Interpretation of biomechanical simulations of normal and chaotic vocal fold oscillations with empirical eigenfunctions

David A. Berry, Hanspeter Herzel, Ingo R. Titze, et al.

Citation: The Journal of the Acoustical Society of America **95**, 3595 (1994); doi: 10.1121/1.409875
View online: https://doi.org/10.1121/1.409875
View Table of Contents: https://asa.scitation.org/toc/jas/95/6
Published by the Acoustical Society of America

Interpretation of biomechanical simulations of normal and chaotic vocal fold oscillations with empirical eigenfunctions

David A. Berry
Department of Speech Pathology and Audiology, National Center for Voice and Speech, The University of Iowa, Iowa City, Iowa 52242

Hanspeter Herzel
Technical University Berlin, Institute for Theoretical Physics, Hardenbergstr. 36, D-10623 Berlin, Germany

Ingo R. Titze
Department of Speech Pathology and Audiology, National Center for Voice and Speech, The University of Iowa, Iowa City, Iowa 52242

Katharina Krischer
Fritz Haber Institut der Max Planck Gesellschaft, Faradayweg 4-6, D-14195 Berlin 33, Germany

(Received 7 July 1993; accepted for publication 3 February 1994)

Empirical orthogonal eigenfunctions are extracted from biomechanical simulations of normal and chaotic vocal fold oscillations. For normal phonation, two dominant empirical eigenfunctions capture the vibration patterns of the folds and exhibit a 1:1 entrainment. The eigenfunctions show some correspondence to theoretical low-order normal modes of a simplified, three-dimensional elastic continuum, and to the normal modes of a linearized two-mass model. The eigenfunctions also facilitate a physical interpretation of energy transfer mechanisms in vocal fold dynamics. Subharmonic regimes and chaotic oscillations are observed during simulations of a lax cover, in which case at least three empirical eigenfunctions are necessary to capture the resulting vocal fold oscillations. These chaotic oscillations might be understood in terms of a desynchronization of a few of the low-order modes, and may be related to mechanisms of creaky voice or vocal fry. Furthermore, some of the empirical eigenfunctions captured during complex oscillations correspond to higher-order normal modes described in earlier theoretical work. The empirical eigenfunctions may also be useful in the design of lower-order models (valid over the range for which the empirical eigenfunctions remain more or less constant), and may help facilitate bifurcation analyses of the biomechanical simulation.

PACS numbers: 43.70.Aj, 43.75.Rs

INTRODUCTION

With any model of a physical or physiological process, there is always a trade-off between simplicity and completeness. The model should be simple enough to be useful in conceptualization and prediction, but also complete enough to represent the process accurately.

This certainly applies to vocal fold models. Early one-mass and two-mass models (Flanagan and Landgraf, 1968; Ishizaka and Flanagan, 1972) were simple enough to be described in a few pages of print. They were elegant in that they helped conceptualize the interaction between airflow and tissue movement to produce self-oscillation. But there is considerable doubt that they represented the geometry and the viscoelastic properties of the vocal folds adequately for the study of voice disorders or special vocal qualities. More recent models by Titze and Talkin (1979), and Alipour and Titze (1985a) have enough biomechanical detail to model the three-dimensional layered structure of vocal fold tissue, but a heavy price is paid in terms of mathematical complexity and speed of computation. Furthermore, interpreting the dynamics of such intensive descriptions of the vocal folds can be a formidable task, particularly if irregular, chaotic vibrations occur (Titze *et al.*, 1993).

One way to facilitate the physical interpretation of a vibrating structure is to calculate its principal modes of vibration. Sometimes, even complicated vibration patterns can be explained by a relatively small number of orthogonal modes.

I. MODAL ANALYSIS OF THE VOCAL FOLDS

Modal analysis is a basic technique used to analyze many vibrating structures. Traditionally, it refers to the process of determining the normal (natural) modes and frequencies of a linear (or linearized) system. It is a powerful technique because it provides a framework in which a system can be decomposed into a set of independent vibration patterns, each with a characteristic (although not necessarily unique) frequency. Experimentally, these normal modes/frequencies can be observed immediately after a system is pulse excited, or during a forced, sinusoidal excitation (provided the driving frequency coincides closely enough with one of the systems' natural frequencies). One of the major limitations of this technique is that it is only

 0001-4966/94/95(6)/3595/10/$6.00

valid for linear systems. However, in practice many systems are approximately linear for small-amplitude oscillations.

A. Theoretical normal modes

The concept of normal modes and frequencies is not new to speech science. For example, formants are frequencies which correspond to the normal modes of the vocal tract. They have been discussed extensively in terms of vowel production. The same concepts have not been exploited to the same degree for an understanding of vocal fold movement, although the study of normal modes in vocal fold tissues does have its beginnings. Almost two decades ago, Titze and Strong (1975) theoretically determined normal modes of the vocal folds. By examining a single fold and treating it as a three-dimensional, elastic, compressible medium, and by assuming a rectangular parallelepiped with simple boundary conditions (anterior, posterior, and lateral boundaries fixed; medial, superior, and inferior boundaries free), normal modes were expressed in terms of elementary sines and cosines. For comparison with empirical modes to be shown later, the theoretical modes are reviewed and briefly discussed. The lateral displacements ξ of the x modes are given by

$$\xi(x,y,z,t) = A \exp i\omega_x t \cos\frac{(2n_x-1)\pi x}{2D} \sin\frac{n_y\pi y}{L} \cos\frac{n_z\pi z}{T} \quad (1)$$

and the vertical displacements ζ of the z modes are given by

$$\zeta(x,y,z,t) = B \exp i\omega_z t \cos\frac{(2n_x-1)\pi x}{2D} \sin\frac{n_y\pi y}{L} \cos\frac{n_z\pi z}{T}, \quad (2)$$

where n_x, n_y, and n_z are integers indicating the order of the modes; L, D, and T are the length, depth and thickness of the folds, respectively; A and B are arbitrary constants; and ω_x and ω_z are the radian frequencies of vibration. Any possible y displacements (anterior–posterior direction) are neglected. This is based on experimental evidence that the trajectories of vocal fold fleshpoints are mostly planar (Baer, 1981; Saito *et al.*, 1981; Saito *et al.*, 1985).

In order to distinguish the modes, the order indices (n_x, n_y, and n_z) must be specified and the modes need to be identified as either x or z modes (the assumption of compressible tissue allows the decoupling of such modes). In practice, the n_x index is usually not specified because the standing wave pattern governed by n_x (the first cosine term) is assumed to be constant (e.g., the likelihood of reflections from the fixed lateral boundary is small because of high attenuation in the thyroarytenoid muscle). Thus, following nomenclature introduced previously (Titze and Strong, 1975; Titze, 1976,1988), the modes are designated as either x-$n_y n_z$ or z-$n_y n_z$ modes. Conceptually, the n_y and n_z indices indicate how many half-wavelengths occur along the longitudinal and vertical dimensions, respectively.

A few of the lower-order modes are shown in Fig. 1. Figure 1(a) shows a superior view (upper) and coronal view (lower) of the x-10 mode. From the superior view, the commonly observed oval glottis is displayed. From the coronal view, all the lateral tissue displacements are in phase along the vertical dimension. An x-11 mode is displayed in Fig. 1(b). In the coronal view, the tissue at the top of the folds is 180 deg out of phase with the tissue at the bottom of the folds. Variations of these lowest-order x modes describe some of the most commonly observed vocal fold vibration patterns (Moore and Von Leden, 1958; Hirano, 1975). Indeed, an appropriate combination of these modes is known to be essential for self-oscillation of the folds (Titze, 1988). An x-21 mode is illustrated in Fig. 1(c), and a z-10 mode in Fig. 1(d) (sagittal view on top). These modes are not as easily observed because (1) the superior aspect, which is almost always used in high-speed films and videostroboscopy, is not ideal for viewing z modes (saggital or coronal views would be better), and (2) higher-order modes (such as the x-21 mode) usually have smaller vibrational amplitudes and are often not excited.

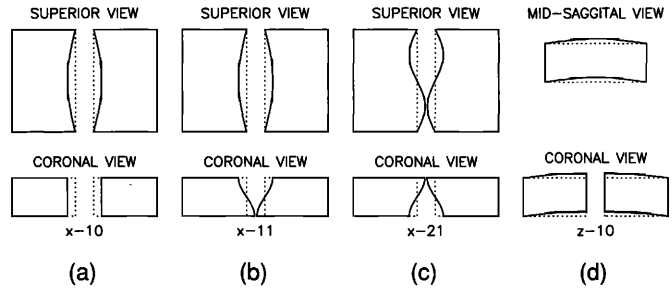

FIG. 1. A few of the low-order, theoretical, normal modes are shown from superior and coronal views: (a) x-10, (b) x-11, (c) x-21. A saggital and coronal view is shown for (d) z-10. An artificial separation of left and right folds is used in order to display the true theoretical modes without deformation from collision.

B. Normal modes in low-order models

Not long after these normal modes were introduced, Titze (1976) claimed that (1) self-oscillation of the vocal folds consists of "approximately linear combinations of the normal modes," and (2) that "self-oscillation... occurs at... one of the natural frequencies of oscillation, usually the lowest." Titze demonstrated the plausibility of these concepts through an analysis of the two-mass model (Ishizaka and Flanagan, 1972). The normal modes of the two-mass model were shown to be analogous to the lowest order x modes of the simplified elastic continuum; that is, the mode where the two masses are in phase is similar to the x-10 mode, and the mode where the two masses are 180 deg out of phase is similar to the x-11 mode. The ability of the two-mass model to self-oscillate can be explained, in large measure, by the existence of these two modes, which facilitate energy transfer from the airflow to the tissue (Stevens, 1977; Broad, 1979; Titze, 1988). For "typical" Ishizaka and Flanagan (1972) parameters, the natural frequencies of the normal modes are 120 and 201 Hz, respectively (Titze, 1976).

Self-oscillation during normal phonation also involves a 1:1 "entrainment" of the modes. Entrainment is a phe-

nomenon in which a nonlinear coupling of system variables causes the natural frequencies of the system to shift so as to be related by an integer ratio. For example, a 1:1 entrainment in the two-mass model means that both modes oscillate at the same frequency. Such an entrainment has been observed over a wide range of parameters in the two-mass model (Herzel *et al.*, 1991). As predicted by Titze (1976), this entrainment occurs at a frequency very close to the natural frequency of the lowest-order mode. Over a certain range of parameters (e.g., those corresponding to low stiffness of the upper mass and weak coupling of the masses), the breakdown or desynchronization of this 1:1 entrainment has also been observed in the two-mass model (Herzel *et al.*, 1991). In such parameter regions, various nonlinear phenomena have been observed including subharmonic regimes, beating-like toroidal oscillations (i.e., low-frequency modulations), and chaotic motion (Herzel *et al.*, 1991).

Subharmonics, low-frequency modulations, and chaos are also commonly observed in patients with vocal disorders (Herzel and Wendler, 1991; Baken, 1991; Herzel *et al.*, in press) and during infant cries (Mende *et al.*, 1990). Consequently, this desynchronization of the modes is believed to be an essential mechanism of many vocal disorders (Titze *et al.*, 1993; Herzel *et al.*, in press).

C. Experimental studies of normal modes

To date, most of the discussion of normal modes in vocal fold tissues has been in a theoretical sense. Direct measurement of the modes has proven problematic, partially because of the small dimensions of the vocal folds (on the order of 1 cm). Traditional modal analysis in which accelerometers are used to trace trajectories at various locations on a structure would undoubtedly yield unsatisfactory results. The number of accelerometers that could be placed on the folds would be limited, the weight of accelerometers might alter the modes significantly, and the ability to firmly attach a device to the elastic tissue of the folds would be limited.

Until direct measurement of the modes becomes more feasible, there are additional theoretical approaches that can be used to investigate the modes, particularly with the help of a biomechanical simulation of vocal fold movement (Alipour and Titze, 1985a). The simulation uses a finite element approach to the solution of viscoelastic waves in a continuum (Titze and Talkin, 1979). A series of experimental studies have been performed to quantify the elastic properties of vocal fold tissues (Alipour and Titze, 1985b), and more work in this area is in progress. Indeed, the development of this simulation has been an effort to integrate many independent measurements and theoretical considerations into one coherent "picture" of vocal fold vibration.

As pointed out earlier, traditional modal analysis is limited by the fact that it is only valid for linear systems. However, there are many nonlinearities associated with the vocal folds. One of these is the nonlinear stress-strain curve of vocal fold tissues. Another is the nonlinear pressure-flow relation in the glottis. A third is the nonlinearity associated with vocal fold collision. While for small transient oscillations these nonlinearities might be neglected, self-sustained oscillation depends critically on at least one nonlinear constitutive equation. Indeed, for many vocal fold configurations linear dynamics is not even approximately true.

However, the method of empirical orthogonal eigenfunctions (Lorenz, 1956) has been used for many years to extract physically meaningful structures from nonlinear systems. For example, Lumley (1967) advocated the technique as a way to extract "coherent structures" from a turbulent flow. In recent years, the method has become a popular technique in a variety of problems in fluid dynamics (Sirovich, 1987; Aubry *et al.*, 1991; Deane *et al.*, 1991; Armbruster *et al.*, 1992). The method is an application of a general technique familiar to many disciplines, and has also been referred to as the singular value decomposition (Golub and Van Loan, 1983), singular spectrum analysis (Vautard *et al.*, 1992), principal-components analysis (Zahorian and Rothenberg, 1981), principal factor analysis (Johnson and Wichern, 1982), the Karhunen-Loève expansion (Fukunaga, 1972), the proper orthogonal decomposition (Lumley, 1967), and the biorthogonal decomposition (Aubry *et al.*, 1991).

Furthermore, Breuer and Sirovich (1991) have recently shown that, for a general class of linear systems, the empirical eigenfunctions actually reduce to the linear normal modes. The ability to extract physically meaningful structures from both linear and nonlinear systems makes the method of empirical orthogonal eigenfunctions a particularly useful tool for analyzing vocal fold movement. In the case of small-amplitude vibrations for which the tissue stress-strain curves are approximately linear, the empirical eigenfunctions should be related to the normal modes of vocal fold tissues. For larger amplitude vibrations for which tissue nonlinearities become important, the eigenfunctions should appear as distortions of the normal modes, i.e., a reflection of the new nonlinear phenomena (Breuer and Sirovich, 1991).

Moreover, the statistical nature of this technique makes it well-suited for the present investigation. In a sense, the method is "blind" to all the complexities of the biomechanical simulation (e.g., nonlinearities in stress-strain curves, complex geometry of the folds, layered tissue structure, tissue incompressibility which induces a coupling between lateral and vertical modes, aerodynamic forces, collision forces, viscous losses). Such complexities forbid an analytical solution of the modes, but present no difficulties for the method of empirical eigenfunctions.

The method of empirical orthogonal eigenfunctions differs from a traditional normal mode analysis in that it does not determine "modes" directly from the equations of motion. Rather, "modes" are determined by statistical correlations of the output variables, i.e., a covariance matrix is generated and eigenvectors are computed. The eigenvectors are orthogonal and are guaranteed to be optimal in the sense that they regenerate the output data with minimum least-square error (for any arbitrary number of eigenvectors). Unlike a normal mode analysis, the method of empirical eigenfunctions does not calculate all the possible

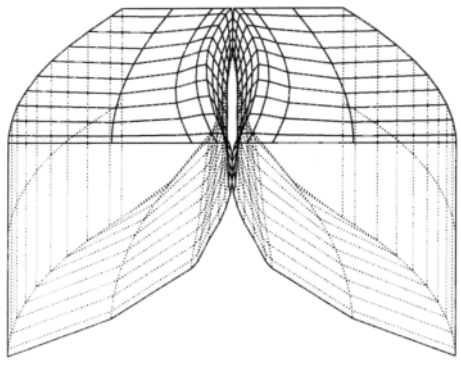

FIG. 2. A view of the biomechanical simulation immediately before glottal closure is shown, with the posterior edge of the folds in the foreground. There are nine layers positioned along the anterior–posterior length.

modes of a system. Rather, it only extracts those "modes" which are excited. For the present investigation, the excited "modes" are the focus of interest, and are used as a tool for interpreting vocal fold dynamics during self-oscillation.

II. PROCEDURES

A. Trajectories from the simulation

Empirical eigenfunctions were calculated based on the output of a biomechanical simulation of vocal fold movement (Alipour and Titze, 1985a). The simulation was run as part of a complete speech synthesis system, including sub- and supraglottal systems. The biomechanical model of the folds consists of nine longitudinal layers as shown in Fig. 2, where the posterior edge of the folds is in the foreground. Anterior and posterior boundaries are fixed. Each layer consists of 32 finite elements (triangles) or 26 nodes (fleshpoints), as shown on the left side of Fig. 3. The elements which correspond to the body (or muscle) are marked "B," the elements which correspond to the ligament are marked "L," and the elements which correspond to the cover (or mucosa) are marked "C." Each of these regions possesses distinct elastic properties. Three nodes per layer are placed on a fixed lateral boundary. Thus, there are 207 nodes per fold (9 layers×23 nodes/layer) which are free to oscillate. As in earlier investigations, lateral and vertical motions are allowed, but no movement along the anterior–posterior direction. With two degrees of freedom per node, there are 414 total degrees of freedom if left and right folds are symmetric, and 828 if asymmetric. Although the simulation is equipped to handle asymmetric folds, all the runs for this analysis employed left–right symmetry.

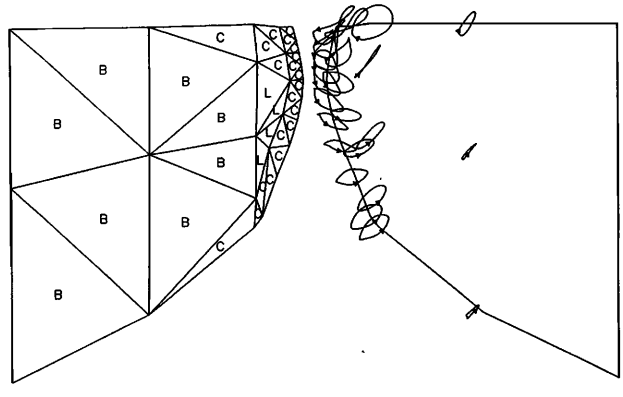

FIG. 3. A coronal view of the fifth longitudinal layer. On the left side of the figure, the 32 elements/layer are displayed and distinguished as corresponding to the body ("B"), the cover ("C"), or the ligament ("L"). On the right side of the figure, trajectories of vocal fold fleshpoints are shown for parameters corresponding to normal phonation.

Nodal trajectories for parameters corresponding to normal phonation are shown on the right side of Fig. 3. The trajectories are taken from the fifth longitudinal layer (the layer midway between anterior–posterior boundaries), which is the layer with the most lateral movement. Qualitative similarities exist between these trajectories and fleshpoint trajectories observed experimentally (Baer, 1981; Saito *et al.*, 1981; Saito *et al.*, 1985). The x and z coordinates from the trajectories of each of the 207 nodes were used as the input for calculating the covariance matrix, and the resulting eigenfunctions. Although the simulation was run at a sampling rate of 20 kHz, the nodal coordinates were only saved at a rate of 5 kHz, which was found to be sufficient. Frequencies above about 1 kHz were essentially nonexistent in the trajectories (measured on a power spectrum, they were at least 40 dB below the strongest frequency).

B. Calculation of empirical eigenfunctions

First, the nodal coordinates R_i were separated into mean and oscillatory components:

$$R_i(t)=\bar{R}_i+r_i(t), \quad i=1,2,...,414, \tag{3}$$

where the bar denotes a mean value. The mean represents the dynamic equilibrium of the system, and the remaining oscillatory component represents the time-varying displacements about this equilibrium. A covariance matrix was generated using the time-varying displacements:

$$S_{ij}=\frac{1}{N}\sum_{k=1}^{N} r_i(t_k)r_j(t_k), \quad i,j=1,2,...,414, \tag{4}$$

where t_k are the discrete times at which the coordinates are sampled, and N is the total number of time samples. The eigenvectors of the covariance matrix correspond to the empirical eigenfunctions. At any time t_k, the nodal displacements may be expressed as a linear combination of the empirical eigenfunctions ϕ_j (Deane *et al.*, 1991):

$$r_i(t_k)=\sum_{j=1}^{414} a_j(t_k)\phi_j(i), \quad i=1,2,...,414, \tag{5}$$

where $\phi_j(i)$ is the ith component of the jth eigenfunction and $a_j(t_k)$ is the temporal coefficient of the jth eigenfunction at time t_k. The temporal coefficients a_j may be computed by projecting the eigenfunctions ϕ_j onto the time-varying displacements:

$$a_j(t_k)=\sum_{i=1}^{414} r_i(t_k)\phi_j(i), \quad j=1,2,...,414. \tag{6}$$

TABLE I. Parameters used in the finite element simulation of "normal" phonation.

lung pressure	0.8 kPa
cricothyroid activity	10%
lateral cricoarytenoid activity	95%
thyroarytenoid activity	50%
transverse Young's modulus of the body	4 kPa
transverse Young's modulus of the cover and ligament	2 kPa
longitudinal shear modulus of the body	12 kPa
longitudinal shear modulus of the cover	10 kPa
longitudinal shear modulus of the ligament	40 kPa
viscosity of the body, cover & ligament	6 poise

The temporal coefficients themselves may be thought of as temporal eigenfunctions, and correspond to the eigenvectors of a temporal correlation matrix which may also be generated from the original data (Sirovich, 1987). The set of temporal eigenfunctions is orthogonal, as is the corresponding set of spatial eigenfunctions. Both sets reveal distinct features of the dynamics of the system. The spatial eigenfunctions (sometimes referred to as "topos," Aubry *et al.*, 1991) reveal topological patterns in the data and are analogous to the normal modes of linear systems. The temporal eigenfunctions (sometimes referred to as "chronos," Aubry *et al.*, 1991) reveal information about possible entrainment of the modes, and capture the frequencies at which the modes oscillate.

Each pair of spatio-temporal eigenfunctions has a corresponding eigenvalue, which quantifies the degree to which the eigenfunctions can regenerate the nodal trajectories (in terms of variance). Often just a few eigenfunction pairs capture the essential dynamics of a system (Deane *et al.*, 1991), which facilitates a reduction of the system as well as a physical interpretation of the dynamics.

All the covariance matrices calculated in this study were generated with one second of stationary output (5000 time frames). Initial transients and other nonstationary segments were not used in calculating the covariance matrices and resulting empirical eigenfunctions. The dominant vibration frequencies of these modes ranged between 80 and 160 Hz, so 80 to 160 cycles were used in calculating the modes.

III. RESULTS AND DISCUSSION

A. Normal phonation

First of all, we consider the results of the analysis for typical parameters corresponding to "normal" phonation (e.g., see Table I). The normalized eigenvalues computed at these parameter values are shown in Table II, in descending order. The far right column shows a cumulative sum of the eigenvalues. From this table, we see that the first eigenfunction explains about 72% of the variance of the nodal trajectories, and the second eigenfunction about 26% of the variance. Together the first two eigenfunctions explain approximately 98% of the variance, suggesting the dominance of just a few primary modes. These results were consistent over a range of subglottal pressures (0.2–1.2 kPa) and elastic constants (0.6 to 2 kPa for the Young's modulus of the cover).

TABLE II. Normalized eigenvalues/variances corresponding to the eigenfunctions of "normal" phonation.

Eigenfunction number	λ_i (%)	Cumulative sum of λ_i (%)
1	72.5	72.5
2	25.2	97.7
3	1.5	99.2
4	0.5	99.7

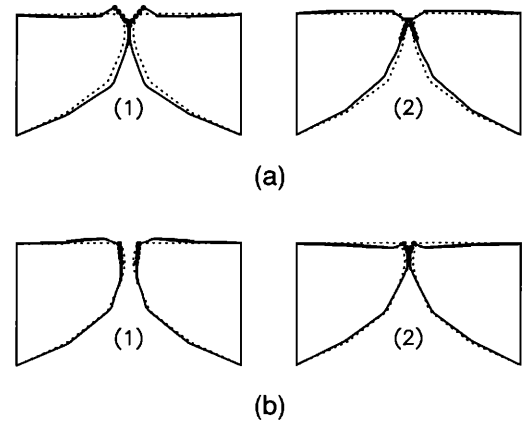

FIG. 4. A coronal view of the two strongest spatial eigenfunctions for normal phonation. The first eigenfunction is shown in (a) and the second eigenfunction is shown in (b). In both cases, frame 1 corresponds to a maximum excursion of the eigenfunction, and frame 2 corresponds to a minimum excursion.

A coronal view of the first eigenfunction is shown in Fig. 4(a). Frames 1 and 2 display maximum and minimum excursions of the eigenfunction, respectively (solid lines). The dotted lines show the mean coordinate values. By examining the motion of the folds near the top of the glottal air passage (e.g., see the top five medial nodes which are indicated on either side), one can note the correspondence of this eigenfunction with the x-11 mode. That is, there is a higher and lower portion of the folds which are 180 deg out of phase. Consequently, this eigenfunction is largely responsible for alternately shaping a divergent (frame 1) and convergent (frame 2) glottis. In addition, there is considerable vertical motion similar to the z-10 mode. This coupling of x and z modes is not surprising given the incompressibility of the tissue (Titze, 1976). Tissue incompressibility implies that the overall tissue volume does not change, so if the folds are compressed laterally, they must bulge out vertically, and vice versa.

A coronal view of the second eigenfunction is shown in Fig. 4(b). Again frames 1 and 2 show maximum and minimum excursions of this eigenfunction. Again, note that near the top of the folds (the region that might approximate a rectangular parallelepiped), this eigenfunction is qualitatively similar to the x-10 mode [Fig. 1(a)], and is largely responsible for the net lateral movement of the folds in this region.

Figure 5 shows the temporal coefficients associated

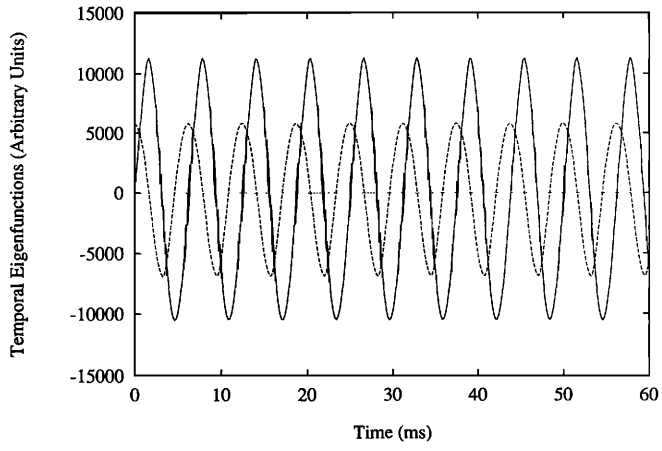

FIG. 5. The two dominant temporal eigenfunctions for parameters corresponding to normal phonation. The solid line corresponds to eigenfunction 1 and the dotted line corresponds to eigenfunction 2.

with each of these eigenfunctions; the solid line displays the temporal coefficients for the first eigenfunction and the dotted line illustrates the time coefficients for the second eigenfunction. The temporal coefficients for both eigenfunctions are nearly sinusoidal with a sine/cosine relationship (mode 1 lags mode 2 by about 90 deg). A simple analysis shows that the modes are synchronized in such a way that energy transfer may occur from the airflow to the tissue, enabling self-oscillation. Specifically, note that the solid line can be expressed as $\sin(t)$. A maximum in this line occurs for a divergent glottis [see Fig. 4(a), frame 1], and a minimum in the solid line occurs for a convergent glottis [see Fig. 4(a), frame 2]. If Bernoulli's law is taken as approximately valid, then the intraglottal pressure will be relatively low for a divergent glottis and relatively high for a convergent glottis (Titze, 1988). As a first-order approximation, one might say that the intraglottal pressure is in phase with $-\sin(t)$.

The dotted line can be expressed as $\cos(t)$. A maximum in this line occurs when the folds are most open [Fig. 4(b), frame 1], and a minimum occurs when the folds are closed [Fig. 4(b), frame 2]. Because this mode roughly corresponds to the net lateral displacement of the tissue, a rough estimate of the net lateral velocity of the tissue is given by the time derivative of $\cos(t)$, or $-\sin(t)$. Thus an examination of the two dominant spatiotemporal eigenfunctions of the biomechanical simulation reveals an in-phase relationship between the intraglottal pressure and the net tissue velocity, which allows energy transfer from the airflow to the tissue.

It is already well known that this condition must be satisfied if self-oscillation of the folds is to occur in the presence of dissipation. However, the important point is that this method of analysis reduced several hundred trajectories of the biomechanical simulation to essentially two modes of vibration. With this reduction, the dynamics of a biomechanical model with many degrees of freedom could be discussed and interpreted as easily as the dynamics of a much more constrained, low-order model. The ability to reduce large amounts of data to essential dynamics will be crucial for understanding more complex output from the biomechanical simulation.

As a word of caution, it should be noted that because the biomechanical simulation was reduced to essentially *two modes* of vibration for parameters corresponding to normal phonation, the biomechanical model was in no way reduced to a two-mass model. Even for normal phonation, the most dominant mode of the biomechanical simulation was not simply a lower-order x mode such as might be captured by a two-mass model, but an x mode coupled with a z mode. Furthermore, although the modes of a two-mass model may be qualitatively similar to the lower-order modes of a simplified elastic continuum, it is questionable whether two bar-shaped masses can adequately capture the smoothly varying shape of the glottis. The discontinuities introduced by such gross spatial discretization would likely have an adverse effect on synthesis.

Moreover, the biomechanical simulation has hundreds of degrees of freedom which allow it to be excited into many modes of vibration not possible for the two-mass model. The fact that just a few of the lower-order modes are excited for a range of parameters corresponding to normal phonation is to be expected and might even be viewed as one validation of the biomechanical simulation. For other parameter configurations, additional modes are excited in the simulation. The study of these modes may yield additional insights into vocal fold dynamics, and may have relevance for an understanding of voice disorders.

B. Chaotic oscillations

Next, empirical eigenfunctions are calculated when the transverse Young's modulus of the cover (E_c) is decreased in the biomechanical simulation. This parameter change simulates a lax cover, and may have physiological relevance to vocal fry or creaky voice (Scherer, 1989). As a preliminary analysis, the acoustic output of the simulation was observed as E_c was gradually lowered from its initial value of 2 kPa. No unusual behavior was noticed until E_c reached values below 0.6 kPa, at which point the signal became irregular, and perceptually rough. At 0.4 kPa, the signal became regular again, but with a doubling of the original period (an "octave jump"), which appeared as alternating high and low amplitudes in the acoustic output. Such phenomena (e.g., irregular oscillations, low frequencies, and alternating high/low amplitudes) are characteristic of the acoustic output of creaky voice (Hollien and Michel, 1968). Listening to the acoustic output also gave the perception of creaky voice.

A spectral bifurcation diagram (e.g., Lauterborn, 1986) is shown in Fig. 6, where E_c is slowly varied from 0.35 to 0.65 kPa. From left to right, one views transitions from a subharmonic regime to chaos to the periodic regime characteristic of normal phonation. This figure shows striking similarities to spectrograms of newborn cries (Mende *et al.*, 1990) and to acoustic cavitation experiments (Lauterborn, 1986). A more complete bifurcation analysis of this region will be treated in a forthcoming paper.

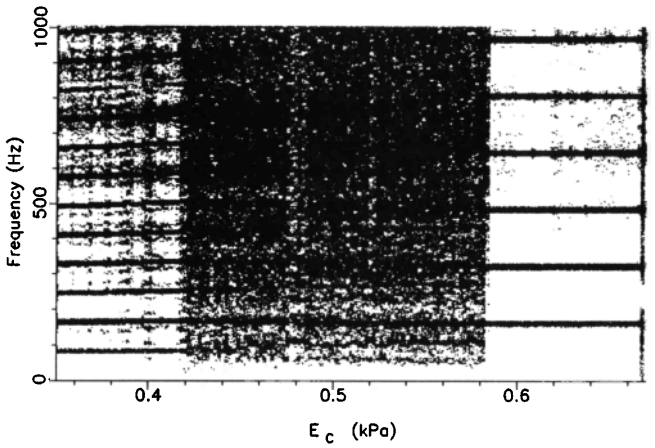

FIG. 6. A spectral bifurcation diagram as E_c is slowly varied from 0.35 to 0.65 kPa, in increments of 0.01 kPa every 400 ms. Transitions from a subharmonic regime to chaos to periodic motion are displayed.

It should be noted that, so far, none of the standard routes to chaos (period doublings, intermittency, secondary Hopf bifurcations, etc.) have been observed in this model. Rather, we have observed abrupt transitions (and hysteresis) which are based on *global* bifurcations rather than *local* bifurcations. Further studies with additional parameter variations will show whether some of the more traditional routes to chaos can be identified.

In the present investigation, empirical eigenfunctions are determined at $E_c=0.4$ kPa and $E_c=0.5$ kPa. Table III shows the eigenvalues for both parameter configurations. At $E_c=0.4$ kPa, four eigenfunctions are needed to describe the nodal trajectories in as much detail (in terms of variance) as the first two eigenfunctions at $E_c=2$ kPa (see Table II). At $E_c=0.5$ kPa, additional eigenfunctions are needed. However, even for the complicated, nonperiodic behavior at $E_c=0.5$ kPa, relatively few eigenfunctions are needed to capture most of the variance of vocal fold dynamics. Out of 414 possible eigenfunctions, only six are needed to describe the motion in considerable detail.

Moreover, the first three spatial eigenfunctions at $E_c=0.5$ kPa are essentially equivalent to the first three spatial eigenfunctions at $E_c=0.4$ kPa. This can be observed by taking the dot product of the first three spatial eigenfunctions from each parameter configuration, as shown in Table IV. This illustrates that the oscillations from both configurations can be described as superpositions of nearly

TABLE III. Normalized eigenvalues/variances corresponding to eigenfunctions calculated for $E_c=0.4$ kPa and 0.5 kPa.

	$E_c=0.4$ kPa		$E_c=0.5$ kPa	
Eigenfunction number	λ_i (%)	Cumulative sum of λ_i (%)	λ_i (%)	Cumulative sum of λ_i (%)
1	43.9	43.9	45.6	45.6
2	30.9	74.8	27.0	72.6
3	16.1	91.0	12.5	85.1
4	7.1	98.1	5.2	90.3
5	0.7	98.8	3.1	93.4
6	0.5	99.3	1.6	95.0

TABLE IV. Dot product of first three spatial eigenfunctions of $E_c=0.4$ and 0.5 kPa.

Eigenfunction number	$\langle \phi_{i,EC=0.4} \mid \phi_{i,EC=0.5} \rangle$
1	0.913
2	0.883
3	0.901

the same underlying spatial eigenfunctions, even though the overall vibration patterns are quite distinct. Indeed, across the range of E_c where these bifurcations occur (0.3 kPa $< E_c <$ 0.57 kPa), the first three spatial eigenfunctions remain nearly constant.

The essential difference in the system at $E_c=0.4$ kPa and $E_c=0.5$ kPa is revealed by the temporal eigenfunctions, as illustrated in Fig. 7. The temporal eigenfunctions of $E_c=0.4$ kPa are entrained and nearly periodic, while the temporal eigenfunctions of $E_c=0.5$ kPa are nonperiodic and not entrained. Because the spatial eigenfunctions of the two configurations are nearly the same, it suggests that the essential difference between the two configurations is a desynchronization of the "modes." Thus in some cases, it is possible that the origin of chaos in vocal fold vibrations may be related to a desynchronization of a few of the low-order "modes" of vibration. Although this hypothesis needs further substantiation, it suggests an analogy to coupled oscillators (or circle maps), which are known to exhibit coexistence of limit cycles (and chaos) and, hence, sudden jumps and hysteresis. Such analogies to the theory of coupled oscillators merit further studies.

Furthermore, the fact that this simulation of partial differential equations (PDE's) can be projected onto just a few eigenfunctions is reminiscent of the findings of Saltzman (1962) and Lorenz (1963) in relation to Bernard convection. In their studies, it was found that close to the onset of convection there were only a few dominant modes, which led to the derivation of the celebrated Lorenz equations (1963). Whereas Lorenz employed a trigonometric expansion, in this study empirical eigenfunctions might be appropriate to reduce the original PDE's to a small set of ordinary differential equations (ODE's). Admittedly, such a reduction cannot be done in general because the spatial eigenfunctions may vary as parameter values are changed. However, over limited parameter regions where the spatial eigenfunctions remain more or less constant (such as in the example given above), a reduction is at least representative of the original PDE's. Such reductions may be useful for the design of lower-order models (which can nevertheless simulate various vocal qualities), and may help facilitate bifurcation analyses over specific parameter regions of the model (Deane *et al.*, 1991).

The first three spatial eigenfunctions of $E_c=0.5$ kPa are shown in Fig. 8. In this case, it may not be possible to claim a definite relationship between the empirical eigenfunctions and the normal modes of the simplified folds. Indeed, many factors (e.g., tissue incompressibility, complex geometry, nonlinearities) may cause significant deformations in the modes of vibrations. Nevertheless, the

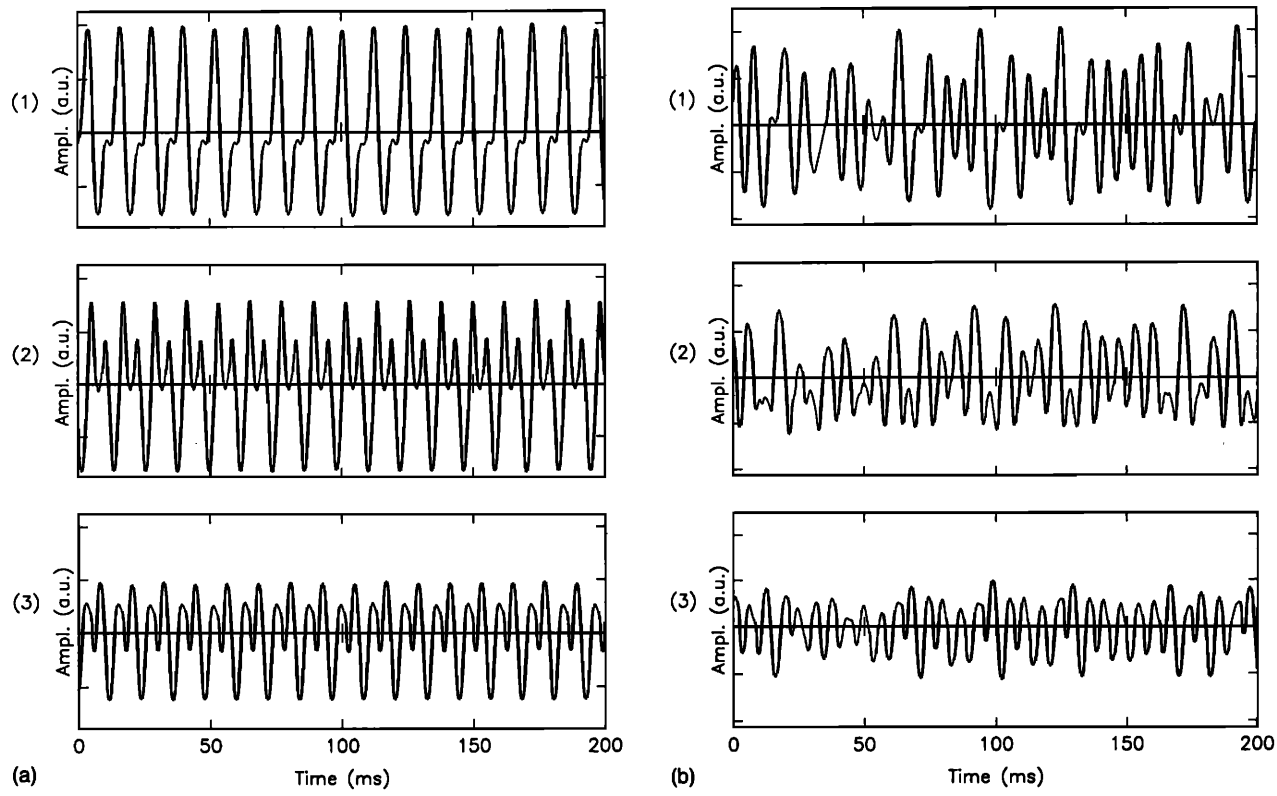

FIG. 7. The first three temporal eigenfunctions for E_c=0.4 kPa (a) and E_c=0.5 kPa (b).

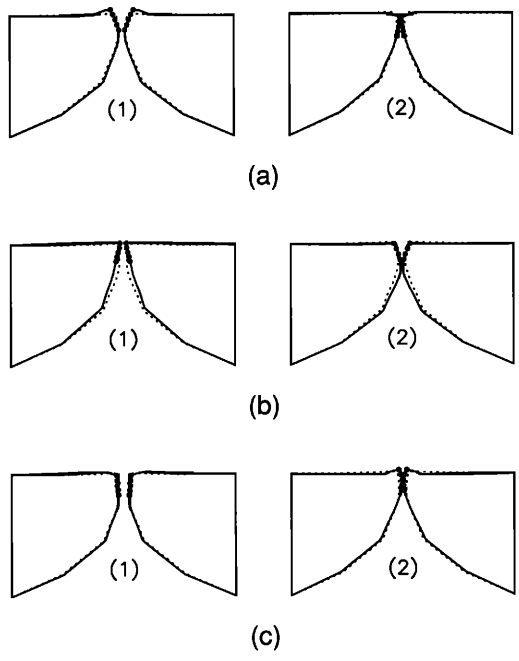

FIG. 8. Coronal views of the first three spatial eigenfunctions at E_c=0.5 kPa. Eigenfunction 1 shows some resemblance to a z-10 mode, eigenfunction 2 to an x-11 mode, and eigenfunction 3 to an x-10 mode. Coronal views are shown as in Fig. 4.

eigenfunctions appear to be manifestations of simple, low-order modes. For example, the first eigenfunction [Fig. 8(a)] shows some resemblance to a z-10 mode, the second eigenfunction [Fig. 8(b)] to an x-11 mode, and the third eigenfunction [Fig. 8(c)] to an x-10 mode.

In addition, the sixth eigenfunction for E_c=0.5 kPa is analogous to a higher-order normal mode (e.g., the x-21 mode). Although not as commonly observed, these higher-order modes have been viewed occasionally with high-speed cinematography (Rubin and Hirt, 1960). Figure 9 shows a superior view of this eigenfunction. This eigenfunction did not appear, at least as clearly, in the more stable oscillations corresponding to E_c=0.4 kPa and E_c=2 kPa. This may be related to the fact that this is an unstable eigenfunction, and is thus usually only excited during more unstable, nonperiodic vibrations. Even in the complex oscillations from which this eigenfunction was extracted, the higher-order eigenfunction was so weak that it could not be visually detected in the overall vibration pattern.

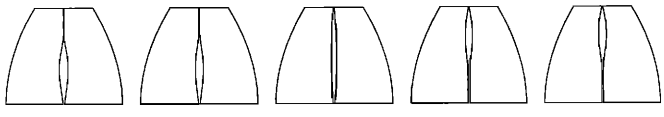

FIG. 9. A superior view of the sixth spatial eigenfunction for E_c=0.5 kPa. This eigenfunction is similar to an x-21 mode. To illustrate this eigenfunction, a series of five sequential snapshots are shown from left to right.

IV. CONCLUSIONS

As is well known from high-speed films, stroboscopy, and sophisticated models, vocal fold vibrations exhibit complex three-dimensional patterns. However, normal phonation produces fairly periodic acoustic output. These observations may be explained by the fact that only a few modes are excited and all the modes are entrained. This concept has been substantiated through examining the empirical eigenfunctions of a biomechanical simulation of vocal fold vibrations during self-oscillation. Even though hundreds of degrees of freedom exist, two eigenfunctions explain 98% of the variance of the nodal trajectories. By viewing the finite element simulation as a superposition of just two dominant eigenfunctions, an interpretation of the mechanism of self-oscillation of the folds is facilitated. Furthermore, the calculated eigenfunctions are qualitatively similar to low-order normal modes predicted in earlier theoretical work.

The technique of empirical eigenfunctions is also useful for describing irregular oscillations related to rough voice. Changing parameters of the biomechanical model leads to subharmonic regimes and chaos. However, despite complex motion, a relatively small number of eigenfunctions captures the essential dynamics of the folds. In addition, some of the more subtle dynamics captured by "weaker" eigenfunctions correspond with higher-order normal modes.

Over a specified parameter region (related to a lax cover), the spatial eigenfunctions remained nearly constant, while the presence of chaos was manifested by a desynchronization of the temporal eigenfunctions, suggesting that in some instances the appearance of chaos in vocal fold oscillations may be related to the desynchronization of a few of the low-order "modes" of vibration. Although not fully substantiated, this hypothesis suggests another analogy between vocal fold vibrations and coupled oscillators which merits further investigation.

In general, the method of empirical eigenfunctions enhances the study of vocal fold dynamics by allowing the principal modes of vibration to be extracted during self-oscillation, despite inherent nonlinearities.

ACKNOWLEDGMENTS

This work was supported by Grant No. P60 00976 from the National Institute on Deafness and Other Communication Disorders. The authors also thank I. G. Kevrekidis for introducing them to the method of empirical eigenfunctions.

Alipour-Haghighi, F., and Titze, I. R. (**1985a**). "Simulation of particle trajectories of vocal fold tissue during phonation," in *Vocal Fold Physiology: Biomechanics, Acoustics, and Phonatory Control,* edited by I. R. Titze and R. C. Scherer (Denver Center for the Performing Arts, Denver, CO).

Alipour-Haghighi, F., and Titze, I. R. (**1985b**). "Viscoelastic modelling of canine vocalis muscle in relaxation," J. Acoust. Soc. Am. **78**, 1939–1943.

Alipour-Haghighi, F., and Titze, I. R. (**1991**). "Elastic models of vocal fold tissues," J. Acoust. Soc. Am. **90**, 1326–1331.

Armbruster, D., Heiland, R., Kostelich, E. J., and Nicolaenko, B. (**1992**). "Phase-space analysis of bursting behavior in Kolmogorov flow," Physica D **58**, 392–401.

Aubry, N., Guyonnet, R., and Lima, R. (**1991**). "Spatiotemporal analysis of complex signals: theory and applications," J. Stat. Phys. **64**, 683–739.

Baer, T. (**1981**). "Investigation of the phonatory mechanism," ASHA Rep. (11), 38–46.

Baken, R. J. (**1991**). "Géométrie fractale et évaluation de la voix: Application préliminaire á la dysphonie," Bull. d'Audiophonogie. Ann. Sc. Univ. Franche-Comté 7(5–6), 731–749.

Breuer, K. S., and Sirovich, L. (**1991**). "The use of the Karhunen-Loève procedure for calculation of linear eigenfunctions," J. Comput. Phys. **96**, 277–296.

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).

Deane, A. E., Kevrekidis, I. G., Karniadakis, G. E., and Orszag, S. A. (**1991**). "Low-dimensional models for complex geometry flows: application to grooved channels and circular cylinders," Phys. Fluids A **3**, 2337–2354.

Flanagan, J. L., and Landgraf, L. (**1968**). Self-oscillating source for vocal tract synthesizers," IEEE Trans. Audio Electroacoust. **AU-16**, 57–64.

Fukunaga, K. (**1972**). *Introduction to Statistical Pattern Recognition* (Academic, New York).

Golub, G. H., and Van Loan, C. F. (**1983**). *Matrix Computations* (Johns Hopkins U.P., Baltimore, MD).

Herzel, H., Steinecke, I., Mende, W. and Wermke, K. (**1991**). "Chaos and bifurcations during voiced speech," in *Complexity, Chaos, and Biological Evolution,* edited by E. Mosekilde (Plenum, New York), pp. 41–50.

Herzel, H., and Wendler, J. (**1991**). "Evidence of chaos in phonatory samples," in *Proceedings of EUROSPEECH 91* (European Speech Communication Association, Genova, Italy).

Herzel, H., Berry, D. A., Titze, I. R., and Saleh, M. (**in press**). "Analysis of vocal disorders with methods from nonlinear dynamics," J. Speech. Hear. Res.

Hirano, M. (**1975**). "Phonosurgery: Basic and Clinical Investigations," Official Rep. 76th Annual Convention Oto-Rhino-Laryngo. Soc., Kurume, Japan.

Hollien, H., and Michel, J. (**1968**). Vocal fry as a phonational register. J. Speech. Hear. Res. **11**, 600–604.

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**, 1233–1267.

Johnson, R. A., and Wichern, D. W. (**1982**). *Applied Multivariate Statistical Analysis* (Prentice-Hall, Englewood Cliffs, NJ).

Lauterborn, W. (**1986**). "Acoustic turbulence," in *Frontiers in Physical Acoustics,* edited by D. Sette, (North-Holland, Amsterdam), Italian Physical Society, pp. 123–144.

Lorenz, E. N. (**1956**). Tech. Rep. 1, Statistical Forecasting Program, Department of Meteorology, Massachusetts Institute of Technology.

Lorenz, E. N. (**1963**). "Deterministic nonperiodic flow," J. Atmos. Sci. **20**, 130–141.

Lumley, J. L. (**1967**). "The structure of inhomogeneous turbulent flows," in *Atmosphere Turbulence and Radio Wave Propagation,* edited by A. M. Yaglom and V. I. Tatarski (Nauka, Moscow), pp. 166–178.

Mende, W., Herzel, H. and Wermke, K. (**1990**). "Bifurcations and chaos in newborn cries," Phys. Lett. A **145**, 418–424.

Moore, G., and Von Leden, H. (**1958**). "Dynamic variations of the vibratory pattern in the normal larynx," Folia Phoniatr. **10**, 205–238.

Rubin, H. J., and Hirt, C. C. (**1960**). "The falsetto: a high-speed cinematographic study," Laryngoscope **70**, 1305–1324.

Saito, S., Fukuda, H., Isogai, Y., and Ono, H. (**1981**). "X-ray stroboscopy," in *Vocal Fold Physiology,* edited by K. Stevens (University of Tokyo Press, Tokyo), pp. 95–106.

Saito, S., Fukuda, H., Kitahira, S., Isogai, Y., Tsuzuki, T., Muta, H., Takyama, E., Fujika, T., Kokawa, N., and Makino, K. (**1985**). "Pellet tracking in the vocal fold while phonating: experimental study using canine larynges with muscle activity," in *Vocal Fold Physiology,* edited by I. R. Titze and R. C. Scherer (Denver Center for the Performing Arts, Denver, CO).

Saltzman, B. (**1962**). "Finite amplitude free convection as an initial value problem," J. Atmos. Sci. **19**, 329–341.

Scherer, R. (**1989**). "Physiology of creaky voice and vocal fry," J. Acoust. Soc. Am. Suppl. 1 **86**, S25.
Sirovich, L. (**1987**). "Turbulence and the dynamics of coherent structures," Parts I–III, Q. Appl. Math. XLV, 561–590.
Stevens, K. N. (**1977**). "Physics of laryngeal behaviour and larynx modes," Phonetica **34**, 264–279.
Titze, I. R., and Strong, W. (**1975**). "Normal modes in vocal fold tissues," J. Acoust. Soc. Am. **57**, 736–744.
Titze, I. R. (**1976**). "On the mechanics of vocal-fold vibration," J. Acoust. Soc. Am. **60**, 1366–1380.
Titze, I. R., and Talkin, D. T. (**1979**). "A theoretical study of the effects of various laryngeal configurations on the acoustics of phonation," J. Acoust. Soc. Am. **66**, 60–74.
Titze, I. R. (**1988**). "The physics of small-amplitude oscillation of the vocal folds," J. Acoust. Soc. Am. **83**, 1536–1552.
Titze, I. R., Baken, R., and Herzel, H. (**1993**). "Evidence of chaos in vocal fold vibration. in *Vocal Fold Physiology: Frontiers in Basic Science*, edited by I. Titze (Singular Group, San Diego, CA).
Vautard, R., Yiou, P., and Ghil, M. (**1992**). "Singular-spectrum analysis: a toolkit for short, noisy chaotic signals," Physica D **58**, 95–126.
Zahorian, S. A., and Rothenberg, M. (**1981**). "Principal-components analysis for low-redundancy encoding of speech spectra," J. Acoust. Soc. Am. **69**, 832–845.