

Aerodynamically and acoustically driven modes of vibration in a physical model of the vocal folds

Zhaoyan Zhang, Juergen Neubauer, and David A. Berry

UCLA School of Medicine, 31-24 Rehabilitation Center, 1000 Veteran Avenue, Los Angeles, California 90095-1794

(Received 23 February 2006; revised 15 August 2006; accepted 16 August 2006)

In a single-layered, isotropic, physical model of the vocal folds, distinct phonation types were identified based on the medial surface dynamics of the vocal fold. For acoustically driven phonation, a single, in-phase, x -10 like eigenmode captured the essential dynamics, and coupled with one of the acoustic resonances of the subglottal tract. Thus, the fundamental frequency appeared to be determined primarily by a subglottal acoustic resonance. In contrast, aerodynamically driven phonation did not naturally appear in the single-layered model, but was facilitated by the introduction of a vertical constraint. For this phonation type, fundamental frequency was relatively independent of the acoustic resonances, and two eigenmodes were required to capture the essential dynamics of the vocal fold, including an out-of-phase x -11 like eigenmode and an in-phase x -10 like eigenmode, as described in earlier theoretical work. The two eigenmodes entrained to the same frequency, and were decoupled from subglottal acoustic resonances. With this independence from the acoustic resonances, vocal fold dynamics appeared to be determined primarily by near-field, fluid-structure interactions. © 2006 Acoustical Society of America.

[DOI: 10.1121/1.2354025]

PACS number(s): 43.70.Aj, 43.70.Bk [BHS]

Pages: 2841–2849

I. INTRODUCTION

A primary prerequisite for self-sustained oscillation of the vocal folds is that the net transfer of energy from the airflow to the tissue be sufficient to overcome frictional forces (Ishizaka and Matsudaira, 1972; Stevens, 1977; Broad, 1979; Titze, 1988). In order for this transfer of energy to occur, the tissue velocity must be approximately in-phase with the driving forces. Titze (1988) argued that this condition may be facilitated by (1) an inertive vocal tract, (2) the superposition of two or more lower-order eigenmodes which propagate a mucosal wave up the medial surface of the vocal fold, or (3) some combination of the previous two conditions.

For a one-mass model, or any model which does *not* propagate a mucosal wave, it has been shown that self-sustained oscillations can only be achieved when the vocal folds entrain with one of the acoustic resonances of the *vocal tract*, a condition which is facilitated by an *inertive supra-glottal tract* (Flanagan and Landgraf, 1968; Ishizaka and Flanagan, 1972; Titze, 1988). Alternately, it has been shown that self-sustained oscillations of a one-mass model may also be achieved when the vocal folds entrain with one of the acoustic resonances of the *subglottal tract*, a condition which is facilitated by a *compliant subglottal tract* (Fletcher, 1993; Zhang *et al.*, 2006). Because of the strong dependence on the acoustic resonances of the vocal system, we refer to such vibrations as *acoustically driven modes of phonation*. As a rule, such modes of phonation are *undesirable* for voice and speech because they tend to be plagued by frequency jumps and involuntary voice breaks, as the vocal folds entrain to distinct acoustic resonances of sub- or supraglottal tracts.

In contrast, the two-mass model (Ishizaka and Flanagan, 1972), or any model which propagates a mucosal wave, tends to oscillate independently of the acoustic resonances. Instead, its dynamics are governed primarily by a near-field fluid-structure interaction. We refer to such vibrations as *aerodynamically driven* modes of phonation. In this mode of phonation, two or more eigenmodes of the vocal fold tissues (Fig. 1) entrain or phase lock so as to create favorable aerodynamic conditions for phonation. Usually the strongest eigenmode creates an alternating convergent/divergent glottis near the top of the glottal airway [Fig. 1(c)], which makes it intimately associated with the glottal aerodynamics, e.g., a convergent glottis is associated with a relatively high intraglottal pressure, and a divergent glottis is associated with a relatively low intraglottal pressure (Titze, 1988). In the literature, this eigenmode has been referred to as an x -11 mode [which according to the nomenclature of Titze and Strong (1975), Berry *et al.* (1994), Berry and Titze (1996) is in the form of x - n_y - n_z , where x refers to a medial-lateral mode of vibration, and n_y and n_z represent the number of half-wavelengths in the anterior-posterior and inferior-superior directions, respectively]. The second mode, usually capturing an in-phase medial-lateral motion of a rectangular glottis near the top of the glottal airway [Fig. 1(b)], governs the net lateral tissue velocity in this region, and modulates the glottal airflow. In the literature, this eigenmode has been referred to as an x -10 mode. Using the method of empirical eigenfunctions (EEFs), Berry *et al.* (1994) investigated the details of this theory with regard to a finite element model of vocal fold vibration. In that investigation, the tissue dynamics were explained by two dominant eigenmodes. The two eigenmodes exhibited a 1:1 entrainment or phase locking such that the intraglottal pressure (governed by an x -11-like mode)

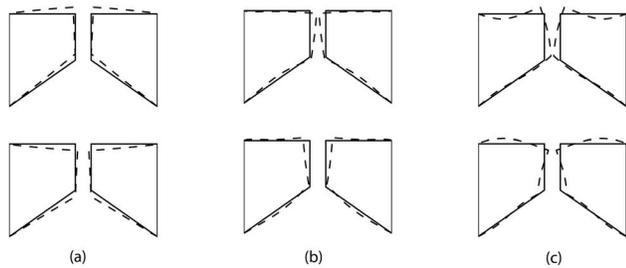


FIG. 1. The three lowest-order eigenmodes of a two-dimensional vocal foldlike structure: (a) eigenmode 1, analogous to the z -10 mode; (b) eigenmode 2, analogous to the x -10 mode, describes in-phase motion along the vertical extent of the medial surface; (c) eigenmode 3, analogous to the x -11 mode, describes out-of-phase motion along the vertical extent of the medial surface. Left and right folds are shown to give an indication of the glottal geometry produced by the corresponding mode. Upper and lower frames indicate extreme positions of eigenmodes, spaced 180° apart in a vibratory cycle. Solid lines indicate equilibrium positions.

was in phase with the net lateral tissue velocity (governed by an x -10-like mode), thus facilitating energy transfer from the airflow to the tissue.

More recently, Thomson *et al.* (2005) examined these same concepts in the laboratory using a physical model of vocal fold vibration, and a numerical simulation. Based on the results of the numerical simulation (which was driven by an imposed unsteady subglottal pressure as measured from the physical model), they investigated aerodynamic energy transfer mechanisms. In particular, they reported further substantiation of “the hypothesis that a cyclic variation of the orifice profile from a convergent to a divergent shape leads to a temporal asymmetry in the average wall pressure, which is the key factor for the achievement of self-sustained oscillation.” However, using the same physical model with a slightly lower Young’s modulus, Zhang *et al.* (2006) showed that the physical model usually exhibited self-oscillation *only when entrained to the acoustic resonances of the subglottal tract*, suggesting that the model exhibited *acoustically driven* rather than *aerodynamically driven* modes of phonation. Indeed, Thomson *et al.* (2005) disclosed that the vibration frequency observed in their study was near the half-wavelength resonance of the subglottal tube. Thus, for aerodynamically driven modes of phonation, it would appear that energy transfer mechanisms remain relatively unexplored in the physical model.

Using the physical model of Thomson *et al.* (2005), the purpose of the present study was to (1) attempt to induce aerodynamically driven phonation in the physical model, and (2) to distinguish *aerodynamically driven* and *acoustically driven* modes of phonation based on the medial surface dynamics of the physical model, over a range of subglottal pressures and tracheal lengths. Historically, a variety of studies have suggested that acoustically-driven modes of phonation are analogous to one-mass models of vocal fold vibration (Flanagan and Landgraf, 1968; Ishizaka and Flanagan, 1972; Titze, 1988; Fletcher, 1993), and that aerodynamically-driven modes of phonation are analogous to any tissue model which propagates a mucosal wave, usually through the superposition of two or more, lower-order eigenmodes (Titze and Strong, 1975; Titze, 1988; Berry *et al.*, 1994; Berry,

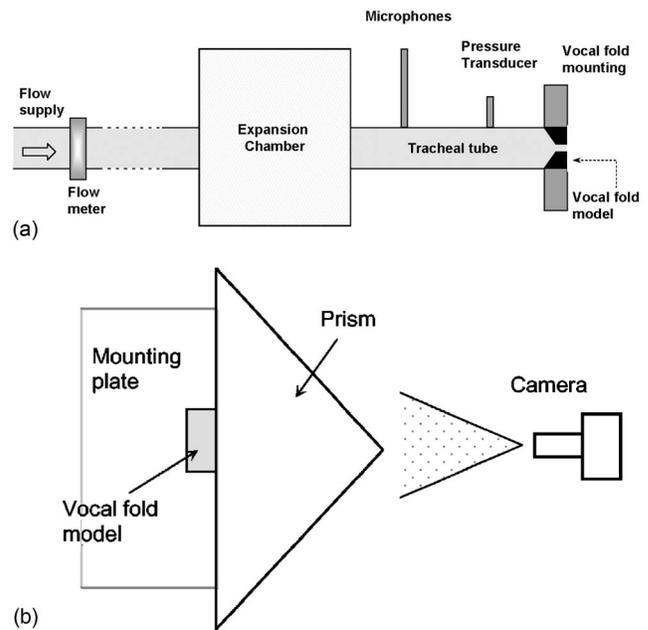


FIG. 2. (a) Schematic of the experimental setup, and (b) the hemimodel configuration.

2001; Doellinger *et al.* 2005a, 2005b). Consequently, we hypothesize that aerodynamically and acoustically driven modes of phonation may be distinguished based on the medial surface dynamics of the vocal fold, and by their degree of dependence on the acoustic resonances of the vocal system. In particular, we will seek to substantiate and elucidate this hypothesis with quantitative, experimental data from the physical model.

II. METHOD

A physical replica of the human vocal system [Fig. 2(a)] was constructed using a rubber model of the vocal folds and a uniform (PVC) tracheal tube (2.54-cm inner diameter) connected upstream to an expansion chamber, simulating the subglottal system. The rubber model was a single-layered, cover-only, isotropic, physical model of the vocal folds (Thomson *et al.*, 2005; Zhang *et al.*, 2006). Using a mold, it was created with a two-component liquid polymer solution mixed with a liquid flexibilizer solution. For additional details regarding the fabrication and dynamical properties of the physical model, please see the original papers. Each vocal fold model 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 [Fig. 3(a)]. The inferior side of each vocal fold had an entrance convergence angle of approximately 60° measured from the inferior-superior axis, yielding an inferior-superior vocal fold thickness of approximately 5.4 mm. Using a 5544 Instron Testing System, the stress-strain relationship of the artificial vocal fold tissue was measured to be nearly linear in the strain range of 0–20%. Using the stress-strain data, the Young’s modulus was calculated to be approximately 11 kPa across this range. The density was 997 kg/m^3 . The vocal folds were glued into a rectangular groove on the medial surface of two

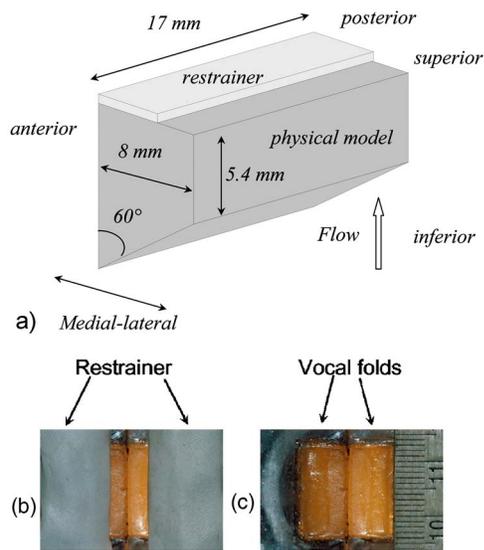


FIG. 3. (a) A sketch of the physical model of the vocal fold, and a superior view of the model both (b) with, and (c) without restrainers.

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.

Different lengths of the PVC tracheal tube were used in this study, ranging from 11 to 120 cm. No efforts were made to smooth the transition from the tracheal tube to the physical model, although such an abrupt transition may have generated undesired perturbations in the airflow. The geometry of the transition region was kept unchanged so that its effects remained constant during the course of the experiments. The expansion chamber had an inner cross-sectional area of 23.5×25.4 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) have shown that the cross-sectional area function from the trachea and the primary bronchi to the lungs increases 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 air flow 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 in a separate study (Zhang *et al.*, 2006) and were found to be similar to an ideal open-ended termination of the tracheal tube for frequencies above approximately 50 Hz. Admittedly, this subglottal system does not exactly reproduce the human subglottal system, particularly the lossy compliance of the lungs. However, the relative simplicity of the system and its acoustical similarity to the human lungs yielded a reliable, controllable model of the subglottal acoustic system, which was essential for the present study.

A hemi-model procedure was used to investigate the medial surface dynamics of the vocal fold model, a technique which has been previously implemented for a variety of other laboratory experiments, including other physical mod-

els of the folds (Titze *et al.*, 1995; Chan *et al.*, 1997), excised larynx experiments (Jiang and Titze, 1993; Berry *et al.*, 2001; Doellinger *et al.*, 2005b), and the *in vivo* canine laryngeal model (Doellinger *et al.*, 2005a). Figure 2(b) shows a schematic of the hemi-model setup. In this experimental configuration, 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 was imaged using a high-speed digital camera (Fastcam-Ultima APX, Photron Unlimited, Inc.). A frame rate of 2000 Hz 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 dynamics.

In particular, time-series cross-correlation analysis was performed on the medial surface images using the image-processing package DaVis (LaVision Inc.). A multipass algorithm was used in the cross-correlation analysis. Initial estimates of the displacements were made using a relatively large interrogation window. Next, these estimates were used as input for subsequent analyses which used smaller interrogation window sizes. For each camera view, the correlation analysis yielded the medial surface displacements in image coordinates. Using an appropriate mapping, the three-dimensional, physical coordinates of the displacements were derived from the two sets of image coordinates. The mapping function was determined through means of a calibration process (Hartley and Zisserman, 2000), using a calibration target with known three-dimensional coordinates.

In this case, the calibration target (LaVision Inc., Type no. 2.5) was 25×25 mm with a 2-mm spacing between calibration points in the inferior-superior and anterior-posterior directions, and 0.5-mm spacing between calibration points in the medial-lateral direction. Due to the periodic motion of the medial surface, the correlation analysis was performed with respect to the first frame. Therefore, a Lagrangian displacement field was obtained as a function of the reference coordinates in the first frame. Ultimately, displacements were computed over a total medial surface area of 16×16 mm, with a grid spacing of 0.22×0.22 mm in the inferior-superior and anterior-posterior directions, respectively. In contrast to previous methods of tracking vocal fold displacements which were semiautomatic (Berry *et al.*, 2001; Doellinger *et al.*, 2005b), the present technique was fully automated.

The subglottal pressure in the tracheal tube was monitored using a probe microphone (B&K 4182), 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 transglottal pressure was measured using a pressure transducer (Baratron type 220D). Because no vocal tract was used in this study, the subglottal pressure was equivalent to the transglottal pressure. The volumetric flow rate through the orifice was measured using a precision mass-flow meter (MKS type 558A) at the inlet to the setup. During the experiments, the flow rate was increased from zero to a certain

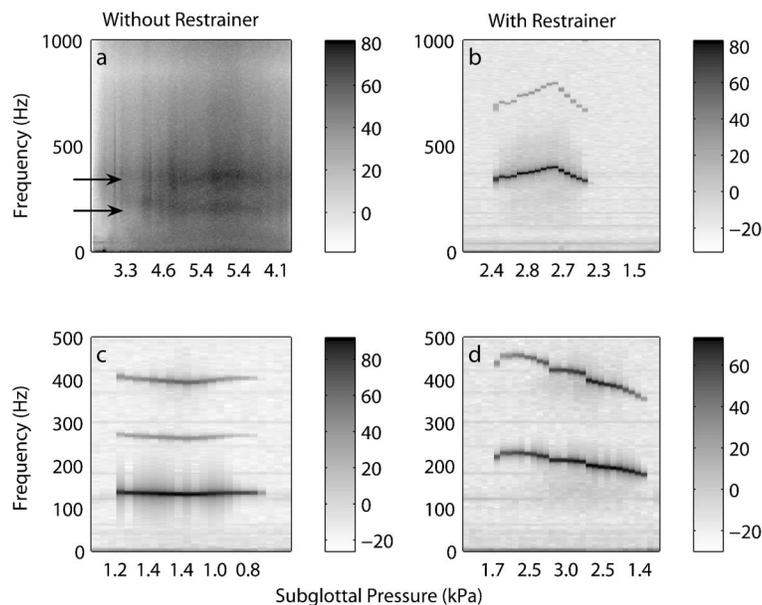


FIG. 4. Sound pressure power spectra as a function of increasing and decreasing subglottal pressure in the hemimodel configuration for: (a) a tracheal tube of length 17.1 cm, without restrainer; (b) a tracheal tube of length 17.1 cm, with restrainer; (c) a tracheal tube of length 60 cm, without restrainer; (d) a tracheal tube of length 60 cm, with restrainer.

maximum value in discrete increments, and then decreased back to zero in discrete decrements (Zhang *et al.*, 2006). At each step, measurement was delayed for an interval of 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. Analog-to-digital conversion of the output signals was performed using a United Electronic Industries Powerdaq board, with 16 bit resolution over a ± 10 V measurement range at a sampling rate of 50 kHz.

III. RESULTS

In a companion paper (Zhang *et al.*, 2006), the vibrations of the physical model were studied as a function of tracheal length and subglottal pressure. In that study, the tracheal length was varied from 11 to 325 cm, and subglottal pressure was varied between 0 and 6 kPa. Despite the desire to generate both aerodynamically and acoustically driven modes of vibration, over the range of parameters investigated (0–6 kPa for the subglottal pressure and 11–325 cm for the tracheal length), only acoustically driven modes of vibration were produced by the model, i.e., phonation only occurred when the phonation frequency entrained with one of the subglottal acoustic resonances, and showed little variation with subglottal pressure. Furthermore, as a function of tracheal length, the vocal folds exhibited bifurcations or abrupt jumps to distinct subglottal resonances. Finally, the acoustically driven modes did not occur for typical human tracheal lengths (i.e., approximately 17 cm). Instead, such modes appeared only for tracheal lengths of 30 cm or more. It was also noted that the acoustically driven modes of vibration exhibited significant vertical (inferior-superior) vibrations similar to the eigenmode in Fig. 1(a), a vibration pattern which easily coupled with the subglottal acoustics.

Thus, over the original range of parameters investigated, aerodynamically driven modes of vibration were not observed. In this study, in a further effort to induce aerodynamically driven phonation, several physical constraints

were applied to the vocal fold model to reduce the large vertical vibrations characteristic of the acoustically driven modes of phonation, and thereby reduce source-tract coupling. Indeed, it was hoped that the restriction of the large vertical vibrations would suppress the vertical motion and facilitate the emergence of an x -11 eigenmode and mucosal wave propagation, and thereby promote an aerodynamically driven mode of phonation (Titze, 1988). The following constraint produced the desired effect: rubber restrainers were laterally positioned over the superior surface of the vocal fold model [Figs. 3(a) and 3(b)]. The thickness of the restrainers was chosen so that the restrainers were thin enough to avoid disturbance of the near-field flow (approximately 2.5 mm), yet stiff enough to restrain the vertical motion of the vocal fold body. The lateral-medial extent to which the restrainers covered the vocal fold was adjustable. The lateral distance of the uncovered superior portion of the vocal fold roughly approximated the effective lateral depth of vibration of the physical model.

Experiments were conducted on both full and hemimodel configurations. Both configurations produced qualitatively similar vibration patterns. The vibration frequencies between configurations generally differed by a few hertz. Also, phonation threshold pressures were slightly higher (0.1–0.2 kPa) in the hemimodel than in the full model. However, as expected, the threshold airflow in the hemimodel was about half that of the full model. Because of our focus on medial surface dynamics, only results from the hemimodel experiments are discussed below.

A. Phonation frequency

Figures 4(a) and 4(b) compare the power spectra of the subglottal acoustic pressure as a function of the mean subglottal pressure for cases with and without restrainers, for a tracheal tube length of 17.1 cm, which is a typical estimate of the length of the human trachea (Flanagan, 1958; Ishizaka *et al.*, 1976). The subglottal resonance for this tracheal length could be measured directly as dominant energy peaks in the

acoustic power spectra before onset, or estimated from linear acoustic theory using the measured reflection factor of the subglottal system in Sec. II. In this case, the first subglottal resonance occurred around 350 Hz. The power spectra for an unrestrained fold are shown in Fig. 4(a), and the power spectra for a restrained fold are shown in Fig. 4(b). Note that, during the experiments, the mean subglottal pressure was first increased to a maximum value and then decreased back to zero. For the unrestrained fold, phonation did not occur. In this case, two broad horizontal lines can be seen in Fig. 4(a), which correspond to broadband flow noise around the subglottal acoustic resonance of approximately 350 Hz, and the would-be phonation frequency of approximately 210 Hz (small-amplitude vibrations were observed at this latter frequency at higher subglottal pressures than those originally investigated in this study [up to about 7 kPa]). Figure 4(b) shows the restrained case with an effective lateral depth of vibration of 2.2 mm (compared with an effective lateral depth of vibration of 8.0 mm when no restrainer was applied). Phonation onset occurred at a subglottal pressure of approximately 2 kPa. After onset, phonation frequency appeared to vary continuously as a function of subglottal pressure. Although close to the acoustic resonance, the fundamental frequency clearly exhibited variations which were independent of the subglottal resonance.

Figures 4(c) and 4(d) show the acoustic power spectra for unrestrained and restrained folds, respectively, when a longer tracheal tube of approximately 60 cm was utilized. The first subglottal resonance was about 134 Hz. Without the application of the restrainer, phonation was observed. The vocal fold entrained to a subglottal acoustic resonance, and the fundamental frequency exhibited weak dependence on the subglottal pressure [Fig. 4(c)]. Clearly, the physical model vibrated in an acoustically driven mode. However, with the restrainer (with an effective lateral depth of vibration of 5.4 mm), the physical model vibrated independently of the subglottal acoustic resonance (about 134 Hz), and appeared to vary continuously as a function of subglottal pressure in the range of 200–250 Hz, as shown in Fig. 4(d).

Shown in Fig. 5 are the medial surface trajectories of one coronal slice of the medial surface, midway between anterior and posterior extremes of the same physical model with and without restrainers. Without restrainers, large vertical vibrations were observed in the acoustically driven mode of vibration. However, this vertical motion was suppressed when restrainers were applied. Note that lip vibrations in brass instruments exhibit similar vibratory phenomena. For example, the unrestrained vibrations observed in this study roughly correspond to the “swinging door” motion of the vibrating lips, and the vertically restrained vibrations roughly correspond to the “sliding door” motion (Adachi and Sato, 1996; Copley and Strong, 1996).

The fact that the physical model vibrated independently of the subglottal acoustic resonances illustrates that the restrained physical model vibrated in the aerodynamically driven modes of vibration. This suggests that the application of vertical constraints may facilitate the excitation of aerodynamically driven modes of vibration, perhaps by enhancing the near-field fluid-structure energy transfer. The results

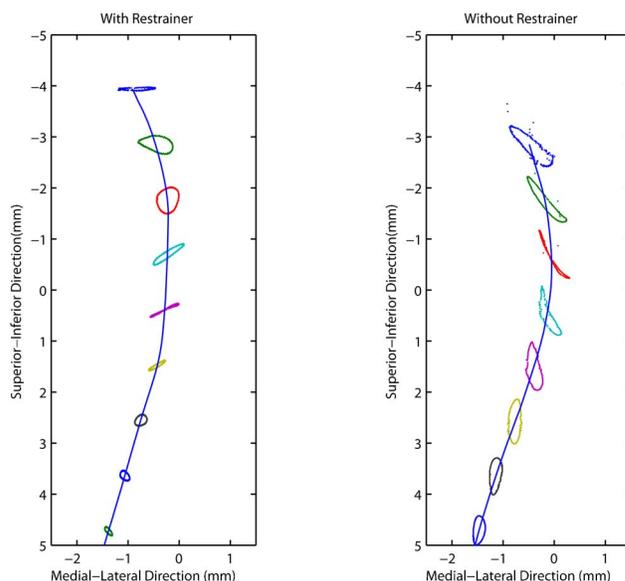


FIG. 5. Anterior view of medial surface trajectories for one coronal slice of the medial surface of the vocal fold, midway between anterior and posterior extremes, for cases with (left) and without (right) restrainers. Because the vibrational amplitudes from the case with restrainers were relatively small, they have been amplified by a factor of three. For a clearer illustration, only trajectories from every fifth grid point are shown along the inferior-superior length. The tracheal tube length was 60 cm.

also suggest that the suppression of vertical motion reduced the coupling of the tissue vibrations with the subglottal acoustics.

B. Spatiotemporal analysis

Figure 6 is a spatiotemporal representation of the medial-lateral component of motion for one coronal slice of the medial surface of the vocal fold model, midway between anterior and posterior extremes. For each subplot, the x axis corresponds to time and the y axis corresponds to the inferior-superior direction. The superior edge of the vocal

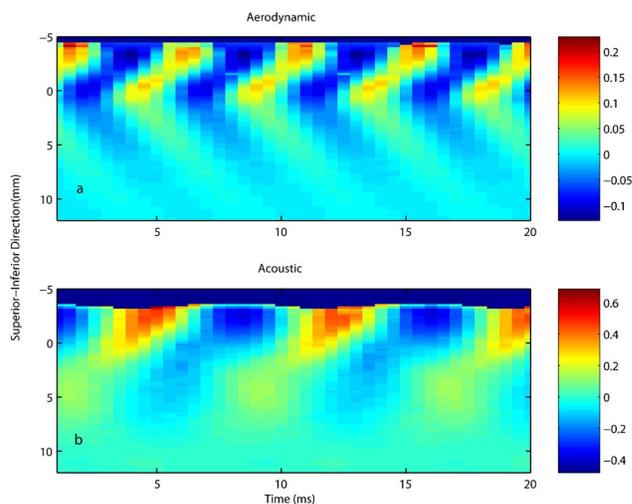


FIG. 6. (Color online) Spatiotemporal plot of the medial-lateral component of the medial surface displacements for one coronal slice of the physical model, for both aerodynamically driven (top) and acoustically driven (bottom) modes of phonation. The tracheal tube length was 60 cm.

fold model was at about -4.5 mm and -3 mm in the superior-inferior direction for the aerodynamically driven mode and acoustically driven mode, respectively. For both modes of phonation, displacement amplitudes were greatest near the superior edge of the vocal fold. However, the vibration patterns were quite distinct. For the acoustically driven mode of phonation [Fig. 6(b)], a traveling wave was initiated about 10-mm below the superior edge, and traveled superiorly with time until it reached the superior edge of the vocal fold, where it exhibited its greatest amplitudes.

In contrast, for the aerodynamically driven mode of phonation [Fig. 6(a)], a forward or superiorly traveling wave and a backward or inferiorly traveling wave were initiated at about 5-mm below the superior surface. Although the backward-traveling wave had smaller amplitudes than the forward-traveling wave, it constituted a characteristic feature of the overall wave structure. The forward traveling wave proceeded until it reached the superior edge of the fold. The more complicated wave structure associated with the aerodynamic mode of phonation was further confirmed through means of a principal component analysis (PCA), as discussed below.

C. Empirical eigenfunctions

Empirical eigenfunctions (EEFs) were extracted from the three-dimensional displacement field of the entire medial surface of the vocal fold using the method of principal components analysis (PCA) (Berry *et al.*, 1994). Prior to analysis, the mean component of the displacement data was calculated and subtracted from the displacement data. Using the method of PCA, the three-dimensional displacement data over the medial surface $D(x, y, z, t)$ was decomposed into a linear combination of empirical eigenfunctions

$$D(x, y, z, t) = \sum_{k=1}^N \sigma_k \Phi_k(t) \Psi_k(x, y, z), \quad (1)$$

where the normalized scalar function $\Phi_k(t)$ and normalized vector function $\Psi_k(x, y, z)$ were the k th temporal and spatial empirical eigenfunctions, σ_k was the modal scaling factor, and N was the total number of time samples. The temporal eigenfunctions may be also thought of as the temporal coefficients of the spatial eigenfunctions. The total energy associated with the vibration was given by

$$E_{\text{tot}} = \left\langle \int |D(x, y, z, t)|^2 dx dy dz \right\rangle_t = \sum_{k=1}^N \sigma_k^2, \quad (2)$$

where the angle brackets denoted a time average. The relative weight of the contribution of each eigenfunction toward the total energy (or percentage variance) can be calculated as the ratio of corresponding σ_k^2 and the total energy E_{tot} .

Figure 7 shows the first two temporal eigenfunctions for both modes of phonation. The percentage variance (or relative weight) of the displacement data captured by the first four EEFs, and the vibration frequencies of the temporal eigenfunctions, are shown in Table I. The temporal eigenfunctions of the first four EEFs were harmonic functions with a dominant single-frequency component, with the first

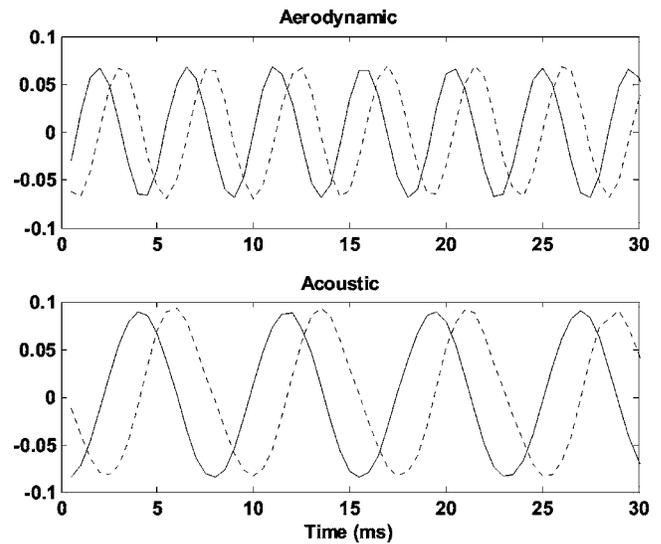


FIG. 7. The first two temporal empirical eigenfunctions for aerodynamically driven (top) and acoustically driven (bottom) modes of phonation. The tracheal tube length was 60 cm. — : first temporal eigenfunction; - - - : second temporal eigenfunction.

two EEFs entrained to the fundamental frequency and the next two EEFs entrained to the second harmonic.

For both modes of phonation, Fig. 8 shows the normalized, medial-lateral component of the first two spatial eigenfunctions. From a coronal aspect, Fig. 9 illustrates medial-lateral and inferior-superior components of these same eigenfunctions, for one coronal slice of the medial surface, midway between anterior and posterior extremes. For the acoustically driven mode of phonation, considerable vertical motion was present (see Fig. 5), which was captured by the first eigenfunction (see Fig. 9). The medial-lateral component of the first eigenfunction was similar to an x -10 eigenmode (Titze and Strong, 1975; Berry *et al.*, 1994; Berry and Titze, 1996), in which the medial-lateral motion near the top of the glottal airway was in phase. This eigenfunction captured 85% of the variance of the medial surface displacements. Although the second eigenfunction was similar to an x -11 eigenmode and alternately shaped a convergent/divergent glottis (Titze and Strong, 1975; Berry *et al.*, 1994; Berry and Titze, 1996), this eigenfunction captured only 11.1% of the variance, and was approximately 7.6 times less energetic than the first eigenfunction. Other higher-order eigenfunctions captured 1.5% or less of the variance of the medial surface displacements. Therefore, in an acoustically driven mode of phonation, the vocal fold oscillations were captured primarily by one eigenfunction, with the medial-

TABLE I. Percentage variances (weights) and frequencies of the first four EEFs.

	Aerodynamic		Acoustic	
	Weight (%)	Frequency(Hz)	Weight (%)	Frequency(Hz)
EEF1	65.3	217	84.8	131
EEF2	31.3	217	11.1	131
EEF3	1.0	434	1.5	263
EEF4	0.5	434	1.1	263

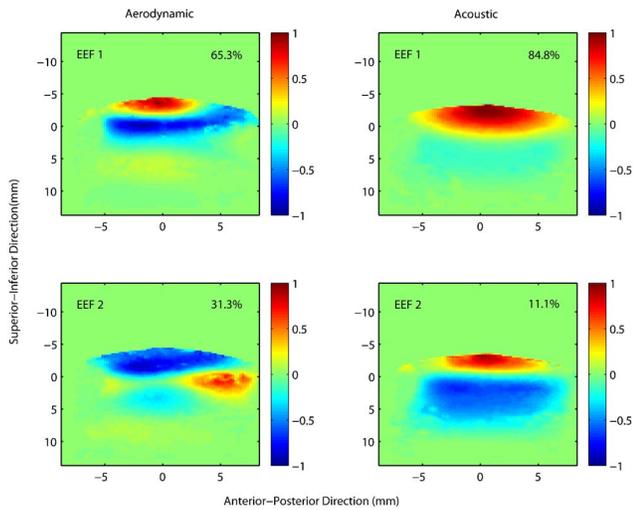


FIG. 8. (Color online) Normalized medial-lateral components of the first two spatial eigenfunctions for aerodynamically driven (left) and acoustically driven (right) modes of phonation. The tracheal tube length was 60 cm.

lateral component corresponding to an x -10 mode and the inferior-superior component corresponding to a z -10 mode.

In contrast, for the aerodynamically driven mode of phonation, the medial surface displacements were more complex, capturing two relatively strong eigenfunctions. The first eigenfunction (Figs. 8 and 9) was similar to an x -11 eigenmode (alternately shaping a convergent/divergent glottis) and captured 65% of the variance of the medial surface displacements. However, near the top of the glottal airway, the second eigenfunction was similar to an x -10 mode, and captured 31% of the variance of the medial surface displacements. Thus, for this aerodynamically driven mode of phonation, at least two eigenfunctions were required to reconstruct the primary dynamics of the vocal fold. Moreover, the two eigenfunctions entrained to the same frequency, which was independent of the subglottal acoustic resonance. This superposition of two eigenfunctions in the aerodynamically driven mode of phonation, with a phase difference determined by the temporal eigenfunctions, generated a traveling wave along the medial surface. This may explain the more complicated wave structure shown in Fig. 6(a), as compared to Fig. 6(b). Indeed, for the acoustically driven mode of phonation, only one eigenfunction was required to capture the essential dynamics of the vocal fold, resulting in a less-complicated wave structure, and a necessary dependence on and coupling with a nearby acoustic resonance.

IV. DISCUSSION

In this study, an attempt was made to distinguish aerodynamically and acoustically driven modes of vibration in a physical model based on the medial surface dynamics of the vocal fold, and the relative independence of the source and the acoustic resonator. The physical model was a single-layered, isotropic model of vocal fold vibration (Thomson *et al.*, 2005). Initially, only acoustically driven modes of phonation were observed in the physical model. For acoustically driven modes of phonation: (1) a single, dominant, x -10 like eigenmode captured the essential dynamics, explaining 85%

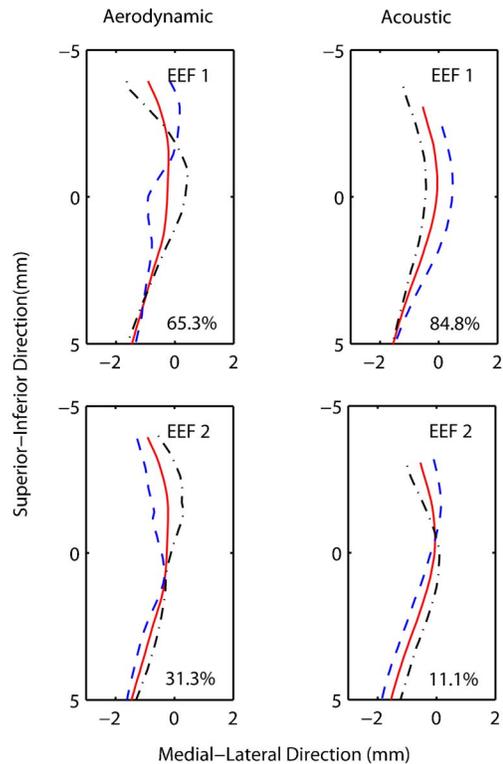


FIG. 9. (Color online) Medial-lateral and superior-inferior components of the first two spatial eigenfunctions for one coronal slice of the medial surface of the vocal fold, midway between anterior and posterior extremes, for both aerodynamically driven (left) and acoustically driven (right) modes of phonation. Maximum (dash-dotted) and minimum (dashed) projections of the eigenfunctions are superimposed on the mean surface projection (solid). The tracheal tube length was 60 cm.

of the variance of the medial surface displacements; (2) the eigenfunction always coupled with one of the acoustic resonances of the subglottal tract; and (3) phonation occurred only for relatively long tracheal lengths of 30 cm or more. In contrast, aerodynamically driven modes of phonation were observed only when a physical constraint restricted the vertical vibrations of the folds. For aerodynamically driven modes of phonation: (1) at least two eigenmodes were required to capture the essential dynamics, including an x -11-like eigenmode which explained 65% of the variance of the medial surface displacements, and an x -10-like eigenmode which explained 31% of the variance; (2) the two eigenmodes entrained to the same frequency and were successfully decoupled from subglottal acoustic resonances; and (3) self-oscillation was achieved for typical, human, tracheal lengths of approximately 17 cm.

The idea that human phonation, particularly the singing voice, may be influenced by sub- or supraglottal acoustic resonances is not new to speech science. Indeed, 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, 1968; Large, 1972; Austin and Titze, 1997). Similarly, Mergell and Herzel (1997) showed an example of high-frequency female singing, in which one vocal fold entrained with the first formant frequency, after the two folds had desynchronized due to left-right asymmetries.

Titze (1988) stated, "...the vocal instrument operates with varying degrees of source-resonator coupling. When coupling is weak...the larynx is able to control frequency, intensity, and the glottal source spectrum rather independently. Decoupling preserves constancy of phonation and seems to be desirable." The results of the present investigation demonstrate both weak and strong source-resonator coupling. The acoustically driven mode of phonation corresponds to strong source-resonator coupling, and the aerodynamically driven mode of phonation corresponds to weak source-resonator coupling. Although the physical model (Thomson *et al.*, 2005) had a propensity to vibrate in an acoustically driven mode of phonation, we were able to induce an aerodynamically driven mode of phonation by applying an appropriate vertical restraint. In the normal human condition, epilaryngeal manipulation through means of the ventricular folds may be able to provide a natural vertical restraint on the true vocal folds. Simultaneous contraction and stiffening of the thyroarytenoid muscle may also cause a natural vertical restraint.

In retrospect, after performing the experiment, it was realized the Young's modulus of 11 kPa utilized for the vocal fold model was relatively stiff, especially for the vocal fold cover, which may have discouraged the propagation of the mucosal wave. By lowering the value of the Young's modulus, post-experiment studies revealed a whole continuum of source-resonator coupling possibilities, with the examples in this investigation representing the extremes of both strong source-resonator coupling (e.g., acoustically driven modes of phonation) and weak source-resonator coupling (e.g., aerodynamically driven modes of phonation). Further details regarding the continuum of source-resonator coupling possibilities will be the subject of future work.

While the effectiveness of the vertical constraint in facilitating aerodynamically driven modes of phonation is not fully understood, it has potential applications in the area of phonosurgery which might merit further investigation. For example, despite the existence of a relatively stiff mucosal cover in the current study (both body and cover had a Young's modulus of approximately 11 kPa), a vertical constraint facilitated the emergence of an x -11 eigenmode, and the propagation of a mucosal wave. Whether induced by aging, scarring, or some other means, a stiffened mucosal cover is clinically known to be a major deterrent to mucosal wave propagation. Thus, it may be that the introduction of a vertical constraint through phonosurgical means could have potential benefit for patients with a stiffened vocal fold cover. Similarly, the introduction of a vertical constraint may also facilitate the propagation of a mucosal wave in patients with a relatively lax vocal fold body, such as might be caused by RLN (recurrent laryngeal nerve) paralysis.

Although no supraglottal vocal tracts were included in the current experiments, we would expect the supraglottal vocal tract to exhibit effects similar to that of the subglottal system, especially near phonation onset. Moreover, the inclusion of a supraglottal vocal tract would introduce yet another resonator, further complicating the interactions of the coupled system. This topic will be explored in future studies.

While medial surface dynamics were a useful tool to distinguish aerodynamically and acoustically driven modes of phonation in this study, such dynamics may not be a useful tool to distinguish these modes of phonation in other contexts, since medial surface dynamics are rarely reported. A few exceptions are Berry *et al.* (1994), Berry *et al.* (2001), and Doellinger *et al.*, (2005a, 2005b). In nearly all of these studies, the x -11 eigenmode was the strongest EEF, as might be expected in an aerodynamically driven mode of phonation. However, in Berry *et al.* (2001), the x -10 eigenmode was stronger, which was indicative of strong source-resonator coupling. Indeed, in a related conference proceeding, Berry *et al.* (1999) indicated that they did observe involuntary frequency jumps, which may be another indication of source-tract coupling.

Indeed, a variety of commonly observed vocal phenomena may help to distinguish the degree of source-resonator coupling when access to medial surface dynamics is not available. For example, continuous control of fundamental frequency as a function of vocal fold tension or subglottal pressure may be suggestive of aerodynamic phonation (e.g., independence of source and resonator). On the other hand, involuntary frequency jumps or "voice breaks" may be suggestive of source-resonator coupling. Because such jumps may also be caused by the intrinsic dynamics of the vocal folds (e.g., different eigenmodes within the folds may be excited), voice breaks do not guarantee the existence of source-tract coupling. However, when voice breaks occur, source-tract coupling is highly probable, and must not be ruled out.

V. CONCLUSION

Using a physical model of vocal fold vibration, aerodynamically and acoustically driven modes of vibration were distinguished based on the medial surface dynamics of the vocal fold, and the relative independence of source and acoustic resonator. The single-layered, cover-only, isotropic physical model was dynamically similar to a one-mass model, and its medial-lateral vibration could be described efficiently by an x -10 eigenmode. The physical model had a propensity to vibrate in an acoustically driven mode of phonation, i.e., vocal fold vibration did not occur unless it was coupled to an acoustic resonance of the vocal system. However, an aerodynamically driven mode of phonation could be facilitated by introducing a vertical constraint, and or by lowering the stiffness of the vocal fold. This mode of phonation captured two relatively strong eigenmodes (e.g., x -11 and x -10), exhibited a clear mucosal wave, and was vibrated independently of acoustic resonances. Because of the increased control over fundamental frequency and register, aerodynamically driven modes of phonation are generally deemed preferable to acoustically driven modes of phonation in both speech and singing.

ACKNOWLEDGMENTS

This study was supported by Research Grant No. R01 DC03072 from the National Institute on Deafness and Other Communication Disorders, the National Institutes of Health.

The authors also thank Dr. Ingo R. Titze and other anonymous reviewers for their help in improving an earlier version of this manuscript.

- Adachi, S., and Sato, M. (1996). "Trumpet sound simulation using a two-dimensional lip vibration model," *J. Acoust. Soc. Am.* **99**(2), 1200–1209.
- Austin, S. F., and Titze, I. R. (1997). "The effect of subglottal resonance upon vocal fold vibration," *J. Voice* **11**, 391–402.
- Berry, D. A. (2001). "Mechanisms of modal and nonmodal phonation," *J. Phonetics* **29**, 431–450.
- Berry, D. A., Herzel, H., Titze, I. R., and Krischer, K. (1994). "Interpretation of biomechanical simulations of normal and chaotic vocal fold oscillations with empirical eigenfunctions," *J. Acoust. Soc. Am.* **95**, 3595–3604.
- Berry, D. A., Montequin, D. W., and Tayama, N. (2001). "High-speed digital imaging of the medial surface of the vocal folds," *J. Acoust. Soc. Am.* **110**, 2539–2547.
- Berry, D. A., and Titze, I. R. (1996). "Normal modes in a continuum model of vocal fold tissues," *J. Acoust. Soc. Am.* **100**, 3345.
- Berry, D. A., Titze, I. R., and Herzel, H. (1999). "Empirical eigenfunctions obtained from highspeed imaging of the vocal folds," *J. Acoust. Soc. Am.* **105**(2), 1304.
- Broad, D. (1979). "The new theories of vocal fold vibration," in *Speech and Language: Advances in Basic Research and Practice*, edited by N. Lass (Academic, New York).
- Chan, R. W., Titze, I. R., and Titze, M. R. (1997). "Further studies of phonation threshold pressure in a physical model of the vocal fold mucosa," *J. Acoust. Soc. Am.* **101**, 3722–3727.
- Copley, D. C., and Strong, W. J. (1996). "A stroboscopic study of lip vibrations in a trombone," *J. Acoust. Soc. Am.* **99**(2), 1219–1226.
- Doellinger, M., Berry, D. A., and Berke, G. S. (2005a). "Medial surface dynamics of an *in vivo* canine vocal fold during phonation," *J. Acoust. Soc. Am.* **117**, 3174–3183.
- Doellinger, M., Tayama, N., and Berry, D. A. (2005b). "Empirical eigenfunctions and medial surface dynamics of a human vocal fold," *Methods Inf. Med.* **44**, 384–391.
- Flanagan, J. L. (1958). "Some properties of the glottal sound source," *J. Speech Hear. Res.* **1**, 99–116.
- Flanagan, J. L., and Landgraf, L. (1968). "Self-oscillating source for vocal tract synthesizers," *IEEE Trans. Audio Electroacoust.* **AU-16**, 57–64.
- Fletcher, N. H. (1993). "Autonomous vibration of simple pressure-controlled valves in gas flows," *J. Acoust. Soc. Am.* **93**, 2172–2180.
- Hartley, R., and Zisserman, A. (2000). *Multiple View Geometry in Computer Vision* (Cambridge University Press, UK).
- Ishizaka, K., and Flanagan, J. L. (1972). "Synthesis of voiced sounds from a two-mass model of the vocal cords," *Bell Syst. Tech. J.* **51**, 1233–1268.
- Ishizaka, K., and Matsudaira, M. (1972). "Fluid mechanical considerations of Vocal Cord Vibration," Monogra. 8, Speech Commun. Res. Lab., Santa Barbara, CA.
- 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 Titze, I. R. (1993). "A methodological study of hemilaryngeal phonation," *Laryngoscope* **103**, 872–882.
- Large, J. (1972). "Towards an integrated physiologic-acoustic theory of vocal registers," *NATS Bull.* **29**, 18–25.
- Mergell, P., and Herzel, H. (1997). "Modelling 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 Registers Voice* **23**, 21–31.
- Stevens, K. N. (1977). "Physics of laryngeal behavior and larynx modes," *Phonetica* **34**, 264–279.
- 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.
- Titze, I. R., and Strong, W. J. (1975). "Normal modes in vocal fold tissues," *J. Acoust. Soc. Am.* **57**, 736–744.
- Titze, I. R. (1988). "The physics of small-amplitude oscillation of the vocal folds," *J. Acoust. Soc. Am.* **83**(4), 1536–1552.
- 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.
- van den Berg, J. W. (1968). "Register problems," *Ann. N.Y. Acad. Sci.* **155**, 129–135.
- Vennard, W. (1967). *Singing ... the Mechanisms and the Technique*, revised ed. 2 (Carl Fisher, Inc., New York).
- Zhang, Z., Neubauer, J., and Berry, D. A. (2006). "The influence of subglottal acoustics on laboratory models of phonation," *J. Acoust. Soc. Am.* **120** (3), 1558–1569.