls affect the formant frequencies and also lead to appearance of a low-frequency acoustic-mechanical resonance (Radolf, Laukkanen, Horáček, & Liu, 2014; Radolf et al., 2016; Horáček et al., 2017a).

For simplicity, we use F_1 to refer both to the first formant frequencies measured in the model and in humans and to the first computed acoustic resonance frequencies.

METHODS

Physical model

A model of the *human lungs*, which includes the splitting of the airways up to fourth-order branching, was built in the subglottic part of the experimental facility (see Figure 1 and more in Horáček, Radolf, Bula, & Košina, 2017b). The air flowed through the model of the lungs to the *trachea*, which was modeled by a metal tube. The total length of the trachea was 23 cm, and the inner diameter was 18 mm. The experiments were performed with a 1:1 scaled three-layered *vocal fold* model with a total length of 20 mm, vertical thickness 10.3 mm and width 8 mm (Horáček et al., 2016b; Horáček et al., 2017b). The middle cross-section of the vocal fold model was based on CT measurements of a female subject (Hampala et al., 2015). The layers were formed by a silicon wedge modeling the vocal fold body inside the VFs, a 2-3 mm thick water layer modeling the lamina propria, and a silicon cover with the thickness slightly below 2 mm. Filling with water lowers the F_0 of the VF model to ca. 80 Hz and the phonation threshold flow (PTF; i.e., the lowest flow rate to start sustained phonation) to ca. 0.02 l/s, covering also the lowest limits of human phonation (see, e.g., Baken & Orlikoff, 2000).

The fundamental phonation frequency was preset by an initial pre-stressing of the VFs. This was accomplished by a slight variation of the distance between the frames enclosing the VFs, followed by changing the hydrostatic pressure inside the VFs. After filling the "lamina propria" layer in the VFs with water and completing the tuning procedure, the model was fixed in a plexiglass frame and connected to the model of the subglottal spaces on one side and to the model of the VT on the other side.

The geometrical configuration of the plexiglass model of the *vocal tract* corresponded to the Czech vowel [u:]. It consisted of 46 circular cross-sectional areas, perpendicular to the midline along the VT from the VFs up to the lips. These areas were obtained from magnetic resonance images (MRI) recorded during the phonation of a 35-year-old male subject (Vampola, Horáček, & Švec, 2008). The total length of the VT model was 18.7 cm.

Set-up and measurements on the physical model

Measurements were performed for: a) simulated phonation on the vowel [u:], b) phonation with [u:]-vowel-shaped VT attached to a glass resonance tube (length 27 cm, inner diameter 7.8 mm) ending in air, c) phonation with [u:]-vowel-shaped VT attached to the same tube ending in water, and d) blowing through the setup mentioned in c) but without the vocal folds vibrating (no phonation). The (d) condition was included to compare subglottic pressure, oral (back) pressure, and water bubbling frequency with and without phonation. Such a comparison may reveal more on the interplay between different types of oscillations of the air pressure occurring simultaneously in the VT during phonation into water.

In the experiments with phonation, flow rates ranged between the PTF and the maximum rate beyond which the vibration amplitudes would endanger the material coherence of the artificial VFs. The glottal width between the VFs was preset to zero and not changed. In the experiment without phonation, the VFs were abducted by deflation, thus simulating the breathing position of the VFs. In this (d) condition the glottal width was preset to ca.1 mm.

The following parameters were measured: 1) mean airflow, Q; 2) phonation threshold flow, PTF; 3) mean subglottic air pressure, P_{sub}; 4) phonation threshold pressure, PTP (i.e. the lowest P_{sub} to start and sustain phonation); 5) peak-to-peak variation in $P_{sub}(t)$; 6) frequency of the lowest subglottic resonance; 7) mean oral air pressure, P_{oral}; 8) peak-to-peak variation in P_{oral}; 9) frequency F_1 of the lowest oral resonance (formant); 10) 3 dB bandwidth B_1 of the lowest oral resonance, 11) fundamental frequency F_0 of phonation (measured from P_{oral} variation); 12) water bubbling frequency, F_b (measured from P_{oral} variation); and 13) variation of glottal area, $GA(t)$. The time derivative of glottal area variation dGA/dt (representing the speed of glottal opening and closing) was calculated by the method of central differences (Mathews, & Fink, 2004). Parameters 1–13 were measured because the VF vibrations and thus glottal area variation are dependent on aerodynamic variables below and above the glottis. The time variation in the pressures, in turn, is related to supra- and subglottic resonances.

A schema of the measurement set-up is shown in Figure 1. The VFs were excited by airflow coming from a regulated central pressure supply. *The mean airflow rate* was measured by a float flowmeter (EMKO type DF3-09K5), and by an orifice flow meter with a pressure sensor (LP1000 DRUCK model LPX1812-C1S) using an analogue output. *The mean subglottic and oral pressures* were measured by integrated pressure semiconductor sensors (NXP Freescale MPXV5010GC6U) mounted in the walls of the trachea and oral cavity models. The mean subglottic pressure at the entrance to the VFs was registered by a digital manometer (Greisinger Electronic, GDH07AN) connected by a short compliant tube to the subglottic space of the glottal cavity. A second digital manometer, a Greisinger Electronic GMH3151, measured the mean pressure of the water inside the VFs, which was adjusted with a water-filled syringe.

The fluctuations of the subglottic and oral pressures were measured by a miniature microphone (Bruel & Kjaer 4138, range 6.5 Hz–140 kHz) and by a special microphone probe (B&K 4182, range 1 Hz–20 kHz), respectively. The spectra of the pressure signals were calculated by an in-house developed program in Matlab, and the formant frequencies were estimated from the peaks of the averaged spectral data (Horáček et al., 2017a).

All the measured signals were simultaneously sampled at a frequency of 16.4 kHz and registered by a B&K PULSE type 3560 C measurement system with type 7537A and 3109 Input/Output Controller Modules controlled by a personal computer (PC I) equipped with SW PULSE LabShop Version 10.

A high-speed CCD NanoSense Mk. III camera (maximum resolution 1280 x 1024 pixels) with a Nikon AF micro Nikkor 60mm zoom lens was included in the measurement set-up for the *analysis of the VFs' vibration*. The camera was positioned at a 90-degree bend of the trachea model where a glass window was installed; this enabled the viewing of the VFs' vibration from the subglottal side. A personal computer (PC II) was used for recording the VFs' vibration. During image recording at a frequency of

3000 frames/second, three intensive LED lights (2 x 13 W+1 x 35 W) were focused on the vibrating VFs. The images recorded by the camera were synchronized with the time records of the pressure signals.

Figure 1. Somewhere here.

Computational analysis of the measured relations between $P_{trans}(t)$ and $GA(t)$

The relationship of $P_{trans}(t)$ vs. $GA(t)$ measured during regular periodic self-sustained VF oscillations can be visualized as a closed, clockwise-oriented cyclic curve (loop). Similar graphs showing the relationship between the subglottic pressure and the glottal opening time variation were introduced by Mišun, Švancara, & Vašek (2011). They measured these relationships on artificial VFs for various self-oscillating regimes without any VT. In the present study, we use similar graphs to compare phonation on a vowel with phonation through the resonance tube with the distal end in air and in water. By computing the average area inside the loop, we get an estimate of a part of the average mechanical work done by the airflow per one cycle of the VFs' vibration (i.e. part of the flow energy needed to excite the self-sustained VF oscillations). This can be used as a measure of vocal efficiency. We note that the acoustic energy is not included in our considerations, because it was not possible to measure the airflow volume velocity waveform $Q(t)$ synchronously with the pressure signals.

The normalized area ΔA inside the graph (loop) was computed by an in-house developed program in Matlab that uses the following mathematical formula for the counter-integral along the boundary ∂A of the area ΔA:

$$\Delta A = \oint_{\partial \Delta A} F_{eq}\, \mathrm{d}GO_{eq} = \frac{1}{n}\int_0^{nT_0} F_{eq}\left(t\right)\frac{dGO_{eq}\left(t\right)}{dt}\,\mathrm{dt} = T\,\frac{1}{n}\int_0^{nT_0} P_{trans}\left(t\right)\frac{dGA\left(t\right)}{dt}\,\mathrm{dt}\,, \tag{1}$$

where $F_{eq}\left(t\right) = P_{trans}(t)\,T\,l$ is a roughly estimated force loading the VF surface approximated by the area $T\;l$; $T \cong 10\,\mathrm{mm}$ and l=20 mm are the VFs' thickness and length, respectively; $GO_{eq}\left(t\right) = GA\left(t\right)/\,l$ is an equivalent glottal width; $GA\left(t\right)$ and $P_{trans}(t)$ are the measured glottal area and transglottic pressure, respectively; and n is the number of considered oscillation periods $T_0 = 1/F_0$. In the present study, n=9. The higher is the area ΔA, measured in joules, the higher is the average work done by the airflow per one oscillation cycle. The transglottic pressure was substituted by a quasi transglottic pressure $P_{trans}(t)=P_{sub}(t)-P_{oral}(t+t_0)$, where the measured oral pressure signal was shifted by the time $t_0=\Delta L/c_0$ corresponding to the time delay of sound propagation between the vocal folds and the pressure sensor positioned in the mouth cavity; ΔL=153 mm is the distance between the vocal folds and the pressure sensor in the mouth and c_0=346 m/s is the speed of sound.

Computational modeling of the acoustic resonances

The VT of the physical model used in this study was made of hard plexiglass. The human VT, instead, has softer, yielding walls. In the present study, computer modeling was used to calculate what the acoustic resonance frequencies of the physical model would be if the vocal tract wall were softer, resembling that of the human vocal tract. The effects of the yielding walls in the human VT have been estimated by computer modeling for the tube in air (see Story et al., 2000; Radolf et al., 2016) and for the tube in water (see Horáček et al., 2017a).

In the computer modeling included in the present study, the following input parameters were considered for air: density ρ_0=1.2 kgm^{-3}; speed of sound c_0=353 m/s (corresponds to 36°C) for a human VT and c_0=346 m/s (corresponds to 24°C) for the VT model with hard walls used in the experiments; a dynamic viscosity μ=1.8 10^{-5} Pa s; and ρ_o=998 kgm^{-3} and c_o=1500 m/s for water at the tube's end.

The following parameters for the mechanical system representing the yielding walls were calculated from the data published by Liljencrants (1985, Table 4.1): the eigenfrequency $\omega_w = 2\pi.66$ rad/s, the damping ratio of the yielding wall $\zeta_w =$ 0.99, and the mass $m_w = 1400\, S_0^2 = 0.2$ gram, which corresponds to the mass of the yielding wall. The area S_1=3.80 cm^2 was considered for the first acoustic element at the glottis.

RESULTS

Effect of tube phonation on formant frequencies

Figure 2 illustrates the formants for all three cases studied with the physical model. It is possible to see that phonation into a tube decreased the first formant frequency F_1 (compared to vowel) both when the outer end of the tube was in air and when it was in water (Figure 2b and c, respectively). The effect was much stronger when the tube was in water. For phonation on vowel [u:] the first formant frequency was $F_1 \cong 315$ Hz. For phonation through the tube in air it decreased to $F_1 \cong 105$ Hz, which was near the fundamental frequency of the model ($F_0 = 90$ Hz). For phonation through the tube with the distal end submerged 10 cm in water the first formant frequency was $F_1 \cong 28$ Hz, which in turn was near the water bubbling frequency ($F_b \cong 22$ Hz).

Table 1 shows the formant frequencies obtained with the physical model in comparison to those obtained in computational modeling for the hard-walled and soft-walled vocal tract and to earlier results obtained for humans. The measured formant frequencies are in good agreement with the computational results for the hard-walled vocal tract. The first computed acoustic resonance frequency was $F_1 \cong 97$ for the VT prolonged by the tube in air, and $F_1 \cong 28$ Hz for the VT prolonged by the tube with the distal end submerged 10 cm in water. Considering the yielding VT walls, the computation resulted in $F_1 \cong 181$ Hz for the VT prolonged by the tube in air, and the acoustic-mechanical resonance $F_{a\text{-}m}$ appeared at 27 Hz. For the VT prolonged by the tube with the distal end in water, F_1 was 143 Hz and $F_{a\text{-}m}$ was 9 Hz. The resonance frequencies computed with yielding VT walls are close to those previously measured in humans. The 3dB frequency bandwidths B_1 measured for the first formant of the vowel decreased from 220 Hz to 99 Hz for the tube in air and to ca. 25 Hz for the tube in water. We remark, that for the phonations through the tube, the differences between the excitation frequency F_0 or F_b and the first formant frequencies F_1 are much smaller than the formant bandwidths, see Figure 2 and Table 1.

Figure 2d shows that the measured first subglottal resonance frequency $F_{sub1} \cong 720$ Hz of the subglottic cavities was much higher than the first formant frequencies F_1 in the supraglottal cavities (Fig. 2 a-c), and therefore F_{sub1} did not interfere with the phenomena studied here. The F_{sub1} was practically the same for all the sample types studied.

Figure 2. Somewhere here.

Table 1**.** Somewhere here.

Effect of tube phonation on aerodynamic variables

The phonation threshold flow (PTF) substantially decreased for phonation through the tube into air and into water compared to phonation on the vowel [u:]. Therefore the flow rate intervals were different for phonation on [u:] (Q=0.08–0.25 l/s), and on [u:] with the VT prolonged by the resonance tube with the distal end in air and in water (Q=0.02–0.10 l/s; see Figure 3). *Similarly as for PTF, the subglottic pressure needed for phonation through the tube in air (phonation threshold pressure; PTP) was substantially lower than for phonation on [u:].* For the tube in water, the PTP≅1.4 kPa was naturally higher than for the tube in air, being approximately the same as for phonation on [u:], see Figure 3.

The measured mean subglottic pressure values P_{sub}=0.4–2.4 kPa were within a physiologically relevant range for human voice production (Baken & Orlikoff, 2000). P_{sub} for blowing into water was lower than for phonation through the tube into water because no flow energy was needed for VFs' vibration. The mean and peak-to-peak values of the subglottic pressure increased in an approximately linear manner with the flow rate Q. The peak-to-peak amplitudes of the subglottic pressure $P_{sub\_p\text{-}t\text{-}p}$ remained in approximately the same range for all considered cases. The relative amplitudes of $P_{sub\_p\text{-}t\text{-}p}$ related to the mean subglottic pressure $P_{sub\_p\text{-}t\text{-}p}/P_{sub}$ increased in the all studied cases, with the flow rate from about 6–16% at the PTF to ca. 45–65% for the maximum flow rate used in the experiment. The values of $P_{sub\_p\text{-}t\text{-}p}$ for blowing into water were much lower than for phonation into water.

Similar trends were visible for the measured mean oral pressure P_{oral} and peak-to-peak values of the oral pressure $P_{oral\_p\text{-}t\text{-}p}$ related to Q. The mean oral pressure for the tube with the distal end submerged 10 cm in water was much higher than for phonation into air due to the hydrostatic pressure. *The* *magnitudes of $P_{oral\_p\text{-}t\text{-}p}$ for phonation through the tube into air and into water were high* and almost the same. This high magnitude of the peak-to-peak oscillation of oral pressure results from the fact that F_1 decreased close to F_0 for the tube in air and close to the water bubbling frequency F_b for the tube in water (see Table 1). *The fluctuations of the oral pressure in the case of blowing were much smaller*, because they were excited only by water bubbling and not supported by the acoustic resonance at the frequency F_1 that increased above 180 Hz, corresponding to the open VT end at the abducted VFs. Therefore, the massage effect for the VFs caused by blowing into water would be much smaller than for phonation into water.

Figure 3. Somewhere here.

Effect of tube phonation on glottal area variation

Figure 4 shows the results for the maximum glottal area (max$GA(t)$) and the maximum glottal area derivative (max($-dGA/dt$)) evaluated in the closing phase of the glottis. The latter quantity, which corresponds to the maximum glottal closing velocity, is sometimes called the maximum area declination rate (MADR; see Titze (2006a). The values of max$GA(t)$ and max($-dGA/dt$) increased nearly linearly with the flow rate Q. Considering the same flow rate, e.g., 0.04 l/s, the values of max$GA(t)$ and max($-dGA/dt$) for phonation through the tube into water were higher than for phonation into air. The magnitudes of max$GA(t)$ for the tubes were higher than for the vowel, while max($-dGA/dt$) , i.e. MADR, was roughly in the same range for the tubes and the vowel.

Figure 4. Somewhere here.

Effects of flow rate on fundamental and water bubbling frequencies

Figure 5 shows the measured fundamental frequency of phonation F_0 and the bubbling frequency F_b as functions of flow rate Q. In the considered range of Q, the fundamental frequency decreased from about $F_0 \cong 110\text{–}113$ Hz for phonation on [u:] to $F_0 \cong 90\text{–}94$ Hz for phonation through the tube into air and further to $F_0 \cong 75\text{–}86$ Hz for phonation through the tube into water. In the case of no phonation, the bubbling frequency increased quickly with Q from zero to the approximate constant $F_b \cong 20$ Hz. The bubbling frequency for phonation through the tube into water was $F_b \cong 19–23$ Hz for all Q values. This indicates that the production of bubbles is not significantly influenced by phonation.

Figure 5. Somewhere here.

Measured relationships $P_{trans}(t)$ vs. $GA(t)$

The relationships $P_{trans}(t)$ vs. $GA(t)$ constructed from the measurements on the physical model clearly distinguish the vowel phonation from both therapy methods and uncover the physical background of the methods (see Figures 6–9). Such relationships are difficult to obtain from measurements in humans.

Figure 6 shows the results for phonation on the vowel [u:] for one period of self-sustained VF vibration. The cyclic graphs ($P_{trans}(t)$ vs. $GA(t)$ and P_{trans}(t) vs. d$GA(t)$/dt) create certain loops, which are oriented in the clockwise direction, as marked by the arrows for increasing time. They start in an open glottis position (at the time instant marked 1). Then, the glottal area GA increases up to the maximum (at the time instant marked 5). Thereafter, during the glottal closing phase, the glottal area decreases to the time instant 7, where is the minimum transglottic pressure, and then through the time instant 13, which corresponds to the MADR, down to zero (at the time instants 15 and 16). After the VFs' collision, the pressure $P_{trans}(t)$ increases up to the maximum just before the opening phase of the glottis. The area ΔA=0.031 mJ inside the loop $P_{trans}(t)$ vs. $GA(t)$ computed according to equation (1) is a measure of the part of the airflow energy consumed by the self-sustained VF oscillations.

Figure 6. Somewhere here.

Similarly, Figure 7 shows an example of the relationships $P_{trans}(t)$ vs. $GA(t)$ and P_{trans}(t) vs. d$GA(t)$/dt measured for phonation into the resonance tube with the distal end in air. The loops start in the opening phase of the glottis (at the time instant marked 1) and continue through the time instant 3 to the minimum of $P_{trans}(t)$ at the time instant 5 and then to the maximum glottal opening (marked as 11). However, the closing phase is clearly different from the previous case for phonation on the vowel [u:], because the part of the loop $P_{trans}(t)$ vs. $GA(t)$ between the time instants 1–18 is oriented in a counter-clockwise direction. This means a negative contribution to the work ΔA in this phase of the oscillation period. After time instant 18, the closing phase continues in the clockwise-oriented part of this loop through the time instant 19 up to the complete glottal closure (marked 21) and then to the end of the period (at the time instant marked 33). The maximum transglottic pressure is substantially delayed after the moment of glottal closure and the minimum of $P_{trans}(t)$ is negative, unlike in the phonation on [u:].

The computed total magnitude of the area ΔA=-0.054 mJ inside the complete loop is negative, because the VFs' vibration for phonation through a tube into air is supported by the first acoustic resonance thanks to a near coincidence of the formant frequency $F_1 \cong 105$ Hz (see Figure 2 and Table 1) with the fundamental frequency $F_0 \cong 94$ Hz (see Figure 5). *This effect, when the frequency F_0 is close to F_1, results in less flow energy being needed for the self-sustained VF oscillations. This demonstrates the principle of vocal exercising and therapy with the resonance tube in air, because this effect makes the phonation easier.* Because of this, lower subglottic pressure and a lower flow rate are needed for the self-sustained VF oscillations. Comparing Figure 7 with Figure 6, it can be concluded that the mean transglottic pressure $P_{trans} \cong 0.89$ kPa, and the mean flow rate Q=0.08 l/s are much lower for phonation into the tube than for phonation on the vowel [u:], where the mean transglottic pressure for self-sustained VF oscillations was about 1.95 kPa and flow rate equaled 0.2 l/s.

The loop for the derivative of the glottal area (reflecting glottal closing speed) is shown in the lower panel of Figure 7. The minimum derivative of the glottal area (MADR), about $dGA/dt \cong -7500$ mm^2/s, occurs at the time instant 18 before the contact of the VFs at the time instant 21, where dGA/dt=0. Values of MADR are considered to be roughly proportional to the impact stress loading the VFs during the collision. However, following the graph for the derivative of the glottal area between time instants 18 and 21 in Figure 7, it can be clearly seen that in reality *the glottal closing speed is noticeably reduced just before the VFs' contact. Thus, MADR is not an adequate measure of impact stress, whereas the glottal area derivative just prior to glottal closing is.*

Figure 7. Somewhere here.

Figure 8 presents an example of the relationships $P_{trans}(t)$ vs. $GA(t)$ and P_{trans}(t) vs. d$GA(t)$/dt for phonation through the resonance tube with the distal end 10 cm in water for the same mean flow rate Q=0.08 l/s as in the previous case for phonation through the tube into air. The measured loops are very similar as in the previous case. This suggests that the effect of the near coincidence of the formant frequency $F_1 \cong 26$ Hz and the bubbling frequency $F_b \cong 23$ Hz is similarly strong as in the case of the near coincidence of F_0 and F_1 for phonation through the tube into air, where the difference F_1–F_0 was 11 Hz. The mean flow rate is again much lower than the flow rate Q=0.25 l/s for the phonation on the vowel [u:]; furthermore, as in the previous case, the mean transglottic pressure $P_{trans} \cong 0.78$ kPa is lower than the mean $P_{trans} \cong 1.95$ kPa for phonation on the vowel [u:].

Figure 8. Somewhere here.

Figure 9 demonstrates that phonation into water is more complicated due to the bubbling process, which creates waves on the water's surface. When the variation of the glottal area, the derivative of the glottal area, transglottic pressure, and the constructed loops are observed for a longer time, the irregularities in amplitudes of all quantities caused by the water bubbling are evident. The areas inside the loops vary irregularly from one oscillation cycle to the next. Therefore, when studying phonation through a tube into water, the energy transfer between the flow and the VFs needs to be investigated for a longer time, taking into account more cycles in order to obtain an average value of the area ΔA. This procedure resulted in $\Delta A \cong$-0.041 mJ which is quantitatively comparable value with that obtained for phonation through the tube into air, where the behavior of the system is periodic with only small disturbances. *The principle of supporting the self-sustained VF oscillations by the effect of the near coincidence of the formant frequency F_1 with the fundamental frequency F_0 or bubbling frequency F_b is valid in both cases of phonation into the tube.*

Figure 9. Somewhere here.

This principle can be documented by the flow power $\Delta FP = \Delta A\ F_0$ [W] that was computed from all the loops measured within the airflow rate interval Q=0.02–0.25 l/s considered in this study. Figure 10 shows the measured power ΔFP as a function of the input steady aerodynamic flow power defined as the mean subglottic pressure P_{sub} multiplied by mean flow rate Q (Schutte, 1980). The power ΔFP measured for the vowel phonation approximates the part of the flow power consumed by the self-oscillating vocal folds. If ΔFP for phonation with the tube is negative, it implicates that the part of the flow power consumed by the self-oscillating vocal folds is smaller.

The flow power ΔFP, evaluated from the area inside the loops, is the highest for phonation on the vowel [u:], for which $\Delta FP \cong 0.8$–7 mW, while in the case of phonation through the tubes ΔFP is lower and mostly negative with a minimum $\Delta FP \cong$-8 mW for the tube phonation into air, where the difference F_1-F_0=14-15 Hz between the formant frequency F_1=105 Hz and the frequency F_0=91 Hz was minimal (see Figures 2b and 5, and Table 1). The flow power ΔFP for the tube phonation into water is negative having the minimum $\Delta FP \cong$-6 mW at the highest input flow power used in the experiments for tube phonation into water where the difference between the formant frequency F_1=28 Hz and the bubbling frequency F_b was ca. 4-8 Hz (see Figures 2c and 5, and Table 1).

The experiments on vowel phonation and on tube phonations were performed in only slightly overlapping intervals of steady input flow power, because of substantial lowering of the PTF caused by the tube. Thus in only one case shown in Figure 10, it was approximately possible to compare directly the power ΔFP for phonation on vowel with that for phonation into water. The steady input flow power for the vowel was 159 mW and for the tube in water it was 150 mW, which are close values. The power ΔFP decreased from 0.8 mW for vowel to -4.2 mW for tube in water. This shows that 5 mW of flow power needed for vocal folds' self-sustained vibration was saved.

Figure 10. Somewhere here.

DISCUSSION

An important finding in the present study is that for phonation through the resonance tube in air, not only F_1 decreased, but the fundamental phonation frequency F_0 also decreased slightly below F_1 (see Table 1). Similarly, for phonation through the tube into water, F_1 decreased further, slightly above the water bubbling frequency F_b. When F_1 is close to F_0 or F_b, it assists the VFs' vibration, which enables easier phonation utilizing the resonance effect of the acoustic system. In practice, it is possible *to tune the system in order to gain maximum support for phonation* when phonating through the tube either into air or into water. In the former case it can be accomplished by changing F_0, while in the latter case by changing F_b.

The relationships $P_{trans}(t)$ vs. $GA(t)$ constructed for vowel [u:] resulted in clockwise-oriented cyclic graphs (loops), similar to those measured by Mišun et al. (2011) for artificial VFs without a VT, while the loops obtained for both therapy methods are more complicated. They show that less airflow energy is needed for phonation because the first acoustic resonance at the frequency F_1 is closely above the excitation frequency F_0 or F_b. This effect reduces the airflow energy needed for the VFs' vibration and the excitation of acoustic waves.

In principle the same situation happens, if the fundamental frequency F_0 of ordinary vowel phonation gets close to the first formant frequency F_1 of the vocal tract and increases the inertance of the supraglottic spaces (Titze, 2001).

Comparison of results from physical modeling, computer modeling, and measurements in humans

The experimental results presented in this paper for phonation through tubes into air and into water can be compared to measurements on humans published by Horáček et al. (2017a) for a male subject and by Radolf et al. (2014) for a female subject. The main difference between the results is caused by the existence of the low frequency acoustic-mechanical resonance in the human VT, whereas in the physical model with hard VT walls, such a resonance does not exist.

The pressure and frequency data measured here on the physical model are compared with measurements on humans in Table 1, where the data from the model are carefully chosen according to the same magnitudes of the mean subglottic pressure P_{sub} as was measured in humans. From the measurement using the model, the values of the mean oral pressure P_{oral}, peak-to-peak oral pressure

$P_{oral\_p\text{-}t\text{-}p}$, the water bubbling frequency F_b, and the fundamental frequencies were interpolated from the data given by the graphs shown in Figures 3 and 5, considering the same subglottic pressure P_{sub} as measured in humans. The acoustic resonance frequencies computed for the mathematical models with yielding and hard walls are also included in Table 1 for comparison with the measured data.

Good agreement was found between the measurements on human subjects and the experiments with the physical model for the mean oral pressure P_{oral}. However, the peak-to-peak fluctuations of the oral pressure $P_{oral\_p\text{-}t\text{-}p}$ measured in the model are far higher than in humans. The reason for this is that the yielding walls of the human VT may cause a higher damping of the pressure fluctuations, and that the difference between the first resonance frequency and the pressure fluctuation frequency, was smaller in the model.

Really, for phonation through the tube with the distal end in air, the differences F_1-F_0≅90 Hz for the male and F_1-F_0≅34 Hz for the female are evidently higher than the differences F_1-F_0≅14–15 Hz for the measurement with the VT model (see Table 1). Therefore, the acoustic-structural coupling was stronger in the VT model, because the excitation frequency F_0 was nearer to the formant frequency F_1, and consequently the amplitudes of oral pressure fluctuations were higher in the model than in humans. The acoustic-mechanical resonance at the computed frequency $F_{a\text{-}m}$≅27 Hz is not so important in vocal exercises using the resonance tube with the distal end in air, because the frequency $F_{a\text{-}m}$ is considerably below the lower limit of F_0 for the ordinary human voice (at least when vocal fry is not considered). However, at least in theory, the mechanical resonance associated with the yielding walls may raise the computed value of F_1 for the hard VT from 97 Hz to 181 Hz, which comes close to F_1=190 Hz measured in humans (see Table 1). This in turn would be beneficial, especially for the female speaking voice, since F_0 could be easily set close to F_1.

In phonation into water the $F_{a\text{-}m}$ becomes important. For the model, the VF vibrations are supported by water bubbling, because the difference F_1-F_b ≅4–8 Hz between the first formant frequency and the bubbling frequency is very small (see Table 1), while the fundamental phonation frequency F_0≅79–82

Hz is far from the acoustic resonance at $F_1 \cong 28$ Hz. We remark that the vibration amplitude of the VF in the vertical direction should be essential at the bubbling frequency. It should be proved in a next study, e.g. using laser vibrometry.

In humans, the bubbling frequency $F_b \cong 11.5$–14 Hz was considerably lower than in the model, where $F_b \cong 20$–24 Hz (see Table 1). The computed frequency of the acoustic-mechanical resonance $F_{a\text{-}m} \cong 9$ Hz is very close to the bubbling frequency F_b measured in humans. This means that in humans, the acoustic-mechanical resonance can be excited by the water bubbling because the difference of only 2–5 Hz can be theoretically estimated between $F_{a\text{-}m}$ and F_b (see Table 1). Moreover, in the case of a female subject, also the first acoustic resonance at the frequency $F_1 \cong 150$ Hz was excited by the very close fundamental frequency $F_0 \cong 149$ Hz, which is also close to the acoustic resonance frequency $F_1 \cong 143$ Hz computed for the VT with the yielding walls (see Table 1). This means that for the female subject, the effect of the water voice therapy can be doubled, because F_b can be close to $F_{a\text{-}m}$, and at the same time F_0 is close to F_1. It is worth noting that according to the measurements on a human subject the amplitudes of the bubbling induced vibrations of the larynx were higher than those caused by the fundamental frequency of phonation (Laukkanen et al., 2018).

From a comparison of the results obtained from the model with hard VT walls and the measurements made in humans, it can be concluded that the effects of tube phonation in humans and in the model are the same, but the situation is more complicated in humans due to the existence of the low frequency acoustic-mechanical resonance. The results show that optimized conditions for a maximal utilization of the resonances for supporting phonation into the tubes could be found. It would be necessary to tune the system. Tuning can be accomplished by changing several parameters:

- the inner diameter and length of the tube following the theoretical results (Horáček et al., 2017a), since the tube length and diameter affect the first resonance;
- the fundamental frequency of phonation, within physiological limits, for males and females;
- the airflow rate which changes the bubbling frequency, as shown in Figure 5.

The most problematic seems to be to change the bubbling frequency F_b to a wider frequency range covering the acoustic-mechanical resonance of the VT (ca. 9 Hz), because an exact theoretical modeling of water bubbling is excluded and only special experiments for chosen tubes can be carried out, as recently presented in Wistbacka et al. (2017), where for a hard-walled VT, the F_b increased from ca. 6 Hz at $Q \cong 0.001$ l/s to ca. 21 Hz at $Q \cong 0.04$ l/s. It is known that the bubbling frequency increases with the flow rate very quickly from zero for low flow rates to a limit value over which is not possible to cross (see Figure 5 and Davidson & Amick, 1956).

General discussion

Our modeling results are in line with the computations of Story et al. (2000) for humans when it comes to the lowering of F_1 with an artificial prolongation of the VT by a tube in air. Our F_1 values were lower when a hard-walled VT was considered. However, when the values were calculated considering a yielding-walled VT, as Story et al. (2000) did, our results (F_1=181 Hz) for tube in air are roughly similar to the results by Story et al. for a 30 cm tube with a ca. 1 cm inner diameter (F_1 about 210 Hz). Their results for the bilabial plosive [b] (F_1 about 170 Hz,) in turn, resemble our results for the tube in water (F_1=143 Hz; see Table 1). The differences can be caused by the different VT and slightly different tube length and inner diameter. As far as we know, the present study and the study by Horáček et al. (2017a) are the first to calculate F_1 for a water resistance exercise showing the effect of the low frequency acoustic-mechanical resonance in the VT.

The P_{sub} and airflow rates considered in the physical modeling were within the range reported for humans (Baken & Orlikoff, 2000), and they remained in the same region as the one that has been found in humans during tube exercises (Tyrmi et al., 2017; Radolf et al., 2014). The bubbling frequency F_b=19–24 Hz in phonation through a tube into water was close to the 21 Hz reported by Wistbacka et al. (2017) for a hard-walled VT. Granqvist et al. (2015) measured F_b in humans in the range 10.5–12.7 Hz and Tyrmi et al. (2017) in the range 15.8-16.5. The results agree with our finding that the yielding walls in the human VT lower the bubbling frequency (see Table 1).

The analysis of the measured relationships $P_{trans}(t)$ vs. $GA(t)$ showed the principle of supporting the self-sustained VF oscillations by a near coincidence of the first resonance frequency with the fundamental or bubbling frequency in both cases of phonation into the tube. The measured relationships showed that when F_1 was closer to F_0 in the tube phonation, less airflow energy was needed for the VFs' vibration. Thus, the results are in line with the computer modeling results presented by Titze (2006) and Titze & Laukkanen (2007), suggesting improved vocal efficiency with the tubes.

The results of our study are also in line with the computer modeling results of Fant & Lin (1987) and Titze (1988) as well as the results of Rothenberg (1986) for a female singer, which showed that the condition where F_0 and F_1 are close improves the VFs' vibration. A further study is warranted to quantify, how much the flow energy reduction and thus improvement of VF vibration depends on the difference between the fundamental phonation frequency and the first resonance frequency for phonation through the tube into air, and on the difference between the bubbling frequency and the first acoustic (or the acoustic-mechanical) resonance frequency of the vocal tract for phonation through the tube into water.

Hypothetically, it is possible to obtain an efficiency maximum in tube therapy by varying (a) the fundamental phonation frequency, (b) the lowest acoustic resonance frequencies using different tube lengths and inner diameters, and (c) the bubbling frequency (through changing the airflow rate). A question of interest is whether the enhancement of VF vibration may increase the mechanical loading imposed on the VFs. According to the results of the present study, for phonation through the tube into air or into water, PTP, PFT and P_{trans} decreased, while MADR did not increase, which suggests that the mechanical loading of the VFs did not increase. These results are in line with the previous modelling study that showed that the impact stress was even smaller during water resistance therapy compared to vowel phonation (Horacek et al. 2018).

On the other hand, Guzman et al. (2017) have found in a high-speed study with humans that in some cases of phonation through a tube into water (especially when the immersion depth is deep, 10-18 cm), there is an increased amplitude-to-VF length ratio (substitute of maximum glottal amplitude), closing quotient (substitute of MADR), and glottal spectral flatness; all of these changes suggesting increased impact stress in phonation. Similarly, the high-speed results reported by Laukkanen et al. (2007) showed a trend that in phonation with a longer tube compared to a shorter one, the P_{sub} was higher and glottal open time shorter (i.e. closed time longer). These observations seem to imply increased phonatory effort as a compensation for increased supraglottic resistance. In the present study, the method used was modeling. It is possible that in some cases humans increase laryngeal adduction and P_{sub}, so that the impact stress may rise somewhat; on the other hand, the human high-speed imaging data did not include data for P_{sub} and airflow. The results of the present study suggest that maximum glottal area and the area derivative need a reference to P_{sub} and airflow for drawing conclusions regarding impact stress. Furthermore, as our results show, maximum area declination rate is not a correct parameter for estimating the impact stress; instead, the area declination rate just prior to vocal fold collision should be studied in future.

In any case, during tube voice therapy, the subject him-/herself should avoid phonation that may potentially be harmful due to exaggerated enhancement of VF vibration near the first resonance or in water resistance therapy when there is a double effect near the first two resonances (F_1 and F_{a-m}). Such a situation is likely to be felt as unpleasantly strong vibrations in the vocal tract and larynx. This can be expected especially for loud phonation with high subglottic pressure.

CONCLUSIONS

The main findings from the modeling of the voice therapy based on phonation through a tube into air and into water can be summarized as follows.

1) Glottal area variation measured simultaneously with the variation of the transglottic pressure showed that the airflow energy needed for VF vibration and the excitation of acoustic waves can be

noticeably reduced by phonation through a resonance tube into air and into water compared to phonation on the vowel [u:]. In both voice therapy methods, part of the airflow energy required for phonation is substituted utilizing the first acoustic or acoustic mechanical resonance.

2) The basic principle in vocal exercises with a resonance tube with the distal end in air and in water (water resistance therapy) is the same. For phonation into air, the fundamental phonation frequency excites the acoustic resonance at the first formant frequency, and for phonation into water, the bubbling frequency excites a low frequency acoustic mechanical resonance in the human VT with yielding walls.

3) The maximum glottal area time derivative during the glottal closure (so-called maximum area declination rate; MADR) is considered a measure of the maximum impact stress between the colliding VFs. However, it was found in the present study that the glottal closing speed was noticeably reduced just before the VFs' collision. Therefore, MADR does not seem to be an adequate parameter for impact stress estimation. The glottal closing speed should be investigated in detail in a future study.

Acknowledgments

This study was supported by a grant from the Czech Science Foundation: No. 16-01246S "Computational and experimental modelling of self-induced vibrations of vocal folds and influence of their impairments on human voice."

References

Amarante Andrade, P., Wistbacka, G., Larsson, H., Södersten, M., Hammarberg, B., Švec J. G., & Granqvist S. (2016). The Flow and Pressure Relationships in Different Tubes Commonly Used for Semi-occluded Vocal Tract Exercises. *Journal of Voice*, 30(1), 36–41.

Baken, R. J., & Orlikoff, R. F. (2000). *Clinical measurement of speech and voice* (2nd ed.). San Diego, CA: Singular/Thomson Learning.

Carroll, L. M., & Sataloff, R. T. (1991). The singing voice. In R. T. Sataloff (Ed.): *Professional voice: The science and art of clinical care* (pp. 381-383). New York, NY: Raven Press Ltd.

Davidson, L., & Amick, E. H. JR. (1956). Formation of gas bubbles at horizontal orifices. *A.I.CH.E. Journal*, 2(3), 337-342.

Fant, G., & Lin, Q. (1987) Glottal source –vocal tract acoustic interaction. *Quarterly Progress and Status Report* STL-QPSR, 1, 13-27.

Granqvist, S., Simberg, S., Hertegård, S., Holmqvist, S., Larsson, H., Lindestad, P. A., Södersten, M., & Hammarberg, B. (2015). Resonance tube phonation in water: High-speed imaging electrographic and oral pressure observations of vocal fold vibrations – a pilot study. *Logopedics Phoniatrics Vocology*, 40, 113-121.

Guzman, M., Laukkanen, A. M., Traser, L., Geneid, A., Richter, B., Muñoz, D.,& Echternach, M. (2017). Effect of water resistance therapy on vocal fold vibration: A high speed registration study. *Logopedics Phoniatrics Vocology*, 42(3), 99-107.

Hampala, V., Laukkanen, A. M., Guzman, M., Horáček, J., & Švec, J. G. (2015). Vocal fold adjustment caused by phonation into a tube: A double-case study using computed tomography. *Journal of Voice*, 29(6), 733-742.

Horáček, J., Radolf, V., Bula, V., & Laukkanen, A. M. (2014). Air-pressure, vocal folds vibration and acoustic characteristics of phonation during vocal exercising. - Part 2: Measurement on a physical model. *Engineering Mechanics*, 21(3), 193-200.

Horáček, J., Bula, V., Košina, J., & Radolf, V. (2016a). Phonation characteristics of self-oscillating vocal folds replica with and without the model of the human vocal tract. In I. Zolotarev & V. Radolf (Eds.) *Engineering Mechanics 2016* (pp. 214-217), Prague: Institute of Thermomechanics of the CAS, v.v.i., 2016, http://www.engmech.cz/2016/im/im/page/proc. ISBN 978-80-87012-59-8, ISSN 1805-8248.

Horáček, J., Bula, V., Radolf, V., Vampola, T., & Dušková, M. (2016b). Development of self-oscillating human vocal folds prosthesis. *Procedia Engineering*, 144, 867-874. ISSN 1877-7058.

Horáček, J., Radolf, V., & Laukkanen, A. M. (2017a). Low frequency mechanical resonance of the vocal tract in vocal exercises that apply tubes. *Biomedical Signal Processing and Control*, 37, 39–49.

Horáček, J., Radolf, V., Bula, V., & Košina, J. (2017b). Experimental modelling of phonation using artificial models of human vocal folds and vocal tracts. In V. Fuis (Ed.), *Engineering Mechanics 2017* (pp. 352-385), Brno: University of Technology, 2017, http://www.engmech.cz/2017/im/im/page/proc, ISBN 978-80-214-54972-2, ISSN 1805-8248.

Horáček, J., Radolf, V., & Laukkanen, A. M. (2018) Impact stress in water resistance voice therapy: A physical modeling study, *Journal of Voice*, Available online 5 June 2018 https://doi.org/10.1016/j.jvoice.2018.01.025.

Laukkanen, A. M. (1992). About the so called "resonance tubes" used in Finnish voice training practice. An electroglottographic and acoustic investigation on the effects of this method on the voice quality of subjects with normal voice. *Scandinavian Journal of Logopedics and Phoniatrics* 17, 151-161.

Laukkanen, A. M., Pulakka, H., Alku, P., Vilkman, E., Hertegård, S., Lindestad, P. Å, Larsson, H., & Granqvist, S. (2007). High-speed registration of phonation-related glottal area variation during artificial lengthening of the vocal tract. *Logopedics Phoniatrics Vocology*, 32, 157-164.

Laukkanen, A. M., Geneid, A., Bula, V., Radolf, V., Horáček, J., Ikavalko, T., Kukkonen, T., Kankare, E., & Tyrmi, J. (2018). How much loading does water resistance voice therapy impose on the vocal folds? An experimental human study. Journal of Voice, Vol.&&, No.&&, pp.1−11, https://doi.org/10.1016/j.jvoice.2018.10.011.

Liljencrants, J. (1985). *Speech Synthesis with a Reflection-Type Line Analog* (Doctoral dissertation). Royal Institute of Technology, Stockholm. (Google).

Mathews, J. H., & Fink, K. K. (2004). *Numerical methods using matlab* (4th ed.). Prentice-Hall Inc. Upper Saddle River, New Jersey, USA.

Mišun, V., Švancara, P., & Vašek, M. (2011). Experimental analysis of the characteristics of artificial vocal folds. *Journal of Voice*, 25(3), 308-318.

Mori, H., Ohsawa, H., Tanaka, T.H., Taniwaki, E., Leisman, G., & Nishijo, K. (2004). Effect of massage on blood flow and muscle fatigue following isometric lumbar exercise. *Medical Science Monitor*, 10(5), CR173–8.

Radolf, V., Laukkanen, A. M., Horáček, J., & Liu, D. (2014). Air-pressure, vocal fold vibration and acoustic characteristics of phonation during vocal exercising. - Part 1: Measurement in vivo. *Engineering Mechanics*, 21(1) 53-59. ISSN 1802-1484.

Radolf, V., Horáček, J., Dlask, P., Otčenášek, Z., Geneid, A., & Laukkanen, A. M. (2016). Measurement and mathematical simulation of acoustic characteristics of an artificially lengthened vocal tract. *Journal of Sound and Vibration*, 366, 556-570.

Rothenberg, M. (1981). An interactive model for the voice source. *Quarterly Progress and Status Report* STL-QPSR, 4.

Rothenberg, M. (1986). Cosi fan tutte and what it means or nonlinear source-tract acoustic interaction in the soprano voice and some implications for the definition of vocal efficiency. In T. Baer, C. Sasaki & K. S. Harris (Eds.), *Vocal fold physiology: laryngeal function in phonation and respiration* (pp.254-263). Colle-eHill Press, San Diego.

Rothenberg, M. (1988). Acoustic reinforcement of vocal fold vibratory behavior in singing. In: O. Fujimura (Ed.) *Vocal fold physiology, volume 2. Vocal physiology, Voice production, Mechanisms and Functions* (pp.379-89). Raven Press, New York.

Schutte, H. (1980). *The efficiency of voice production*. Groningen: State University Hospital.

Sihvo, M. (2006). Terve ääni, äänen hoidon A B C. [Healthy Voice. The A B C for Voice Care], Helsinki: Kirjapaja.

Sihvo, M. (2017). History of the Lax Vox – tube exercise. LAP Lambert Academic Publishing.

Sihvo, M., & Denizoglu, I. (2007). Lax Vox: Voice Therapy Technique. Turkey: AD Izmir;. http://laxvox.com/documents/LAXVOX%20handouts.pdf.

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

Sovijärvi, A. (1969). Nya metoder vid behandling av röstrubbningar. [New methods for treating voice disorders], *Nordisk Tidskrift för Tale of Stemme* 3,121–131.

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

Titze, I. R. (1988). A framework for the study of vocal registers. *Journal of Voice*, 2(3), 183–194.

Titze, I. R. (2001). Acoustic interpretation of resonant voice. Journal of Voice, 15(4), 519–528.

Titze, I. R., Finnegan, E. M., Laukkanen, A. M., & Jaiswal, S. (2002). Raising lung pressure and pitch in vocal warm-ups: The use of flow-resistance straws. *Journal of Singing*, 58, 329-338.

Titze, I. R. (2006a). Theoretical analysis of maximum flow declination rate versus maximum area declination rate in phonation. *Journal of Speech, Language, and Hearing Research*, 49, 439-447.

Titze, I. R. (2006b). Voice training and therapy with a semi-occluded vocal tract: rationale and scientific underpinnings. *Journal of Speech, Language, and Hearing Research*, 49, 448-459.

Titze, I. R., & Laukkanen, A. M. (2007). Can vocal economy in phonation be increased with an artificially lengthened vocal tract? A computer modeling study. *Logopedics Phoniatrics Vocology*, 32(4), 147-156.

Tyrmi, J., Radolf, V., Horáček, J., & Laukkanen, A. M. (2017). Resonance tube or lax vox? *Journal of Voice*, 31(4), 430-437, http://dx.doi.org/10.1016/j.jvoice.2016.10.024.

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.

Wistbacka, G., Amarante Andrade, P., Simberg, S., Hammarberg, B., Södersten, M., Švec, J. G., & Granqvist, S. (2017). Resonance tube phonation in water—the effect of tube diameter and water depth on back pressure and bubble characteristics at different airflows. *Journal of Voice*, 12 p. http://dx.doi.org/10.1016/j.jvoice.2017.04.015.

FIGURES

Figure 1. Schema of the measurement set-up for the experimental study on the model.

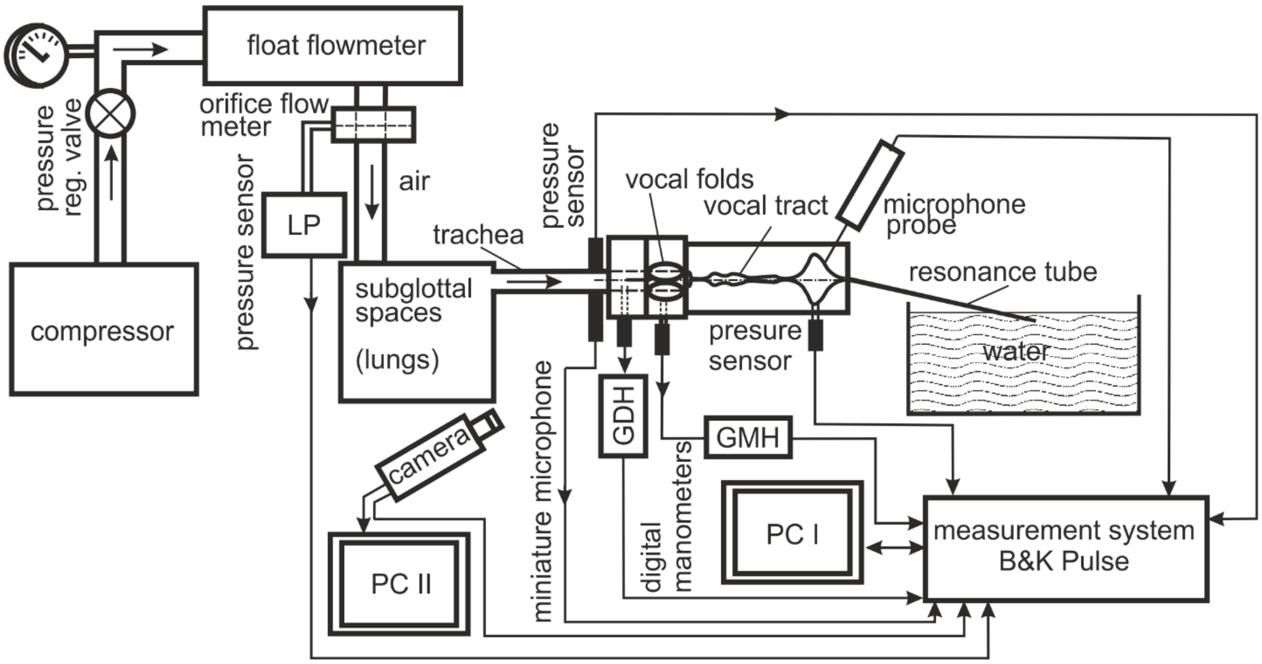

Figure 2. Sound pressure spectrum levels of the measured oral pressure for simulated phonations: a) on the vowel [u:] (F_0=111 Hz, F_1≅315 Hz, Q=0.20 l/s), b) through the tube into air (F_0=90 Hz, F_1≅105 Hz, Q=0.04 l/s), and c) through the tube into water with a detail in the lowest frequency range (F_b=22 Hz, F_1≅28 Hz, F_0=79 Hz, Q=0.04 l/s). Figure 2 d) shows the example of the spectrum of the subglottic pressure (measured for the same phonation through the tube into water, F_{sub1}≅720 Hz). The formant frequencies F_1 were estimated from the peaks of the averaged spectra marked by thick lines.

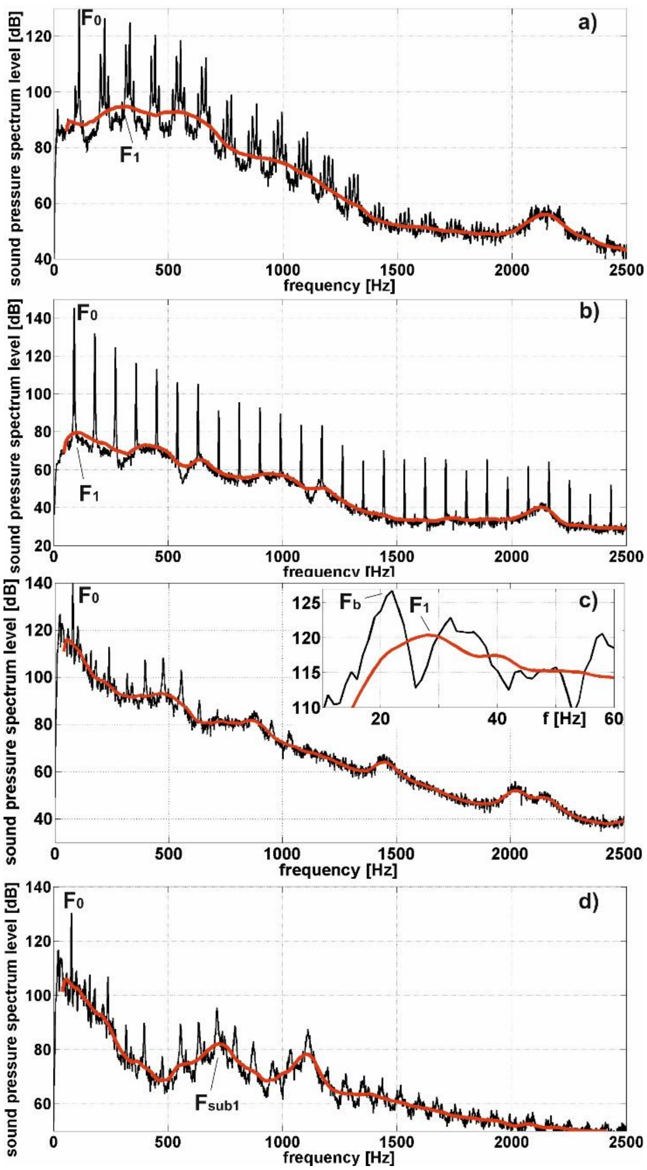

Figure 3. Measured values of: a) mean subglottic pressure (P_{sub}), b) peak-to-peak magnitudes of subglottic pressure ($P_{sub\_p\text{-}t\text{-}p}$), c) mean oral pressure (P_{oral}), and d) peak-to-peak magnitudes of oral pressure ($P_{oral\_p\text{-}t\text{-}p}$) depending on the flow rate Q for: 1) phonation on the vowel [u:], 2) phonation through the tube into air, 3) phonation through the tube into water, and 4) blowing through the tube into water without phonation.

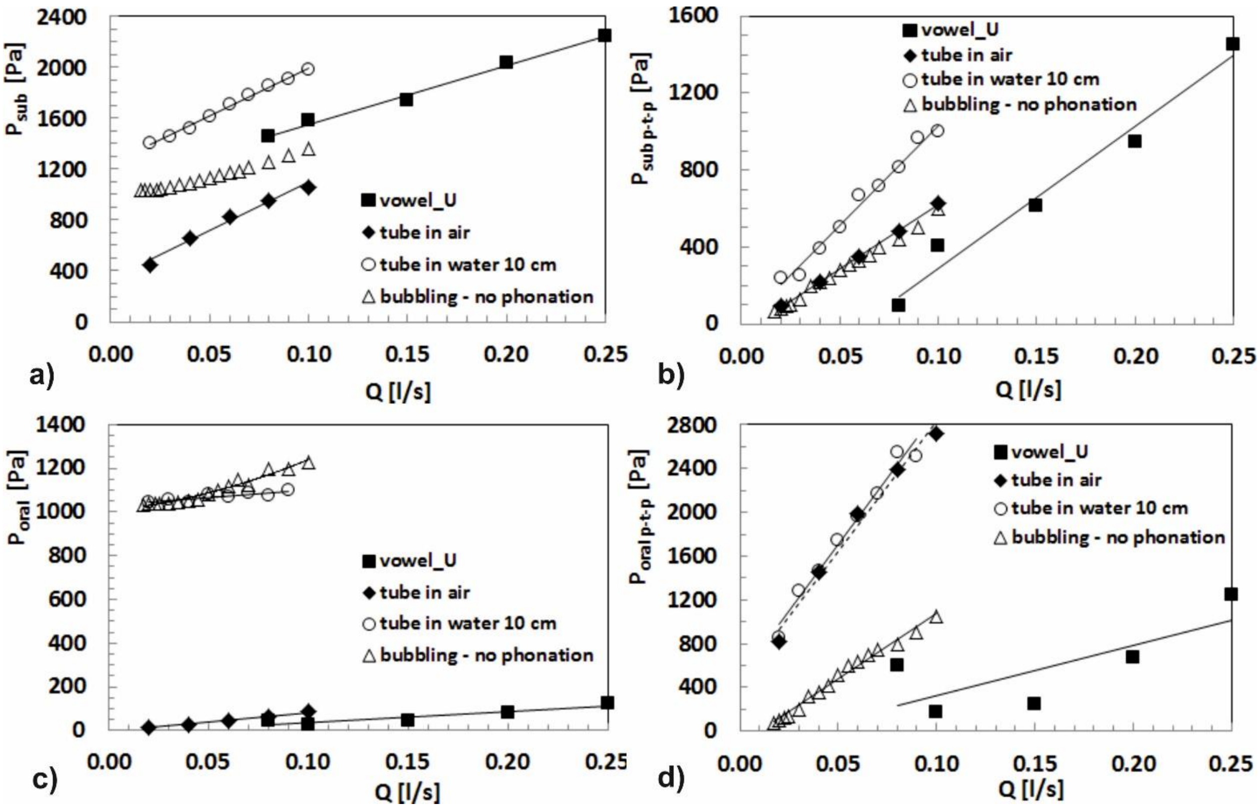

Figure 4. Measured maximum glottal area max$GA(t)$ and maximum glottal area derivative max(-d$GA(t)$/dt) evaluated during the closing phase of the glottis (MADR) as a function of the flow rate Q for: 1) phonation on the vowel [u:], 2) phonation through the tube into air, and 3) phonation through the tube into water.

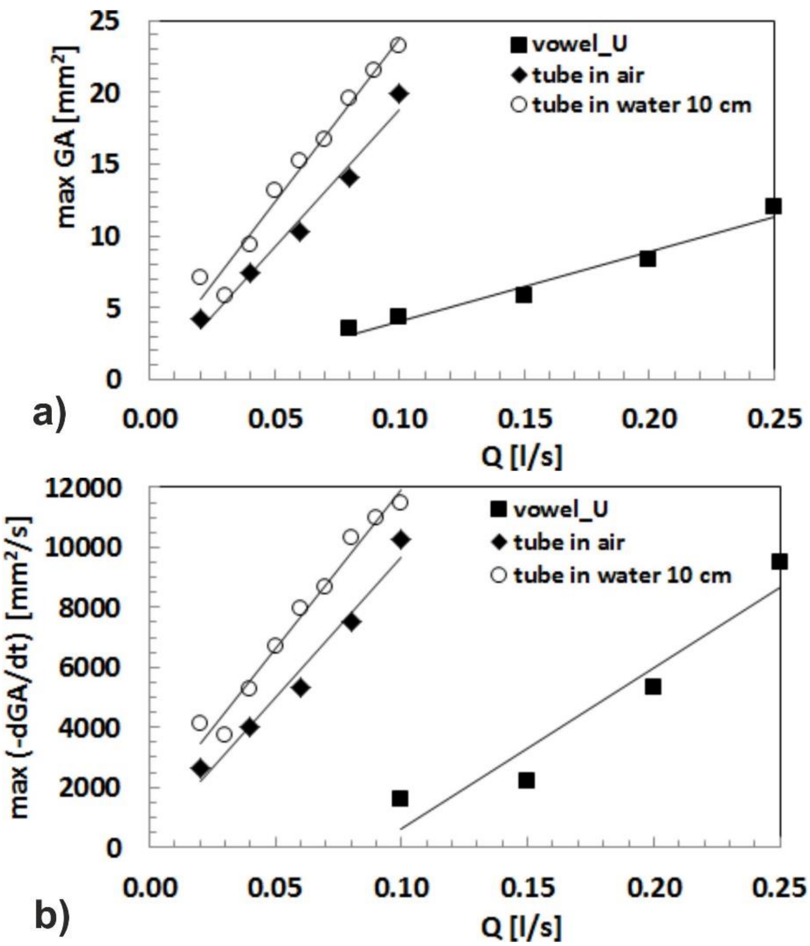

Figure 5. Measured fundamental frequency F_0 and the water bubbling frequency F_b depending on the flow rate Q for: 1) phonation on the vowel [u:], 2) phonation through the tube into air, 3) phonation through the tube into water, and 4) blowing through the tube into water without phonation.

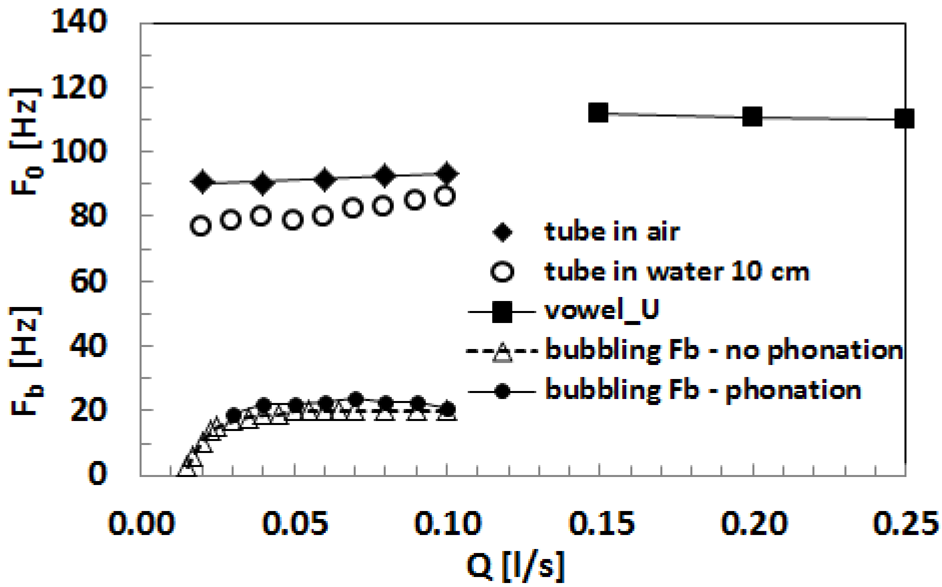

Figure 6. Example of the construction of the relationships $P_{trans}(t)$ vs. $GA(t)$ and $P_{trans}(t)$ vs. d$GA(t)$/dt for phonation on vowel [u:] for one period of the vocal folds' self-oscillation showing: glottal area $GA(t)$ (upper panel) and glottal area derivative d$GA(t)$/dt (2nd panel), transglottic pressure $P_{trans}(t)$ (3rd panel), and the resulting relationships (lower panel) (Q=0.20 l/s, F_0=111 Hz, ΔA=0.031 mJ, ΔFP=ΔA F_0=3.42 mW).). The numbers along the waveform $GA(t)$ indicate the important time instants marked also on the relationships and on the high-speed VF images shown at the bottom.

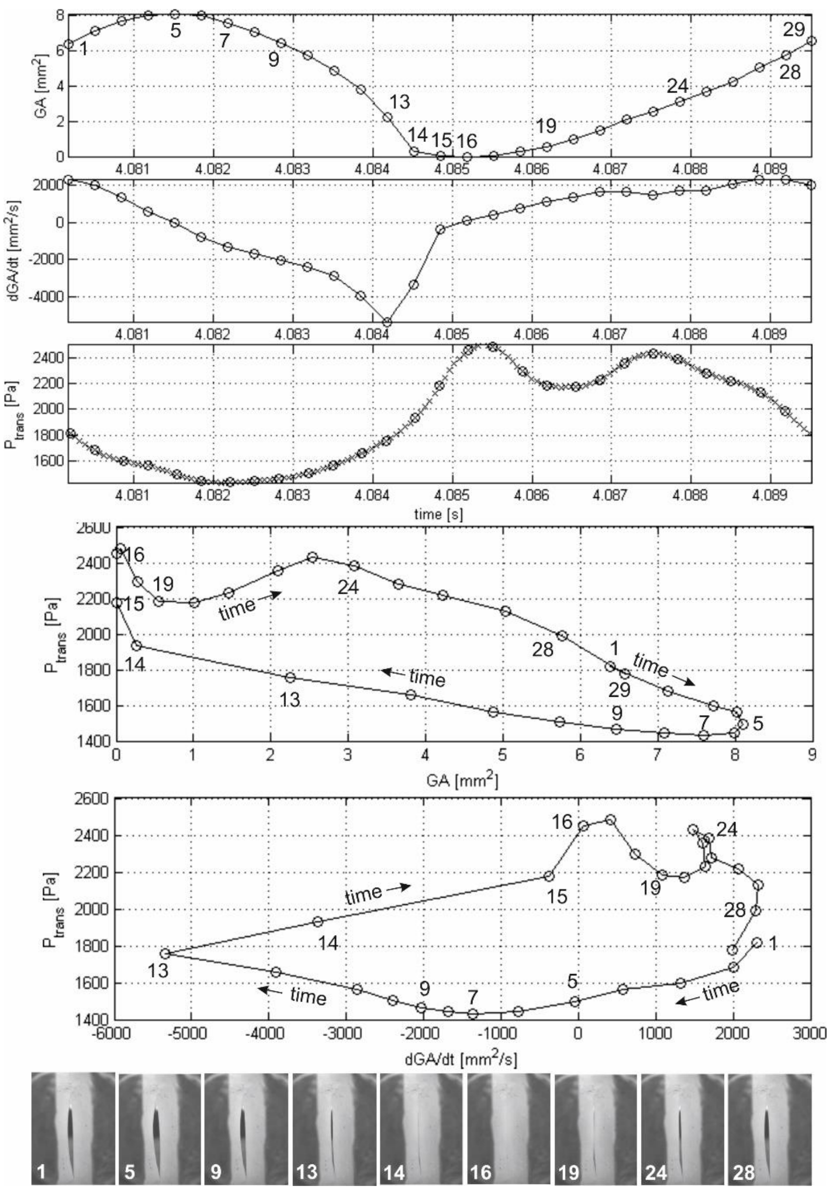

Figure 7. Example of the construction of the relationships $P_{trans}(t)$ vs. $GA(t)$ and $P_{trans}(t)$ vs. d$GA(t)$/dt for phonation through the resonance tube into air for one period of the vocal folds' self-oscillation showing: glottal area $GA(t)$ (upper panel), glottal area derivative d$GA(t)$/dt (2nd panel), transglottic pressure $P_{trans}(t)$ (3rd panel), and the resulting relationships (lower panel), with the images of the vibrating vocal folds also shown at the numbered time instants (Q=0.08 l/s, F_0=94 Hz, ΔA=-0.054 mJ, $\Delta FP=\Delta A\ F_0$ =-5.07 mW). The numbers along the waveform $GA(t)$ indicate the important time instants marked also on the relationships and on the high-speed VF images shown at the bottom.

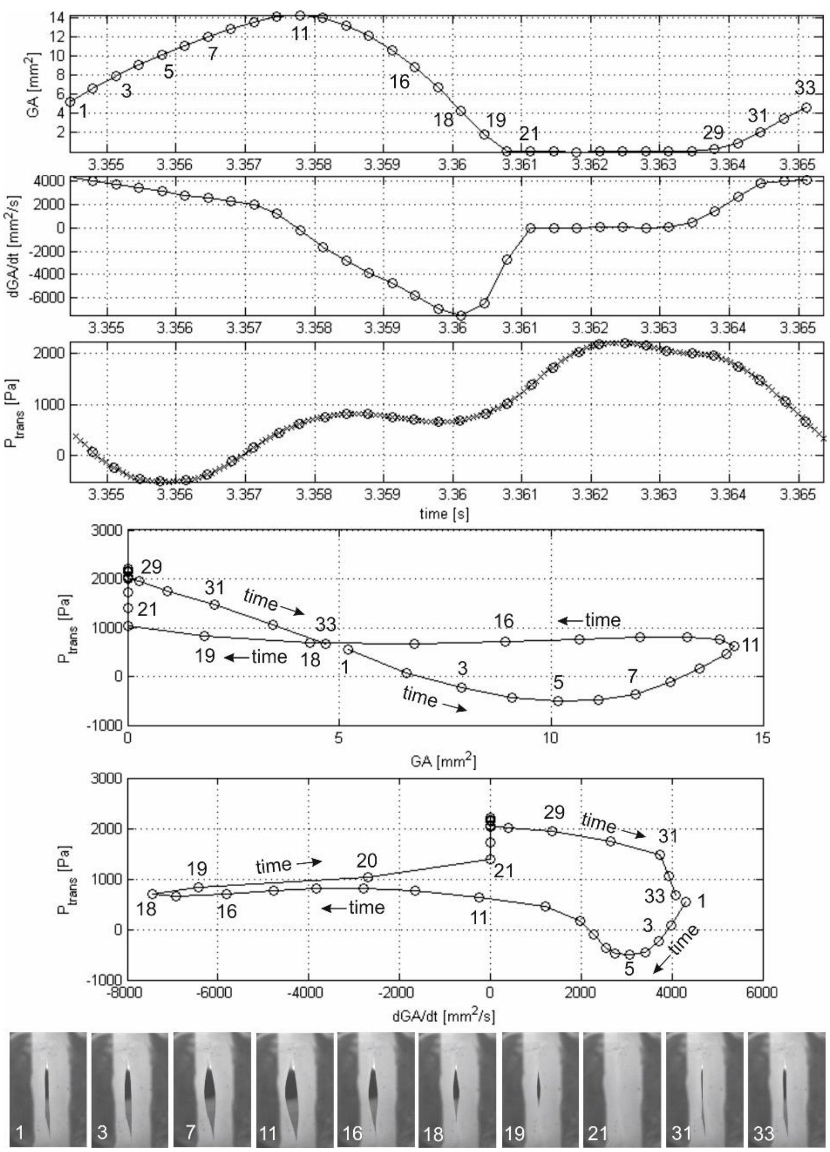

Figure 8. Example of the construction of the relationship $P_{trans}(t)$ vs. $GA(t)$ for phonation on the vowel [u:] when the VT is prolonged by the resonance tube with the distal end submerged 10 cm in water for one period of the vocal folds' self-oscillation showing: glottal area $GA(t)$ (upper panel), transglottic pressure P_{trans} (t) (middle panel), and the resulting relationship P_{trans} (t) vs. $GA(t)$ (lower panel), with the images of the vocal folds' vibration also shown at the numbered time instants (Q=0.08 l/s, F_0=83 Hz, F_b=23 Hz). The numbers along the waveform $GA(t)$ indicate the important time instants marked also on the relationships and on the high-speed VF images shown at the bottom.

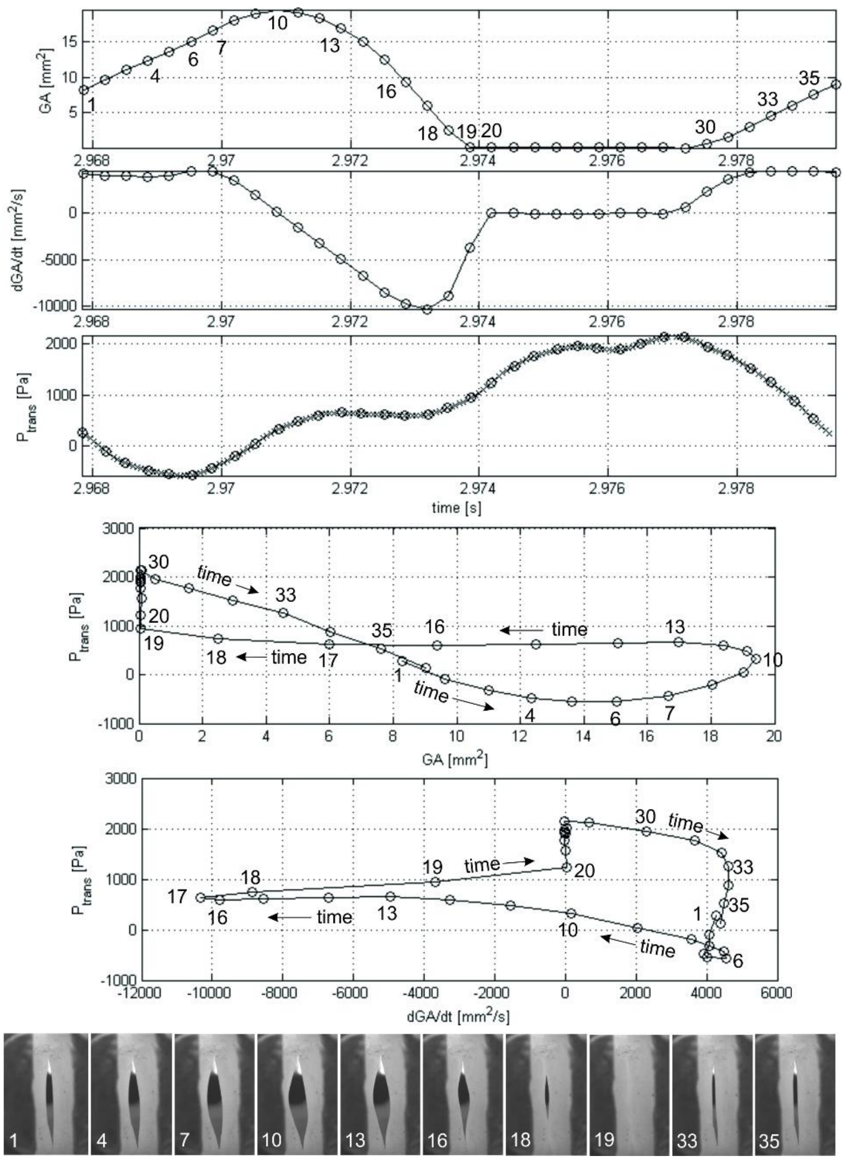

Figure 9. Example of construction of the relationships $P_{trans}(t)$ vs. $GA(t)$ and $P_{trans}(t)$ vs. d$GA(t)$/dt for phonation on the vowel [u:] when the VT is prolonged by the resonance tube with the distal end submerged 10 cm in water, for nine periods of the vocal folds' self-oscillation showing: glottal area $GA(t)$ (upper panel), glottal area derivative d$GA(t)$/dt (2nd panel), transglottic pressure $P_{trans}(t)$ (3rd panel), and the resulting relationships (bottom) (Q=0.08 l/s, F_0=83 Hz, F_b=23 Hz, ΔA= -0.041 mJ, ΔFP=ΔA F_0 =-3.41 mW).

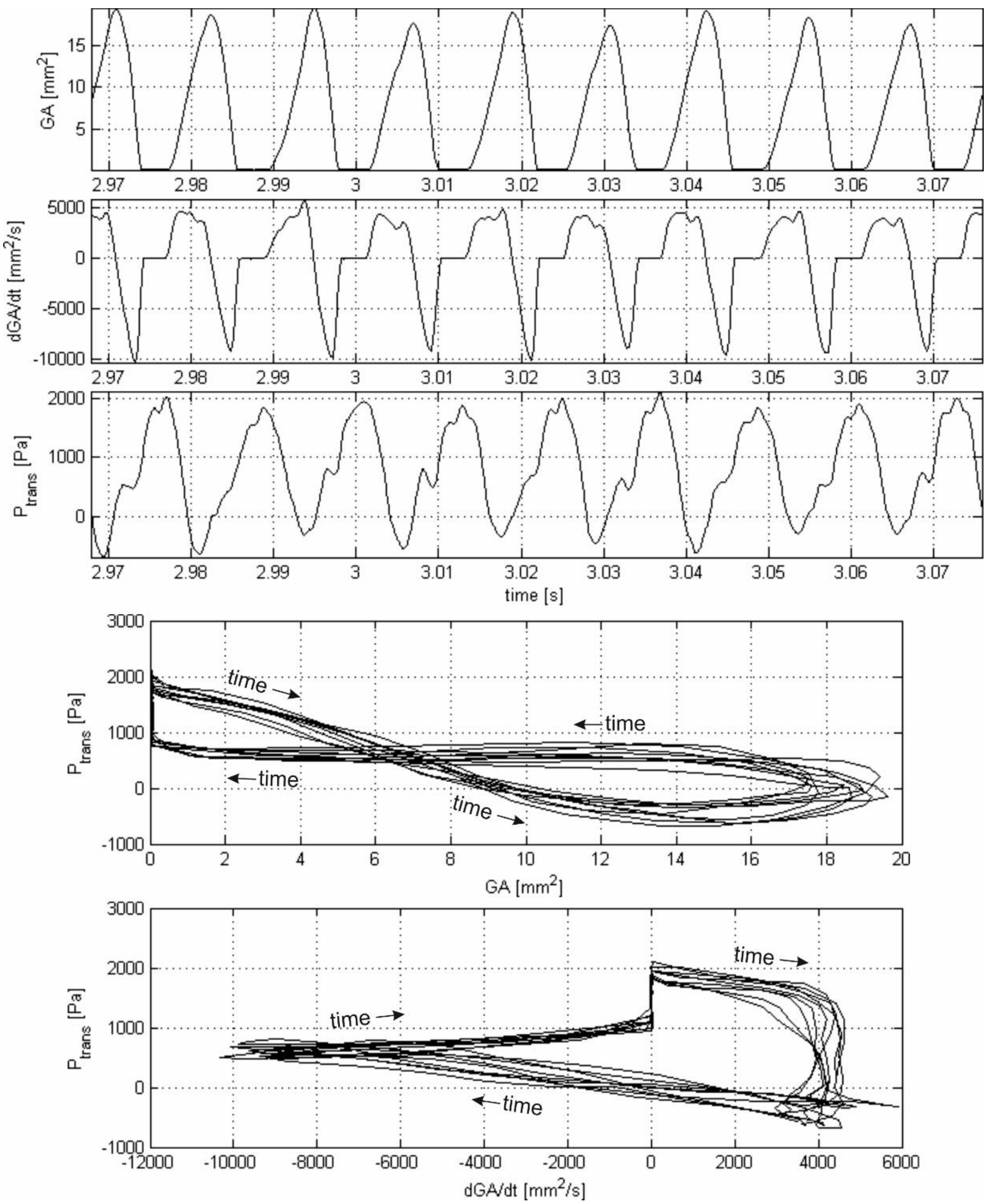

Figure 10. Measured part of the flow power $\Delta FP=\Delta A\ F_0$ consumed by the self-oscillating vocal folds for phonation on the vowel [u:] and its changes caused by phonation through the tube into air and into water in dependence on the input steady flow power defined as mean subglottic pressure times mean airflow rate ($P_{sub}\ Q$).

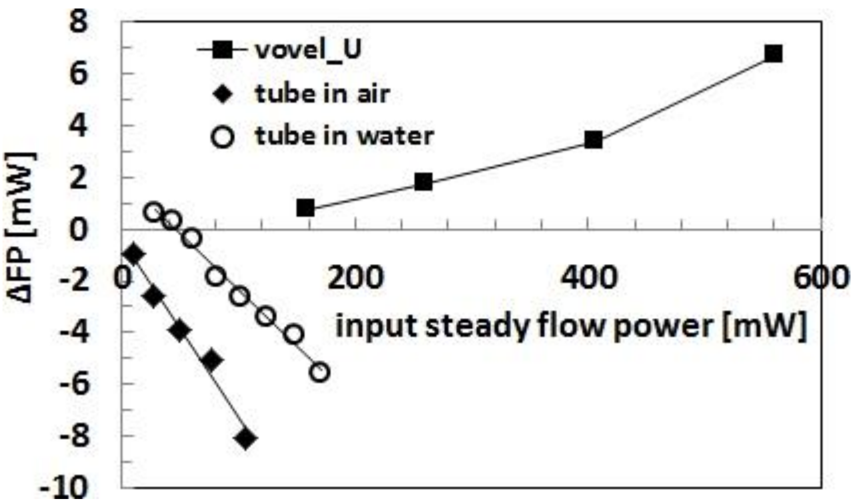

Table 1. Comparison of the mean subglottal pressure (P_{sub}), mean oral pressure (P_{oral}), peak-to-peak oral pressure ($P_{oral\_p\text{-}t\text{-}p}$), frequencies for water bubbling (F_b), fundamental frequencies (F_0), the first formant frequencies (F_1) and 3 dB frequency bandwidths (B_1) for phonations: a) on vowel [u:], b) into the resonance tube in air and c) with the distal end of the tube submerged 10 cm in water, measured on the model with hard vocal tract walls and in human subjects. Measurements are complemented with computational results for F_1 and the acoustic-mechanical resonance frequencies ($F_{a\text{-}m}$) for a vocal tract with hard walls and with yielding walls.

	P_{sub} [kPa]	P_{oral} [Pa]	$P_{oral\_p\text{-}t\text{-}p}$ [Pa]	F_b [Hz]	F_0 [Hz]	$F_{a\text{-}m}$ [Hz]	F_1 [Hz]	B_1 [Hz]	F_1-F_0 [Hz]	F_1-F_b [Hz]	F_b-$F_{a\text{-}m}$ [Hz]
Phonation on vowel [u:]											
RESULTS FROM THE PRESENT STUDY♦)	1.45-2.24	44-123	170-1240	/	111	/	316	220	205	/	/
Computer model with hard VT	/	/	/	/	/	/	333		/	/	/
Computer model with yielding VT	/	/	/	/	/	31	410		/	/	/
RESULTS FOR HUMANS											
Female subject**)	0.710	220	257	/	164	x	360		196	/	/
Phonation into the resonance tube in air											
RESULTS FROM THE PRESENT STUDY											
Physical model with hard VT***)	0.600	28	1284	/	90	/	105	99	15	/	/
	0.746	44	1740	/	91	/	105		14	/	/
Computer model with hard VT	/	/	/	/	/	/	97		/	/	/
Computer model with yielding VT	/	/	/	/	/	27	181		/	/	/
RESULTS FOR HUMANS											
Male subject*)	0.600	25	210	/	100	x	190		90	/	/
Female subject**)	0.746	47	429	/	156	x	190		34	/	/
Phonation into the resonance tube into water (10 cm)											
RESULTS FROM THE PRESENT STUDY											
Physical model with hard VT***)	1.800	1082	2271	24	82	/	28	25	-54	4	/
	1.480	1054	1166	20	79	/	28		-51	8	/
Computer model with hard VT	/	/	/	/	/	/	28		/	/	/
Computer model with yielding VT	/	/	/	/	/	9	143		/	/	≅2-5♦♦)
RESULTS FOR HUMANS											
Male subject*)	1.800	1019	391	11.5	98	x	179		81	168	x
Female subject**)	1.480	920	551	14	149	x	≅150		≅1	136	x

*) Horáček et al. (2017a, Table 3); **) Radolf et al. (2014, Tables 2 and 3 – normal phonation); ♦) see Figure 3;
***) measured data interpolated from the graphs in Figures 2, 3, and 5; ♦♦) related to results for humans; x… undefinable values.