The influence of subglottal acoustics on laboratory models of phonation

Zhaoyan Zhang,[a] Juergen Neubauer, and David A. Berry
UCLA School of Medicine, 31-24 Rehabilitation Center, 1000 Veteran Avenue, Los Angeles, California 90095-1794

(Received 13 May 2005; revised 20 June 2006; accepted 20 June 2006)

Many previous laboratory investigations of phonation involving physical models, excised larynges, and *in vivo* canine larynges have failed to fully specify the subglottal system. Many of these same studies have reported a variety of nonlinear phenomena, including bifurcations (e.g., various classes of phonation onset and offset, register changes, frequency jumps), subharmonics, and chaos, and attributed such phenomena to the biomechanical properties of the larynx. However, such nonlinear phenomena may also be indicative of strong coupling between the voice source and the subglottal tract. Consequently, in such studies, it has not been clear whether the underlying mechanisms of such nonlinear phenomena were acoustical, biomechanical, or a coupling of the acoustical and biomechanical systems. Using a physical model of vocal fold vibration, and tracheal tube lengths which have been commonly reported in the literature, it is hypothesized and subsequently shown that such nonlinear phenomena may be replicated solely on the basis of laryngeal interactions with the acoustical resonances of the subglottal system. Recommendations are given for ruling out acoustical resonances as the source of nonlinear phenomena in future laboratory studies of phonation. © *2006 Acoustical Society of America.* [DOI: 10.1121/1.2225682]

PACS number(s): 43.70.Aj, 43.70.Bk [AL] Pages: 1558–1569

I. INTRODUCTION

For years, it has been known that the vocal folds may exhibit a wide range of source-resonator coupling with the sub- and supra-glottal systems during different types of phonation. While strong source-resonator coupling is understood to be the norm for many musical instruments (Wilson and Beavers, 1974; Fletcher, 1993), weak source-resonator coupling is generally preferred for the vocal instrument (Titze, 1988a). This is because strong source-resonator coupling tends to be plagued by involuntary voice breaks, or abrupt jumps in fundamental frequency and register. Fortunately, the voice source is relatively independent of acoustic resonances for weak source-resonator coupling, which facilitates continuous, fine control of fundamental frequency and register.

Even for human phonation, source-resonator coupling tends to be strengthened at high frequencies, especially when the fundamental frequency approaches the first formant (Titze, 1988a; Joliveau *et al.*, 2004). Other variables which tend to strengthen source-resonator coupling include a stiff mucosal cover (Titze, 1988a) and a narrow epilarynx tube (Titze and Story, 1995; Mergell and Herzel, 1997). As a general rule, voice breaks and register transitions may also be indicative of strong source-tract coupling. For example, in the literature, a variety of investigators have proposed an intimate connection between subglottal acoustic resonances and vocal registers (Nadoleczny-Millioud and Zimmerman, 1938; Vennard, 1967; Van den Berg, 1968b; Large, 1972; Austin and Titze, 1997).

To reduce the risk of source-resonator coupling in laboratory models of phonation, relatively long subglottal tube lengths have often been employed. For example, in a physical model of vocal fold vibration, Titze *et al.* (1995) utilized a "10-ft-long (305 cm) helical soft-rubber subglottal tube…designed to place subglottal acoustic resonances well below any vibration frequency." Similarly, Alipour and Scherer (2001) used a subglottal tube length of 98 cm, and Thomson *et al.* (2005) used a subglottal tube length of about 180 cm. For computational models of phonation which allow source-tract coupling, a single tube length of about 17–20 cm is generally employed (Flanagan, 1958; Ishizaka *et al.*, 1976). However, it is even more common for numerical simulations to model the subglottal system as a constant pressure source (Ishizaka and Flanagan, 1972), thus precluding any interactions between the voice source and the subglottal system.

In many excised larynx studies, the subglottal tube length has simply been left unspecified (van den Berg and Tan, 1959; van den Berg, 1968a; Berry *et al.*, 1996; Švec *et al.*, 1999; Jiang *et al.*, 2003; Jiang and Zhang, 2005). Consequently, if interactions between the voice source and the subglottal system were to occur, the results could not necessarily be repeated by independent investigators. These same excised larynx studies reported various nonlinear phenomena, including bifurcations (e.g., various classes of phonation onset and offset, register changes, frequency jumps), subharmonics, and chaos—and such phenomena were attributed to the biomechanical properties of the laryngeal system. However, because such phenomena may also be indicative of strong source-resonator coupling, it may be that many of the nonlinear phenomena reported in these laboratory studies of phonation reflected laryngeal coupling with the acoustical

[a]Electronic mail: zyzhang@ucla.edu

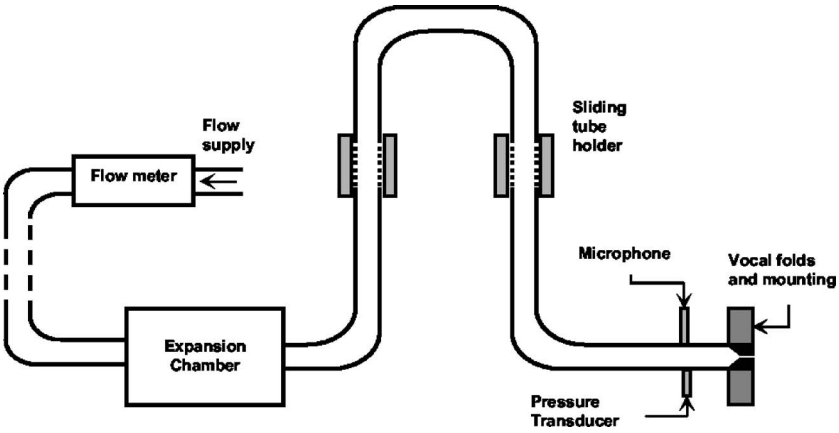

FIG. 1. Schematic of the experimental apparatus.

resonances of the subglottal system more than the biomechanical properties of the larynx. Consequently, in such studies, it was not clear whether the underlying mechanisms of the nonlinear phenomena were acoustical, biomechanical (e.g., cover or body stiffness, vocal fold geometry), or a coupling of the acoustical and biomechanical systems. Moreover, because such laboratory investigations have frequently (1) failed to report the length of the subglottal tube, or (2) utilized subglottal tube lengths considerably longer than that of the human trachea, if acoustical mechanisms were involved, one cannot be certain whether the nonlinear phenomena were reflective of human voice production, at least at the reported, vibrational frequencies. In order to investigate this concern, we hypothesize that many of the nonlinear phenomena reported in previous excised larynx experiments (van den Berg and Tan, 1959; van den Berg, 1968a; Berry *et al.*, 1996; Švec *et al.*, 1999; Jiang *et al.*, 2003; Jiang and Zhang, 2005) may be replicated solely on the basis of laryngeal interactions with the acoustical resonances of the subglottal tube. To explore this hypothesis, we will utilize the physical model of vocal fold vibration recently introduced by Thomson *et al.* (2005) over a range of tracheal tube lengths commonly used in laboratory investigations of phonation.

II. EXPERIMENTAL APPARATUS

A schematic of the experimental apparatus for the phonatory system is shown in Fig. 1. A self-oscillating physical model of the vocal folds (Thomson *et al.*, 2004; 2005) was built and used in this study. This physical model was chosen because of its repeatability and capability of sustaining phonation for long periods of time without significant changes in vibratory behavior. The vocal fold model was molded using a two-component liquid polymer solution mixed with a liquid flexibilizer solution. The stiffness of the physical model could be controlled by varying the mixing ratio of the compound solution. For further details regarding the fabrication and dynamic characteristics of the physical model, the reader is referred to several original papers (Thomson *et al.*, 2004; 2005). The vocal fold model used in this study measured approximately 1 cm in the superior-inferior direction, 1.7 cm in the anterior-posterior direction, and 0.8 cm in the medial-lateral direction. The inferior side of the vocal fold had an entrance convergence angle of 60° from the superior-inferior axis, yielding an inferior-superior thickness of the vocal fold approximately 5.4 mm. The stress-strain relationship of the physical model was measured using the 5544 Instron Testing System (Instron Corp., Norwood, MA). The Young's modulus was estimated to be approximately 11 kPa for strains in the range of 0–20%. The density was about 997 kg/m^3. Two acrylic plates were machined with a rectangular groove on the medial surface of each plate. The vocal fold model was glued into the rectangular grooves on the two 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 vocal fold plates were connected downstream from the subglottal system. In humans, the cross-sectional area of the subglottal system stays fairly constant in the trachea and primary bronchi, and then increases abruptly (Weibel, 1963). Various models of the subglottal system have been developed (Ishizaka *et al.*, 1976; Harper *et al.*, 2001; Van de Plaats *et al.*, 2006). As far as the first subglottal resonance is concerned, the subglottal system can be modeled as a single tube about 17–20 cm long, equivalent to the trachea and the primary bronchi, terminated with the lossy compliance of the lungs (Flanagan, 1958; Ishizaka *et al.*, 1976). In this study, the subglottal system consisted of a uniform rigid tube connected to an expansion chamber, simulating the lung compliance. The expansion chamber was connected to the air supply on the other end. The length of the uniform tube was varied in the range between 17 and 325 cm, encompassing the range typical of humans and those used in previous laboratory experiments (e.g., Titze *et al.*, 1995; Vilain *et al.*, 2003; Ruty *et al.*, 2005). This large range of variation was achieved by using a U-shaped circular copper tube (inner diameter of 2.35 cm), which was connected to a 32-cm-long circular PVC tube (inner diameter of 2.54 cm) on the vocal fold end and the expansion chamber on the air supply side. Tube holders (Fig. 1) were designed to allow smooth sliding of the U-shaped tube while maintaining an airtight seal. The combination of the U-shaped tube and the PVC tube allowed the length of the pseudotrachea (and the subglottal acoustics) to be varied systematically over a large range (60–325 cm). Smaller tracheal lengths were achieved without the U-shaped tube by using a 15-cm-long PVC tube with segmental extensions in 5-cm steps. This extended the range of tracheal lengths to about 17 cm on the low end. The bends of the U-shaped tube may have slightly distorted the incoming flow

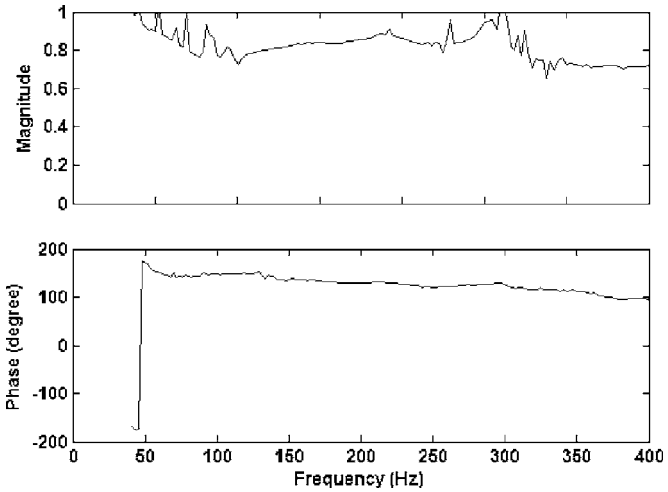

FIG. 2. Magnitude and phase of the reflection factor of the expansion chamber.

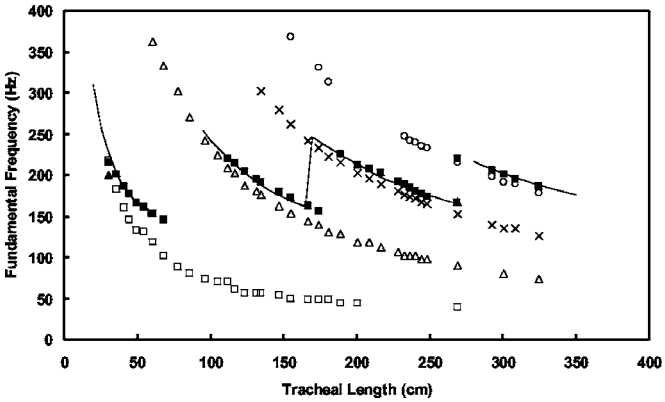

FIG. 3. Measured phonation fundamental frequency as a function of the tracheal length. Closed symbols ■: measured; ▲: second onset frequency measured if more than two frequencies are observed; open symbols: subglottal resonance frequency calculated using the measured reflection factor (□: first resonance; △: second resonance; ×; third resonance; ○; fourth resonance); —: prediction from theory.

distribution inside the tracheal tube. However, this effect was damped by the relatively long PVC tube (more than ten times the tube diameter) downstream, and considered to be negligible.

The expansion chamber was built to simulate the bronchi and the lungs. It had an inner cross-section of 23.5 $\times$25.4 cm and was 50.8 cm long. The inside of the expansion chamber was lined with a 2.54-cm-thick layer of Fiberglass. Cross-sectional area data of the subglottal airway (Ishizaka *et al.*, 1976) showed that the cross-sectional area function from the trachea and the primary bronchi to the lungs increased abruptly by 10–100 times over a distance of about 6 cm. In this setup, the cross-sectional area from the pseudotrachea section to the expansion chamber increased by a factor of about 117. The expansion chamber was connected to the airflow supply through a 15.2-m-long rubber hose, reducing possible flow noise from the air supply. The acoustic characteristics of the expansion chamber and the flow supply were evaluated using the two-microphone method (Seybert and Ross, 1977). Two microphones were mounted along the pseudotracheal tube 5 and 10 cm from the vocal fold plates, respectively. The measured equivalent termination reflection factor is shown in Fig. 2. The magnitude of the reflection factor was very close to one in the frequency range of measurement (40–400 Hz). The phase varied between 100° and 180°. The expansion chamber and upstream system behaved acoustically similar to an ideal open-ended termination to the tracheal tube.

The sound pressure in the tracheal tube was measured using a probe microphone (B&K 4182, Denmark), which was mounted flush with the inner wall of the tracheal tube, 5 cm upstream from the vocal fold plates. A pressure tap was also mounted flush with the inner wall of the tracheal tube, 2 cm upstream from the vocal fold plates. The time-averaged subglottal pressure was measured using a pressure transducer (Baratron type 220D, MKS Instruments, Inc., Andover, MA). The volumetric flow rate through the orifice was measured using a precision mass-flow meter (MKS type 558A, MKS Instruments, Inc., Andover, MA) at the inlet of the setup. Analog-to-digital conversion of the output signals was performed using a United Electronic Industries Powerdaq board (model No. PD2-MFS-8-500/16), with 16 bit resolution over a ±10 V measurement range at a sampling rate of 50 kHz.

During the experiments, the flow rate was increased from zero to a certain maximum value in discrete increments, and then decreased back to zero in discrete decrements. The maximum value of the flow rate varied, being slightly higher than the onset flow rate if phonation onset was observed or around 1800 ml/s when no obvious onset was observed. The increment in flow rate was comparatively large at low flow rates and became smaller when phonation onset was imminent. At each step, measurement was delayed for an interval (about 4–5 s) after the flow rate change, allowing the flow field to stabilize. Sound pressure inside the subglottal tube, flow rate, and subglottal pressure were recorded for a 2-s period. The procedure was repeated over a 17- to 325-cm range of tracheal tube lengths, as previously noted. For some cases, the vibration of the physical model was recorded from a superior view using a high-speed, digital camera (Fastcam-Ultima APX, Photron USA, Inc., San Diego, CA) at a frame rate of 2000 fps and at an image resolution of 1024$\times$1024 pixels.

III. RESULTS

A. Fundamental frequency

The fundamental frequency of vocal fold vibration at onset was measured from the spectra of the sound pressure signal and is shown in Fig. 3 as a function of the tracheal length. The variation of the fundamental frequency with the tracheal length showed a quasi-cyclic pattern, demonstrating a marked dependence on subglottal acoustics (see the Appendix), rather than an exclusive dependence on material properties of the physical model. The fundamental frequency decreased with increasing tracheal length within each cyclic range of the tracheal length in Fig. 3, and then jumped to the highest value of the next cyclic range. Also shown in Fig. 3 is the prediction from a simplified theoretical model of subglottal acoustics (see the Appendix) using the measured reflection factor and model constants [Eq. (A12)]. The theory

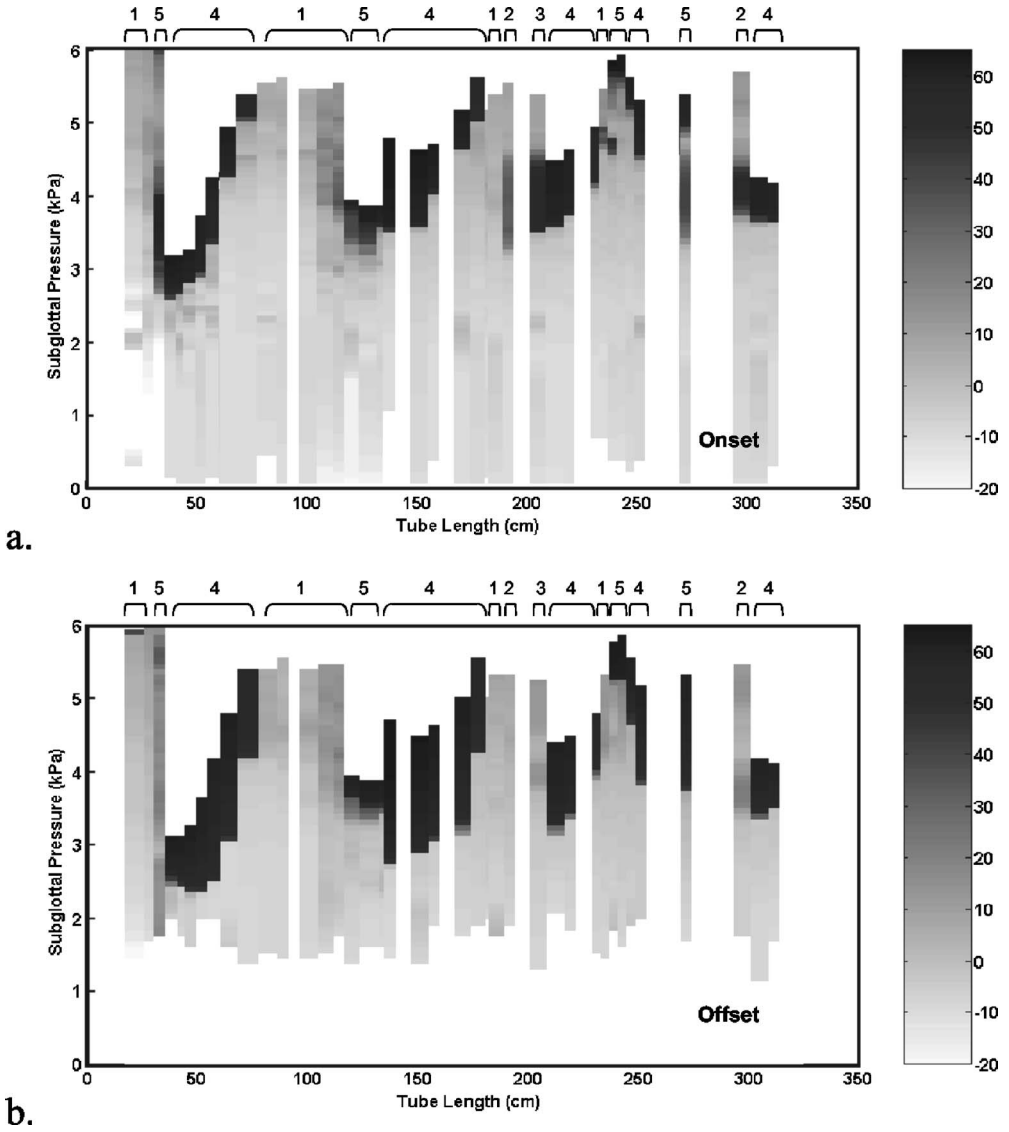

FIG. 4. Sound pressure amplitude (dB) in the subglottal tube as a function of the tracheal length and mean subglottal pressure in (a) increasing-flow and (b) decreasing-flow steps. Regions of various onset regimes are labeled for: 1. irregular phonation; 2. weak supercritical; 3 weak subcritical; 4. strong subcritical; 5. mixed types. The sound pressure amplitude in dB is represented by the gray scale.

and measurements were comparable, except for slight mismatches in the regions of frequency jumps. Note that no phonation was observed in the experiments (Figs. 3 and 9) over certain ranges of the tracheal length (e.g., 70–100 cm, and below 30 cm) in the flow rate range of the experiments. This is discussed further in Sec. IV. As the tracheal length increased, the variation in fundamental frequency over one cyclic range of the tracheal length decreased gradually from 72 Hz (over roughly a 30- to 70-cm range in tracheal length) to 55 Hz (over roughly a 190- to 270-cm range in tracheal length), which is consistent with the acoustic theory presented in the Appendix. The resonance frequencies of the subglottal system were calculated from the measured reflection factor of the expansion chamber and are also shown in Fig. 3. Clearly, the fundamental frequency closely approximated the subglottal acoustic resonance within each cyclic range of the tracheal length. Generally the fundamental frequency was slightly higher than the acoustic resonance frequency it followed. The difference gradually increased with tracheal length within each cyclic range of the tracheal length, but decreased with tracheal length between cyclic ranges. After onset, the fundamental frequency generally stayed constant and varied only slightly with the subglottal pressure.

B. Phonation onset quality

Figure 4 shows the sound pressure amplitudes for different values of the tracheal length and the mean subglottal pressure during increasing- (onset) and decreasing-flow (offset) experiments. The sound pressure amplitudes were estimated as the standard deviation of the recorded sound pressure data. Phonation onset and offset were identified by a significant change in the sound pressure amplitude between two consecutive steps (or a significant change of gray scale in Fig. 4).

The onset and offset behavior (related to the Hopf bifurcation from nonlinear systems theory [Berge *et al.*, 1984])

demonstrated a quasicyclic pattern based on the subglottal acoustics. In particular, alternating weak and strong onset regions appeared as the tracheal length increased. Four types of phonation onset were identified: irregular, weak supercritical, weak subcritical, and strong subcritical.

1. Irregular onset

Regions existed in Fig. 4 in which no apparent phonation onset was observed. An example is shown in Fig. 5(a), which was measured for a tracheal length of 182 cm. The left figure shows a bifurcation diagram in which the sound pressure amplitude is plotted against the mean subglottal pressure for both increasing- and decreasing-flow experiments. There was no clear phonation onset for this case. Only a very weak onset was observed in the increasing stage, as indicated by the small hump (around 4 kPa) in the bifurcation diagram. The right figure shows the corresponding power spectra of the measured sound pressure at each step, with subglottal pressure increased from zero to an upper limit and decreased back to zero. The power spectra showed that the energy inside the subglottal tube distributed mostly around the acoustic resonance frequency of the subglottal system. Time series data of the sound pressure showed a periodic component at 220 Hz, superimposed with intermittent low-frequency modulations, and chaos. The periodic component was so weak that the sound field was dominated by such instabilities.

2. Strong subcritical onset

Strong phonation onset observed in the experiments was normally subcritical (an abrupt increase in oscillation amplitude). One example is shown in Fig. 5(d) measured at a tracheal length of 208.4 cm. The sound pressure amplitude increased abruptly at a critical subglottal pressure [around 3.5 kPa in Fig. 5(d)]. After onset the sound pressure amplitude continued to increase with the subglottal pressure in the flow range of the experiments. In the decreasing-flow stage, the sound pressure amplitude decreased following a slight different curve from that in the increasing-flow stage. At a critical (offset) subglottal pressure slightly lower than the onset subglottal pressure, the sound pressure amplitude decreased abruptly to a very small value and phonation abruptly ceased. Power spectra during phonation showed that both the fundamental and harmonics were strongly excited. The onset pattern was repeatable whether immediately following a previous measurement, or after a period of rest.

3. Weak supercritical and subcritical onset

In between regions of irregular onset and strong onset in Fig. 4, two transitional types of onset existed. The onsets of these types were typically weak in strength and existed only for a certain limited range of subglottal pressure, i.e., there was a phonation offset even in the increasing-flow stage. Unlike the strong onset, the weak onset can be either supercritical (a gradual increase in glottis oscillation amplitude) or subcritical. Figure 5(b) shows one case of a weak and supercritical onset, obtained using a tube length of 188.7 cm. The difference between this type and the irregular onset is that there was a clear phonation onset above a critical subglottal pressure (or onset pressure). Time series sound pressure data at the onset showed a clear dominant periodic component with slight modulations and noise. The modulations disappeared as subglottal pressure continued to increase. Beyond the onset pressure, the sound pressure amplitude gradually increased with increasing subglottal pressure, indicating a supercritical (or soft) onset. The onset was weak and phonation was sustainable for only a narrow range of subglottal pressure. As subglottal pressure further increased, the sound pressure amplitude saturated and then gradually decreased to phonation offset. Although there was an apparent phonation onset, the onset strength was much weaker than that in the strong onset case. The maximum sound pressure was at least ten times smaller than that in the strong onset case. In the sound pressure spectra, only the fundamental frequency was strongly excited during the weak onset. The harmonics were hardly excited. It was observed in this study that the weak type of onset was always accompanied by weakly excited harmonics. In contrast to the strong onset case, there was no phonation in the decreasing-flow stage.

A case of weak subcritical onset is shown in Fig. 5(c), which was obtained for a tracheal length of 200.7 cm. This case was similar to the previous case [Fig. 5(b)], except that the onset was of the subcritical type. There was a jump in the sound pressure amplitude at phonation onset. As in the previous case, phonation was sustained only for a narrow range of subglottal pressure in the increasing-flow stage. As the subglottal pressure increased further, the sound pressure amplitude first saturated and then gradually decreased to phonation offset. The maximum sound pressure in this case was much larger than that in the previous case, but still smaller than that in the strong onset case. Sound pressure spectra during the onset showed that harmonics (particularly the first and third) were weakly excited. As in the weak supercritical case, no apparent phonation was observed in the decreasing-flow stage.

For both of the two weak onset cases, in an immediate repetition of the experiment, the bifurcation pattern (onset and offset) was qualitatively repeatable but with a much reduced saturation amplitude of the sound pressure. Phonation offset occurred at a lower value of subglottal pressure and the range of subglottal pressure for onset was reduced. The complete bifurcation pattern was repeatable only after a period of rest.

4. Coexistence of multiple onsets

More than one type of onset was observed at a same tracheal length. Figure 6 shows a case in which a supercritical onset was immediately followed by a subcritical onset. The tracheal length was 123.5 cm. The sound pressure spectra for these two onsets show that the difference lies in the

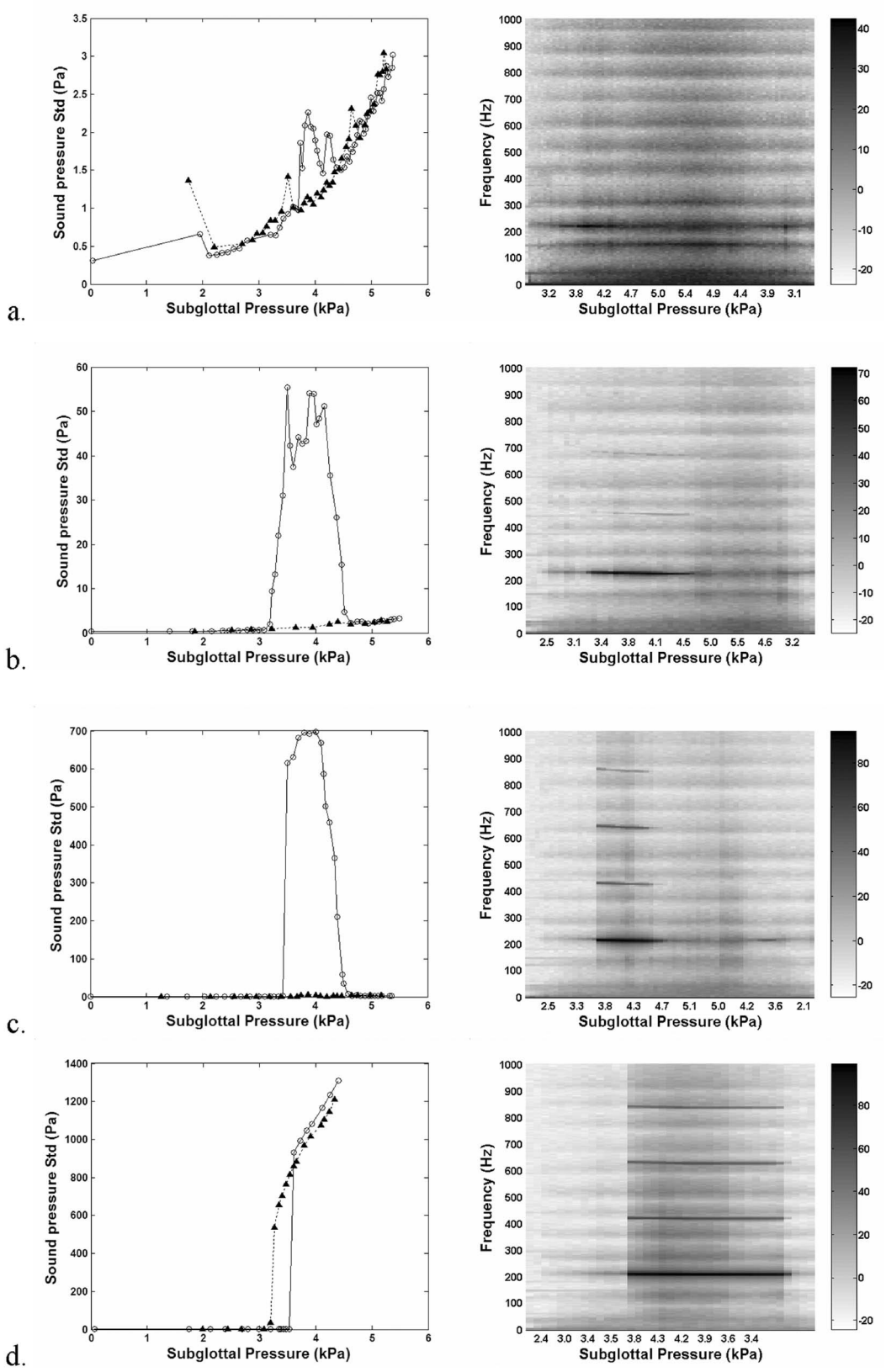

FIG. 5. Sound pressure amplitude (left) and power spectra (right) for increasing and decreasing subglottal pressure at four typical onset types: (a) irregular; (b) weak supercritical; (c) weak subcritical; and (d) strong subcritical. ○: increasing subglottal pressure; ▲: decreasing subglottal pressure. The power spectral amplitude in dB is represented by the gray scale.

excitation of the harmonics. The abrupt transition at the subcritical onset was accompanied by a boost in the strength of the harmonics, especially the second harmonic. A similar supercritical and subcritical behavior was observed in the decreasing-flow stage, in which the sound pressure amplitude first decreased abruptly to an intermediate value and then decreased gradually to phonation offset.

Figure 7 shows a case (with a tracheal length of 269 cm)

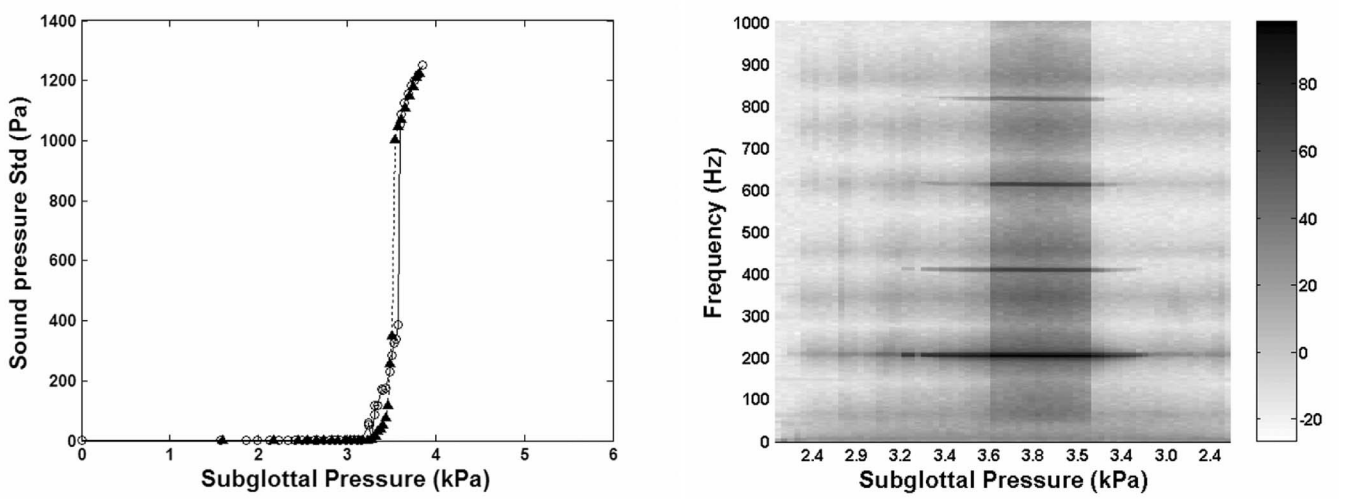

FIG. 6. Sound pressure amplitude (left) and power spectra (right) for increasing and decreasing subglottal pressure. Tracheal length $L=123.5$ cm. Two types of onset can be observed with a supercritical onset followed by a subcritical onset. ○: increasing subglottal pressure; ▲: decreasing subglottal pressure. The power spectral amplitude in dB is represented by the gray scale.

in which onsets of two different fundamental frequencies were observed. In this figure, the following phenomena were noted: a register change, a frequency jump, and a distinct entrainment to the subglottal acoustics. The onset at lower subglottal pressure (around 3.3 kPa) was weak and supercritical, with a fundamental frequency around 220 Hz. The sound pressure amplitudes increased slowly with the subglottal pressure after onset and then decreased gradually to an offset (around 4.5 kPa). The harmonics were weakly excited, characteristic of a weak onset. As subglottal pressure continued to increase, a second onset occurred at a subglottal pressure of about 4.9 kPa. The second onset was strong and subcritical, with a lower fundamental frequency around 168 Hz. Harmonics were strongly excited during the second phonation. The two fundamental frequencies corresponded roughly to the third and fourth resonance frequency of the subglottal system (Fig. 3). Superior views of the physical vocal fold model were recorded using a high-speed camera during the two onsets. The time history of a medial-lateral line taken from the center of the vocal folds (known as a spatio-temporal plot) is shown in Fig. 8 for a few oscillation cycles. The vocal folds during the first phonation demonstrated a falsetto-like, high-frequency, and low-amplitude vibration pattern, while the second phonation showed a more chest-like vibration with much larger amplitude at a lower frequency.

As discussed in Sec. V of the Appendix, more than one solution may exist for the fundamental frequency of the coupled laryngeal-subglottal system. However, only the solution with the largest negative damping is singled out and amplified. The two-onset case shown in Fig. 7 was in the region where frequency jump occurred in Fig. 3. In this region, several competing solutions were likely to have similar total damping, therefore similar onset threshold pressure. Even a slight increase in the subglottal pressure could trigger the system to favor one over another. It is possible that the

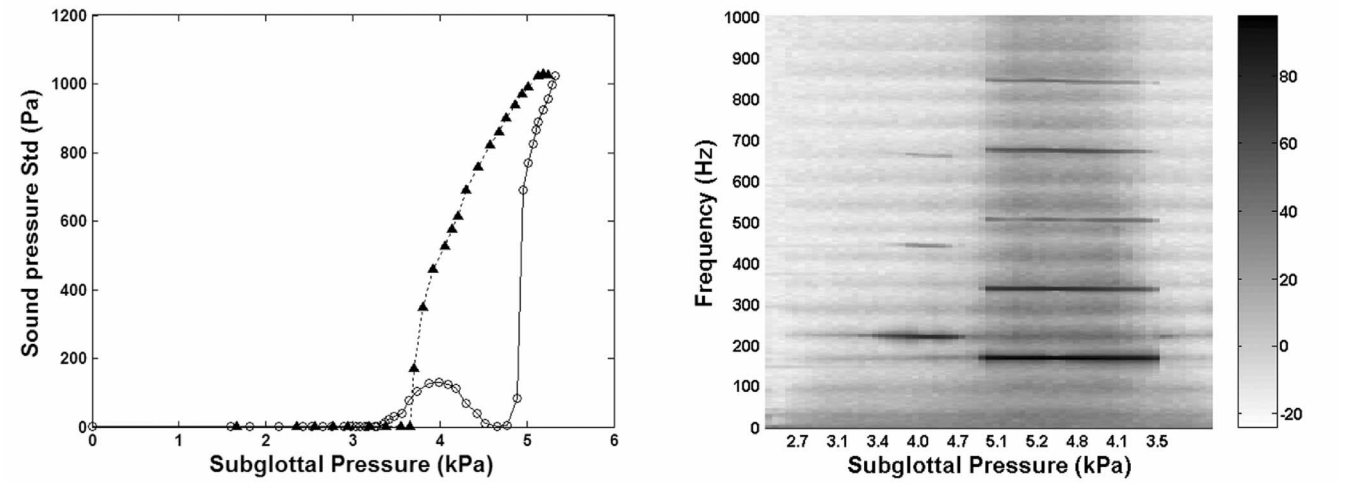

FIG. 7. Sound pressure amplitude (left) and power spectra (right) for increasing and decreasing subglottal pressure. Tracheal length $L=269$ cm. Two types of onset with different fundamental frequencies can be observed with a supercritical onset followed by a subcritical onset, ○: increasing subglottal pressure; ▲: decreasing subglottal pressure. The power spectral amplitude in dB is represented by the gray scale.

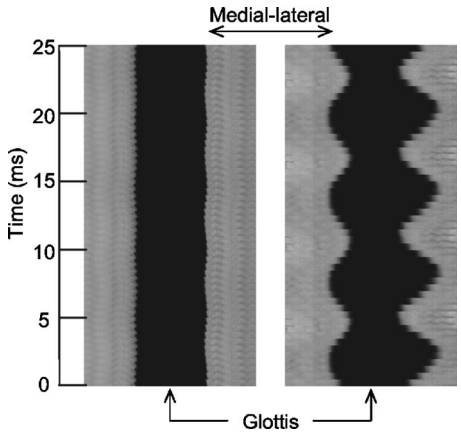

FIG. 8. Time history of a medial-lateral strip of the vocal folds at the center during the two types of onset shown in Fig. 7. Left figure: onset with high F_0: right figure: onset with low F_0. Note that there is a phase difference between the left and right vocal folds.

first onset in Fig. 7 had a lower threshold subglottal pressure but a weaker strength. As subglottal pressure further increased the first one was less preferred. Above the threshold pressure for the second type, the second onset became the favored and dominated with a much stronger onset.

C. Subglottal pressure at onset and offset

Onset and offset subglottal pressures were measured for cases with clear onset and offset. The onset and offset pressures were defined as the subglottal pressure at which sudden change in sound pressure amplitude first occurred. For supercritical onset, it was difficult to identify the exact conditions at which phonation onset and offset occurred. In these cases, the onset (offset) subglottal pressure was estimated as the subglottal pressure at which there was a considerable increase (decrease) in the slope in the plot of the sound pressure amplitude versus subglottal pressure. No offset pressure was measured for weak onsets as there was normally no phonation in the decreasing-flow stages.

The measured onset and offset pressures are shown in Fig. 9. Both the onset and offset subglottal pressures showed cyclic behavior in their variation with the tracheal length.

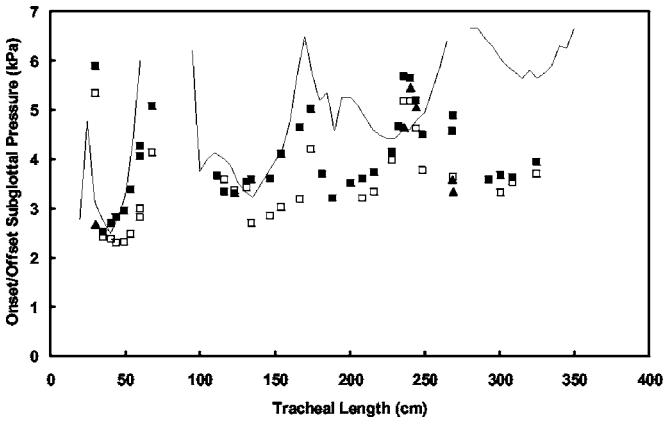

FIG. 9. Measured phonation onset (closed symbols) and offset (open symbols) threshold subglottal pressure as a function of the tracheal length.—: predicted onset threshold pressure from theory. ■: major onset; ▲: weak or supercritical onset for cases where more than one onset is observed.

Within each cyclic range of the tracheal length, the onset pressure slightly decreased and then increased with the tracheal length. The minimum value of the onset subglottal pressure within each cyclic range slightly increased with tracheal length. Theoretically, this behavior could be reproduced by including acoustic losses inside the tracheal tube in the model (see the Appendix). The offset pressure followed a similar trend as the onset pressure. Generally the offset subglottal pressure was lower than the onset subglottal pressure. Within each cyclic range of the tracheal length, the difference between onset and offset pressures was largest at the local minimum of the onset subglottal pressure, and gradually decreased as the tube length increased.

Also shown in Fig. 9 is the predicted onset pressure from the theory in the Appendix. The predicted onset pressure was overall slightly higher than the measured values. The theoretical model in this study only focused on the effects of the subglottal system. Accurate prediction of phonation onset pressure would require other factors to be included, which are beyond the scope of this study. The predicted and measured values exhibited the same general trends over the first two cyclic ranges of the tracheal length (up to a tracheal length of about 170 cm). There was a slight mismatch between measurement and prediction in the tracheal length range of the third cyclic range of the onset pressure. The highest onset pressure within the third cyclic range in experiments occurred at a tracheal length around 240 cm, while the predicted tracheal length was around 270 cm. In this region (230–270 cm), experiments exhibited mixed types of onsets and onsets of distinct frequencies. It is possible that two neighboring cyclic ranges overlapped for large values of the tracheal length, which may have distorted the onset pattern.

The variation of the phonation onset pressure with tracheal length was related to that of the onset strength. Generally low onset pressure corresponded to the regions of strong onset, while high onset pressure occurred at regions of weak and irregular onset types. Comparison of Figs. 3 and 4 shows that regions of irregular onset occurred roughly at fundamental frequencies above a threshold value (about 215 Hz), while the strong onset occurred at fundamental frequencies below this threshold value. As discussed in the Appendix, the phonation onset condition [Eq. (A11)] tended to restrict the fundamental frequency to a limited range. In this range the total system damping was negative. This may correspond to the strong onset case. For F_0 above this range, the total system damping was positive and no phonation or irregular phonation would occur. At the upper bound of this range, the total system damping was very small, either positive or negative, and the system is at a critical state. This seems to correspond to the regions of weak onset. As the theoretical model in the Appendixwas based on linear theory, it could not predict all the onset types encompassed by nonlinear systems theory. Nevertheless, it is reasonable to suppose that total system damping may have played a role in determining onset quality.

IV. DISCUSSION

Many previous laboratory investigations of phonation involving physical models, excised larynges, and *in vivo* canine larynges have failed to fully specify the subglottal system. Many of these same studies have reported a variety of nonlinear phenomena, including bifurcations (e.g., various classes of phonation onset and offset, register changes, frequency jumps), subharmonics, and chaos—and attributed such phenomena to the biomechanical properties of the larynx. However, such nonlinear phenomena may also be indicative of strong coupling between the voice source and the subglottal tract. Consequently, in such studies, it is not clear whether the underlying mechanisms of such nonlinear phenomena are acoustical, biomechanical, or a coupling of the acoustical and biomechanical systems. Moreover, because such laboratory investigations have frequently failed to report the length of the subglottal tube, or utilized subglottal tube lengths considerably longer than that of the human trachea, if acoustical mechanisms were involved, one cannot be certain whether the nonlinear phenomena were reflective of human voice production, at least at the reported vibrational frequencies. Using a physical model of vocal fold vibration, it is shown that such nonlinear phenomena may be replicated solely on the basis of laryngeal interactions with the acoustical resonances of the subglottal tube, using tracheal tube lengths which have been commonly reported in the literature, and no changes in the biomechanical properties of the laryngeal model.

In these experiments, no phonation was observed for short tracheal lengths (17–30 cm). In subsequent studies, however, phonation was achieved with the same physical model for short tracheal lengths when the flow rate was increased beyond the 1800 ml/s levels of flow rate utilized in this study. As might be expected, lower phonation threshold pressures have also been achieved at these short tracheal tube lengths when the vocal fold stiffness was lowered. Nevertheless, for longer tracheal tube lengths, phonation onset was strongly influenced by the subglottal acoustics in a cyclic manner, as previously shown. Moreover, for these longer tracheal tube lengths, after onset occurred, the secondary instabilities such as period-doubling, biphonation, and frequency jumps appeared more frequently.

In addition, this study has confirmed that the constructive or destructive interference of the subglottal acoustics can result in changes of the vibratory modes of the vocal folds, which Titze (1988b) suggested as a possible explanation for register change. For singing, in which the subglottal resonances are close to the pitch frequency, the subglottal influence is likely to play a more important role. Furthermore, the vocal tract, another acoustic resonator above the glottis, is expected to have similar influence on the vocal fold vibration. It has been shown that singers can tune the vocal tract to benefit from the pitch frequency being close to a formant (Joliveau *et al.*, 2004). The interplay between the subglottal system and the vocal tract will be addressed in a future study.

The results presented in this paper indicate that laboratory experiments involving excised larynges or physical models should be performed with a detailed understanding of the subglottal system. Previous excised-larynx experiments have reported a variety of onset conditions and instabilities, without including a detailed description of the subglottal system utilized. Accurate interpretation of the results would be difficult without a good understanding of the characteristics of the subglottal system. Indeed, without appropriate details regarding the subglottal system, it would also be impossible to compare results across laboratories. For example, consider the experimental setup used in this study. With a slight change of the tracheal length from 170 to 220 cm, the onset pattern changed significantly from irregular, weak supercritical, to strong subcritical. Comparison of phonation threshold pressure across these different cases would be meaningless.

The following conclusions are reached: (1) the only way to rule out source-tract interactions in laboratory experiments of phonation is through the construction of an anechoic termination, which is a time-consuming and difficult task (see Zhang *et al.*, 2002); (2) if an anechoic termination is not utilized, source-tract interactions are always a possibility, so the subglottal system must be fully specified, and subglottal resonances identified; (3) when register jumps and other nonlinear phenomena occur, one must rule out the possibility of source-tract interactions before concluding that such phenomena are solely a reflection of the laryngeal biomechanics, (4) in case investigators want to study source-tract interactions, to help ensure their possible relevance to human phonation, the tracheal tube length should include physiological lengths (approximately 12–17 cm).

V. CONCLUSIONS

Many previous laboratory studies of human phonation have utilized tracheal lengths considerably longer than found in the human system. Other studies have simply not reported the tracheal dimensions. However, this study has shown the potential for strong interactions between laryngeal dynamics and subglottal acoustics. This concern becomes amplified when one is studying nonlinear vocal phenomena and attempting to relate such phenomena to laryngeal biomechanics. As illustrated in this study, this task cannot be performed without ruling out the possible influence of subglottal acoustics.

ACKNOWLEDGMENTS

This investigation was supported by Research Grant Nos. R01 DC03072 and R01 DC004688 from the National Institute on Deafness and Other Communication Disorders, the National Institutes of Health. The authors also acknowledge the assistance of Dr. Scott L. Thomson and Dr. Luc G. Mongeau in providing a prototype vocal fold model and for their suggestions in constructing the vocal fold models.

APPENDIX: DETAILS OF THE SIMPLIFIED THEORETICAL MODEL

Patterned after the models of Gupta *et al.* (1973) and Wilson and Beavers (1974), in the following a one-dimensional, theoretical model of the subglottal acoustics is presented which captures the vocal phenomena in this laboratory investigation. We assumed a steady laminar flow

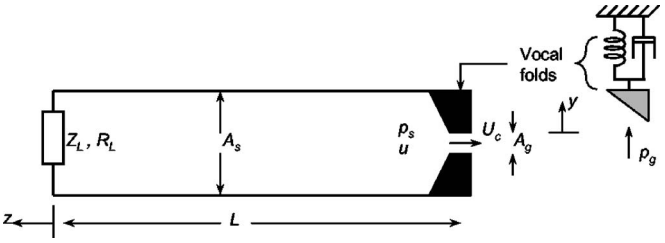

FIG. 10. Simplified model of the vocal folds and subglottal system. The vocal folds are modeled as a symmetric spring-mass-damping system.

through the glottis prior to phonation onset. Stability of the system under small-amplitude, single-frequency disturbances in the subglottal system was considered.

1. The vocal fold model

The vocal fold was modeled as a symmetric one-dimensional spring-mass-damping system (Fig. 10). Consider a small-amplitude simple harmonic oscillation of the vocal folds $y=Ye^{i\omega t}$, where y is the displacement of the vocal fold mass from its equilibrium position, Y is the complex displacement amplitude, and ω is the complex angular frequency. For small-amplitude oscillations the nonlinearities in the coupled system can be neglected so that all dynamic variables are simple harmonic functions of the same frequency.

The driving pressure applied on the vocal fold surface, p_g, can be obtained by integrating the pressure profile within the glottis. It can be related to the subglottal pressure p_s as follows:

$$p_g = C_g p_s, \tag{A1}$$

where the coupling factor C_g accounts for the dependence of the mean pressure p_g on the resting glottal geometry. The equation of motion for vocal fold oscillation is

$$M\ddot{y} + B\dot{y} + Ky = C_g p_s', \tag{A2}$$

where p_s' is the fluctuating component of the subglottal pressure p_s, M, K, and B are the effective lumped mass, stiffness, and damping per unit area of the vocal folds, respectively.

2. Pressure-flow relationship through the glottis

For an open prephonatory glottis and small-amplitude oscillation, the unsteady flow through the glottis can be accurately modeled using a quasisteady approximation (Zhang *et al.*, 2002). The jet centerline velocity u_c is related to the subglottal pressure p_s by

$$p_s = \tfrac{1}{2}\rho u_c^2. \tag{A3}$$

Equation (A3) can be linearized by decomposing the pressure and flow fields into a mean and a fluctuating component and neglecting higher-order terms:

$$p_{s0} = \tfrac{1}{2}\rho u_{c0}^2, \quad p_s' = \rho u_{c0} u_c', \tag{A4}$$

where (p_{s0}, u_{c0}) denote the mean components, and (p_s', u_c') denote the fluctuating components. Similarly, linearization of the continuity equation gives, neglecting density fluctuations,

$$u_s' A_s = C_d(A_{g0} u_c' + u_{c0} A_g'), \tag{A5}$$

where C_d is the orifice discharge coefficient, and A_g is the minimum cross-sectional area of the glottal channel. The fluctuating glottal area A_g' is related to the vocal fold displacement by

$$A_g' = 2l_g y = 2l_g Y e^{i\omega t}, \tag{A6}$$

where l_g is the length of the glottis.

3. Subglottal system

The subglottal system consists of a uniform tube of length L and cross-sectional area A_s simulating the trachea (Fig. 10). The bronchi and the lungs are modeled as an equivalent acoustic impedance Z_L (or reflection factor R_L). The input impedance of the subglottal system to the glottis, Z_s, is (Kinsler *et al.*, 2000):

$$Z_s = \frac{P_{z=-L}}{U_{z=-L}} = \rho c \frac{e^{-ik(-L)} + R_L e^{ik(-L)}}{e^{-ik(-L)} - R_L e^{ik(-L)}}, \tag{A7}$$

where ρ is the air density, c is the speed of sound, and P and U are the complex acoustic pressure and velocity inside the subglottal tube, respectively. The complex wave number $k = 2\pi f/c - i\alpha$, where α is the absorption coefficient which accounts for the acoustic loss inside the trachea.

4. Equation of motion of the coupled system

The subglottal acoustic pressure and velocity are related to the fluctuating subglottal pressure and flow velocity by

$$p_s' = P_{z=-L} e^{i\omega t}, \quad u_s' = U_{z=-L} e^{i\omega t}. \tag{A8}$$

Combining Eqs. (A4)–(A8) yields

$$p_s' = Z_a y = \frac{2Z_s C_d u_{c0} l_g}{\left(1 - \dfrac{Z_s C_d A_{g0}}{\rho u_{c0} A_s}\right) A_s} y, \tag{A9}$$

where Z_a is the effective loading of the subglottal system to the vocal folds. Equation (A9) gives an explicit relationship between the driving oscillating pressure in the subglottal tube and the resulting oscillation of the vocal folds. In general, the effective loading Z_a is a function of tracheal length, oscillating frequency, subglottal pressure, and coupling between the tracheal and the glottis A_{g0}/A_s. The characteristic equation of the coupled system can be obtained from Eqs. (A2) and (A9):

$$[-\omega^2 M + i\omega(B - C_g X_a/\omega) + K - C_g R_a] Y = 0, \tag{A10}$$

where R_a and X_a are the real and imaginary part of Z_a, respectively.

Equation (A10) shows that the influence of the subglottal acoustics on the vocal fold oscillation is twofold. First, the subglottal acoustics introduces an effective damping $-C_g X_a/\omega$ into the coupled system. This effective damping may be positive or negative, depending on the oscillating frequency and tracheal length, and may increase or decrease the phonation threshold pressure. The subglottal acoustics

also introduces an effective stiffness $-C_g R_a$. This will lead to cyclic variations of the fundamental frequency with the tracheal length.

Equation (A10) can be solved numerically for the complex angular frequency $\omega=\omega_r+i\omega_i$ as a function of the tracheal length and the subglottal pressure. The fundamental frequency F_0 can then be determined from the real part of ω. For sustained oscillations of the vocal folds, the fundamental frequency also has to satisfy the net energy transfer condition, i.e., the total system damping needs to be zero or negative:

$$B - C_g X_a/\omega_r \leqslant 0, \tag{A11}$$

where the equality is reached at phonation onset. Phonation threshold pressure can be determined to be the minimum subglottal pressure for which Eq. (A11) was satisfied. Note that when the inequality in Eq. (A11) is satisfied, the vocal fold displacement will increase exponentially. The small-amplitude oscillation assumption will not be valid and this model is not applicable anymore.

5. Open-ended ideal tracheal tube

As an example, Eq. (A10) was solved for the case of an ideal open-ended and lossless trachea as a function of kL, where k is the wave number and L the tracheal length. Typical results are shown in Fig. 11 for a mean subglottal pressure 1.6 kPa. We used the following values for the model constants, which were adapted from the two-mass model (Ishizaka and Flanagan, 1972):

$$M = 2.656 \text{ kg m}^{-2}, \quad K = 1.904 \times 10^6 \text{ N m}^{-3},$$

$$B = 1600 \text{ N s m}^{-3},$$

$$l_g = 1.4 \times 10^{-2} \text{ m}, \quad A_{g0} = 6 \times 10^{-6} \text{ m}^2,$$

$$A_t = 6.45 \times 10^{-4} \text{ m}^2,$$

$$C_d = 1, \quad C_g = 0.5. \tag{A12}$$

The fundamental frequency showed a periodic variation with kL. Over most of the range of kL, the fundamental frequency stayed close to the natural frequency of the lumped vocal fold model

$$\omega_0 = \sqrt{K/M - (B/2M)^2}, \tag{A13}$$

which was obtained when kL was integer multiples of π. The subglottal influence on the fundamental frequency dominated only at values of kL close to the quarter-wavelength resonances of the subglottal system. In this range the fundamental frequency closely approximated the quarter-wavelength resonance of the subglottal system. The maximum fundamental frequency was obtained when kL was odd multiples of $\pi/2$:

$$\omega_{\max} = \sqrt{K/M - (B/2M)^2 + C_g l_g P_{s0}/MA_{g0}}, \tag{A14}$$

which increases with the mean subglottal pressure. In between the two extreme cases, the fundamental frequency varied smoothly between ω_0 and $\omega_{\max}$.

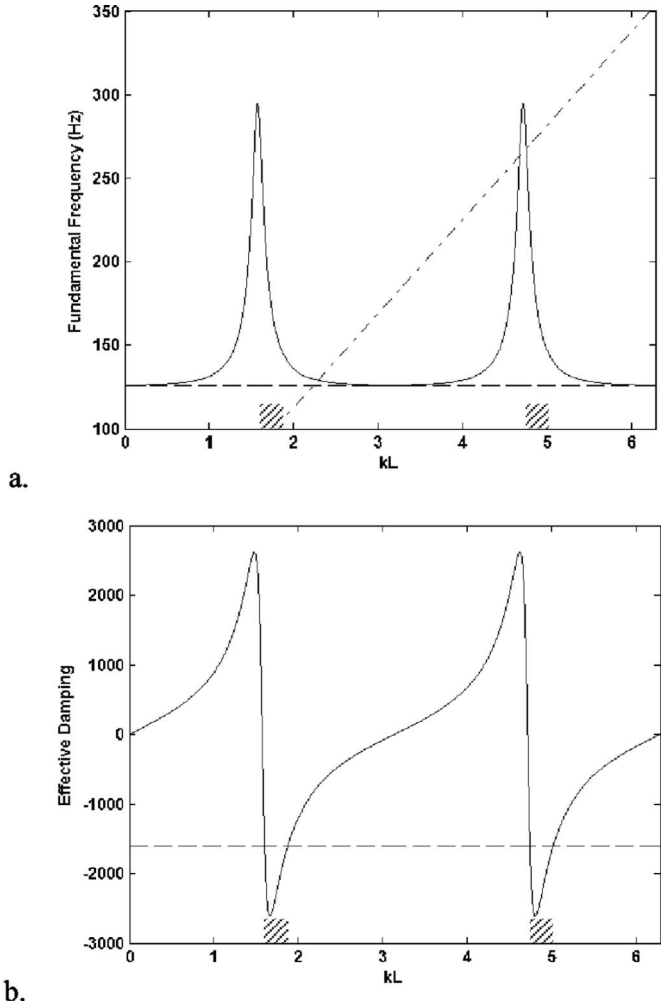

FIG. 11. (a) Fundamental frequency of vocal fold vibration and (b) corresponding effective damping induced by the subglottal system as a function of the product of wave number and tracheal length for an open-ended and lossless trachea tube. —: solution curve; – - –: (a) natural frequency of the vocal folds and (b) critical effective damping at which the equality in Eq. (A11) is reached; –·–: $f=(c/2\pi L)\cdot kL$ for $L=100$ cm; shaded area: regions possible for phonation onset.

For self-sustained oscillations of the vocal folds to occur, Eq. (A11) had to be satisfied, i.e., the total damping of the coupled system had to be zero (at onset) or negative. This requirement limited eligible kL values to series of narrow bands at kL values slightly higher than the quarter-wavelength resonances of the subglottal system, as indicated by the shaded areas in Fig. 11(b). The fundamental frequency at a certain tracheal length L could be obtained by drawing a straight line from the origin with a slope of $c/(2\pi L)$ in Fig. 11(a). The intersection of this line with the solution curve gave the fundamental frequency. For a certain length of the trachea, one or multiple solutions for the fundamental frequency were possible [e.g., three solutions are possible for $L=100$ cm, as shown in Fig. 11(a)]. A unique set of solutions of the fundamental frequency may be obtained by choosing the one with the largest negative value of the effective damping.

Alipour, F., and Scherer, R. C. (**2001**). "Effects of oscillation of a mechanical hemilarynx model on mean transglottal pressures and flows," J.

Acoust. Soc. Am. **110**, 1562–1569.
Austin, S. F., and Titze, I. R. (**1997**). "The effect of subglottal resonance upon vocal fold vibration," J. Voice **11**, 391–402.
Berge, P., Pomeau, Y., and Vidal, C. (**1984**). *Order Within Chaos* (Hermann and Wiley, Paris, France).
Berry, D. A., Herzel, H., Titze, I. R., and Story, B. H. (**1996**). "Bifurcations in excised larynx experiments," J. Voice **10**, 129–138.
Flanagan, J. L. (**1958**). "Some properties of the glottal sound source," J. Speech Hear. Res. **1**, 99–116.
Fletcher, N. H. (**1993**). "Autonomous vibration of simple pressure-controlled valves in gas flows," J. Acoust. Soc. Am. **93**, 2172–2180.
Gupta, V., Wilson, T. A., and Beavers, G. S. (**1973**). "A model for vocal chord excitation," J. Acoust. Soc. Am. **54**, 1607–1617.
Harper, P., Kraman, S. S., Pasterkamp, H., and Wodicka, G. R. (**2001**). "An acoustic model of the respiratory tract," IEEE Trans. Biomed. Eng. **48**, 543–550.
Ishizaka, K., and Flanagan, J. L. (**1972**). "Synthesis of voiced sounds from a two-mass model of the vocal chords," Bell Syst. Tech. J. **51**, 1233–1267.
Ishizaka, K., Matsudaira, M., and Kaneko, T. (**1976**). "Input acoustic-impedance measurement of the subglottal system," J. Acoust. Soc. Am. **60**, 190–197.
Jiang, J. J., and Zhang, Y. (**2005**). "Spatiotemporal chaos in excised larynx experiments," Phys. Rev. E **72**, 035201–1:4.
Jiang, J. J., Zhang, Y., and Ford, C. N. (**2003**). "Nonlinear dynamics of phonations in excised larynx experiments," J. Acoust. Soc. Am. **114**, 2198–2205.
Joliveau, E., Smith, J., and Wolfe, J. (**2004**). "Vocal tract resonances in singing: The soprano voice," J. Acoust. Soc. Am. **116**, 2434–2439.
Kinsler, L. E., Frey, A. R., Coppens, A. B., and Sanders, J. V. (**2000**). *Fundamentals of Acoustics* (Wiley, New York).
Large, J. (**1972**). "Towards an integrated physiologic-acoustic theory of vocal registers," NATS Bull. **29**, 18–25.
Mergell, P., and Herzel, H. (**1997**). "Modeling biphonation – The role of the vocal tract," Speech Commun. **22**, 141–154.
Nadoleczny-Millioud, M., and Zimmerman, R. (**1938**). "Categories et registres de la voix," Names and Registers of the Voice **23**, 21–31.
Ruty, N., Van Hirtum, A., Pelorson, X., Lopez, I., and Hirschberg, A. (**2005**). "A mechanical experimental setup to simulate vocal folds vibrations. Preliminary results," ZAS Papers in Linguistics **40**, 161–175.
Seybert, A. F., and Ross, D. F. (**1977**). "Experimental determination of acoustic properties using a two-microphone random-excitation technique," J. Acoust. Soc. Am. **61**, 1362–1370.
Švec, J. G., Schutte, H. K., and Miller, D. G. (**1999**). "On pitch jumps between chest and falsetto registers in voice: Data from living and excised human larynges," J. Acoust. Soc. Am. **106**, 1523–1531.
Thomson, S. L., Mongeau, L., and Frankel, S. H. (**2005**). "Aerodynamic transfer of energy to the vocal folds," J. Acoust. Soc. Am. **118**, 1689–1700.
Thomson, S. L., Mongeau, L., Frankel, S. H., Neubauer, J., and Berry, D. A. (**2004**). "Self-oscillating laryngeal models for vocal fold research," *Proceedings of the Eighth International Conference on Flow-Induced Vibrations*, Paris, France, Vol. **2**, pp. 137–142.
Titze, I. R. (**1988a**). "The physics of small-amplitude oscillation of the vocal folds," J. Acoust. Soc. Am. **83**(4), 1536–1552.
Titze, I. R. (**1988b**). "A framework for the study of vocal registers," J. Voice **2**, 183–194.
Titze, I. R., Schmidt, S. S., and Titze, M. R. (**1995**). "Phonation threshold pressure in a physical model of the vocal fold mucosa," J. Acoust. Soc. Am. **97**, 3080–3084.
Titze, I. R., and Story, B. H. (**1995**). "Acoustic interactions of the voice source with the lower vocal tract," J. Acoust. Soc. Am. **101**, 2234–2243.
van den Berg, J. W. (**1968a**). "Sound production in isolated human larynges," Ann. N.Y. Acad. Sci. **155**, 18–27.
van den Berg, J. W. (**1968b**). "Register problems," Ann. N.Y. Acad. Sci. **155**, 129–134.
van den Berg, J. W., and Tan, T. S. (**1959**). "Results of experiments with human larynxes," Pract. Otorhinolaryngol. (Basel) **21**, 425–450.
Van der Plaats, A., Schutte, H. K., van der Eerden, F. J. M., de Vries, M. P., Mahieu, H. F., and Verkerke, G. J. (**2006**). "An *in-vitro* test set-up for evaluation of a voice- producing element under physiologic acoustic conditions," Ann. Biomed. Eng. **34**, 893–900.
Vennard, W. (**1967**). "*Singing…the Mechanisms and the Technique*," Revised 2nd ed. (Fisher, New York).
Vilain, C. E., Pelorson, X., Hirschberg, A., Le Marrec, L., Op't Root, W., and Willems, J. (**2003**). "Contribution to the physical modeling of the lips. Influence of the mechanical boundary conditions," Acta Acoustica United with Acustica **89**, 882–887.
Weibel, E. R. (**1963**). *Morphometry of the Human Lung* (Springer, Berlin).
Wilson, T. A., and Beavers, G. S. (**1974**). "Operating modes of the clarinet," J. Acoust. Soc. Am. **56**(2), 653–658.
Zhang, Z., Mongeau, L., and Frankel, S. H. (**2002**). "Experimental verification of the quasi-steady approximation for aerodynamic sound generation by pulsating jets in tubes," J. Acoust. Soc. Am. **112**(4), 1652–1663.