Influence of vocal fold stiffness and acoustic loading on flow-induced vibration of a single-layer vocal fold model

Zhaoyan Zhang*, Juergen Neubauer, David A. Berry

UCLA School of Medicine, 31-24 Rehabilitation Center, 1000 Veteran Ave., Los Angeles, CA 90095-1794, USA

Abstract

The flow-induced vibrations of a single-layer vocal fold model were investigated as a function of vocal fold stiffness, and subglottal and supraglottal acoustic loading. Previously, it was reported that the single-layer vocal fold model failed to vibrate when short, clinically relevant tracheal tubes were used. Moreover, it was reported that the model had a propensity to be acoustically driven, and aerodynamically driven vibration was observed only when a vertical restraint was applied superiorly to the vocal folds. However, in this study involving a wider range of source/tract conditions, the previous conclusions were shown to apply only for the special case of a stiff vocal fold model, for which self-oscillation occurred only when the vocal fold vibration synchronized to either a subglottal or supraglottal resonance. For a more general case, when vocal fold stiffness was decreased, the model did exhibit self-oscillation at short tracheal tubes, and no vertical restraint was needed to induce aerodynamically driven phonation. Nevertheless, the vocal fold vibration transitioned from aerodynamically driven to acoustically driven vibration when one of the subglottal resonance frequencies approximated one of the natural frequencies of the vocal folds. In this region, strong superior–inferior vibrations were observed, the phonation threshold pressure was significantly reduced, and the phonation onset frequency was heavily influenced by the dominant acoustic resonance. For acoustically driven phonation, a compliant subglottal system always lowered phonation threshold. However, an inertive vocal tract could either increase or decrease phonation threshold pressure, depending on the phonation frequency.
© 2008 Elsevier Ltd. All rights reserved.

1. Introduction

This paper presents an experimental study of the flow-induced vibrations of a vocal fold-shaped compliant constriction in a pipe flow, under the influence of subglottal (upstream) and supraglottal (downstream) acoustic loading. Understanding the underlying physics of such flow-induced vibrations in channel flows has important applications in human phonation and animal vocalization. Similar phenomena also occur in musical instruments (e.g., brass and woodwind instruments [1]).

In general, self-sustained vocal fold oscillations are initiated by a complex fluid–structure–acoustic interaction, which results in a net energy transfer from the flow into the vocal folds to overcome dissipative

---

*Corresponding author. Tel.: +1 310 206 4324.

E-mail address: zyzhang@ucla.edu (Z. Zhang).

losses and sustain structural oscillations. There are two major energy transfer mechanisms. The first mechanism is related to a near field fluid–structure interaction (FSI), in which two structural eigenmodes synchronize and coalesce into a coupled-mode flutter [2–6]. The coupling strength generally increases with increasing flow velocity and decreasing stiffness of the compliant structure [6]. The second mechanism is related to an acoustics–structure interaction, or the coupling of one structural mode to one of the acoustic resonances of the upstream or downstream system [7,8], as is common for some musical instruments. Depending on a variety of system variables, the relative strength of the FSI and the acoustics–structure interaction in the coupled system may change.

In voice production, the synchronization of two or more structural modes due to a near field FSI is considered the primary mechanism of normal phonation [6,9,10]. Synchronization of two eigenmodes to the same frequency but different phases allows the flow pressure to have an in-phase component with the motion of the vocal fold surface, therefore transferring energy from the airflow to the vocal fold tissue [6]. In a continuum model of the vocal folds, the synchronization of two or more structural modes may lead to a wave-like motion (termed a mucosal wave) along the medial surface. The phonation frequency (fundamental frequency or pitch) is controlled primarily by the vocal fold biomechanics rather than by the upstream or downstream acoustic environment. In a previous study [11], this type of phonation was referred to as an aerodynamically driven mode of phonation, as contrasted to an acoustically driven mode of phonation in which the energy transfer was provided by acoustic-structure coupling to the subglottal acoustics.

The acoustics–structure interaction, or source–resonator coupling, is generally understood to play a minor role in normal phonation. The source–resonator coupling is weak as the phonation frequency is normally lower than both the subglottal and supraglottal resonance frequencies. Under this condition, the vocal tract introduces an inertive acoustic load to the vocal folds, which has been shown theoretically to lower phonation threshold pressure (mean subglottal pressure at phonation onset) [9]. Strong source–resonator coupling may occur in pathological phonation, or in singing when the fundamental frequency approaches the first formant [12–15], or for large amplitude vocal fold vibration in which nonlinear effects are significant. For example, it was shown that synchronization of the vocal fold to subglottal acoustics may lead to various irregular vibration patterns (subharmonics and biphonation) [16]. In addition, Titze [17] proposed that the constructive or destructive interference of the subglottal acoustics with vocal fold vibration may be a mechanism of register changes (a sudden change in the vibratory pattern of the vocal folds).

Previously, the influence of acoustics on the self-oscillation of a single-layer rubber vocal fold model was investigated by systematically varying the subglottal tract length [8]. Self-oscillation was observed only when the vocal fold vibration entrained with one of the subglottal resonances. The resulting vocal fold vibration exhibited primarily a single-mass motion. No oscillation was observed for short tracheal tube lengths typical of humans. Aerodynamically driven mode of phonation was observed only when a vertical restrainer was applied to the superior surface of the vocal fold model [11]. High-speed imaging of the medial surface dynamics showed that the aerodynamically driven mode of phonation featured strong excitation of high-order out-of-phase transverse motion (perpendicular to the flow direction) of the medial surface. This contrasts with a dominantly in-phase vertical motion (along the flow direction) [11] for an acoustically driven mode of phonation.

However, because a relatively stiff vocal fold model was used in these two previous studies, the near field FSI may have been considerably weaker than in a more typical situation in which the vocal folds were more pliable. Reducing vocal fold stiffness is expected to enhance the FSI, and thus facilitate aerodynamically driven phonation at short, clinically relevant tracheal tube lengths, without the application of vertical restrainers. The interaction of these two mechanisms of self-oscillation (the acoustics–structure interaction and the FSI), when both are simultaneously present, was not investigated in these two previous studies, due to the weak FSI in a stiff vocal fold model. Furthermore, no vocal tract tube was included in these previous studies [8,11], and the interaction between the subglottal and supraglottal acoustics was therefore not investigated either.

In this study, the interplay between two self-oscillating mechanisms (either a FSI and an acoustics–structure interaction, or two different acoustics–structure interactions) is investigated in a single-layer rubber vocal fold model. The relative strength of the near field FSI and the acoustics–structure interaction can be varied by changing either the vocal fold biomechanics or the sub- or supraglottal system. In this study, the near field

fluid–structure coupling strength is controlled by varying the stiffness (Young's modulus) of the vocal fold model, which was shown to affect the phonation threshold pressure associated with the FSI [6]. The subglottal and supraglottal acoustical loading is controlled by using straight tubes of uniform cross-section and varying the length of corresponding tubes, as in previous work [8]. The goal of this study is not to reproduce the exact physiological geometries or the resonance structures of the human subglottal and supraglottal tracts. Instead, by varying both the vocal fold properties and the sub- or supraglottal acoustics, we systematically scan a wide range of conditions with various degrees of relative strength between the FSI and acoustics–structure interaction, in an effort to quantify the interplay between the two primary mechanisms of self-oscillation. Although tube lengths other than those physiologically possible were used, the relative strengths between the resulting FSI and acoustics–structure interaction, or specifically, the ratio between the natural frequencies of the vocal folds and the acoustical resonance frequencies of the sub- and supraglottal system, are still within the physiological range of human phonation (including both speech and singing). This study also provides experimental data for validation of theoretical phonation models. As a first step, this study focuses only on the system behavior before and around phonation onset. The post-onset behavior of the coupled system will be the topic of future investigations.

2. Method

The experimental setup (Fig. 1) was the same as used in previous studies [8,12], which was designed to simulate the human subglottal system and the vocal folds. For details of the experimental apparatus readers are referred to the previous studies. The subglottal system consisted of an expansion chamber (inner cross-sectional area of 23.5×25.4 cm and 50.8 cm long), simulating the lungs, and a uniform circular PVC tube (inner diameter of 2.54 cm) of variable length, simulating the tracheal tube. The expansion chamber was connected to the air flow supply through a 15.2-m-long rubber hose, reducing possible flow noise from the air supply. Acoustic reflection factor measurements [8] showed that, at low frequencies, the expansion chamber and the upstream flow supply could be treated approximately as an ideal open-ended termination of the uniform tracheal tube. The combination of the expansion chamber and the variable-length uniform tracheal tube provided a simple yet controllable acoustical environment, which was essential for the present study. The vocal tract was simulated in this study using a straight circular PVC tube (inner diameter of 2.54 cm) of variable tube length.

The glottis was formed by a single-layer, rubber model of the vocal folds, which was installed in between the tracheal tube and the vocal tract tube [8,18]. Each vocal fold model measured approximately 1 cm in the superior–inferior (streamwise) direction, 1.7 cm in the anterior–posterior (spanwise) direction, and 0.8 cm in the medial–lateral (cross-stream) direction (Fig. 2). The inferior (upstream) side of each vocal fold had an entrance convergence angle of approximately $60°$ measured from the superior–inferior (streamwise) axis, yielding a vocal fold thickness (in the superior–inferior direction) of approximately 5.4 mm. The physical models were made with a two-component liquid polymer solution mixed with a liquid flexibilizer solution [18]. The stiffness or Young's modulus of the model was controlled by the ratio of the flexibilizer solution used in the compound mixture. The Young's moduli of the physical models were measured using the indentation method with a 1 mm-diameter cylindrical rod and a 1-mm indentation depth [19]. The vocal folds were glued
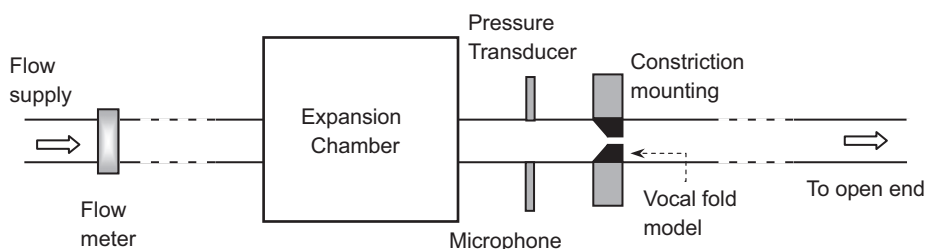

Fig. 1. Schematic of the experimental setup.

FSI: fluid-structure interaction interface;
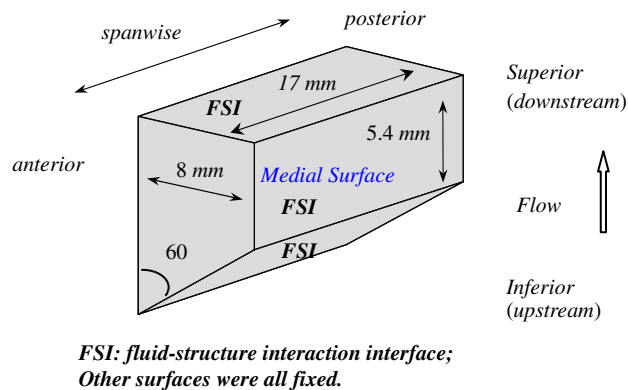Other surfaces were all fixed.

Fig. 2. Sketch of the vocal fold model.

into a rectangular groove on the medial surface of two acrylic plates. The medial surfaces of the two folds were positioned to be in contact so that the glottis was closed when no airflow was applied.

The acoustic pressures in the tracheal tube and the vocal tract were monitored using two probe microphones (B&K 4182), which were mounted flush with the inner wall of the tracheal tube and vocal tract, each 5 cm away from the vocal fold plates. Pressure taps were also mounted flush with the inner wall of the tracheal tube and the vocal tract, 2 cm from the vocal fold plates. The time-averaged transglottal pressure was measured using a pressure transducer (Baratron type 220D). The volumetric flow rate through the orifice was measured using a precision mass-flow meter (MKS type 558A) at the inlet to the setup. Analog-to-digital conversion of the output signals was performed at a sampling rate of 50 kHz.

For each vocal fold model and experimental configuration, the flow rate was increased to a certain upper limit in discrete increments, and then decreased back to zero. At each step, sound pressure, flow rate and subglottal pressure were measured. Phonation frequency at onset and subglottal pressure were then extracted in a same procedure as described in [8].

In some cases, a hemi-model procedure [11] was used to measure the medial surface motion of the vocal fold model. One of the vocal fold plates was removed and replaced by a glass prism. The prism provided two distinct views of the medial surface of the vocal fold, which were imaged using a high-speed digital camera (Fastcam-Ultima APX, Photron Unlimited, Inc.). A frame rate of 2000 Hz for the camera was used with a spatial resolution of 1024×1024 pixels per image. Prior to imaging, graphite powder was sprinkled on the medial surface of the vocal fold to form random dot patterns. In the post-processing stage, such patterns facilitated cross-correlation analysis to compute the medial surface displacements. The time-series cross-correlation analysis was performed on the medial surface images using the image-processing package DaVis (LaVision Inc.). For further details of the hemi-model experimental setup and the cross-correlation analysis process please refer to [11]. In this study, displacements were computed over the vocal fold medial surface with a grid spacing of 0.22×0.22 mm in the inferior–superior and anterior–posterior directions, respectively.

3. Results

3.1. Acoustics–structure interaction

Results are first discussed for a stiff vocal fold model with a Young's modulus of about 11.1 kPa. The relatively large Young's modulus raised the onset threshold of the coupled-mode flutter [6], and phonation only occurred when the vocal fold vibration was synchronized to one of the acoustical resonances of the sub- or supra-glottal system. Thus, the acoustics–structure interaction was isolated and studied.

Fig. 3 shows the phonation frequency (symbol ×) at phonation onset, F_0, and phonation threshold subglottal pressure, P_t, as a function of the tracheal tube length, L_{up}. No vocal tract tube was attached in this case. Also shown in Fig. 3a is the subglottal resonance frequency, as calculated from the measured reflection coefficients data of the subglottal expansion chamber and assuming plane wave propagation in the tracheal
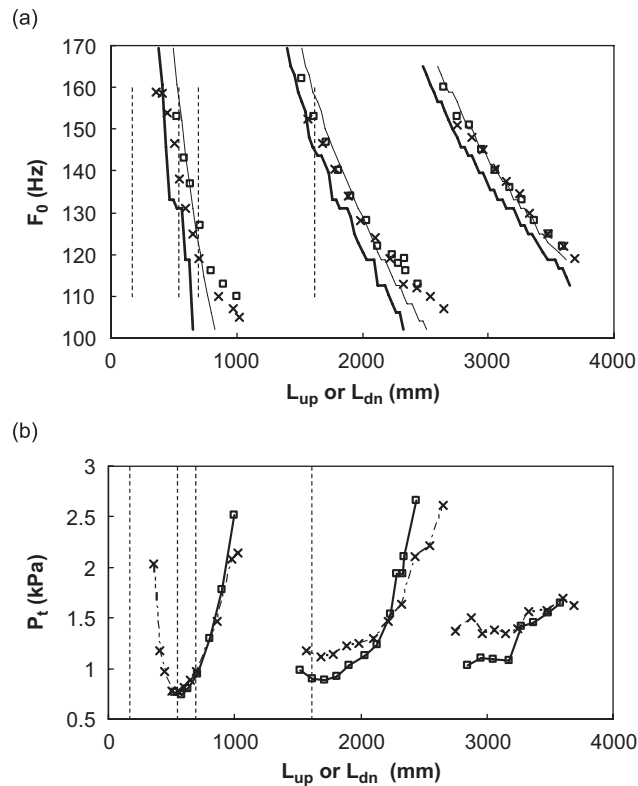
Fig. 3. (a) Phonation frequency F_0 at onset and (b) phonation threshold pressure as a function of tracheal tube length (×) or vocal tract tube length (□). For the former case, no vocal tract was included and the tracheal tube length was varied. For the latter case the vocal tract tube was varied and the tracheal tube length was fixed at 17.2 cm. The two solid lines are the resonance frequency of the subglottal system (thick solid line) and the vocal tract (thin solid line). The four vertical lines indicate the tracheal tube lengths used for the four cases of Fig. 4 as discussed in Section 3.2.

tube and infinite glottal impedance. Both the F_0 and P_t showed cyclic variations with the subglottal tube length. Within each cycle, the phonation frequency F_0 closely approximated the subglottal resonances, suggesting an acoustically driven mode of vibration. The phonation frequency was generally slightly higher than the subglottal resonance, suggesting the vocal fold self-oscillation was sustained by a compliant subglottal system. These observations are consistent with those in [8], which showed that this vocal fold oscillation could be described effectively by a single-mass model.

Fig. 3 also shows the phonation frequency at onset and the phonation threshold pressure (symbol □) as a function of the vocal tract tube length, for a subglottal tube of length 17.2 cm. The corresponding vocal tract resonance frequency is also shown in the figure. The vocal tract resonance was calculated by assuming plane-wave propagation in a uniform tube terminated by an infinite impedance at one end and at the other end by a radiation impedance of $Z = 0.25(ka)^2 + j0.6ka$, where k and a are the wavenumber and radius of the vocal tract tube, respectively. Due to the high stiffness and the short tracheal tube, neither the FSI nor the subglottal acoustics was strong enough to sustain self-oscillation. Phonation only occurred when the vocal fold vibration synchronized with one of the vocal tract resonances. A similar cyclic pattern in the phonation onset frequency and threshold pressure was observed. Within each cycle, the phonation frequency also closely approximated the vocal tract acoustics. However, as the phonation frequency decreased within the same cycle, the phonation frequency gradually changed from below to above the corresponding vocal tract resonance frequency. This suggests that the vocal fold self-oscillation was sustained by an inertive vocal tract at high phonation frequencies, and a compliant vocal tract at low phonation frequencies. This is in contrast to the subglottal-controlled self-oscillation in which a compliant subglottal system consistently assisted phonation onset at all observed phonation frequencies.

3.2. Interaction between subglottal and supraglottal acoustics– structure interaction

Experiments were repeated to scan the phonation frequency at onset and threshold pressure as a function of vocal tract tube length, for fixed subglottal tube lengths of 54.9 cm (squares), 70.0 cm (triangles), and 162.2 cm (circles). The results are shown in Fig. 4 in a normalized format: the frequency F_0 and the vocal tract tube length L_{dn} were normalized by $c/2\pi L_{up}$, and $c/2\pi F_0$, respectively. For comparison, the result for a short subglottal tube (17.2 cm) as shown in Fig. 3 was also shown in Fig. 4 (diamonds). Fig. 4 clearly shows the competition of two acoustics–structure interactions for dominance. When subglottal acoustics dominated, the resulting phonation frequency lay on horizontal lines corresponding to the subglottal resonances, in this case $kL_{up} \approx \pi/2$, $3\pi/2$. When the vocal tract dominated, the resulting phonation frequency lay on vertical lines corresponding to the vocal tract resonances.

In regions where it deviated from the subglottally controlled values, the phonation frequency F_0 showed different variation trends for the three cases shown in Fig. 4. For subglottal tube lengths $L_{up} = 54.9$ cm and 70.0 cm, F_0 abruptly increased and then gradually decreased as the vocal tract tube length was increased. In the supraglottal-dominated regions F_0 was consistently higher than that in the subglottally dominated regions. In contrast, for a subglottal tube length of 162.2 cm, F_0 exhibited an opposite variation pattern with the vocal tract tube length: F_0 gradually decreased with increasing vocal tract length in the supraglottally dominated regions until it abruptly increased back to its value in the subglottally controlled regions. In the supraglottal-dominated regions F_0 was consistently lower than that in the subglottally dominated regions.

The different behaviors in F_0 variation can be explained by the interaction between two coexisting acoustics–structure interaction mechanisms competing for dominance over the vocal fold dynamics. While these two acoustics–structure interactions interacted with each other, each had a distinct eigenfrequency.
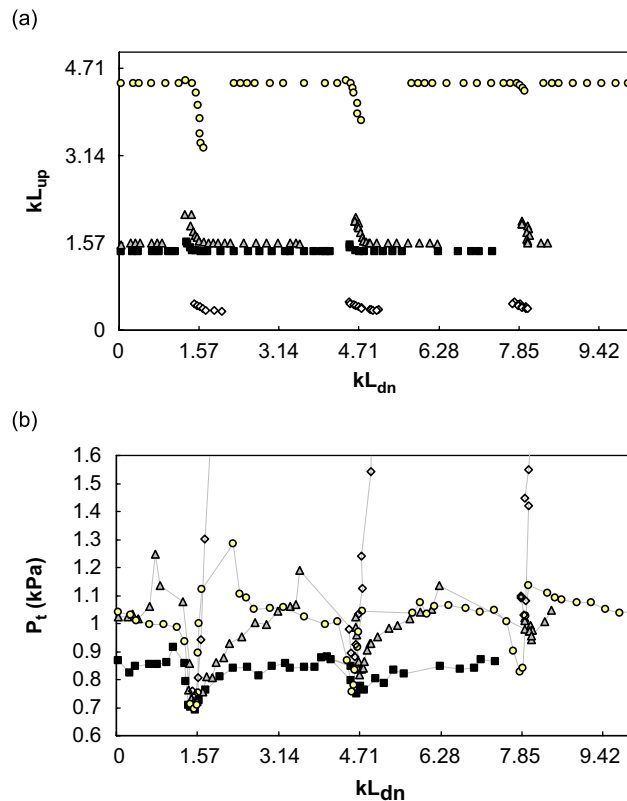

Fig. 4. (a) Normalized phonation frequency at onset $kL_{up} = 2\pi F_0 L_{up}/c$ and (b) phonation threshold pressure as a function of normalized vocal tract length $kL_{dn} = 2\pi F_0 L_{dn}/c$. The tracheal tube lengths used in the four cases were 17.2 cm ($\diamond$), 54.9 cm ($\blacksquare$), 70.0 cm ($\blacktriangle$), and 162.2 cm ($\bullet$). The four cases are also indicated in Fig. 3 by vertical lines.

The phonation frequency was determined by the acoustics–structure interaction which yielded the lower phonation threshold pressure. More specifically, for a given tracheal tube length with a threshold pressure $P_{t,\text{up}}$ associated with the subglottal acoustics–structure interaction, the vocal fold vibration synchronized to the vocal tract resonance only when the supraglottal acoustics–structure interaction provided a lower threshold pressure, or $P_{t,\text{dn}} < P_{t,\text{up}}$. Refer to Fig. 3b, this condition was satisfied first at a point on the left branch within each cycle, which generally corresponded to a high F_0 value. Depending on whether the tracheal tube length was located on the left or right branch of the P_t curve in Fig. 3b, this F_0 was either greater than (for the cases of $L_{\text{up}} = 54.9$ and 70.0 cm) or similar to (for $L_{\text{up}} = 162.2$ cm) the subglottally controlled F_0 value. Accordingly, the phonation frequency either first abruptly increased before decreasing gradually to the subglottal-controlled value (for $L_{\text{up}} = 54.9$ and 70.0 cm, Fig. 4a), or first gradually decreased before it abruptly increased back to its original value (for $L_{\text{up}} = 162.2$ cm, Fig. 4a).

In all cases, the phonation threshold pressure was generally lowered in regions where kL_{dn} values were close to a vocal tract resonance. However, outside these regions or at off-resonance conditions, the phonation threshold pressure as a function of vocal tract length showed different variation patterns, depending on the subglottal tube length. For the two cases in which the phonation frequency abruptly increased ($L_{\text{up}} = 54.9$ and 70.0 cm), for values of kL_{dn} below each vocal tract resonance, the phonation threshold pressure first increased and then decreased with the vocal tract length. Note that, in this region before the frequency jump, the phonation frequency was lower than the vocal tract resonance frequency. Therefore, Fig. 4b shows that an inertive vocal tract actually increased phonation threshold pressure, in contrast to previous predictions [9]. In the region of vocal tract dominance, the phonation threshold pressure first decreased rapidly and then increased slowly. After the phonation frequency changed back to the subglottally controlled values, the phonation threshold pressure did not immediately return to its corresponding value but only gradually increased back. The pattern was somewhat reversed in the case $L_{\text{up}} = 162.2$ cm. As the vocal tract tube length was increased, the phonation threshold pressure first gradually decreased. When the phonation frequency began to decrease, phonation threshold pressure continued to decrease but at a larger rate, similar to the previous two cases, but then increased rapidly to a large value. After the phonation frequency returned to its original value, the phonation threshold pressure then gradually decreased towards its original value.

This difference in P_t variation is consistent with the observations in Fig. 3a that the vocal tract plays different roles at different phonation frequencies. At high frequencies, an inertive (compliant) vocal tract facilitates (inhibits) self-oscillation, while at low frequencies, an inertive (compliant) vocal tract inhibits (facilitates) self-oscillation. For the case of $L_{\text{up}} = 70.0$ cm, the subglottally controlled vocal fold vibrated at a relatively low frequency (Fig. 3a, symbols $\times$), at which a compliant vocal tract assisted phonation onset, as shown in Fig. 3a (symbols $\square$ and the thin line). Therefore, the vocal tract introduced a positive damping to the coupled system for a phonation frequency below the vocal tract resonance (an inertive vocal tract), and a negative damping for phonation frequencies above the vocal tract resonances (a compliant vocal tract). Maximum positive and negative dampings were reached at a phonation frequency just below and above the corresponding vocal tract resonance, respectively (see, e.g., Fig. 11b of [8]). This variation pattern of the negative damping is consistent with the variation pattern of the phonation threshold pressure in Fig. 4b. For the case of $L_{\text{up}} = 162.2$ cm, the subglottally controlled vocal fold vibrated at a relatively high frequency (Fig. 3a, symbols $\times$), at which an inertive rather than a compliant vocal tract assisted phonation onset, as shown in Fig. 3a (symbols $\square$ and the thin line). The corresponding damping was therefore negative (positive) at phonation frequencies below (above) the vocal tract resonance, which is consistent with Fig. 4b. The case of $L_{\text{up}} = 54.9$ cm seems to be in between these two cases, with both a smaller P_t variation amplitude and a reduced range of kL_{dn} in which F_0 was influenced.

3.3. Fluid–structure interaction (FSI)

When the stiffness of the vocal fold model was lowered, phonation threshold pressure of the coupled-mode flutter was also reduced [6]. Self-oscillation of the vocal fold model occurred naturally at conditions of weak acoustical influence, e.g., for a short tracheal tube with a relatively high first resonance frequency and no vocal tract. This contrasts with previous studies in which a stiffer single-layer model was used, which failed to yield any self-sustained oscillations for tracheal lengths in the range 17–30 cm [8]. In this study, because the effective

negative damping due to subglottal acoustics was relatively small for these short tracheal tubes (see Fig. 11 of [8]), presumably the self-sustained vibrations were induced primarily by the near field FSI (see further discussion in Section 3.4).

Fig. 5a shows the phonation frequency at onset as a function of the Young's modulus of the vocal fold model, for both the full-model and hemi-model experiments. The subglottal tube length was about 11 cm and no vocal tract was attached. Phonation onset frequency generally increased with the Young's modulus. For the same vocal fold model, the phonation frequency at onset was similar for the full-model and the hemi-model configurations.

A previous study [6] showed that the phonation frequency at onset was generally close to the natural frequencies of the two interacting eigenmodes (e.g., the second and third eigenmode in [6]). For comparison, the natural frequency of the lowest structural eigenmode (generally the second eigenmode) with a dominantly in-phase medial–lateral motion along the medial surface of the vocal fold model is also shown in Fig. 5a. Two estimates of the in vacuo eigenfrequencies are shown in the figure: one was estimated using a two-dimensional plane-strain model and the other using a three-dimensional model with the same boundary conditions as in the mounted vocal fold models. The natural frequencies were calculated using the commercial finite-element software COMSOL. The comparison between the in vacuo eigenfrequencies and the measured phonation frequency at onset shows a positive correlation, which confirms that the phonation frequency at onset was strongly influenced by the structural dynamics [6].

Fig. 5b shows the phonation threshold pressure as a function of the Young's modulus of the vocal fold models, for both full- and hemi-model configurations. The phonation threshold pressure generally increased with increasing Young's modulus of the model, except for the two models with similar Young's moduli
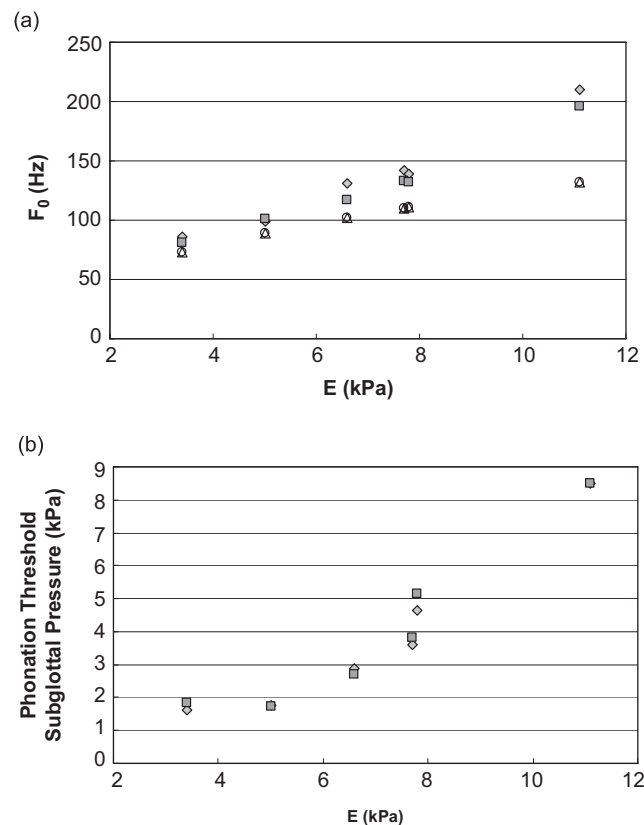

Fig. 5. (a) Phonation frequency F_0 and the in vacuo eigenfrequencies, and (b) phonation threshold pressure as a function of model Young's modulus. Tracheal tube length $L_{up} = 11$ cm. ♦: results from full-model experiments; ■: results from hemi-model experiments; △: second in vacuo eigenfrequency of a two-dimensional plane strain vocal fold model; ○: corresponding in vacuo eigenfrequency in a three-dimensional vocal fold model.

(around 7.7 kPa). For these two models, the phonation threshold pressure was different by about 1 kPa while they had similar values for the phonation frequency. The reason for this difference is not clear. Possible factors include, among others, variability in the mounting of different physical models, different prephonatory openings, and different material damping, which were not measured in this study.

3.4. Interaction between FSI and acoustics– structure interaction

Fig. 6 shows the phonation frequency at phonation onset, F_0, and phonation threshold subglottal pressure, P_t, as a function of tracheal tube length, L_{up}, for physical models with different Young's moduli. A hemi-model configuration was used, and no vocal tract was attached. The data were obtained using physical models discussed above as well as others whose Young's moduli were not measured. Although the Young's moduli of some vocal fold models were unknown, Fig. 6 showed that the phonation threshold frequencies and threshold pressures of different models collapsed onto one curve when properly normalized. Ideally, we would normalize the phonation frequency and threshold pressure by the corresponding values when no acoustic influence is present (e.g., in an anechoic subglottal system). In this study, these normalization values were approximated by the phonation frequency and the threshold subglottal pressure at short tracheal tube lengths (as reported in Section 3.3), $F_{0,ae}$ and $P_{t,ae}$, respectively. For short subglottal tubes it is expected that the acoustic influence would be minimal and therefore the values of phonation frequency and phonation threshold pressure at very
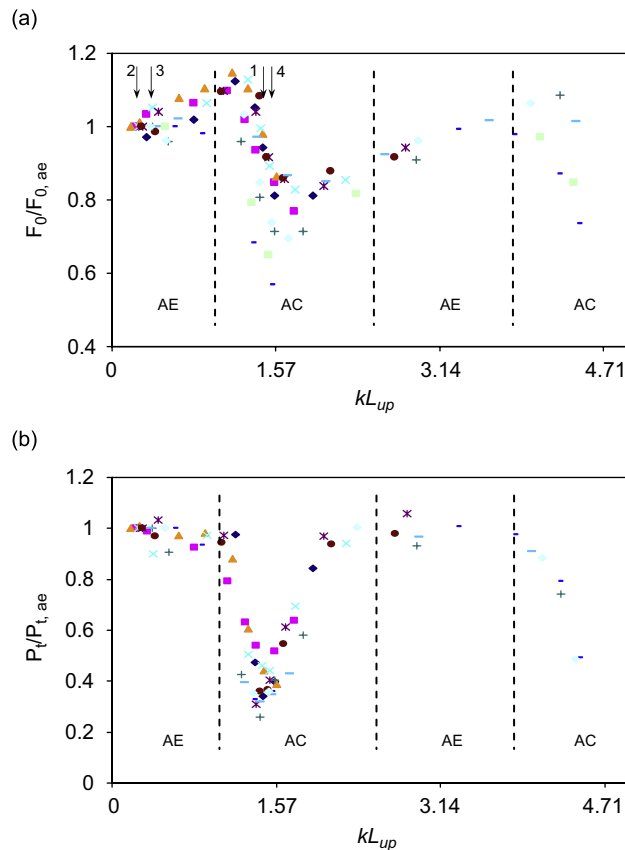

Fig. 6. (a) Normalized phonation frequency $F_0/F_{0,ae}$ and b) phonation threshold pressure $P_t/P_{t,ae}$ as a function of the normalized tube length $kL_{up} = (2\pi F_0/c)L_{up}$ for vocal fold models of different Young's modulus (as denoted by different symbols). The labels 'AE' and 'AC' denote roughly the regions of aerodynamically driven and acoustically driven modes of phonation, respectively. The arrows and numbers in the figure denote the kL values for the four cases discussed in Section 3.4. Cases 1 and 4 are in the region of strong influence of subglottal acoustics, whereas cases 2 and 4 are in a region with weak subglottal acoustic influence.

short tube lengths approach the values in aerodynamically driven mode of phonation. The tracheal tube length was normalized by the wavelength at the phonation frequency as $kL_{up} = (2\pi F_0/c)L_{up}$.

Two different regions can be identified in Fig. 6. For values of kL_{up} around odd integer multiples of $\pi/2$, which correspond to the quarter-wavelength resonances of the subglottal system, the vibration of the physical models were clearly influenced by the subglottal acoustics (labeled as 'AC' in Fig. 6). For example, as kL_{up} increased from slightly below to slightly above $\pi/2$, the normalized phonation frequency first increased gradually, then decreased rapidly toward a minimum at around $kL_{up} = \pi/2$, and then increased gradually toward one. In this region, the phonation frequency was entrained to the first subglottal resonance frequency, as $kL \approx \pi/2$. The normalized phonation threshold pressure in this region also decreased first from and then increased back toward one. This region begins when the normalized phonation frequency starts to deviate from one. In other words, the influence of subglottal acoustics becomes significant when the aerodynamically driven phonation frequency, $F_{0,ae}$, approaches one of the subglottal resonances.

In regions of minimal acoustical influence, labeled as 'AE' in Fig. 6, the normalized phonation threshold pressure stays approximately at one, indicating weak influence of subglottal acoustics on phonation threshold pressure. This is consistent with the theoretical prediction that the effective negative damping introduced by the subglottal acoustics is significant only at or around subglottal resonances (Fig. 11b in Ref. [8]). In these regions, the phonation frequency increased slightly with increasing kL_{up} value until it entrained with the acoustic resonances and began to decrease.

Although both the FSI and the acoustics–structure interaction are integral components of a complex fluid–structure–acoustics interaction, they can be viewed as two co-existing self-oscillating mechanisms competing for dominance. From this perspective, Fig. 6 also indicates that entrainment to subglottal acoustics (in terms of phonation frequency at onset and threshold pressure) depends on the relative values of the phonation threshold pressures associated with the aerodynamically driven and acoustically driven modes of vibration. Specifically, the acoustically driven mode of vibration was induced only when its phonation threshold pressure was lower than that of the aerodynamically driven vibration. This is reflected in Fig. 6b as a dip around $\pi/2$ for the phonation threshold pressure data.

The same experiments were repeated in a full-model configuration. Similar entrainment patterns as shown in Fig. 6 were obtained for both phonation frequency and threshold pressure. The phonation frequencies for the same physical model in full- and hemi-model configurations generally differed by a few Hertz. The phonation threshold pressure was generally slightly higher in the hemi-model configuration than in the full model. However, the phonation threshold flow rate was about half of that in the full-model configuration.

To compare the vocal fold vibration pattern in different regions of acoustical influence, the medial surface vibration dynamics for four cases in Fig. 6 was measured using a hemi-model configuration [11]. Table 1 shows the experimental parameters of these four cases. The corresponding kL_{up} values for these four cases are labeled in Fig. 6a. Roughly, cases 1 and 4 fell into the region of strong influence of subglottal acoustics, while cases 2 and 3 were in a region with weak acoustical influence.

Fig. 7 shows, for the four cases, the medial surface trajectories in a coronal plane, midway between anterior and posterior extremes. The corresponding spatiotemporal plots are shown in Fig. 8, for both the medial–lateral and superior–inferior components of the surface displacement. The most significant difference is that, on the superior part of the medial surface, the superior–inferior (vertical direction in Fig. 7) and medial–lateral (horizontal direction in Fig. 7) components of the vocal fold vibration were of roughly equal

Table 1
Vocal fold Young's modulus E, subglottal tube length L_{up}, phonation frequency F_0, subglottal pressure P_s, and flow rate Q for the four cases in Section 3.4. Also shown is the ratio $Q/P_s^{1/2}$ as an estimation of the glottal opening area.

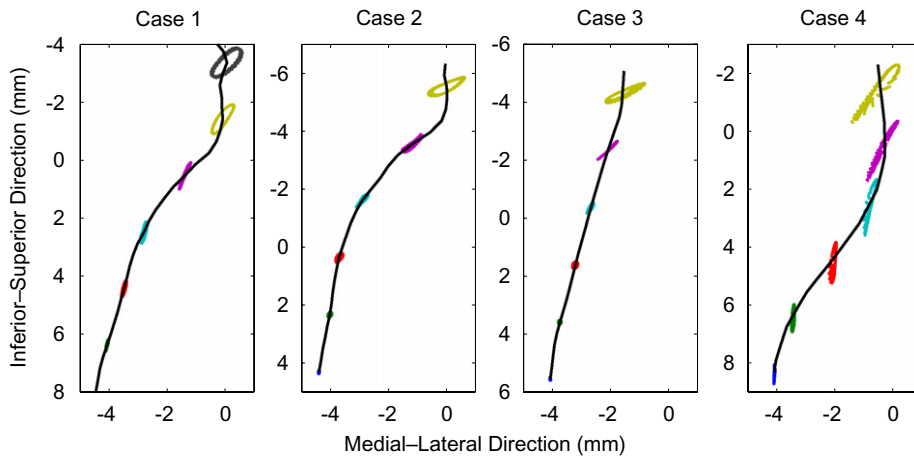| Case	E (kPa)	L_{up} (cm)	F_0 (Hz)	P_s (kPa)	Q (ml/s)	$Q/P_s^{1/2}$
1	3.4	118.1	69.8	0.6	249.1	321.5
2	3.4	17.5	83	1.95	1132	810.7
3	7.8	17.1	129.4	4.68	1238.7	572.8
4	7.8	79.5	106.4	1.72	303.3	231.3

Fig. 7. Anterior view of medial surface trajectories in a coronal plane, midway between anterior and posterior extremes, for four cases labeled in Fig. 6a and Table 1. For a clearer illustration, only trajectories from every fifth grid point are shown along the inferior–superior length.

amplitude in cases 1 and 4, which contrasts with much smaller superior–inferior components in cases 2 and 3 (see both Figs. 7 and 8). For all cases, both the superior–inferior and the medial–lateral motion decreased in amplitude toward the inferior part of the vocal fold model. However, in cases 1 and 4, the superior–inferior motion decreased at a smaller rate than the medial–lateral motion so that the motion of the inferior part of the vocal fold model was dominated by the inferior–superior motion.

Fig. 8 further shows that, in the region of considerable motion (the superior portion of the medial surface), the vocal fold model in case 4 moved almost at the same phase: the maximum displacement (both components) occurred at the same time for different locations along the medial surface in Fig. 8. In case 1, a slight phase difference in the superior–inferior component along the inferior part of the surface was observed. For all four cases, a slight phase difference in the medial–lateral motion can be observed. For cases 2 and 3 with reduced acoustic influence, the region of significant superior–inferior motion was also confined to the superior portion of the vocal fold.

These features suggest that the first in vacuo eigenmode of the one-layered vocal fold structure, which features an in-phase inferior–superior motion [11], was strongly excited when the vocal fold vibrated under strong acoustic influence. Previous studies [8,11] have shown that this eigenmode has a tendency to entrain with subglottal acoustics, probably because its in-phase superior–inferior motion couples well with the plane acoustic wave propagation in the subglottal system.

Preliminary experiments with excised human larynges also confirmed the pattern shown in Fig. 6. Fig. 9 shows the results from an excised larynx experiment using a similar subglottal system but without any vocal tract tube. The phonation frequency at onset with adducted vocal folds (narrow glottal opening at rest) stayed constant with increasing tracheal tube length until it approached one of the subglottal resonance frequencies, after which it entrained to the subglottal resonance. However, for the same range of tracheal tube lengths, this pattern was not observed for more abducted vocal folds (large glottal opening), for which the glottal opening area was comparable with the tracheal cross-sectional area. This is probably due to that, for large glottal opening, the subglottal system may have changed from a quarter-wavelength resonator to a half-wavelength resonator, leading to an increased first subglottal resonance and reduced coupling between the vocal folds and the subglottal acoustics.

4. Discussion

In a previous study using a similar vocal fold model [11], acoustically and aerodynamically driven modes of vibration were differentiated based on their medial surface dynamics. In that study, the aerodynamically driven vibration was induced by applying a vertical restrainer to the superior surface of the vocal fold model,
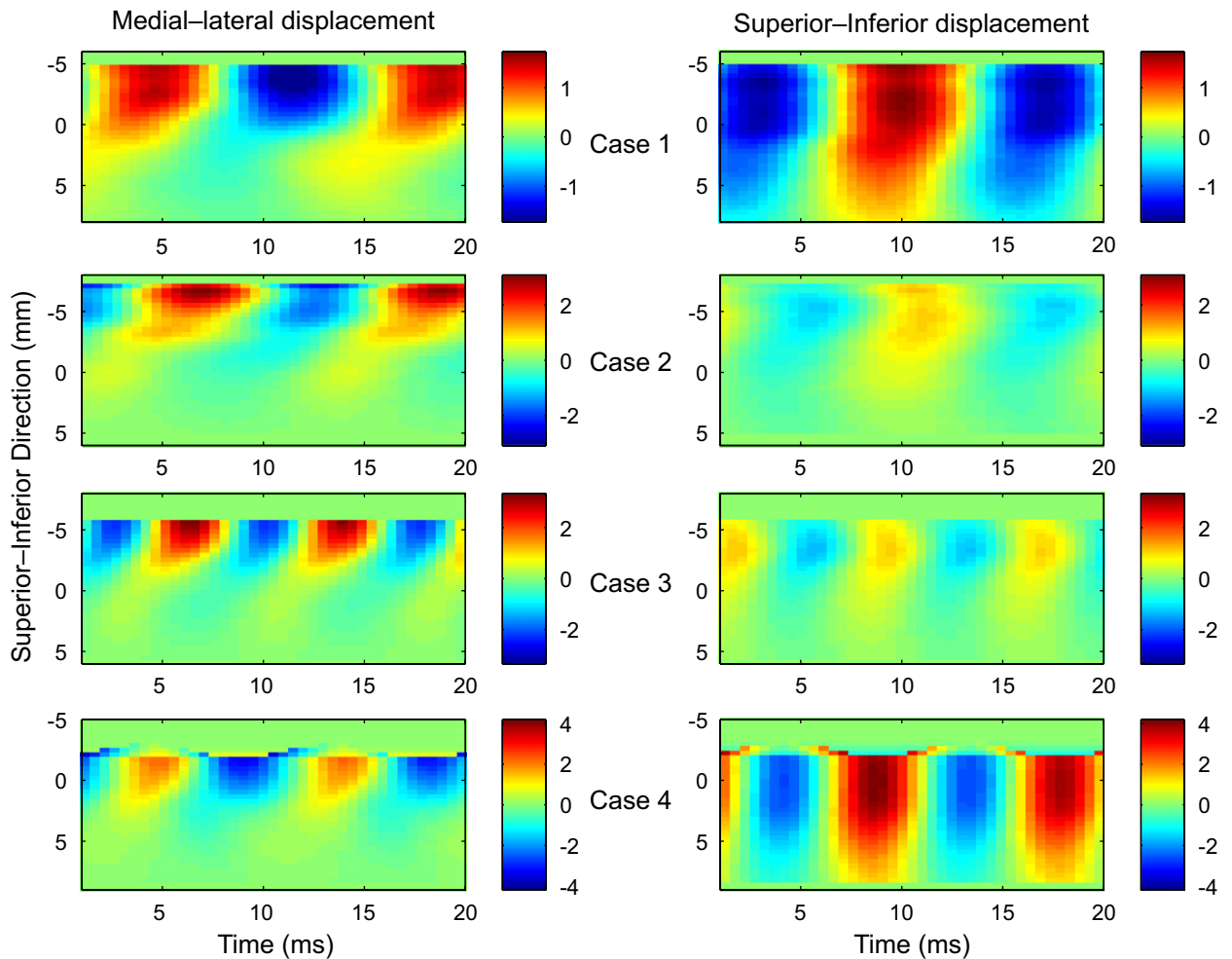
Fig. 8. The spatiotemporal plot of the medial–lateral (left column) and superior–inferior (right column) components of the surface displacement for one coronal slice of the medial surface of the vocal fold model for four cases discussed in Fig. 7 and Section 3.4. For each case, the two displacement components were normalized by a common factor, their maximum value along the slice and over time.

which presumably restricted the superior–inferior motion and reduced source–tract coupling. In this study, the aerodynamically driven mode of vibration was induced by lowering the vocal fold stiffness, which enhanced the near field FSI. Both types of aerodynamically driven modes of vibration yielded reduced amplitudes of superior–inferior motion (compare Fig. 7 of this study to Fig. 5 in [11]). However, in this study the aerodynamically driven vibration exhibited a phase difference in the medial–lateral motion along the superior–inferior direction (Fig. 8) that was much smaller than that in the restrained vocal fold model of [11] (Fig. 6 of Ref. [11]). A wave-like motion along the medial surface was clearly observed in the aerodynamically driven vibration of the restrained model, which was less obvious in cases 2 and 3 of this study. Perceptually, the subglottal acoustic pressure for the aerodynamically driven cases in the unrestrained models sounded more breathy than both the acoustically driven cases of this study and the aerodynamically driven cases in restrained models. This is probably due to a larger prephonatory glottal opening (as estimated by $Q/\sqrt{P_s}$ in Table 1) in the unrestrained aerodynamically driven cases: the flow rates (above 1000 ml/s, Table 1) in the unrestrained aerodynamically driven cases were much larger than those (around 300 ml/s) in both the acoustically driven cases of this study and the restrained cases in [11], although the subglottal pressures were comparable (around 2 kPa).

The difference in vibration patterns between these two types of aerodynamically driven vibrations is probably related to differences in the natural modes of the two vocal fold structures. For an unrestrained single-layer vocal
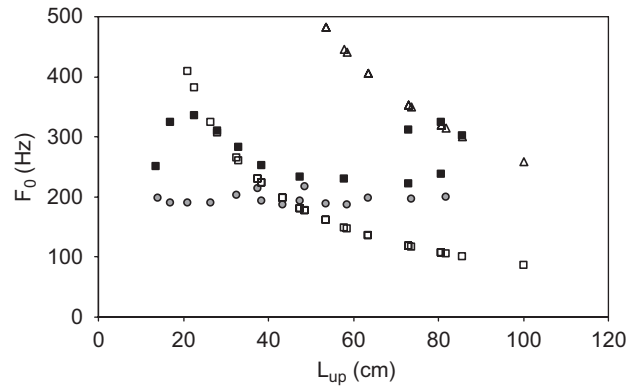
Fig. 9. Phonation frequency F_0 at onset as a function of the tracheal tube length in an excised-larynx experiment. ●: the excised larynx was not adducted; ■: the excised larynx was adducted; □: first resonance of the subglottal system calculated assuming infinite glottal impedance; △: second resonance of the subglottal system calculated assuming infinite glottal impedance. The self-oscillation of the adducted excised larynx showed a similar entrainment pattern as in Fig. 6. However, this entrainment pattern was not observed when the excised larynx was not adducted and had a relatively large glottal opening which presumably reduced source–tract coupling.

fold model, as used in this study and [8], the first eigenmode is dominated by an in-phase superior–inferior motion, i.e., the whole vocal fold body moves mostly in one direction. Applying a vertical restrainer restricted this superior–inferior motion and the vocal fold model was forced to vibrate in a higher-order mode, leading to a more wave-like motion as observed in [11]. The restraining effect may also lead to a smaller prephonatory glottal opening. Similarly, we would expect that a multi-layer vocal fold structure, such as the human vocal folds, would exhibit a vibration pattern different from that of an unrestrained single-layer model, but probably similar to that of the restrained single-layer model as in [11]. This issue will be addressed in future studies.

In this study we showed that the acoustically driven mode of self-oscillation was sustained by an inertive vocal tract at high phonation frequencies and by a compliant vocal tract at low phonation frequencies (Fig. 3a). Consistently, an inertive vocal tract could either increase or decrease phonation threshold, depending on the phonation frequency (Fig. 4b). This finding contradicts previous theoretical findings that an inertive vocal tract lowers phonation threshold pressure [9]. A possible explanation for this discrepancy is that the vocal fold model may vibrate in different patterns at low and high phonation frequencies. It is possible that, due to the excitation of higher-order eigenmodes of the vocal fold structure, the vocal fold vibration exhibited more medial–lateral motion and less superior–inferior motion at high frequencies than that at low frequencies. Such change in the vibration pattern was observed in trombone playing [20], which showed that the player's lips vibrated with a more pronounced longitudinal motion (in flow direction) at low frequencies and more pronounced transverse motion (perpendicular to flow direction) at high frequencies. If such frequency-dependent vibrational change did occur in our model, it may help to explain the different roles of the vocal tract which we observed. The medial–lateral (transverse) motion is driven by the Bernoulli's pressure, while the superior–inferior motion is driven by the difference between the subglottal and supraglottal pressure. Considering the vocal fold geometry, which had a convergent entrance, an increase in the subglottal acoustic pressure opens the glottis no matter which motion dominates. Therefore, it can be shown that a compliant subglottal system always facilitates phonation onset by lowering the phonation threshold pressure [7,8]. However, an increase in the supraglottal acoustic pressure has an opposite influence on the two components: it opens the glottis if the vocal fold vibrates in the medial–lateral direction but may close the glottis for a dominating superior–inferior motion. Using a similar theoretical approach as in [7,8], Adachi and Sato [21] showed that an inertive vocal tract facilitates self-oscillations in the medial–lateral direction, while a compliant vocal tract facilitates self-oscillations in the superior–inferior direction. Therefore, the vocal tract acoustics may have different overall influence on phonation onset depending on which motion is more dominant. If the vocal fold vibrated only in the medial–lateral direction, as assumed in [9], an inertive vocal tract would then always lower phonation threshold pressure, as concluded in [9]. Due to limitations in our present experimental setup, it was not possible to image the vocal fold vibration pattern when a vocal tract was attached. Future experiments will be designed to verify this hypothesis.

Although the subglottal and supraglottal acoustics was controlled by varying the length of the corresponding tube in this study, the same general conclusion can be extended to cases when the subglottal or supraglottal acoustical loading is varied by other means, and strong source–tract interaction is expected to occur when one of the natural frequencies of the vocal folds approaches one of the resonance frequency of the sub- or supraglottal acoustics. These other means include, for example, changing the shape of the vocal tract to vary formants [13], or changing the geometry in the region where the narrow glottis transitions to the trachea and the pharynx (e.g., changing the glottal opening area relative to the cross-sectional area of the subglottal or supraglottal tube [22], ventricular cavity and other supraglottal structures). The results from the excised larynx experiments of this study showed that geometric details in the region close to the vocal folds can significantly affect source–tract coupling. On the other hand, the natural frequency of the vocal folds can be varied significantly through laryngeal muscle activation. For example, although in speech the pitch is generally lower than the first resonance frequency of either the subglottal or supraglottal system, in singing the pitch can be dramatically increased to above 1 kHz (e.g., in soprano), which is way above the first subglottal resonance. This would change the ratio between the natural frequency of the vocal folds and the acoustical resonance frequencies of the subglottal or supraglottal resonators. Our results show that strong source–tract interaction can occur when this frequency ratio approaches one. It has been reported that singers can tune their vocal tract to benefit from a strong source–tract interaction due to the match between a formant and the pitch [13]. Finally, the result of this study (the interaction between two self-oscillating mechanisms) can be also extended to the case of infant phonation or bird vocalization, if the structural oscillation frequencies and the acoustical resonance frequencies are properly normalized.

5. Conclusions

The interaction of two mechanisms of self-oscillation (the fluid–structure interaction and acoustics–structure interaction due to either the subglottal or supraglottal acoustics) was experimentally investigated using a single-layer isotropic vocal fold model. The conclusions of this study are as follows:

1. Decreasing vocal fold stiffness in the single-layer model enhanced the near field fluid–structure interaction, and aerodynamically driven phonation occurred naturally even for short clinically relevant tracheal tube lengths, for which phonation was either not observed or observed only with the application of a vertical restraint in previous studies [8,11].
2. Significant influence of the subglottal acoustics on the vibration pattern was observed for values of kL_{up} around the corresponding resonance of the subglottal tract, where k is the acoustic wavenumber and L_{up} is the tracheal tube length. In this region, the phonation threshold pressure was reduced, and the phonation frequency at onset was significantly influenced by the dominant acoustic resonances.
3. Under strong coupling to sub- or supra-glottal acoustics, the vocal fold exhibited a dominantly in-phase motion with a strong inferior–superior component.
4. For a stiff vocal fold model, self-oscillation occurred only as the vocal fold vibration synchronized to one of the acoustic resonances. The phonation frequency entrained to either a subglottal or supraglottal resonance, depending on which acoustics–structure interaction yielded the lower threshold pressure.
5. An inertive vocal tract does not necessarily lower phonation threshold pressure.

Acknowledgments

References

[1] J.S. Cullen, J. Gilbert, D.M. Campbell, Brass instruments: linear stability analysis and experiments with an artificial mouth, *Acustica* 86 (2000) 704–724.

[2] P.J. Holmes, Bifurcations to divergence and flutter in flow-induced oscillations: a finite dimensional analysis, *Journal of Sound and Vibration* 53 (1977) 471–503.

[3] P.W. Carpenter, A.D. Garrad, The hydrodynamic stability of flow over Kramer-type compliant surfaces. Part 2. Flow-induced surface instabilities, *Journal of Fluid Mechanics* 170 (1986) 199–232.

[4] K. Ishizaka, Equivalent lumped-mass models of vocal fold vibration, in: K.N. Steven, M. Hirano (Eds.), *Vocal Fold Physiology*, University of Tokyo, Tokyo, 1981, pp. 231–244.

[5] K. Ishizaka, Significance of Kaneko's measurement of natural frequencies of the vocal folds, in: O. Fujimara (Ed.), *Vocal Physiology: Voice Production, Mechanisms and Functions*, Raven Press Ltd., New York, 1988, pp. 181–190.

[6] Z. Zhang, J. Neubauer, D. Berry, Physical mechanisms of phonation onset: a linear stability analysis of an aeroelastic continuum model of phonation, *Journal of the Acoustical Society of America* 122 (2007) 2279–2295.

[7] N.H. Fletcher, Autonomous vibration of simple pressure-controlled valves in gas flows, *Journal of the Acoustical Society of America* 93 (1993) 2172–2180.

[8] Z. Zhang, J. Neubauer, D. Berry, The influence of subglottal acoustics on laboratory models of phonation, *Journal of the Acoustical Society of America* 120 (2006) 1558–1569.

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

[10] K. Ishizaka, J.L. Flanagan, Synthesis of voiced sounds from a two-mass model of the vocal cords, *The Bell System Technical Journal* 51 (1972) 1233–1267.

[11] Z. Zhang, J. Neubauer, D. Berry, Aerodynamically and acoustically driven modes of vibration in a physical model of the vocal folds, *Journal of the Acoustical Society of America* 120 (2006) 2841–2849.

[12] H. Hatzikirou, W.T. Fitch, H. Herzel, Voice instabilities due to source–tract interactions, *Acustica* 92 (2006) 468–475.

[13] E. Joliveau, J. Smith, J. Wolfe, Vocal tract resonances in singing: the soprano voice, *Journal of the Acoustical Society of America* 116 (2004) 2434–2439.

[14] P. Mergell, H. Herzel, Modeling biphonation—the role of the vocal tract, *Speech Communication* 22 (1997) 141–154.

[15] M. Zanartu, L. Mongeau, G.R. Wodicka, Influence of acoustic loading on an effective single mass model of the vocal folds, *Journal of the Acoustical Society of America* 121 (2007) 1119–1129.

[16] D. Berry, Z. Zhang, J. Neubauer, Mechanisms of irregular vibration in a physical model of the vocal folds, *Journal of the Acoustical Society of America* 120 (2006) EL36–EL42.

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

[18] S.L. Thomson, L. Mongeau, S.H. Frankel, Aerodynamic transfer of energy to the vocal folds, *Journal of the Acoustical Society of America* 118 (2005) 1689–1700.

[19] J.J. Pawlak, D.S. Keller, Measurement of the local compressive characteristics of polymeric film and web structures using micro-indentation, *Polymer Testing* 22 (2003) 515–528.

[20] D.C. Copley, W.J. Strong, A stroboscopic study of lip vibrations in a trombone, *Journal of the Acoustical Society of America* 99 (1996) 1219–1226.

[21] S. Adachi, M. Sato, Time-domain simulation of sound production in the brass instrument, *Journal of the Acoustical Society of America* 97 (1995) 3850–3861.

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