Experimental modelling and human data of glottal area declination rate for vowel and semi-occluded vocal tract phonation

Jaromír Horáček[1], Vojtěch Radolf [1], Vítězslav Bula[1], Anne-Maria Laukkanen[2]

[1] Institute of Thermomechanics of the Czech Academy of Sciences (CAS), Prague, Dolejškova 1402/5, 182 00 Prague, Czech Republic, https://www.it.cas.cz/en/contacts/
[2] Speech and Voice Research Laboratory, Faculty of Social Sciences, Tampere University, Virta, Åkerlundinkatu 5, 33100 Tampere, Finland
jaromirh@it.cas.cz, radolf@it.cas.cz, bula@it.cas.cz, Anne-Maria.Laukkanen@tuni.fi

Abstract:
Objective: The aim of this study was to investigate how the maximum area declination rate (*MADR*) of the glottis corresponds to the maximum velocity of the self-oscillating vocal folds just before their collision. The maximum velocity before collision is closely related to impact stress in the colliding vocal folds and, thus, it is an important indicator of vocal loading.
Methods: High speed imaging data for a male subject and measurements of the glottal area waveforms performed on a physical 1:1 scaled model for phonation on vowel [u:] were compared with those obtained during a semi-occluded vocal tract which increases voice source-vocal tract interaction. Semi-occlusion was obtained by phonation through a glass resonance tube with the distal end in air or submerged 10 cm in water.
Results: The results show that the area declination rate of the glottis, i.e., closing velocity, before the glottal closure is substantially lower than the *MADR*, and also lower in semi-occluded vocal tract phonation through a tube into air and into water than in vowel phonation.
Conclusions: The results suggest that *MADR* is not a reliable estimate of impact stress, and that impact stress in tube therapy is lower than in vowel phonation.

Keywords: glottis closing velocity; area declination rate of the glottis; *MADR*; impact stress; voice therapy using tubes.

1. Introduction

Periodic mechanical stress is loading the human vocal folds (VFs) particularly due to their collisions during ordinary phonation in chest register where the VFs make full contact and close the glottis. The contact of the vocal folds results in a complicated, generally three-dimensional loading of the VFs tissue by three normal and three shear stresses. These mechanical stresses, when appearing in excessive amounts, are the principal reason for VFs' tissue damage, like nodules or polyps, to appear in the VFs' surface layer, which may cause serious voice problems

[1]. The measure of the loading is the mutual relative speed of the two VFs before their contact. We consider in the study only direct and not oblique impacts of the vocal folds.

Direct impact (contact) stress measurement has been performed on artificial silicon VF models (see e.g. [2-3]) and on excised canine or human larynges (see e.g. [4]) and sometimes even *in vivo* in humans [5], but it is difficult, unreliable and even in the best case not fully correct, because the pressure sensor between the VFs influences the VFs' vibration and the airflow in the glottis.

A noninvasive method to estimate the impact stress (*IS*) is to study the VFs' vibration by high-speed camera and to determine the closing velocity of the glottis. The maximum velocity of the glottal area when the VFs are approaching each other, the *MADR*, has been considered as a measure of the *IS* loading the VFs during collision [6], and thus, an important quantity in studying vocal economy of phonation [7]. Similarly, the maximum velocity of the glottal width during closing has been used for an estimate of *IS* [8].

However, in our two recently published papers, where we studied the glottal width and waveforms of the glottal area variation and their time derivatives *in vivo* [8] and on model [9], we noticed as a side effect that the speed of the glottis closure after *MADR* can be substantially reduced. Therefore, *MADR* does not seem to be an adequate parameter for impact stress estimation; instead, the area declination rate just prior to vocal fold collision should be studied.

We can note that according to the Hertz impact theory [10, 11] the maximum contact (impact) stress *IS* during the VFs' collision is approximately proportional to the velocity V_0 of glottal closing just before the complete closure, in line with the formula:

$$IS \approx const. V_0^{2/5} \quad , \tag{1}$$

where the constant is given by the material properties and the curvature of impacting surfaces. The Hertz theory was used in computer modelling of VFs' self-sustained vibration for estimation of impact stress between the colliding VFs [12]. The results were found to be in good agreement with measurements of *IS* on the artificial silicon VFs [2] as well as with results published on excised larynges [13]. It is known that the Hertz theory is approximately valid only for a very short time after the contact, i.e. before reflections of the stress waves on the edges of the VF influence the stress state in the contact region. Therefore, a relatively good agreement of the Hertz theory with the experimental results is probably caused by high damping of the generated stress waves in the VF tissue.

This study investigates the glottal width and its derivative waveforms measured on a male participant, and thereafter the glottal closing speed is investigated in detail on a physical 1:1 scaled model of phonation using artificial VF and vocal tract (VT).

The waveforms of the glottal area variation, glottal area declination rate, and *MADR* for phonation on vowel [u:] are compared with semi-occluded vocal tract phonation through a glass resonance tube with the distal end in air or submerged 10 cm in water, simulating two widely used voice therapy methods, see e.g., [14-17]. The aim of the study was to investigate how well *MADR* corresponds to the velocity of VFs before collision. Phonation through a tube alters the voice source and vocal tract acoustic-mechanical interaction considerably, and therefore it was chosen for a comparison with vowel phonation.

2. Methods

2.1. Measurement in human

First, we present modified results originating from our previous experimental study of human phonation characteristics during the voice therapy method using phonation into tubes [8]. A 34-year-old male teacher of non-classical singing volunteered as the participant in the experiments. High-speed laryngoscopic data were collected for human VFs' vibration during phonation of a neutral vowel at a comfortable speaking pitch and loudness using a Kay Pentax Color High-Speed Video System 9710 with a resolution of 512x512 pixels. The sampling frequency was 2000 frames/s. The human was recorded with a rigid scope inserted into the pharynx.

The glottal width variation $GW(t)$ values (in pixels) were recalculated to dimensionless quantities by dividing them by the vocal fold width (in pixels). Then, the time derivative of the glottal width variation ($dGW(t)/dt$), representing the velocity of the glottal opening and closing, was in the present paper calculated by the method of backward differences:

$$\frac{dGW(t)}{dt} = \frac{GW(t) - GW(t - \Delta t)}{\Delta t}, \tag{2}$$

where Δt is the time step given by the sampling frequency of the camera, in contrast to the central difference method used in the previous paper of the authors [8]. The backward differences method increased the precision of detecting the speed of the glottis closing $dGW(t)/dt$ just before the complete closure, i.e., at the time instant when $GW(t)$ reaches zero value.

2.2. Measurement on physical model

In order to study the glottal closing velocity in more detail, the measurement in human was followed by experiments performed on a physical model of phonation. The model and a similar measurement setup have been described in a previous study [9]. The air was flowing from a model of the lungs to a trachea model and then through a 1:1-scaled three-layer silicon VFs' model, filled by pressurized water, to a plexiglas hard-walled VT model for vowel [u:]. For changing the vocal tract conditions to resemble voice therapy with tubes, the VT was prolonged by a glass resonance tube (length 27 cm, inner diameter 7.8 mm) with the distal end in air or

submerged 10 cm in water. The controlling parameter was the mean airflow rate Q=0.06-0.09 L/s, measured by a float flowmeter (EMKO type DF3-09K5). The subglottal, supraglottal and oral pressures, P_{sub}, P_{supra} and P_{oral}, were registered with integrated pressure semiconductor sensors (NXP Freescale MPXV5010GC6U) mounted in the walls of the trachea and in the laryngeal and oral cavities of the vocal tract models. The fluctuations of the subglottic and oral pressures were measured by a dynamic semiconductor pressure transducer developed at the Institute of Thermomechanics of the CAS and by a special microphone probe (B&K 4182, range: 1 Hz–20 kHz), respectively. The acoustic signal was registered by the microphone of the sound level meter B&K 2239 positioned at the distance of 30 cm from the outlet of the VT model or from the end of the resonance tube.

All the measured signals were simultaneously sampled at a frequency of 16.4 kHz and registered by a measurement system B&K PULSE (type 3560 C with Input/Output Controller Modules Type 7537A and 3109) controlled by a personal computer equipped by the SW PULSE LabShop Version 10 (Bruel & Kjær Sound & Vibration Measurement, Denmark). 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 [18].

A high-speed CCD camera (NanoSense MkIII, maximum resolution 1280x1024 pixels) with a zoom lens (Nikon ED, AF NIKKOR, 70-300 mm, 1:4-5.6 D) was included in the measurement set up for investigating VFs' vibration. The rate of image recording into a personal computer was 10,000 frames/s with the maximum possible resolution (548x104 pixels). The camera was positioned at a 90° bend of the trachea model where a glass window was installed; this enabled the viewing of the VFs' vibration from the subglottal side. During recording of 1000 images, an LED light (ca 60 W) was focused on the vibrating VFs for improving the quality of the recorded images. The time instants of each image recording were registered in one channel of the system B&K Pulse. The images recorded by the camera were synchronized with the time records of the pressure signals by associating first image with the nearest time instant of the sampled pressure signals. The total recording time of the signals was 4 s for each trial, and it included simultaneous glottal area images recording of the length of 0.1 s.

For the purpose of this paper, the glottal area time variation $GA(t)$ and the time derivative variation $dGA(t)/dt$ were evaluated first from the whole vocal fold lengths, and then in a middle cross-section of the VFs for a bandwidth of 21 pixels, centered at the location of the maximum amplitude of the glottal width, where the edges of the glottis were vibrating approximately in parallel creating a rectangular shape of the area, see Fig. 1. The aim was to detect in more detail

the speed of the glottis closing in the middle cross section of the VFs. The time derivative *dGA*(*t*)*/dt* was computed according to eq. (2) by the method of backward differences.

The middle part of the VFs model was focused since it corresponds to the middle of the membranous part of the human vocal folds, which is usually the critical region for creating vocal fold nodules caused by excessive impact stress, see e.g., [19, 20].

First, the brightness and contrast of the original black and white video images were increased and optimized, so that the black pixels represented the glottal area at the individual times of the video recording, see Fig. 1. Next, the total number of black pixels in the individual columns of the figure was counted and summed across the rows using a MATLAB program written in-house to obtain the time waveform of the glottal area variation *GA*(*t*) in pixels. The area was statically calibrated in horizontal and vertical directions at the level of the glottis.

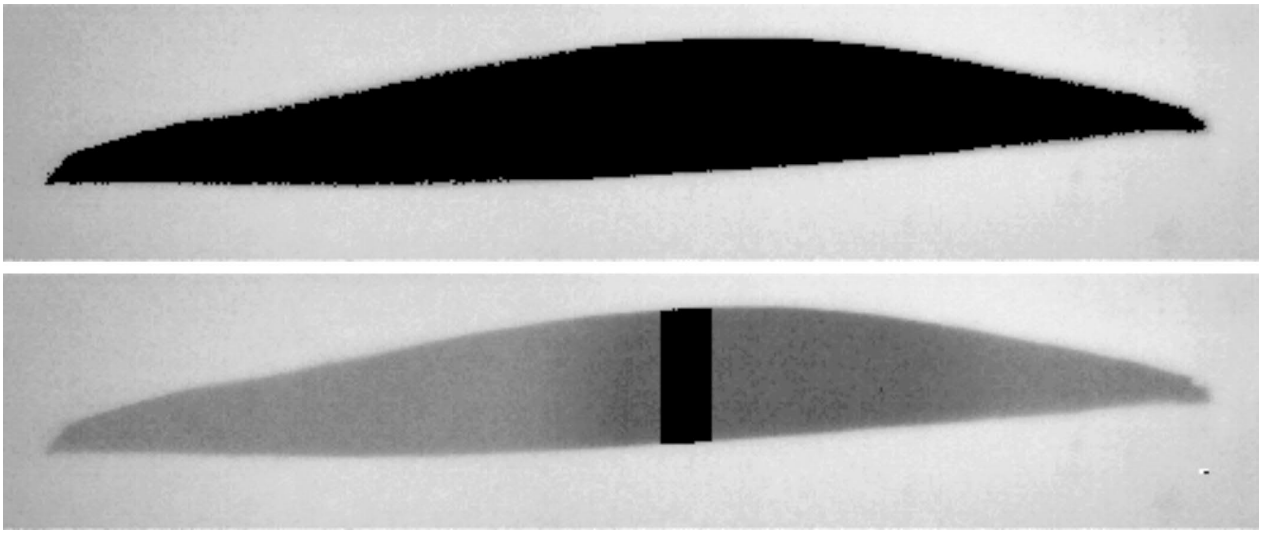

Figure 1. Glottal area regions GA (the black parts) evaluated: a/ in the full VF length (top panel) and b/ in a middle cross-section of the self-oscillating VFs model (bottom).

3. Results

3.1. Glottal area time variation in human

Figure 2 shows the results reconstructed from the measurement of glottal width *GW*(*t*) and the time derivative of the glottal width (*dGW(t)/dt*) waveforms for seven cycles of the VF vibration during phonation of the human subject. By inspecting the first three cycles it can be detected that the maximum velocity of the glottis closing, given by the absolute value of the time derivative of the glottal width, is followed by a smaller glottal width closing velocity $V_0 = \left| dGW\left(t\right)/dt \right|$

measured just before the glottis closure, which according to equation (1) implicates lower impact stress *IS*. The markedly visible decrease of the velocity magnitude is substantially influenced by the inadequate sampling frequency of the high speed camera that was used. In the first cycle, after the maximum velocity of the glottal closing (*MADR*), the velocity V_0 decreases by ca 45% before the collision, in the second cycle it decreases by about 70% and in the third cycle by ca 90%. This indicates that the glottal closing velocity rapidly decreases just before the closure. The results shown in Fig.2 are influenced not only by a low sampling frequency but naturally also by jitter and shimmer known as normal features in human phonation disturbing the perfect periodicity of the measured signals.

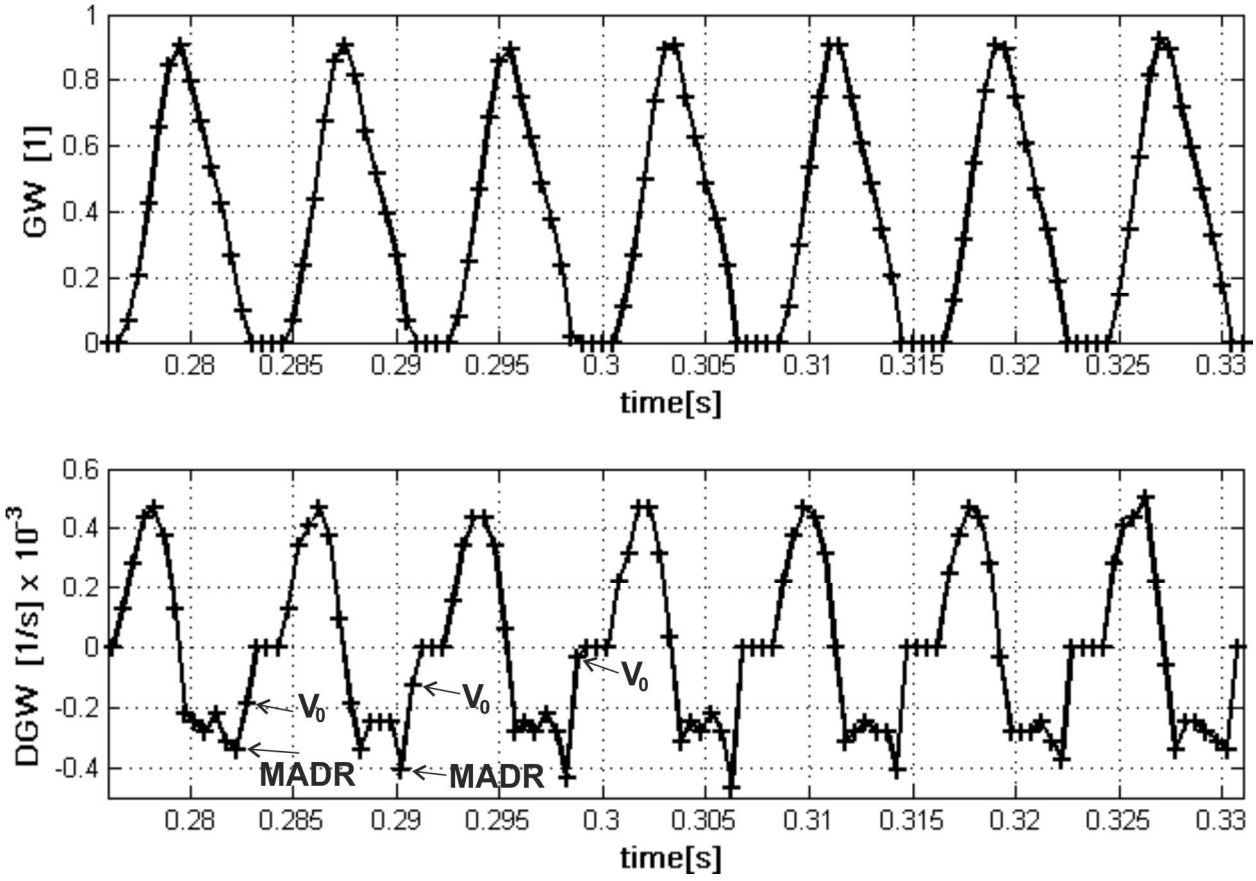

Figure 2. Glottal width variation waveform *GW*(*t*) and its numerical time derivative *DGW*(*t*), obtained from high-speed images for vowel phonation – measurement on human subject for sustained neutral vowel; subglottic pressure $P_{sub}\cong$2.03 kPa, fundamental phonation frequency f_0=126 Hz, high speed video sampling frequency 2 kHz. The results were recomputed from the previous results of the authors published in [8]. The *GW*(*t*) values measured in pixels are normalized to the maximum.

3.2. Results from modelling

3.2.1. Aerodynamic and frequency variables

Figure 3 shows an example of the simultaneously measured waveforms of the oral pressure P_{oral}, supraglottic pressure P_{supra}, subglottic pressure P_{sub} and the calculated waveform of the transglottic pressure P_{trans}, from which the mean and peak-to-peak values of the pressures were computed.

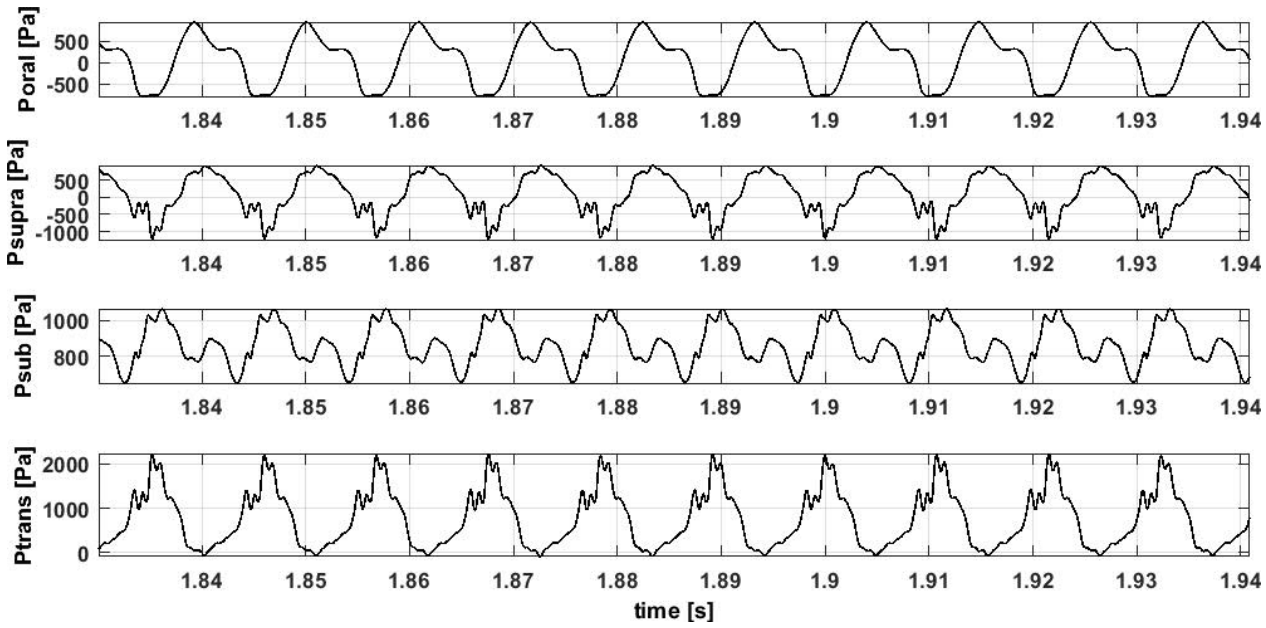

Figure. 3 Measured waveforms of the oral pressure P_{oral}, supraglottic pressure P_{supra}, subglottic pressure P_{sub} and transglottal pressure $P_{trans}(t)= P_{sub}(t) - P_{supra}(t)$ for phonation into the resonance tube with the distal end in air (Q=0.08 L/s, P_{sub}=864 Pa, f_0=93 Hz, first formant frequency F_1=115 Hz, first subglottic formant frequency F_{sub1}=725 Hz).

Figure 4 shows the mean subglottic pressure P_{sub} as a function of the controlled airflow rate Q, which was varied above the phonation threshold flow rate PTF≅0.05 L/s for vowel phonation up to 0.09 L/s, where the VF vibrations were regular for the all three measured configurations, i.e., for vowel [u:], for VT prolonged by tube with the outer end in air or submerged 10 cm in water. Phonation threshold pressure PTP≅727 Pa for the vowel phonation at the PTF was higher than PTP≅579 Pa for the tube phonation with the distal end in air.

Subglottic pressure P_{sub} increased linearly with the flow rate in the all three cases of phonation, for the vowel from ca. 720 Pa at Q=0.06 L/s up to 854 Pa at Q=0.09 L/s, for tube in air from 666 Pa up to 958 Pa, and from 1429 Pa up to 1725 Pa for tube phonation with the distal end in water. All of these quantities are in line with the physiological data for human vowel and tube phonations, see e.g. [21] and [22], respectively.

Because the subglottal pressure variation is an important parameter related to glottis area time variation, Fig. 4 also presents the measured peak-to-peak values of the subglottal pressure $P_{sub\ p\text{-}t\text{-}p}$ in percentages of the mean subglottal pressure P_{sub}. These values also increase with the airflow rate Q approximately linearly for the all three types of phonation. The highest subglottal pressure fluctuations $P_{sub\ p\text{-}t\text{-}p}$ were found for the vowel phonation, as they increased from 56% at Q=0.06 L/s to the maximum 86% at Q=0.09 L/s. This is also in agreement with the physiological data measured for the subglottal pressure variation in humans. The peak-to-peak magnitude of the subglottal pressure and the mean subglottal pressure were measured during normal phonation on vowel [ae:] by Sundberg et al. [23]. Their results show a ratio of 70%, which is in the range of our results measured on the artificial vocal folds. Similarly earlier, Miller and Schutte [24] measured subglottal pressure variation in humans showing this ratio to be about 50% for a continuous phonation on different vowels and about 73% for phonation on vowel [o:].

Figure 4 also presents the peak-to-peak values of the transglottal pressure $P_{trans\ p\text{-}t\text{-}p}$, which represent the pressure drop fluctuations at the glottis during the VF self-oscillations. In the all three cases of phonation, the $P_{trans\ p\text{-}t\text{-}p}$ increases with the flow rate approximately linearly. The maximum fluctuations of the transglottal pressure are observed for the tube with the outer end in air, somewhat smaller fluctuations for the tube with the end in water and markedly smaller fluctuations for the vowel. The higher fluctuations for tube phonations are given by principles of both therapy methods, where part of the airflow energy required for phonation is substituted by the acoustic energy utilizing the first acoustic resonance [9].

The fundamental phonation frequency of the VF self-oscillations was in the range f_0 = 90-93 Hz for all airflow rates in the range Q = 0.06-0.09 L/s and all measured configurations, see Fig. 5. The values for f_0 were thus within physiologically relevant values for male voice production. The water bubbling frequency f_b was between 21-22 Hz. The frequencies f_0 and f_b were determined from the spectra of the oral pressure.

The closed quotient (CQ) was approximately constant with increasing the flow rate, the highest $CQ\cong0.38$ was for the vowel, slightly smaller $CQ\cong0.37$ was found for tube in water and $CQ\cong0.33$ for tube in air.

Measured SPL increased with the flow rate approximately linearly for the all three types of phonation, see Figure 6. SPL for phonation on vowel [u:] was ca 3-6 dB higher than for phonation into tubes. However, the SPL measurement for phonation into tube with distal end in water is influenced by bubbling and its comparison with other two types of phonation is problematic.

Figure 7 shows that the flow resistance for phonation into the tube with the distal end 10 cm in water was substantially higher than for vowel phonation and tube phonation into air.

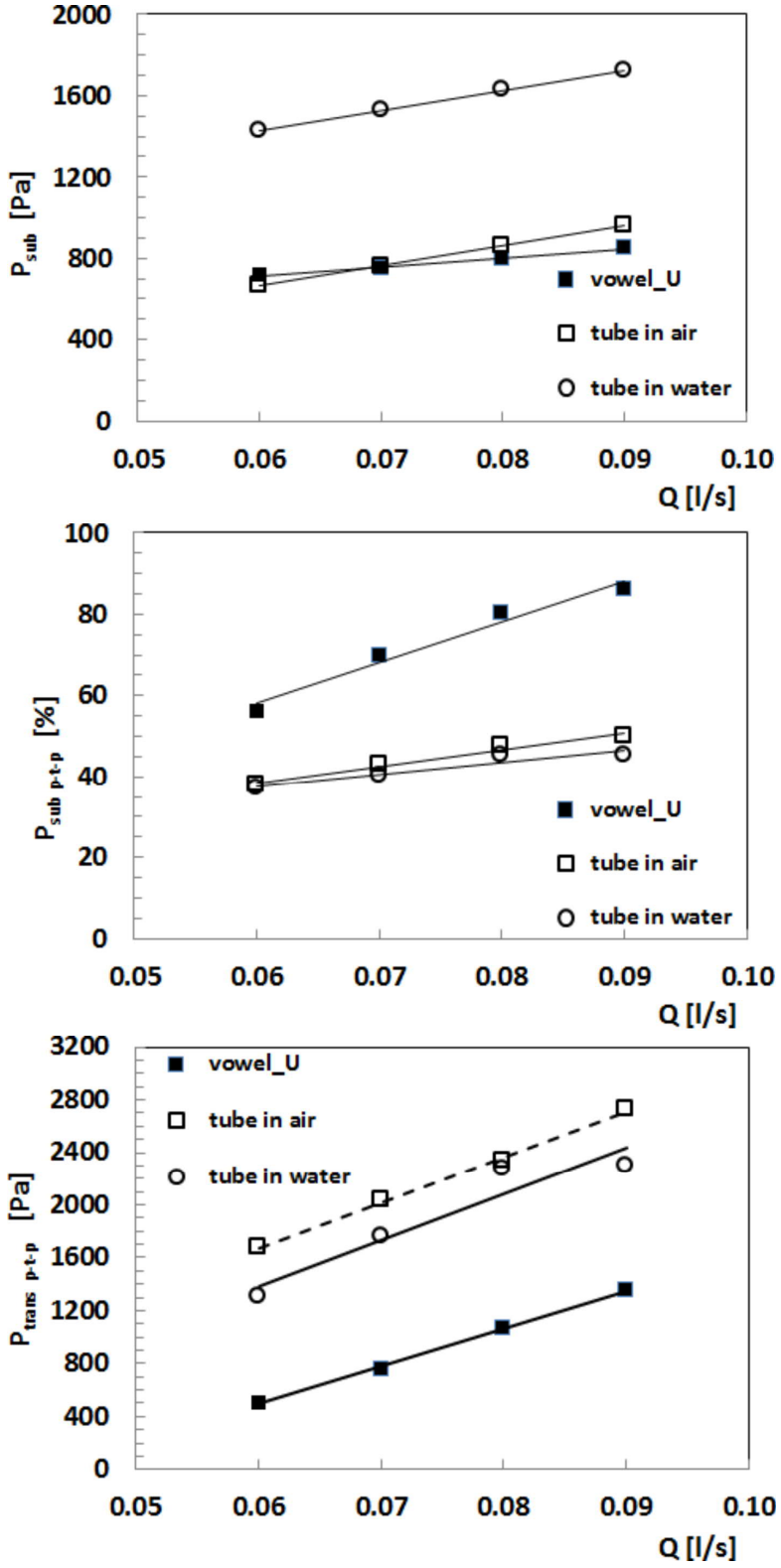

Figure 4. Measured values of mean subglottic pressure (P_{sub}), ratio of peak-to-peak variation of subglottic pressure to mean value of the subglottic pressure (100 $P_{sub\_p\text{-}t\text{-}p}/P_{sub}$), and peak-to-peak variation of transglottic pressure depending on the flow rate Q for simulated phonation with: 1)

[u:]-vowel-shaped vocal tract (VT), 2) [u:]-vowel-shaped VT attached to a glass resonance tube (length 27 cm, inner diameter 7.8 mm) ending in air and 3) [u:]-vowel-shaped VT attached to the same tube ending 10 cm in water.

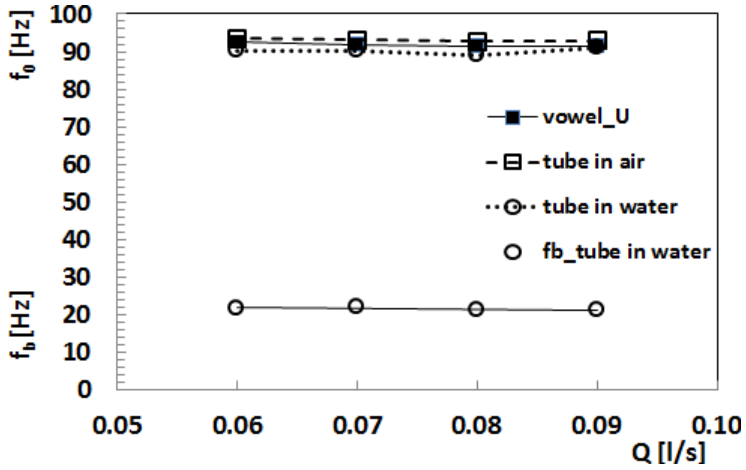

Figure 5. Measured fundamental frequency f_0 and the water bubbling frequency f_b depending on the flow rate Q for simulated phonation with: 1) [u:]-vowel-shaped vocal tract (VT), 2) [u:]-vowel-shaped VT attached to a glass resonance tube ending in air, and 3) [u:]-vowel-shaped VT attached to the same tube ending 10 cm in water.

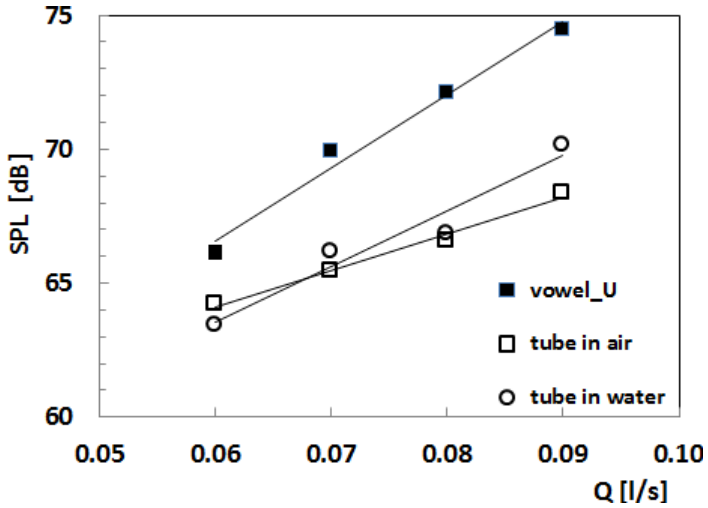

Figure 6. Sound pressure level (SPL) measured at the distance of 30 cm from the VT end or from the end of the resonance tube for the all three types of phonation investigated on the model.

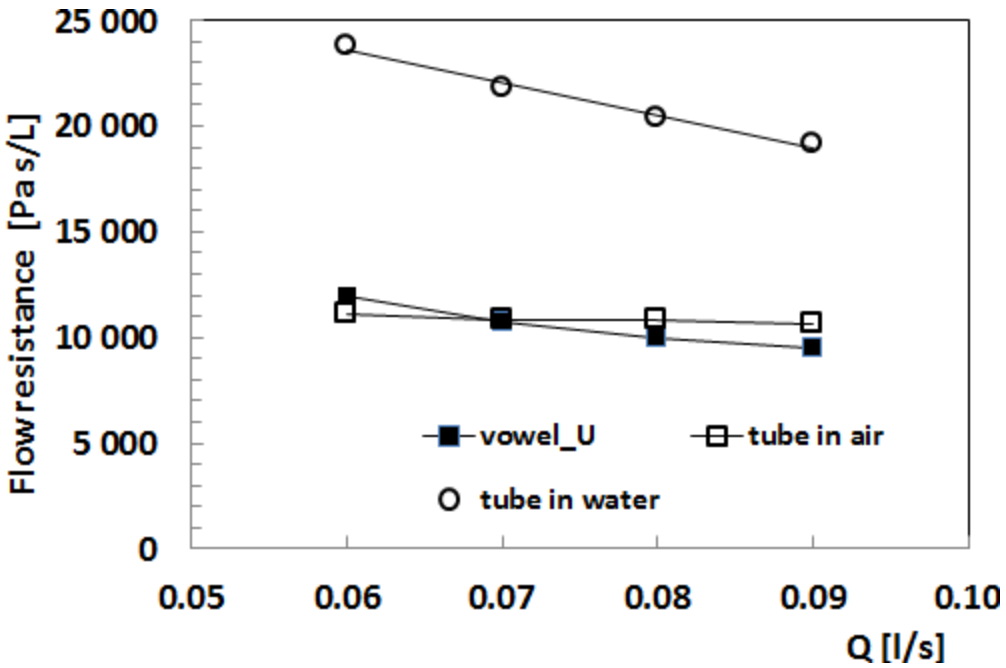

Figure 7. Flow resistance P_{sub}/Q calculated from the measured mean subglottic pressure P_{sub} and mean flow rate Q for three types of phonation investigated on the model.

3.2.2. Glottal area time variation in the model

An example of the glottal area $GA(t)$ and the time derivative of the glottal area $dGA(t)/dt$ variation evaluated for the full VF length is shown in Fig. 8. The upper trace is the glottal area measured by the high speed camera, and the lower trace is the numerical time derivative of the area computed by the backward difference schema, see eq. (2). This means that when the upper trace approaches zero the lower trace gives the speed of glottal area closing before the glottis closure, i.e., at the time of a complete glottal closure, where $GA(t)=0$. The waveforms of the measured glottal area for nine cycles of the VF vibrations show that before each glottal closure the *MADR* value, given by the absolute value of the time derivative of the glottal area $|dGA(t)/dt|$, was followed by lower absolute values of glottal area declination rate before the glottal closure.

The waveform for $dGA(t)/dt$ in Fig. 8 shows that in the first three cycles after *MADR* the speeds $V_0 = |dGA(t)/dt|$ of the glottis closing decreases before the VF collisions to ca 45%, 40% and 38% of the *MADR* magnitudes, in the fourth cycle it decreases to about 26%, in the fifth cycle to ca 20% and in the sixth cycle to ca 15% of the *MADR*. The speeds of V_0 before the VF collisions considered in this study are marked in Fig. 8 by arrows.

Remark. Because the last glottal area derivatives before the complete glottis closure are sometimes very close to zero, see e.g. such small non-zero values of $dGA(t)/dt$ in the first period in Fig. 8, we excluded them from the study, as they could be considered as a measurement error. Instead, we considered as true results only the clearly-detected speeds V_0 of the glottis closing, measured after *MADR* and just before full glottis closure. For example, in the first period in Fig. 8 we took into consideration the speed of glottis closing V_0 = 6 009 mm^2/s equal to ca 45% of the *MADR* value, which is ca 13 362 mm^2/s. It means that in the first period the glottis closing speed was reduced relative to *MADR* by 55 % and the next measured value V_0 = 957 mm^2/s was not taken into account. The percentage decrease related to the *MADR* in each oscillation period was influenced by the sampling frequency of the high speed camera which was synchronized with the pressure signals but not with the glottal area waveforms.

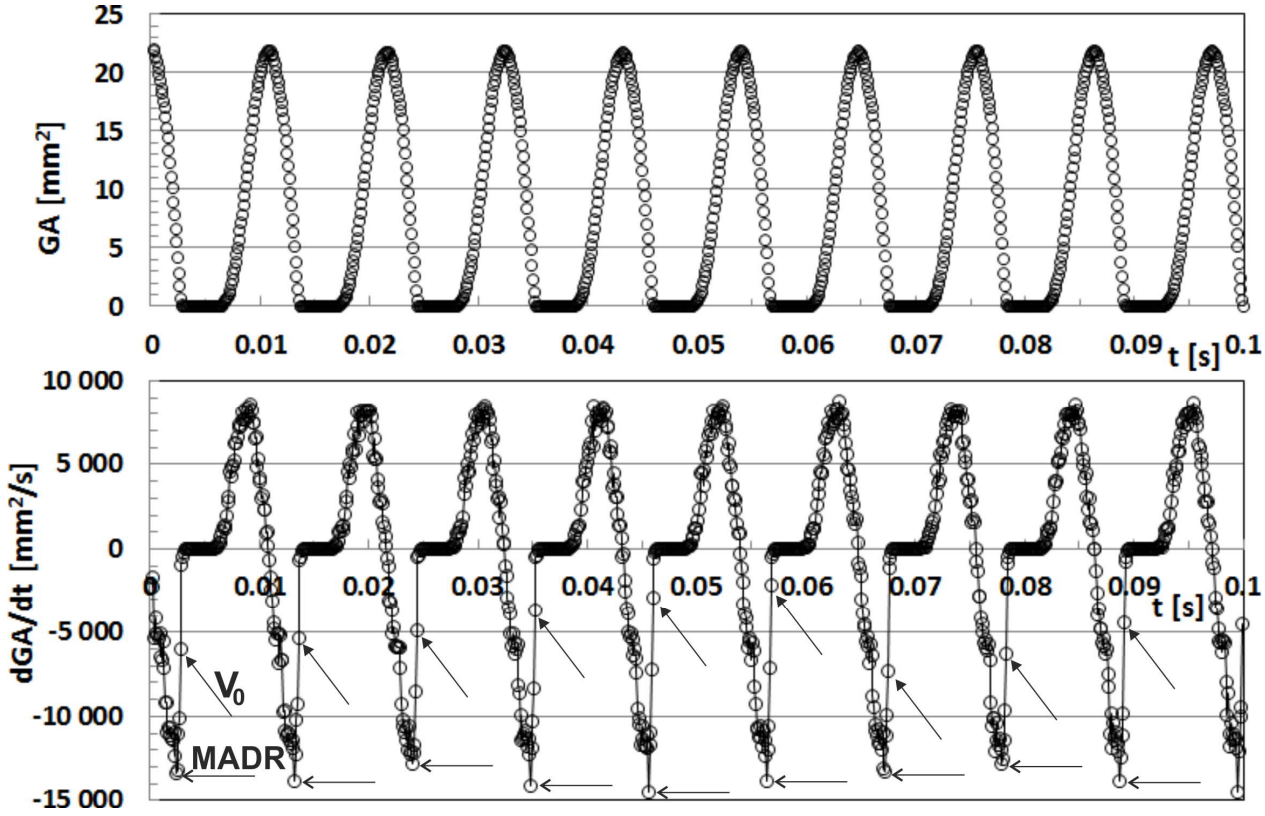

Figure 8. Example of the glottal area waveform (top), and the area derivative waveform (bottom) measured for phonation on tube with the distal end in air (Q=0.08 L/s, P_{sub}=864 Pa, f_0=93 Hz, first formant frequency F_1=115 Hz, first subglottic formant frequency F_{sub1}=725 Hz). The arrows indicate *MADR* and the values of speed $V_0(t_i) = |dGA(t)/dt|$ $(i = 1,2,...,9)$ of the glottis closing measured after $MADR(t_i)$ and just before the glottis closure in each of the nine periods of VFs' vibration. The results correspond to the measurement of the full glottal area as shown in Fig. 1a.

Figure 9 compares the graphs $GA(t)$ and $dGA(t)/dt$ shown in Fig. 8 in more detail for the first and fourth vibration periods of the VFs. The graphs for the glottal area are completed by the measured signal of the subglottic pressure $P_{sub}(t)$. The waveforms of the all three quantities clearly demonstrate the time shift between the samples of the pressure signal and the time instants when the images of glottal area were registered by the high sped camera with the lower sampling frequency than the pressure. Especially, this a kind of a stroboscopic effect influences in each vibration period the differences in the time delay between the last time instant for the $dGA(t)/dt$ registered as the last reliable value for the speed of glottis closing before the complete glottis closure. Figure 9 also demonstrates that at the time instant corresponding to the *MADR* the $P_{sub}(t)$ has a minimum and after that the subglottal pressure starts to increase. This tendency of $P_{sub}(t)$ corresponds to breaking of the glottis closing speed just before the glottis closure.

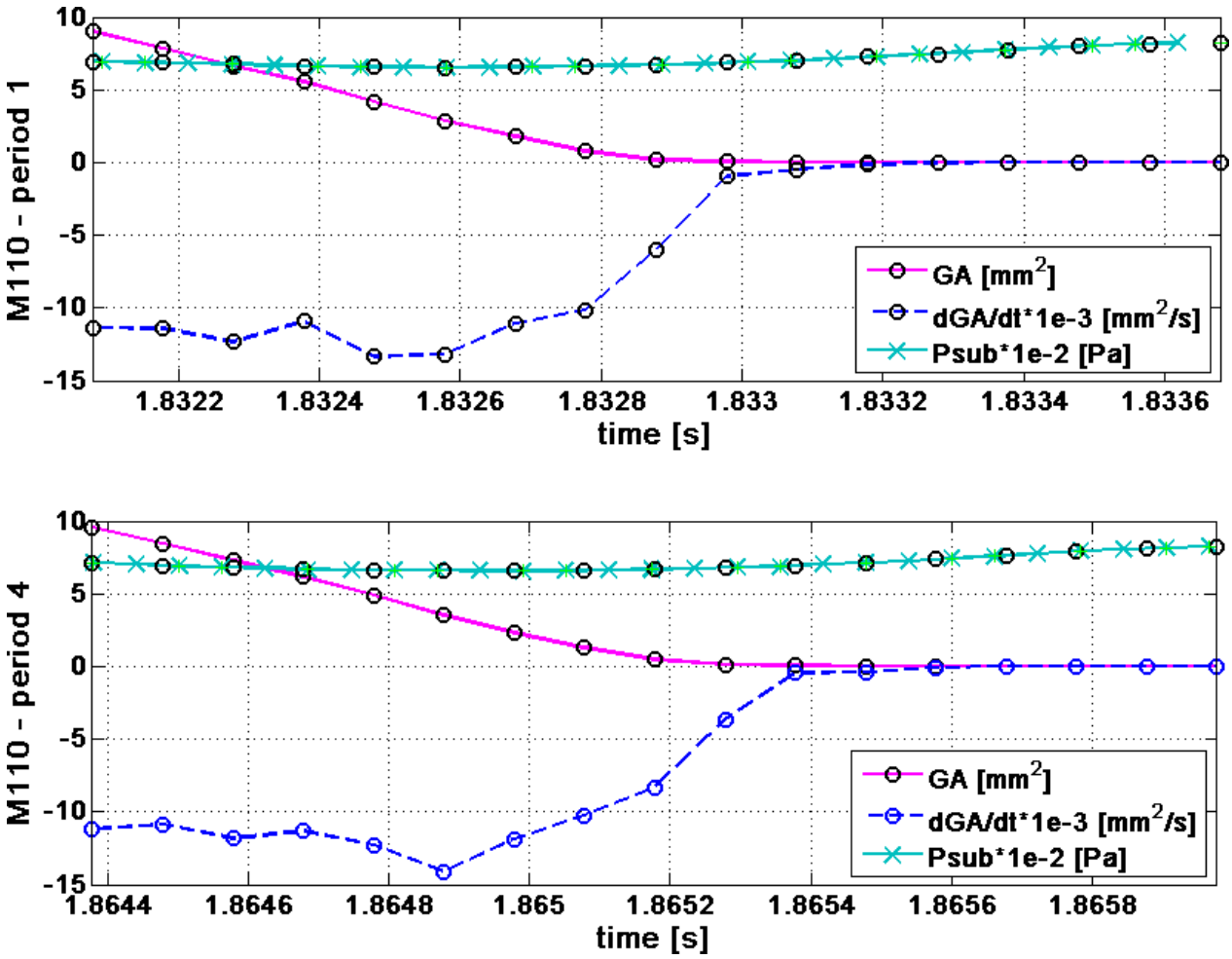

Figure 9. Details of the 1st (upper panel) and 4th (lower panel) periods of the signals $GA(t)$ and $dGA(t)/dt$ shown in Fig. 8 together with the measured signal of the subglottic pressure $P_{sub}(t)$.

Figure 10 shows the measured maxima of the glottis area variation $\max GA(t)$ as a function of the flow rate Q for the all three types of phonation. The maximal amplitudes increase approximately linearly with Q and are in line with the measured peak-to-peak values of the transglottal pressure, see Fig. 4. The glottal area amplitudes are highest for phonation into tube in air and lowest for the vowel phonation.

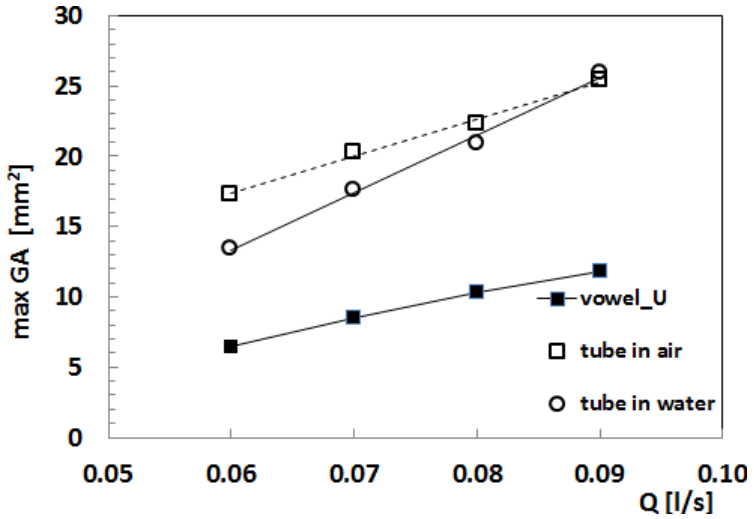

Figure 10. Measured maximum glottal area max$GA(t)$ as a function of the flow rate Q for: 1) phonation on the vowel [u:], 2) phonation through the tube into air, and 3) phonation through the tube into water.

In order to study the relations between *MADR* and the last absolute values of the area derivative $V_0(t_i) = \left| dGA(t)/dt \right|$, that could be reliably evaluated just before the glottal area closure, were averaged from the nine recorded periods of VF vibrations for each flow rate Q, see e.g. the nine values of $MADR(t_i)$ followed by the nine values of $V_0(t_i)$ marked in Fig. 8 by the arrows. Averaged *MADR* and averaged speed of glottal closing just before the glottis closure are defined as follows

$$\overline{MADR} = \left(\sum_{i=1}^{N} MADR(t_i) \right) / N \qquad \text{and} \qquad \overline{V_0} = \left(\sum_{i=1}^{N} V_0(t_i) \right) / N \ , \tag{3}$$

where N is number of vibration periods in each of the twelve trials measured for four chosen airflow rates in the interval $Q \in \langle 0.06, 0.09 \rangle$ L/s and three types of phonations (vowel [u:], tube in air, tube in water).

The results for averaged *MADR* values and averaged speed of glottis closing V_0 are shown in dependence on the airflow rate in Fig. 11. As expected, similarly as the maximum amplitudes of glottal area $\max GA(t)$ in Fig. 10 the *MADR* values in Fig. 11 increase with the flow rate extensively and approximately linearly, while all values of the average speed V_0 are considerably lower and keeping approximately similar level for all flow rates and types of phonation studied.

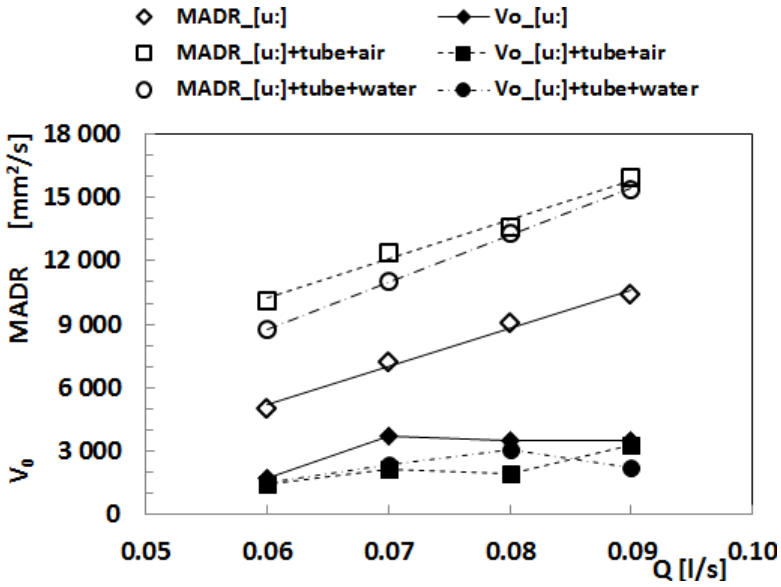

Figure 11. Average values of the glottal area derivative max(-d*GA*(*t*)/dt) during the closing phase of the glottis (*MADR*) and average values of speed V_0 of the glottis closing just before the glottal area closure, evaluated according to eq. (3), as a function of the flow rate *Q* for: 1) phonation on the vowel [u:], 2) phonation through the tube into air, and 3) phonation through the tube into water.

Thereafter, the averaged and median values of *MADR* and V_0 were calculated for each phonation type from all 36 measured *MADR* and V_0 values, i.e. from 9 periods for 4 different flow rates. The final results averaged in this way and denoted by $\overline{\overline{MADR}}$ and $\overline{\overline{V_0}}$ are summarized in Table 1. The highest average $\overline{\overline{MADR}}$ =1.3 10^4 mm^2/s was evaluated for phonation into the tube with the distal end in air. For phonation into the tube with the distal end in water $\overline{\overline{MADR}}$ =1.2 10^4 mm^2/s was slightly smaller, and the lowest $\overline{\overline{MADR}}$ =7.92 10^3 mm^2/s was found for vowel phonation. However, the average speed $\overline{\overline{V_0}}$ of glottis closing just before the glottal closure was in all cases clearly smaller than $\overline{\overline{MADR}}$. In phonation through the tube into air $\overline{\overline{V_0}}$ was only ca. 17% of $\overline{\overline{MADR}}$ value, in phonation through the tube into water $\overline{\overline{V_0}}$ was ca 19 % of $\overline{\overline{MADR}}$, and in vowel

phonation it was ca. 39% of $\overline{\overline{MADR}}$. Therefore, the highest average speed of glottis closing just before the glottal closure: $\overline{\overline{V_0}} \cong 3.1\ 10^3$ mm^2/s was observed for vowel phonation, and lower values were observed for phonation through the tube with the distal end in air and in water, $\overline{\overline{V_0}} \cong 2.2\text{-}2.3\ 10^3$ mm^2/s. Similar and even more pronounced trends were found for the calculated median values of $\overline{\overline{MADR}}$ and $\overline{\overline{V_0}}$, see Table 1.

Table 1. Average and median values of the *MADR* and of the last absolute values of the area derivative before the glottal closure V_0. These values were obtained for each phonation type from the all glottal area data throughout the VF vibration periods registered by the camera and the measured flow rates *Q*. $R_{average}$ and R_{median} are the ratios $\overline{\overline{V_0}} / \overline{\overline{MADR}}$ of the averaged and median values, respectively.

Phonation type	$\overline{\overline{MADR}}$ [mm^2/s]		$\overline{\overline{V_0}}$ [mm^2/s]		$R_{average}$ [%]	R_{median} [%]
	average	median	average	median		
Vowel [u:]	7.92E+03	8.11E+03	3.11E+03	2.59E+03	39.3	31.9
Vowel [u:]+tube in air	1.30E+04	1.30E+04	2.22E+03	1.60E+03	17.1	12.3
Vowel [u:]+tube in water	1.20E+04	1.21E+04	2.29E+03	2.30E+03	19.1	19.0

Figure 12 introduces an example of the glottal area and the time derivative waveforms and *MADR* and V_0 values measured in the limited middle cross-sectional area of the vocal folds where the colliding VF surfaces are approximately parallel**,** and the glottal area has a rectangular shape (see Fig. 1b). The results confirm the validity of the results presented above for the glottal area measured along the VFs' full length. The decrease of the area declination rate down to V_0 just after MADR and before the VFs' collision is obvious.

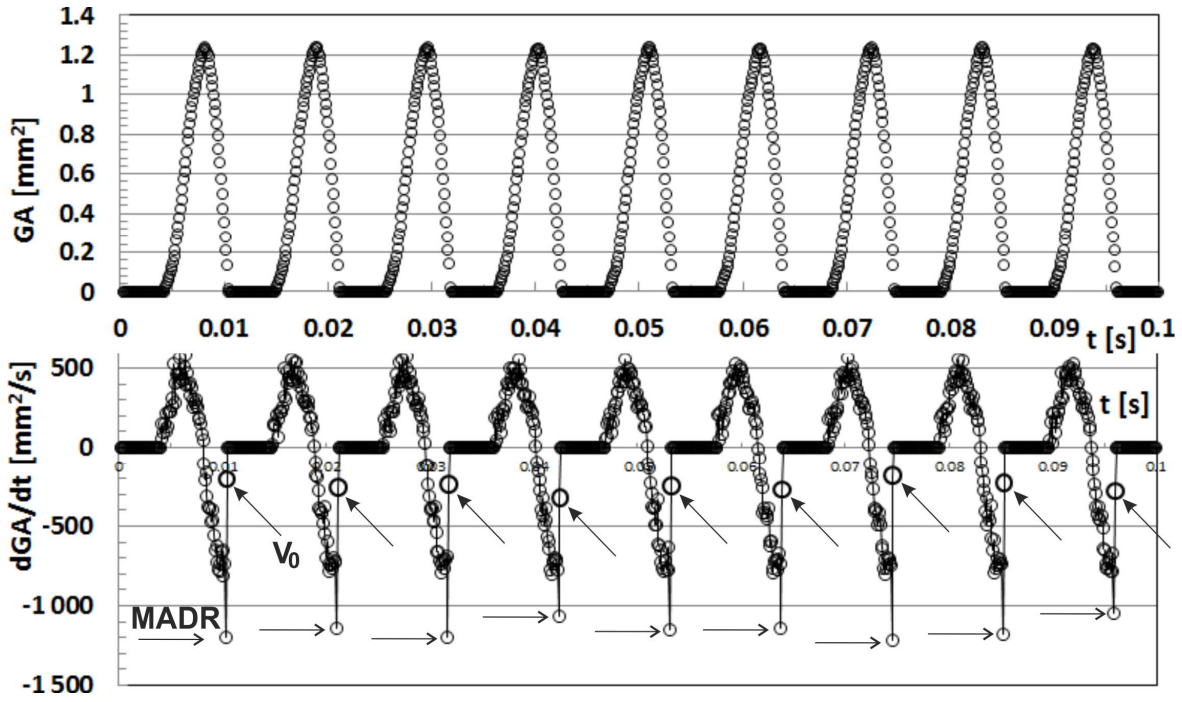

Figure 12. Example of the glottal area waveform (top), and the area derivative waveform (bottom) measured in a middle cross-section of the VFs in the bandwidth of 21 pixels for phonation on tube with the distal end in air (Q=0.06 L/s, P_{sub}=666 Pa, f_0=93 Hz). The arrows indicate *MADR* and the values of speed $V_0(t_i)=\left|dGA(t)/dt\right|$ $(i=1,2,...,9)$ of the glottis closing measured after $MADR(t_i)$ and just before the glottis closure in each of the nine periods of VFs' vibration.

4. Discussion

The results of this study inevitably indicate that *MADR* is not a sufficient quantity for estimating the impact stress during glottal collision. The results obtained confirm the previous random observations of a decrease of the glottal closing speed after *MADR*, see e.g. [8] for measurement *in vivo* and [9] for measurement on model. Moreover, the results for the absolute values of the time derivative of the glottal area measured just before the glottal closure imply that the impact stress for phonation into tubes is not higher than for phonation on vowel [u:], but rather the opposite, which agrees with the results of previous direct impact stress measurements on a similar phonation model [3].

With tubes, there is a stronger breaking effect of closing velocity and the reason is not flow resistance (see Fig. 7) offered by the tube, which is higher with tube in water and about the same with tube in air as in vowel phonation. The decrease of the glottis closing speed just before the complete glottis closure can be partly explained by the increased reactance, because in the

experiment on model for phonation on the resonance tube with distal end in air the fundamental phonation frequency was around $f_0 \cong 93$ Hz, which is close below the first acoustic resonance $F_1 \cong 105$ Hz, and thus less flow energy is needed for VF self-oscillation, see [9]. Similarly, for phonation on the resonance tube with distal end in air the bubbling with frequency $f_b \cong 20\text{-}24$ Hz was close below the first acoustic resonance, which decreased to $F_1 \cong 28$ Hz. As it was shown in [9] these lowest acoustic resonances cause a delay between the maxima of the transglottal pressure and the time instant of the glottis closure, which can influence the speed of glottis closure. However, Table 1 shows also a substantial decrease of the average and median values of MADR down to V_0 values for phonation on vowel, for which the first acoustic resonance $F_1 \cong 315$ Hz is far above $f_0 \cong 93$ Hz. Therefore, the effect of breaking the glottis closing speed just before the complete closure can be predominately caused by the viscosity of the air flowing in a very narrow glottal channel.

The results of the glottal area variation were markedly influenced by the low sampling frequency and low resolution of the high-speed camera used; thus, a further study with a better camera is warranted for a more detailed study of the glottal closing speed which influences the magnitude of the impact stress. Another limitation of the results is that in the experiments it was possible to observe only the lower edge of the vibrating vocal folds, while the vocal folds' collision is a complicated three-dimensional phenomenon. According to a kinematic 3D model developed recently by Smith & Titze [25], eight major VF contact patterns can be identified and characterized by the shape of the flow channel, with the following descriptors assigned: convergent, divergent, convergent-divergent, uniform, split, merged, island, and multichannel.

DeJonckere & Lebacq in a very recent study [26] investigated similar breaking effect of the glottal closing speed for phonation of a male participant in speaking pitch by combined photoglottography and EGG methods. In agreement with our results they found that the VFs collision velocity is ca. 10%-30% lower than the maximum closing velocity for SPL values up to 78 dB. For higher SPL the VFs collision velocity increases, up to 70% of the maximum closing velocity.

We investigated phonation on vowel [u:] since our earlier CT and MRI studies have shown that humans adopt a vocal tract shape for vowel [u:] when phonating into a tube. A closed vowel like [u:] implies a slight semi-occlusion in the vocal tract. Therefore, future studies should compare a semi-occluded vocal tract posture with a clearly open vocal tract, as in the phonation of [a:]. For an open vowel, the airflow resistance is smaller and thus the reduction of the glottal closing velocity just before the full glottis closure could be smaller. Such a study for an open vowel performed on the model would be advantageous because laryngeal high-speed videoendoscopic

observations of open vowels are difficult to carry out *in vivo.* In an open vowel like [a:], the tongue base is lower, which pushes the epiglottis back, thus impairing visibility of the glottis. In studying humans, an important future improvement would also be to develop a reliable calibration tool for glottal measurements. If nasoendoscopic high-speed imaging is further developed, it may allow a deeper visualization of the glottis during different vowels, and with glottal closure measurement, it opens new possibilities for more detailed estimation of impact stress – despite the fact that in imaging there is always the problem that from above one is able to see only the upper part of the vocal folds.

5. Conclusion

The results suggest that the MADR is not a sufficient quantity for estimating impact stress during glottal collision. The glottal area declination rate just prior to vocal folds collision is lower than MADR. Furthermore, in phonation through a tube into air and into water, the glottal area declination rate prior to vocal folds collision is lower than in vowel phonation; thus, the impact stress in tube therapy is lower than in vowel phonation.

The results also show that the most important parameter for maximal glottal area amplitude and the maximal glottal area declination rate is the peak–to–peak amplitude of the transglottal pressure. Such a detailed study carried out on the model would be difficult to perform in humans.

Acknowledgement

The study was supported by a grant from the Czech Science Foundation: No. 19-04477S "Modelling and measurements of fluid-structure-acoustic interactions in biomechanics of human voice production."

References

[1]I.R. Titze, Principles of Voice Production, (2nd printing), Iowa City, IA: National Center for Voice and Speech, 2000.

[2] J. Horáček, V. Bula, V. Radolf, P. Šidlof, Impact stress in a self-oscillating model of human vocal folds, Journal of Vibration Engineering & Technologies 4(3) (2016) 183-190.

[3] J. Horáček, V. Radolf, A.M. Laukkanen, Impact stress in water resistance voice therapy: A physical modeling study, J. Voice 33(4) (2019) 490-496.

[4] J. Jiang, I. Titze, Measurement of vocal fold intraglottal stress and impact stress, J. Voice 8 (1994) 132–144.

[5] K. Verdolini, M.M. Hess, I.R. Titze, W. Bierhals, M. Gross, Investigation of vocal fold impact stress in human subjects, J. Voice 13 (1999) 184-202.

[6] I.R. Titze, Theoretical analysis of maximum flow declination rate versus maximum area declination rate in phonation, J. Speech, Lang. Hear. Res. 49 (2006) 439-447.

[7] I.R. Titze, A.M. Laukkanen, Can vocal economy in phonation be increased with an artificially lengthened vocal tract? A computer modeling study, Logop. Phoniatr. Vocol. 32(4) (2007) 147-156.

[8] A.M. Laukkanen, A. Geneid, V. Bula, V. Radolf, J. Horáček, T. Ikavalko, T. Kukkonen, E. Kankare, J. Tyrmi, How much loading does water resistance voice therapy impose on the vocal folds? An experimental human study, J. Voice (2018) In Press, Available online 22 November 2018, https://doi.org/10.1016/j.jvoice.2018.10.011.

[9] J. Horáček, V. Radolf, A.M. Laukkanen, Experimental and computational modeling of the effects of voice therapy using tubes, J. Speech, Lang. Hear. Res. 62 (2019) 2227–2244.

[10] R. Brepta, M. Prokopec, Wave Propagation and Impacts in Solids (in Czech), Academia, Praha, 1972.

[11] W. Goldsmith, Impact, Arnold Ltd., London, UK, 1960.

[12] J. Horáček, P. Šidlof, J.G. Švec, Numerical simulation of self-oscillations of human vocal folds with Hertz model of impact forces, Journal of Fluids and Structures 20 (2005) 853–869.

[13] J. Horáček, A.M. Laukkanen, P. Šidlof, P. Murphy, J.G. Švec, Comparison of acceleration and impact stress as possible loading factors in phonation. A computer modeling study, Folia Phoniatr. Logop. 61(3) (2009) 137-145.

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

[15] S. Simberg, A. Laine, The resonance tube method in voice therapy: Description and practical implementations, Logop. Phoniatr. Vocol. 32 (2007) 165-170.

[16] G. Wistbacka, J. Sundberg, S. Simberg, Vertical laryngeal position and oral pressure variations during resonance tube phonation in water and in air. A pilot study, Logop. Phoniatr. Vocol., 41(3) (2016) 117-123.

[17] L.S. López, N.P. Marrero, T.R. Rodriguez, Applicability of resonance tube phonation in water for semi-occluded vocal tract voice training and therapy. Folia Phoniatr. Logop. 72 (2020) 22–28.

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

[19] K. Verdolini, C.A. Rosen, R.C. Branski, eds. Vocal fold nodules (nodes, singer's nodes, screamer's nodes, in Classification Manual for Voice Disorders - I. Psychology Press, (2014), pp. 37–40. ISBN 978-1-135-60020-4.

[20] I.R. Titze, Mechanical stress in phonation, J. Voice 8(2) (1994) 99-105.

[21] R.J. Baken, R.F. Orlikoff, Clinical Measurement of Speech and Voice, (2nd ed.), Singular/Thomson Learning, San Diego, CA, 2000.

[22] J. Tyrmi, V. Radolf, J. Horáček, A.M. Laukkanen, Resonance tube or Lax vox?, J. Voice 31(4) (2017) 430–437.

[23] J. Sundberg, R. Scherer, M. Hess, F. Muller, S. Granqvist, Subglottal pressure oscillations accompanying phonation, J. Voice 27(4) (2013) 411-421.

[24] D.G. Miller, H.K. Schutte, Characteristic patterns of sub- and supra-glottal pressure variations within the glottal cycle, in: Lawrence, Van L, ed. Transcripts XIII Symposium Care of the Professional Voice, New York 1984, The Voice Foundation, 1985, pp. 70-75.

[25] S.L. Smith, I.R Titze, Vocal fold contact patterns based on normal modes of vibration, Journal of Biomechanics 73 (2018) 177–184.

[26] P.H. DeJonckere, J. Lebacq, Vocal fold collision speed in vivo: The effect of loudness. Journal of Voice (2020), doi:10.1016/j.jvoice.2020.08.025.

Highlights:

- Glottal closing velocity has been used as an estimate of impact stress in phonation.
- Phonation in physical model and in human were compared for vowel and phonation through a resonance tube with distal end in air or in water.
- Glottal area waveform was studied with high-speed imaging.
- Results show that maximum area declination rate is not a reliable estimate of impact stress; instead, maximum velocity before closure needs to be studied.