Mechanisms of irregular vibration in a physical model of the vocal folds

David A. Berry, Zhaoyan Zhang, Juergen Neubauer, and Anders Lofqvist

Citation: The Journal of the Acoustical Society of America **120**, EL36 (2006); doi: 10.1121/1.2234519
View online: http://dx.doi.org/10.1121/1.2234519
View Table of Contents: http://asa.scitation.org/toc/jas/120/3
Published by the Acoustical Society of America

Mechanisms of irregular vibration in a physical model of the vocal folds

David A. Berry, Zhaoyan Zhang, and Juergen Neubauer
The Laryngeal Dynamics Laboratory, David Geffen School of Medicine at UCLA, 1000 Veteran Avenue, Suite 31-24, Los Angeles, California 90095-1794
daberry@ucla.edu

Abstract: Previous investigations have shown that one mechanism of irregular vocal fold vibration may be a desynchronization of two or more vibratory modes of the vocal fold tissues. In the current investigation, mechanisms of irregular vibration were further examined using a self-oscillating, physical model of vocal fold vibration, a hemi-model methodology, and high-speed, stereoscopic, digital imaging. Using the method of empirical eigenfunctions, a spatiotemporal analysis revealed mechanisms of irregular vibration in subharmonic phonation and biphonation, which were not disclosed in a standard acoustic spectrum.

PACS numbers: 43.70.Aj, 43.70.Bk [AL]
Date Received: May 8, 2006 **Date Accepted:** August 15, 2006

1. Introduction

Over the past decade, the method of empirical eigenfunctions (EEFs) has been used to examine spatiotemporal structures in vocal fold vibration (Berry *et al.*, 1994; Alipour *et al.*, 2000; Svec *et al.*, 2000; Berry *et al.*, 2001; Neubauer *et al.*, 2001; Zhang and Jiang, 2005; Doellinger *et al.*, 2005). Herzel *et al.* (1994) hypothesized that "normal phonation corresponds to an essentially synchronized motion of all vibratory modes. A change of parameters … may lead to a desynchronization of certain modes resulting in bifurcations and chaos." Berry *et al.* (1994) confirmed this hypothesis using a finite element simulation of vocal fold vibration. In that study, the spatial EEFs (derived from the nodal displacements of the finite element model) were similar for both simple and complex vibrations patterns, and the desynchronization of the EEFs was the primary mechanism of irregular vibration. Neubauer *et al.* (2001) further substantiated this hypothesis using high-speed digital imaging of vocal fold vibration in human subjects. Desynchronization of anterior-posterior EEFs (derived from glottal contours) distinguished two classes of biphonation: one based on left-right asymmetries, and one based on anterior-posterior asymmetries. In the current study, this desynchronization hypothesis will be subjected to more stringent testing through the use of a self-oscillating, physical model of vocal fold vibration. That is, the laboratory investigations will not be limited by the mathematical simplifications of the previous theoretical studies. Moreover, in contrast to the clinical study which only permitted vocal fold imaging from a superior aspect, the laboratory investigations will yield a systematic study of the medial surface dynamics of the vocal folds. In particular, medial surface dynamics will be extracted from the model using a hemi-model methodology and high-speed, stereoscopic, digital imaging. As a function of subglottal pressure, both simple and complex oscillatory regimes will be imaged, and medial surface coordinates will be extracted. Using the method of EEFs, a spatiotemporal analysis of the medial surface dynamics will evaluate the desynchronization hypothesis.

2. Methods

A physical replica of the human vocal system was constructed using a rubber model of the vocal folds and a uniform (PVC) tracheal tube (with an inner diameter of 2.54 cm and a length of approximately 60 cm) connected upstream to an expansion chamber, simulating the subglottal system. The expansion chamber had an inner cross-section 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. 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. Because the isotropic, single-layered, physical model of the vocal fold has been described previously (Thomson *et al.*, 2005), few details will be given here. However, in the current study, the Young's modulus was approximately 3 kPa for strains in the range of 0–20%, as compared to 13.7 kPa in the previous study. As described elsewhere (Berry *et al.*, 2001; Doellinger *et al.*, 2005), the medial surface dynamics of the folds were imaged using a hemi-model methodology, stereoscopic imaging (through means of a glass prism), and 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, a multipass, time-series cross-correlation analysis was performed on the medial surface images using the image-processing package DaVis (LaVision Inc.). In contrast to previous methods for tracking vocal fold displacements which were semiautomatic (Berry *et al.*, 2001; Doellinger *et al.*, 2005), the DaVis software package was fully automated. The lack of automation in previous imaging studies precluded the study of irregular vibrations.

The subglottal acoustic 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. The mean transglottal pressure was measured using a pressure transducer (Baratron type 220D), which was also mounted flush with the inner wall of the tracheal tube, 2 cm upstream from the vocal fold plates. Analog-to-digital conversion of the signal 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.

3. Results and discussion

During the experiment, the subglottal pressure was increased from zero to a maximum value in discrete increments, and then decreased back to zero in discrete decrements. At each step, measurement was delayed for an interval of approximately 4–5 s after the subglottal pressure changed, allowing the flow field to stabilize. The power spectra of the microphone signal at each of these constant subglottal pressure values are summarized in Fig. 1 . Inspection of this figure suggests that several distinct oscillatory regimes appeared in the physical model: "1" denotes biphonation at a subglottal pressure of approximately 2.1 kPa, "2" denotes a quintupling bifurcation (creating a period-5 subharmonic) at a subglottal pressure of approximately 1.8 kPa, and "3" denotes periodic phonation, or a limit cycle, at a subglottal pressure of approximately 0.8 kPa.

The acoustic spectra for each of these vibratory regimes are shown in Fig. 2 . In Fig. 2(a), the acoustic spectrum for periodic phonation is shown with a fundamental frequency of approximately 200 Hz, and two overtones. In Fig. 2(b), the acoustic spectrum of a subharmonic phonation is shown, with a fundamental frequency of approximately 231 Hz, and subharmonics occurring at integer multiples of approximately 46.2 Hz, corresponding to a period-5 or quintupling of the fundamental frequency. The entrainment, or integer ratio, of these two dominant frequencies is captured by the discrete, finite number of spectral lines in Fig. 2(b). On the other hand, the dense spectral line structure in Fig. 2(c) suggests a lack of entrainment, or a noninteger ratio, of two independent frequencies.

Using the method of EEFs, a spatiotemporal analysis was performed on the three distinct phonatory regimes. The EEFs were computed from high-speed recordings over a time window of 150 ms, centered at the three time locations shown in Fig. 1. The EEFs for periodic phonation are illustrated in the two left-most columns of Fig. 3 . The first column illustrates the spatial EEFs (only the medial-lateral component is shown for comparison with lateral modes of vibration introduced previously in the literature). The second column illustrates the spectra of the corresponding temporal EEFs. Spatial EEFs are often considered the buildings blocks of

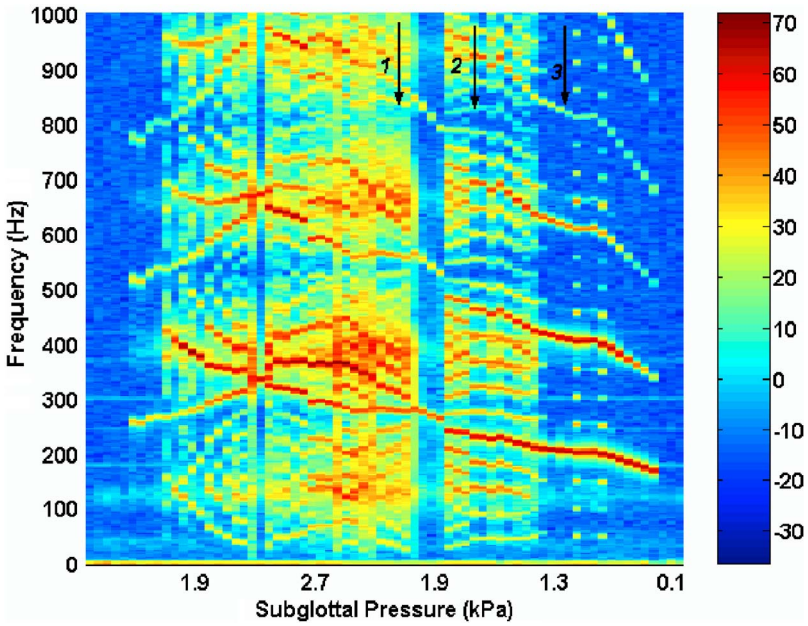

Fig. 1. (Color online) A spectral bifurcation diagram produced by the physical model of vocal fold vibration [in SPL (dB)]. The numbers 1, 2, and 3, denote the instances at which biphonation, subharmonic phonation (quintupling), and periodic phonation occurred, respectively.

vibration. For example, for a linear system, the spatial EEFs would correspond to the normal modes of the system. In column 1, the first spatial EEF (e.g., illustrated in the top subplot) appears to be qualitatively similar to a x-11 mode (according to nomenclature used previously by Titze and Strong (1975), and Berry and Titze (1996)), in which superior and inferior portions of the medial surface vibrate 180 deg out-of-phase with each other. These out-of-phase regions are illustrated by the adjacent red and blue regions found between 0 and 5 mm along the x axis (anterior-posterior direction) and near 0 mm along the y axis (inferior-superior direction). The second spatial EEF (e.g., the subplot directly below EEF1) appears qualitatively similar to a x-10 mode, which captures the in-phase lateral vibrations of the fold (as marked by the blue region, between 0 and 5 mm on the x axis). As shown in the second column, the corresponding temporal EEFs entrained at a fundamental frequency of 200 Hz. As shown in Table 1 these first

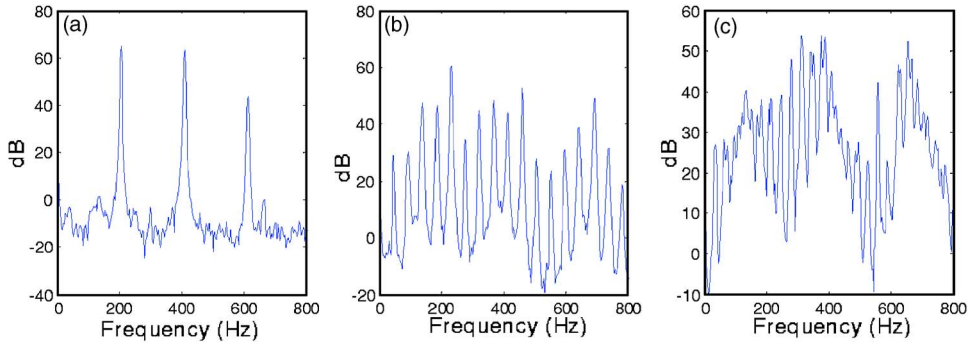

Fig. 2. (Color online) Acoustic spectra corresponding to the numbers 3, 2, and 1 in Fig. 1, which represent the following phonation types, respectively: (a) periodic phonation; (b) subharmonic phonation (quintupling); and (c) biphonation.

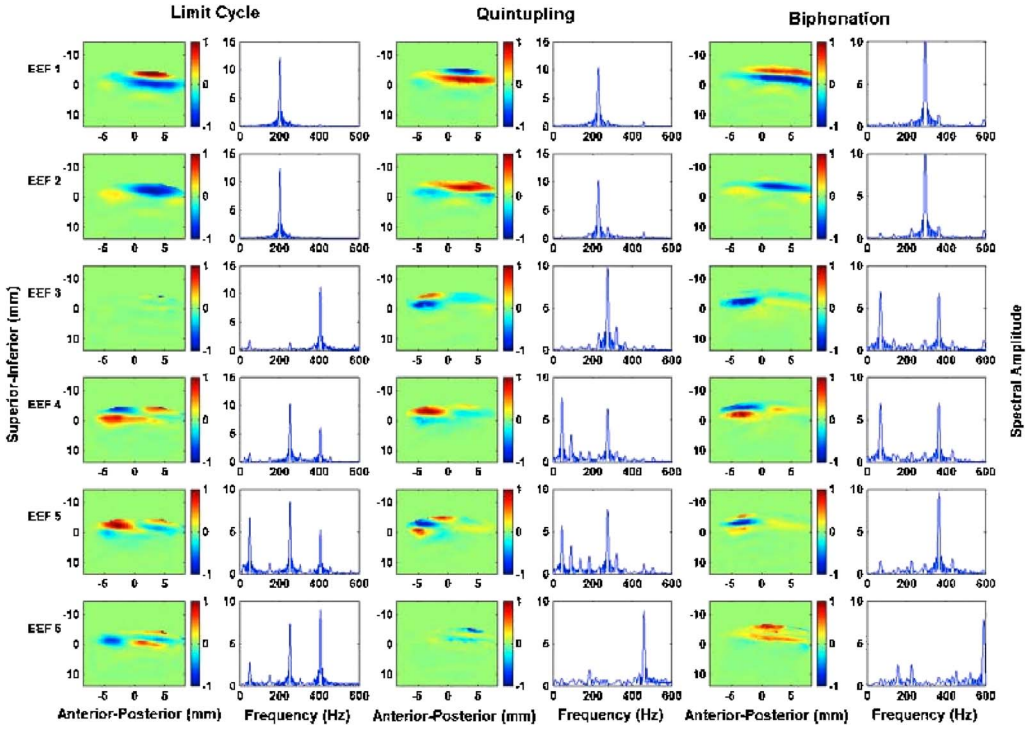

Fig. 3. (Color online) Columns (1), (3), and (5): the normalized medial-lateral components of the first six spatial EEFs for periodic phonation (limit cycle); subharmonic phonation (quintupling); and biphonation, respectively. As shown to the left of the figure, the *y* axis spans the "inferior-superior" direction. Columns (2), (4), and (6): the acoustic spectra of the first six temporal EEFs for periodic phonation (limit cycle); subharmonic phonation (quintupling); and biphonation, respectively. As shown to the right of the figure, the *y* axis captures the "spectral amplitude" (arbitrary units).

two EEFs captured 96.2% of the variance of the medial surface data. Spatial EFFs 3–6 captured higher-order modes in both the anterior-posterior and inferior-superior directions. These higher-order EEFs exhibited both harmonic and nonharmonic frequency components, with respect to the fundamental frequency of 200 Hz.

The EEFs extracted from the subharmonic (quintupling) regime are illustrated in the two central columns of Fig. 3. The first two spatial EEFs entrain at a fundamental frequency $f_1 = 231$ Hz, and are qualitatively similar to the first two EEFs from periodic phonation (the switching of blue and red from superior to inferior positions, as compared to column one, is not significant: Either case demonstrates that the inferior and superior regions are 180 deg out-of-phase with each other). Spatial EEFs 3–6 capture higher-order modes.

Table 1. Percentage variances (or weights) of the first six EEFs.

	Limit cycle weight (%)	Quintupling weight (%)	Biphonation weight (%)
EEF1	58.6	60.5	45.6
EEF2	37.6	27.7	30.1
EEF3	0.76	3.3	6.1
EEF4	0.71	2.3	4.4
EEF5	0.54	1.6	3.9
EEF6	0.42	0.7	1.7

Table 2. Dot products to compare the spatial correlation of the first six EEFs of periodic phonation (P), quintupling (Q), and biphonation (B).

	EEF1	EEF2	EEF3	EEF4	EEF5	EEF6
(P_i, Q_i)	91.2	78.7	4.6	7.6	4.7	41.7
(P_i, B_i)	72.5	66.6	9.0	50.7	23.5	14.6
(Q_i, B_i)	76.8	79.6	41.4	42.1	58.0	19.4

While the EEFs 1–2 entrained with each other, the higher-order EEFs also entrained with the subglottal acoustic resonance f_{sg} of approximately 138 Hz, as estimated by the quarter-length resonance of the subglottal tube. For example, EEF3 (which possessed a dominant frequency component at f_2=277 Hz), exhibited 2:1 entrainment with f_{sg}. EEF4 had significant frequency components at approximately 46, 93, 139, 185, 231, and 277 Hz, etc. All these frequency components were linear combinations of the two primary frequencies, f_1 and f_2 [i.e., $f(n,m)=nf_1+mf_2$, where m and n are integers, e.g., $f(-1,1)=-231$ Hz+277 Hz=46 Hz].

From Table 1, the first two EEFs capture 88.2% of the variance, in comparison to 96.1% of the variance captured by the first two EEFs of periodic phonation. Six EEFs from this quintupling regime were required to capture 96.1% of the variance. In Table 2 , the spatial EEFs from periodic phonation and the quintupling regime are compared using the dot product (the Euclidean scalar product of two vectors). A value of 100% corresponds to identical vectors, or to identical spatial EEFs. A value of 0% corresponds to perpendicular vectors, or to EEFs which are spatially uncorrelated. The dot products of the first three EEFs of the periodic and quintupling regimes were 91.2, 78.7, and 4.6%, denoting a relatively high-spatial correlation between the first two EEFs of the two vibratory regimes. Because EEFs 3 and higher were only weakly excited in the periodic regime, the higher-order EEFs may have captured noise patterns rather than physically meaningful information. Thus, one would not necessarily expect to see a correspondence between the higher-order EEFs of these two oscillation types.

The EEFs extracted from the example of biphonation are illustrated in the two rightmost columns of Fig. 3. The first two spatial EEFs entrain at a fundamental frequency of f_1 =296 Hz, and are qualitatively similar to the first two EEFs from periodic phonation. Spatial EEFs 3–6 capture higher-order modes. EEFs 3–4 have two dominant frequency components at f_2=69 Hz (which exhibited a 1:2 entrainment with the subglottal acoustic resonance of 138 Hz), and 365 Hz. EEF5 also has a frequency component at 365 Hz, and EEF6 has a frequency component at 592 Hz. Thus, again, the higher-order EEFs entrained not only with other tissue modes, but also with the subglottal acoustic resonance. Also, similar to the quintupling regime, all frequency components were linear combinations of f_1 and f_2. Thus, while the central importance of these two frequencies was not apparent in the corresponding acoustic spectra of Fig. 2(c) the spatiotemporal decomposition using EEFs revealed considerable information regarding mechanisms of irregular vibration.

From Table 1, the first two EEFs captured 75.7% of the variance, in comparison to 88.2% for the quintupling regime, and 96.1% for periodic phonation. Six EEFs from this quintupling regime were required to capture 91.8% of the variance. In Table 2, the spatial EEFs from periodic phonation, quintupling, and biphonation regime were compared. For the first spatial EEF, the dot product across the three regimes ranged between 72.5 and 91.2%; and for the second EEF, the dot produced ranged between 66.6 and 79.6%. For the higher-order EEFs the dot products continued to decrease. However, the dot products between the quintupling regime and biphonation (Q_i, B_i) did not diminish as quickly for the higher-order EEFs, perhaps because the higher-order modes were relatively strong and thus essential descriptors of these more complex phonation types, and less significant for periodic phonation.

Dynamical complexity of the vocal fold system not only complicated the entrainment patterns of the EEFs, but also increased the effective number of spatial EEFs more dramatically than observed in a previous theoretical study (Berry *et al.*, 1994). In part, this may have been

because the parameter variation implemented in the current study was subglottal pressure, which naturally excites additional modes when increased. In contrast, the previous study varied mucosal stiffness.

One may argue that the bifurcations observed in this paper may not be related to human phonation because of the relatively long subglottal tube length of 60 cm. However, the entrainment of the vocal folds to acoustic resonances of the supraglottal tract in human subjects has been reported previously (Mergell and Herzel, 1997). Moreover, previous investigators have observed an intimate connection between subglottal acoustic resonances and vocal registers (Austin and Titze, 1997). Also, given the range of first resonances of the human subglottal system reported in the literature (approximately 400–500 Hz, see Ishizaka *et al.*, 1976), some of the frequency components of the EEFs for irregular phonation reported by Neubauer *et al.*, (2001, e.g., in the range of 197–338 Hz), may have been exhibiting 1:2 entrainment patterns with sub- or supra-glottal acoustic resonances. Thus, while future laboratory studies which more accurately model the human subglottal system may be needed to further investigate mechanisms of irregular vibration, the entrainment patterns observed between tissue modes and acoustic resonances in this study appear to be characteristic of source-resonator systems, and are consistent with previous observations in voice literature.

4. Conclusions

Mechanisms of irregular vibrations were explored in a self-oscillating, physical model of vocal fold vibration. Medial surface dynamics of a vocal fold were obtained using a hemi-model methodology together with high-speed, stereoscopic, digital imaging. Using the method of EEFs, a spatiotemporal analysis was performed on the medial surface vibrations of the folds. Mechanisms of irregular vibration for subharmonic phonation (quintupling) and biphonation were revealed by the entrainment patterns of the spatial EEFs. While stronger, lower-order EEFs tended to entrain with each other, weaker, higher-order EEFs also sometimes entrained with subglottal acoustic resonances. These mechanisms of irregular vibration were not disclosed by a spectral analysis of the acoustic signal, but were revealed through a spatiotemporal analysis of vocal fold oscillations using the method of EEFs. Using these same laboratory techniques, future studies are envisioned which will investigate mechanisms of irregular vibration across many phonatory conditions of both the normal and pathological voice. Such studies may employ the following models: multilayered physical models, excised animal larynges, excised human larynges, and *in vivo* animal larynges. While high-speed observation of the medial surface of the vocal folds is not currently available in the clinic, as the temporal and spatial resolution of ultrasonic imaging methods continues to increase, at some point in the future it may be possible to have clinical access to the medial surface of the vocal folds in order to assess mechanisms of irregular vibration. Already, Neubauer *et al.* (2001) have demonstrated that some mechanisms of irregular vocal fold vibration may be investigated clinically using a spatiotemporal decomposition of the superior view provided by high-speed, digital, endoscopic imaging.

Acknowledgment

This investigation was supported by Grant No. R01 DC03072 from NIDCD.

References and links

Austin, S. F. and Titze, I. R. (**1997**). "The effect of subglottal resonance upon vocal fold vibration," J. Voice **11**, 391–402.

Alipour, F., Berry, D. A., and Titze, I. R. (**2000**). "A finite element model of vocal fold vibration," J. Acoust. Soc. Am. **108**, 3003–3012.

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**, 3345–3354.

Doellinger, M., Berry, D. A., and Berke, G. S. (**2005**). "Medial surface dynamics of an *in vivo* canine vocal fold during phonation," J. Acoust. Soc. Am. **117**, 3174–3183.

Herzel, H., Berry, D. A., Titze, I. R., and Saleh, M. M. (**1994**). "Analysis of vocal disorders with methods from nonlinear dynamics," J. Speech Hear. Res. **37**, 1008–1019.
Ishizaka, K., Matsudaira, M., and Kaneko, T. (**1976**). "Input acoustic-impedance measurement of the subglottal system," J. Acoust. Soc. Am. **60**, 190–197.
Mergell, P., and Herzel, H. (**1997**). "Modeling biphonation — The role of the vocal tract," Speech Commun. **22**, 14l–154.
Neubauer, J., Mergell, P., Eysholdt, U., and Herzel, H. (**2001**). "Spatio-temporal analysis of irregular vocal fold oscillations: biphonation due to desynchronization of spatial modes," J. Acoust. Soc. Am. **110**, 3179–3192.
Svec, J. G., Horacek, J., Sram, F., and Vesely, J. (**2000**). "Resonance properties of the vocal folds: in vivo laryngoscopic investigation of the externally excited laryngeal vibrations," J. Acoust. Soc. Am. **108**, 397–407.
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.
Zhang, Y., and Jiang, J. J. (**2005**). "Spatiotemporal chaos in excised larynx vibrations," Phys. Rev. E **72**, Art. No. 035201R, 1–4.

