

Modeling the role of nonhuman vocal membranes in phonation

Patrick Mergell^{a)}

Department of Phoniatrics and Pedaudiology, University Erlangen-Nuremberg, Bohlenplatz 21,
D-91054 Erlangen, Germany

W. Tecumseh Fitch^{b)}

Harvard/MIT Speech and Hearing Sciences, Room 1036 William James Hall, 33 Kirkland Street,
Cambridge, Massachusetts 02138

Hanspeter Herzel^{c)}

Institute for Theoretical Biology, Humboldt University Berlin, Invalidenstr. 43, D-10115, Berlin, Germany

(Received 17 March 1998; accepted for publication 6 November 1998)

Although the mammalian larynx exhibits little structural variation compared to sound-producing organs in other taxa (birds or insects), there are some morphological features which could lead to significant differences in acoustic functioning, such as air sacs and vocal membranes. The vocal membrane (or “vocal lip”) is a thin upward extension of the vocal fold that is present in many bat and primate species. The vocal membrane was modeled as an additional geometrical element in a two-mass model of the larynx. It was found that vocal membranes of an optimal angle and length can substantially lower the subglottal pressure at which phonation is supported, thus increasing vocal efficiency, and that this effect is most pronounced at high frequencies. The implications of this finding are discussed for animals such as bats and primates which are able to produce loud, high-pitched calls. Modeling efforts such as this provide guidance for future empirical investigations of vocal membrane structure and function, can provide insight into the mechanisms of animal communication, and could potentially lead to better understanding of human clinical disorders such as sulcus vocalis. © 1999 Acoustical Society of America. [S0001-4966(99)03602-4]

PACS numbers: 43.80.Ka, 43.70.Aj, 43.64.Bt [FD]

INTRODUCTION

In contrast to the sound-producing organs in other taxa (e.g., insects or birds), the mammalian larynx has a conservative evolutionary history, and shows little qualitative variation from species to species (Negus, 1949; Harrison, 1995; Schön-Ybarra, 1995; Fitch and Hauser, 1995). The gross anatomy (cartilages, muscles, nerve supply) and physiological function of the mammalian larynx appear to be highly constrained by the necessities of airway protection and respiratory function. Despite this, there are a number of interesting differences among mammals in laryngeal anatomy, and the precise effects of such differences on the acoustic output of the larynx remain unknown. For example, the histological fine structure of the vocal folds, particularly the number and distribution of tissue layers, varies from species to species in ways that may have an effect on the vibratory behavior of the fold (Kurita *et al.*, 1983; Hirano, 1991). In nonhuman mammals, two gross morphological features are particularly prominent: air sacs and vocal membranes. A variety of air sacs is seen in a diverse array of mammals; they can be paired or unpaired, large or small, and sub- or supra-glottal (Negus, 1949, gives a review). The precise acoustic function of air sacs is currently unknown (reviewed in Schön-Ybarra, 1995; Gautier, 1971; Fitch and Hauser, 1995).

In this study we focus on the vocal membranes, which are thin upward extensions of the membranous portion of the vocal folds (see Fig. 1). They are variously referred to as “vocal membranes” in bats (Griffin, 1958; Griffiths, 1978; Suthers, 1988; Suthers and Fattu, 1973), “sharp-edged vocal folds” (Hill, 1957; Hill and Booth, 1957; Hast, 1989), or “vocal lips” in primates (Schön-Ybarra, 1995; Brown and Cannito, 1995; Fitch and Hauser, 1995). Here we adopt the term vocal membrane because it is most descriptive, and because vocal lip has an unrelated anatomical meaning in human laryngeal anatomy (Thomas, 1989). They are formed of connective tissue and lack muscle fibers, and can be as thin as a few microns across in bats (Suthers and Fattu, 1973; Suthers, 1988). They are found in many microchiropteran bats and a variety of primate species, including most New World (platyrrhine) monkeys, some Old World (catarrhine) monkeys, including some common laboratory species, and in our closest primate relatives, the apes (Griffin, 1958; Schön-Ybarra, 1995; Starck and Schneider, 1960). We are not aware of a taxonomic review of vocal membranes for all mammals, but they are found in members of the genus *Felis* (the smaller members of the cat family; Hast, 1989), and we have observed them in both llamas and young pigs (W. T. Fitch, personal observation).

Both bat and primate-vocal membranes have been postulated to serve as independent low-mass oscillators (e.g., Griffin, 1958; Griffiths, 1978; Schön-Ybarra, 1995; Fitch and Hauser, 1995), and thus to support the extremely high-

^{a)}Electronic mail: mergell@phoni.med.uni-erlangen.de

^{b)}Electronic mail: tec@wjh.harvard.edu

^{c)}Electronic mail: h.herzel@biologie.hu-berlin.de

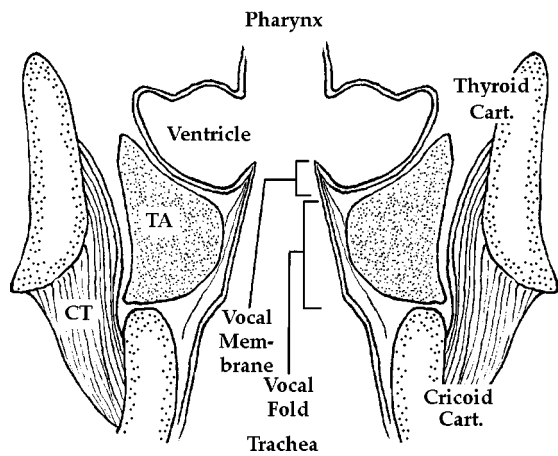


FIG. 1. Schematic coronal section through middle of larynx showing anatomy of the vocal membrane. The membrane projects upward (rostrally) from main body of vocal fold. CT: Cricothyroid muscle. TA: Thyroarytenoid (“vocalis”) muscle.

frequency (sometimes ultrasonic) calls found in many primate species and most echolocating bats. In primates, the vocal lips have been suggested to underlie vibration at multiple frequencies, or biphonation (Kelemen, 1949, 1969; Schön-Ybarra, 1995; Brown and Cannito, 1995). It has been suggested that the nonhuman primate larynx has a broader range but less control of fundamental frequency than the human larynx (Lieberman, 1968; Pressman and Kelemen, 1969), and that this constraint has led to a different pattern of evolution in primate communication (Schön-Ybarra, 1995). However, detailed investigations of the properties and functions of the vocal membranes are lacking, and at present these suggestions remain speculative.

In addition to providing information relevant to animal communication, investigations of the vocal membrane have potential clinical significance. The addition of masses (tumors, polyps, nodules) and/or additional surfaces or flaps (sulcus vocalis, Reinke’s edema) to the vocal folds represent common disorders in the voice clinic. A better understanding of the acoustic manifestations of such features could pave the way for enhanced voice diagnostics based on the acoustic signal. Detailed theoretical and experimental investigations of the vocal membranes could prove extremely useful for such an endeavor (especially since the range of possible experimental manipulations is much greater in animals than in human subjects).

I. POSSIBLE FUNCTIONS OF THE VOCAL MEMBRANES

A number of hypotheses have been advanced as to the function of the vocal membranes. Many researchers have suggested that the stiff, lightweight membranes could vibrate relatively independently from the rest of the vocal fold, and thus result in the production of extremely high-frequency calls, thus giving a wider pitch range (Griffin, 1958; Starck and Schneider, 1960; Schön-Ybarra, 1995; Fitch and Hauser, 1995). We will refer to this as the “increased pitch range” hypothesis.

Another suggestion is that the vocal membranes could result in greater vocal efficiency: the “increased efficiency”

hypothesis. For example, Schön-Ybarra (1995) suggested that the decreased damping that presumably accompanies vocal membranes could result in a decrease in phonation-threshold pressure. Phonation threshold pressure P_s^{th} is an extremely useful parameter for measuring phonatory efficiency. It is defined as the minimum pressure required for self-sustained vocal-fold oscillations. Model calculations by Titze (1988) and Lucero (1993, 1995, 1996) revealed an inverse proportionality between subglottal-threshold pressure and the vertical depth of the vocal folds. Thus, one can expect that vertical extensions of the vocal folds lower the phonation threshold, which is equivalent to an increased phonatory efficiency.

Within the framework of nonlinear dynamics, the phonation threshold can be considered as a supercritical Hopf bifurcation: a sudden change of vocal-fold dynamics from a fixed point (damped oscillation) to a limit-cycle attractor (sustained oscillation) at a critical subglottal pressure. Once the threshold-pressure value is surpassed, the oscillation amplitude increases monotonically with pressure. Hence, lower phonation-threshold pressures typically result in higher oscillation amplitudes for a given lung pressure (Mergell *et al.*, 1998).

A final possible effect of the vocal membranes would be the production of subharmonic or chaotic vibration. Recent research suggests that many broadband, complex acoustic signals produced by animals result from chaotic dynamics in the vocal production mechanism (Wilden *et al.*, 1998). Such dynamics could result from coupling between various oscillatory components within the source, e.g., interactions between the right and left vocal folds, between upper and lower portions of the same fold, or between the air column in the vocal tract and the tissue of the vocal fold, and from turbulence generated at constrictions in the vocal tract. The “chaos hypothesis” holds that one function of the vocal membranes is to support the production of broadband chaotic signals via increased oscillatory coupling. The production of “noisy” signals is held to be intimidating and could be useful in agonistic contexts (Morton, 1977). Alternatively, broadband sounds can provide an accurate outline of the vocal-tract transfer function of the caller, which in turn could carry information about body size, oral configuration, or other features (Fitch, 1997).

II. MODEL

The model explored here is a modification of the original two-mass model by Ishizaka and Flanagan (1972). In this model, the vocal folds are represented by damped mass-spring pairs forming a rectangular-shaped glottal valve. Steinicke and Herzel (1995) simplified this model in order to focus on the basic dynamical features of the larynx: vocal-tract resonances, viscous losses, and nonlinear restoring forces are neglected in this first approximation. For the calculation of the driving aerodynamical force, laminar flow exists below the narrowest constriction in the glottis and a jet separates above that point, giving a pressure drop to atmospheric pressure (gauged to zero pressure). Although more complex models which simulate the precise physiology of

the vocal folds have been created (Titze and Talkin, 1979; Alipour-Haghighi and Titze, 1985), they are computationally expensive, and recent analyses (e.g., Berry *et al.*, 1994) suggest that normal vocal-fold vibration is dominated by the first two low-vibrational modes anyway. These two basic vibration modes are the in-phase and out-of-phase oscillation in the vertical (caudal–cranial) direction; their superposition corresponds to 98% of the variance of all excited vibrational modes found in a highly complex and realistic three-dimensional continuum model (Berry *et al.*, 1994). This suggests that simple low-dimensional models are appropriate for exploratory investigations like ours.

Guided by the anatomical model proposed by Hirano (1974), which subdivides the vocal-fold tissue into a muscle and ligament “body” and a mucosal “cover,” a variety of representations of the vocal membrane could be chosen. The membrane could move independently of the upper mass, or vary its angle of attachment dynamically. Obviously, modeling the vocal membrane as a third low-mass, high-tension oscillator would allow the production of arbitrarily high frequencies; calculations to this effect are given in Griffin (1958). Unfortunately, the critical data on which to base these calculations (the tension on the membranes and the stress–strain relationships within the vocal membrane) are currently unavailable. In the present study, our goal was to build the simplest model which would have some physiological validity, be compatible with anatomical knowledge and still be likely to give appreciable effects, without speculating beyond the available (basically geometric) data. Thus, to the basic two-mass model we added the simplest geometric representation of a vocal membrane: an upward extension of the upper mass, attached to that mass at a fixed angle.

Specifically, we modified the two-mass model of Steinecke and Herzel (1995) by adding a massless rigid reed connected to the upper mass with an angle θ . The vertical depth of the vocal membranes is given by d_3 . Figure 2(a) shows a 3D representation of this model, and Fig. 2(b) gives the corresponding frontal section. The horizontal (dorsal–ventral) length and the vertical depth of the masses are given by l , d_1 , and d_2 , respectively. During vocalization, the areas at the level of the lower-mass pair a_1 , upper-mass pair a_2 , and at the glottal outlet a_{v1} limited by the upper edges of the vocal membranes vary, and interact with aerodynamical forces. The glottal rest areas $a_{0i} = 2lx_{0i}$, where the x_{0i} define the distances of the masses to the glottal symmetry plane prior to oscillation onset [the index (i) distinguishes the lower ($i=1$) and upper ($i=2$) mass pairs]. Besides the parameters introduced above, the set of parameters comprises masses m_i , stiffness coefficients k_i and c_i , damping coefficients r_i , a coupling stiffness k_c , the air density ρ at body temperature, and the subglottal pressure P_s generated by the respiratory system. Since the current model is intended as a general model of the vocal membranes in any species, we did not attempt to scale the model dimensions or parameters to those of (for example) a bat or chimpanzee. Instead, we simply used the “standard parameter set” listed in Table I. (All units are expressed in cm, g, ms, and respective combinations). These parameters were originally developed for the human larynx and have been extensively tested. Although it

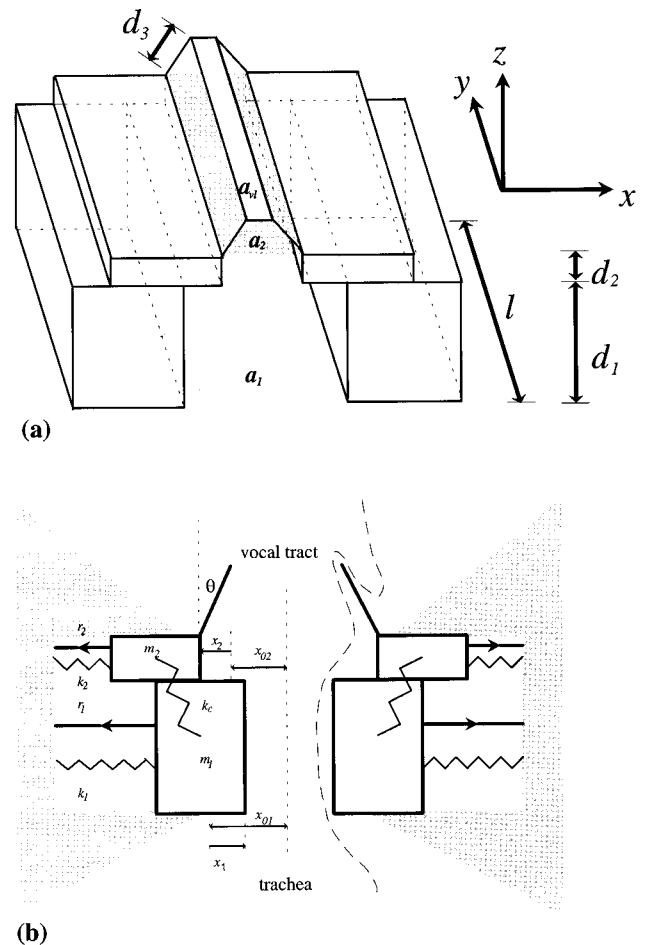


FIG. 2. (a) 3D representation of the two-mass model with vocal membranes. (b) Frontal section of the model. For further explanations, see the text.

would clearly be more appropriate to use values for an organism which has vocal membranes, few if any of the required parameters are available for such species. Thus, it seems safer at present to use parameter values whose effects on phonation are well understood, and whose validity has been subjected to extensive empirical tests. One practical effect of this decision is that the oscillation frequencies observed in this study will not go into the ultrasonic range (as would be the case if we specifically modeled the much smaller vocal folds of bats or small rodents).

TABLE I. Standard parameters of the employed model.

| m_1 | m_2 | k_1 | k_2 | k_c | r_1 | r_2 | c_1 |
|--------|----------|----------|-------|-------|-------|----------|--------|
| 0.125 | 0.025 | 0.08 | 0.008 | 0.025 | 0.02 | 0.02 | $3k_1$ |
| c_2 | a_{01} | a_{02} | d_1 | d_2 | l | ρ | P_s |
| $3k_2$ | 0.05 | 0.05 | 0.25 | 0.05 | 1.4 | 0.001 13 | 0.008 |

Using matrix notation, the motion of the masses can be written as follows:

$$\frac{d}{dt} \begin{pmatrix} x_1 \\ v_1 \\ x_2 \\ v_2 \end{pmatrix} = \begin{pmatrix} 0 & 1 & 0 & 0 \\ -\frac{k_1+k_c}{m_1} & -\frac{r_1}{m_1} & \frac{k_c}{m_1} & 0 \\ 0 & 0 & 0 & 1 \\ \frac{k_c}{m_2} & 0 & -\frac{k_2+k_c}{m_2} & -\frac{r_2}{m_2} \end{pmatrix} \times \begin{pmatrix} x_1 \\ v_1 \\ x_2 \\ v_2 \end{pmatrix} + \begin{pmatrix} 0 \\ I_1(x_1) + F_1(x_1, x_2) \\ 0 \\ I_2(x_2) + F_2(x_1, x_2) \end{pmatrix}. \quad (1)$$

This is a system of four coupled ordinary differential equations. The coordinates x_i are the oscillation amplitudes of the masses and v_i are the corresponding velocities. The right-most term of Eq. (1) contains the nonlinearities, i.e., the driving forces and the vocal-fold contact terms given by

$$I_i(x_i) = -\Theta(-a_i) \frac{c_i}{m_i} \frac{a_i}{2l}, \quad (2)$$

where $\Theta(x) = 1$ for $x > 0$ and $\Theta(x) = 0$ for $x \leq 0$. At vocal-fold collision ($a_i \leq 0$), an additional restoring force controlled by the stiffness coefficients c_i becomes active via the Θ -step function.

We must consider nine different glottal configurations (see Fig. 3) to deduce the aerodynamic forces. We assume jet separation at the minimum glottal cross section, and laminar airflow obeying the Bernoulli law below that point (Story and Titze, 1995; Steinecke and Herzel, 1995). According to Bernoulli's law, total pressure, which is equal to static pressure plus the kinetic energy density, is invariant along a streamline for steady, inviscid flow. Above the jet-separation point, the static pressure drops to zero (atmospheric pressure) and the hydrodynamic energy is carried completely by the airflow. Thus, above the separation point, we write

$$P_s = P(z) + \frac{\rho}{2} \left(\frac{U}{a(z)} \right)^2 = \frac{\rho}{2} \left(\frac{U}{a_{\min}} \right)^2, \quad (3)$$

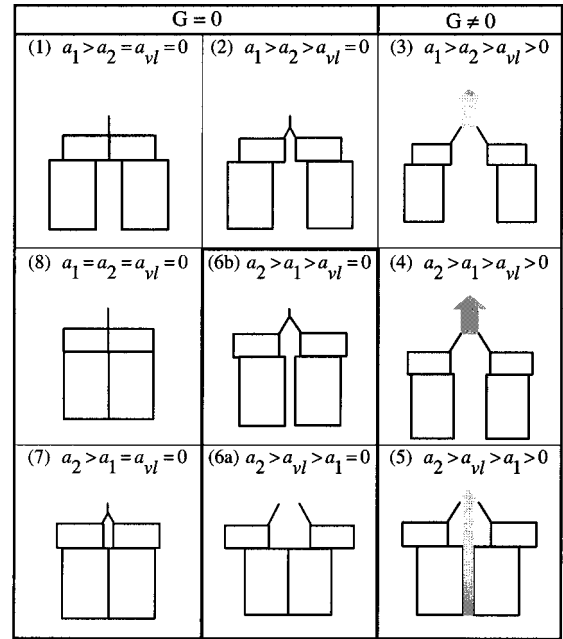


FIG. 3. Glottal configurations during one oscillation cycle. A typical cycle corresponds to clock-wise succession in this diagram: (1) $a_1 > 0 \geq a_2 \geq a_{vl} \rightarrow$ (2) $a_1 > a_2 > 0 \geq a_{vl} \rightarrow$ (3) $a_1 > a_2 > a_{vl} > 0 \rightarrow$ (4) $a_2 > a_1 > a_{vl} > 0 \rightarrow$ (5) $a_2 > a_{vl} > a_1 > 0 \rightarrow$ (6a) $a_2 > a_{vl} > 0 \geq a_1$ or (6b) $a_2 > a_1 > 0 \geq a_{vl} \rightarrow$ (7) $a_2 > 0 \geq a_1 \geq a_{vl} \rightarrow$ (8) $0 \geq a_1 \geq a_2 \geq a_{vl}$. The first two columns show configurations with closed glottis and zero air flow ($U=0$). The jet is indicated by a gray arrow ($U>0$).

where $P(z)$ is the static pressure, U is the airflow, $a(z)$ corresponds to the z -dependent glottal area, and $a_{\min} = \min(a_1, a_2, a_{vl})$. The driving force is a direct consequence of the static pressure

$$dF(z) = P(z) l dz, \quad (4)$$

where ldz is an infinitesimally small segment of the inner vocal-fold surface. Integration over the inner-surface portions yields the aerodynamic-force terms considering the corresponding glottal configurations,

$$F_1(x_1, x_2) = \begin{cases} \frac{ld_1 P_s}{m_1}, & [1,2,6b]. \\ \frac{ld_1 P_s}{m_1} \left[1 - \left(\frac{a_{vl}}{a_1} \right)^2 \right], & [3,4]. \\ 0, & [5,6a,7,8]. \end{cases} \quad (5)$$

$$F_2(x_1, x_2) = \begin{cases} \frac{P_s}{m_2} \left(ld_2 + \frac{1}{2} a_2 \cot \theta \right), & [2,6b]. \\ \frac{ld_2 P_s}{m_2} \left[1 + \frac{d_3}{d_2} \cos \theta - \left(\frac{a_{vm}}{a_2} \right)^2 \left(1 + \frac{a_2}{2 \tan \theta ld_2} \left\{ \frac{1}{1 - \frac{2ld_3 \sin \theta}{a_2}} - 1 \right\} \right) \right], & [3,4,5]. \\ 0, & [1,6a,7,8], \end{cases}$$

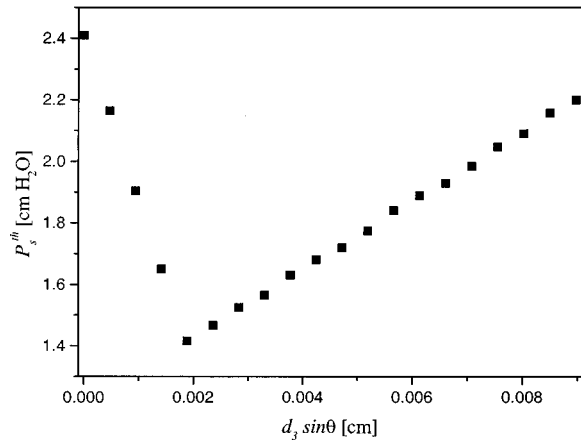


FIG. 4. Hopf bifurcation diagram showing the dependence of the subglottal-threshold pressure P_s^{th} on the vocal membrane overhang $d_3 \sin \theta$. Around $d_3 \sin \theta = 0.0019$ cm, the phonation-threshold pressure exhibits a minimum indicating an optimal membrane geometry ($d_3 \cos \theta = 0.1$ cm, $P_s = 8$ cm H₂O).

taking into account that at collision, the respective aerodynamic driving force drops to zero (Fig. 3, cases [1,6a,7,8]). In the limit $d_3 \rightarrow 0$ (not a straightforward limit case), Eq. (5) reduces to

$$F_1 = l d_1 P_s \left[1 - \left(\frac{U}{a_{\text{min}}} \right)^2 \Theta(a_{\text{min}}) \right] \Theta(a_1),$$

$$F_2 = 0, \quad a_{\text{vl}} = a_2, \quad (6)$$

which corresponds to the force equations in the two-mass model without vocal membranes (Steinecke and Herzel, 1995). Thus, the vocal membranes couple the aerodynamic force directly to the upper-mass pair, whereas in the two-mass model without vocal membranes, the upper masses are driven only indirectly via mechanical coupling to the lower masses.

We explored the dynamic behavior of this model using discrete-time numerical integration of the differential equations described above (using Runge–Kutta). To present the results, we make use of the phonation-threshold concept described above, so we observe the changes in phonation threshold that depend upon variation in some other variable, such as frequency or the angle of the vocal membranes. In nonlinear dynamical terms, these threshold plots are formally identical to bifurcation diagrams. To vary frequency, we used a parameter q to scale the natural frequencies of the model masses $f_i^{\text{nat}} = q f_i^{\text{nat}}$, where $f_i^{\text{nat}} = \sqrt{k_i/m_i}$ are the natural frequencies for standard parameters given in Table I (see Mergell and Herzel, 1997).

III. RESULTS

First, we address the increased efficiency hypothesis. Varying the parameters d_3 and θ , we found that the overhang $d_3 \sin \theta$ is the crucial measure controlling the phonatory effect of the vocal membrane. At fixed vertical projection of the vocal membranes, i.e., $d_3 \cos \theta = 0.1$ cm, we analyzed the behavior of the threshold pressure when changing the overhang (Fig. 4). The resulting dependence of the threshold

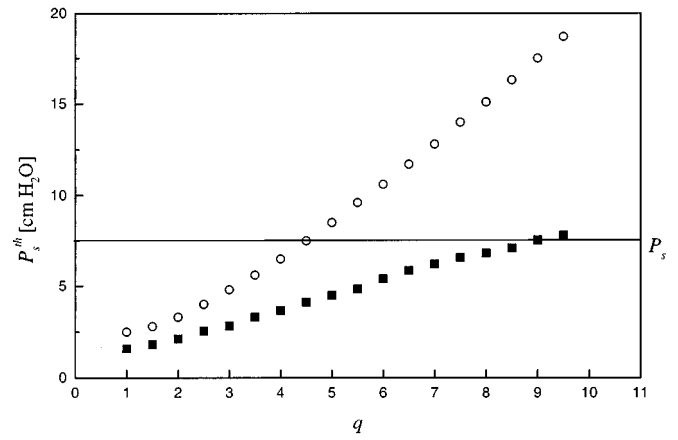


FIG. 5. Frequency-dependent Hopf bifurcation of the two-mass model with (filled squares) and without (circles) vocal membranes. The curves separate the graph plane into phonation and aponia region. Considering a constant lung pressure P_s , the frequency range of the extended model is clearly increased relative to the range of the unmodified model ($d_3 = 0.1$ cm, $\theta = 1.8^\circ$).

pressure on the overhang indicates that there is an optimal membrane geometry: at $d_3 \sin \theta \approx 0.0019$ cm, the threshold curve exhibits a minimum ($d_3 \approx 0.1$ cm, $\theta \approx 1.8^\circ$). These results indicate that for any set of parameters, there will be an optimum angle of attachment for the membranes. In general, this angle will be very small so the optimum orientation of the membranes will be nearly vertical. Hence, the increased vertical dimension does not solely explain the increased phonatory efficiency. It must be adjusted to the inclination angle of the vertical extensions or vice versa. The precise optimum value of θ will depend on a number of other parameters (e.g., actual fold dimensions, natural frequencies of the masses, etc.) which show considerable biological variation. Lacking data to precisely specify these parameters for a given nonhuman species, we will focus on qualitative results, fixing the length and angle of theta to values which seem anatomically reasonable based upon cross-sections of primate larynges (see Fig. 1).

Figure 5 plots phonation-threshold pressure vs natural frequency with and without vocal membranes. These curves separate the q – P_s^{th} -plane into phonation and nonphonating regions for the respective model configurations. The upper boundary of the phonation region is determined by the maximum pressure which can be produced by the respiratory muscles. Assuming a frequency-independent maximum pressure, the maximum achievable frequency is determined by the intersection of the threshold curve with this constant upper boundary. This plot thus defines a total frequency range. In clinical practice, similar voice-range profile diagrams are very useful diagnostics of vocal functioning (Wendler *et al.*, 1996).

At low fundamental frequencies, the effect of the vocal membranes on the threshold pressure is moderate, while at high frequencies, the increase in the phonation region is pronounced. Figure 5 shows that the vocal membrane effects an overall decreased threshold pressure in the investigated frequency range. Assuming constant lung pressure P_s , we find that the intersection point with the threshold curve of the

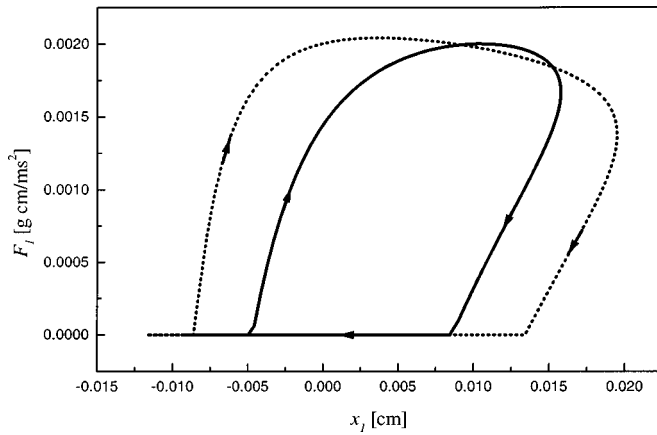


FIG. 6. Energy-balance diagram of the two-mass model with (dotted line) and without (solid line) vocal membranes for $q=3$, $P_s=8$ cm H₂O, $d_3=0.1$ cm, and $\theta=1.8^\circ$. The area enclosed by the F_i-x_i curve which corresponds to the energy pumped into the lower mass during one glottal cycle is increased with addition of vocal membranes. Due to a higher glottal efficiency, the oscillation amplitudes are increased.

model with vocal membrane occurs at a much higher frequency than is the case for the model without the extension. Consequently, the phonation region, and the pitch range, are increased by the vocal membranes. Thus, the related increased pitch range and the increased efficiency hypotheses are both supported by the results of these simulations.

To better illustrate the effects of the vocal membranes on phonation efficiency, Fig. 6 shows how much energy E_1^{in} is pumped from the airflow into one of the lower masses during one glottal cycle. It corresponds to the area which is enclosed by the force-amplitude curve (solid line: model without vocal membranes; dotted line: model with vocal membranes).

$$E_i^{\text{in}} = \oint F_i dx_i = \int_0^T F_i v_i dt, \quad (7)$$

where T is the duration of the oscillation period and the index i again distinguishes the lower ($i=1$) and upper ($i=2$) mass pairs. Equation (7) helps to understand the condition which has to be fulfilled for sustained phonation (see also Titze, 1988). Assuming a time-invariant force, it is clear that $E_i^{\text{in}}=0$ because the velocity v_i oscillates in time and thus contributes to the integration with a positive sign in the same amount as with a negative sign. Consequently, oscillation cannot be supported by a static driving force. A dynamic driving force with an opposite sign compared to the velocity v_1 results in a negative integral: the oscillator is damped by the external force. Clearly, the condition for sustained phonation is a positive integral which can only be obtained with a time-varying driving force which agrees in sign with the oscillator velocity. The optimal situation for phonation is given when the driving force is large during glottal opening (pushing the masses apart) and small (or negative) during closing (not resisting inward movement) (Stevens, 1977).

Since the oscillation amplitude reaches a saturation value after phonation onset, an energy equilibrium is obtained, i.e., all energy pumped into the masses is dissipated. The dissipated energy reads

$$E_i^{\text{out}} = r_i \oint v_i dx_i = r_i \int_0^T v_i^2 dt, \quad (8)$$

where the energy-equilibrium relation can be written as

$$\sum_i E_i^{\text{in}} + E_i^{\text{out}} = 0, \quad (9)$$

The lung is considered to be the energy source providing

$$E_s = P_s \int_0^T U dt = P_s V, \quad (10)$$

where V is the air volume leaving the lung during one glottal cycle. We can now define the glottal efficiency η

$$\eta = 2 \frac{E_1^{\text{in}} + E_2^{\text{in}}}{E_s}, \quad (11)$$

which is the ratio between the energy pumped into the system and the compression work which is done by the respiration muscles (the factor 2 takes into account that energy is transferred into mass pairs). For the parameters $q=3.0$, $\theta=1.8^\circ$, $d_3=0.1$ cm, and $P_s=8$ cm H₂O, we find in the model with vocal membranes an efficiency $\eta_{\text{vm}}=2.82\%$, which is double that of the unmodified model $\eta=1.36\%$. Since the area surrounded by the dotted curve in Fig. 6 is enlarged due to higher peak amplitudes, we can conclude that higher efficiency and an amplitude increase are closely related.

Finally, in order to test the chaos hypothesis, we explored the effects of asymmetries of geometrical and tissue properties on the model, with and without vocal membranes. In our model, the tissue properties of the membranes are simulated with the upper-model masses, whereas the geometrical aspects have been realized by adding the extensions of vertical dimension d_3 . Thus, for the upper-left (l) and the upper-right (r) model mass, we used the relations

$$m_{2l} = m_2 / q_2, \quad m_{2r} = m_2, \quad k_{2l} = q_2 k_2, \quad k_{2r} = k_2, \quad (12)$$

where $q_2=0.2$ is an asymmetry factor changing mass and stiffness coefficients of the left-upper mass. The upper graph of Fig. 7 shows a spectrogram of the glottal flow created by running the model with the pressure curve shown in the lower graph of Fig. 7. At $P_s=13$ cm H₂O, period doubling appears, followed by a cascade of subharmonic-oscillation patterns with increasing subglottal pressure. In contrast, the model version without vocal membranes but with identical asymmetry coefficient q_2 shows regular oscillation in the investigated pressure range (see the middle graph of Fig. 7).

IV. DISCUSSION

The most significant finding of these simulations is that a very simple change in the geometry of the vocal fold can produce major differences in the dynamics of phonation. Several conclusions can be drawn regarding the implications of our model for phonation dynamics in animals which possess vocal membranes. We will focus here on echolocating bats, but similar arguments apply for nonhuman primates and other mammals. In a night of foraging, insectivorous bats spend a considerable amount of energy producing the ultra-

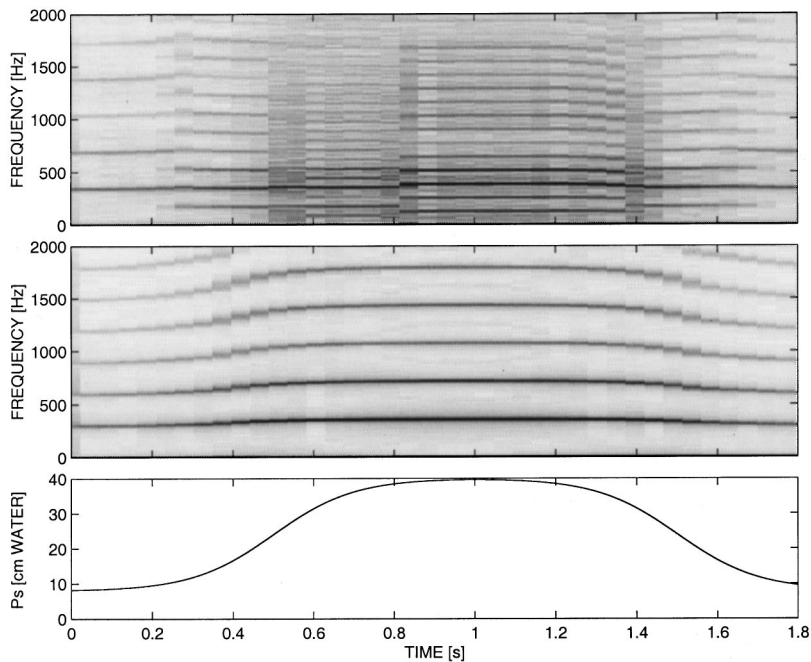


FIG. 7. In the spectrogram of an airflow simulation with asymmetric properties of the upper-mass pair including vocal membranes (upper graph), a cascade of bifurcations sets in at 0.2 s. Increasing subglottal pressure leads to various subharmonic-phonation patterns ($q_2=0.2$, $d_3=0.1$ cm, and $\theta=1.8^\circ$). In contrast, the same conditions in the model without vocal membranes do not produce irregularities, as can be seen in the middle graph. The lower graph shows the pressure variation during the simulations.

sonic pulses which subserve echolocation (Suthers and Fattu, 1973; Hartley and Suthers, 1988). The frequency of the echolocation pulses determines the spatial discrimination of the biosonar system, while pulse intensity controls the range within which the animal can detect its insect prey or other targets (Griffin, 1958; Suthers and Fattu, 1973; Suthers, 1988). Thus, there is a powerful selective pressure for the bat to produce very loud, very high-pitched sounds. For example, *Pteronotus parnellii* produces pulses with a 30-kHz fundamental and most energy at 60 kHz with loudness of 100 dB SPL, and appears to have a vocal efficiency two orders of magnitude greater than humans (Suthers and Fattu, 1973; Suthers, 1988). Unfortunately for the bat, the subglottal pressures necessary to produce these pulses are great enough (20–50 cm H₂O) to potentially counteract pulmonary blood flow and thus impede respiration, and this is “arguably the most important factor ultimately limiting pulse intensity” (Suthers and Fattu, 1973; Suthers, 1988).

Our simulations suggest that the vocal membranes may permit the vocal folds to oscillate at lower subglottal-threshold pressures, and indicate that this effect increases with increasing frequency. Thus, a simple geometrical modification of glottal morphology could enable an organism which habitually makes loud, high-pitched sounds to do so using less energy, or alternatively to increase the loudness and pitch range of sounds produced using a certain subglottal pressure. In other words, the addition of vocal membranes to the standard mammalian vocal folds may not only allow an organism to produce higher-pitched phonation (as previously suggested by many researchers, e.g., Griffin, 1958; Griffiths, 1978; Schön-Ybarra, 1995; Brown and Cannito, 1995; Fitch and Hauser, 1995), but also to produce a given sound louder and more efficiently. While the significance of this effect for echolocating bats is obvious and could be substantial, more subtle benefits could be accrued to any animal producing loud vocalizations for long-distance communication.

To our knowledge, this hypothesis has not been sug-

gested previously, and it indicates that a proper understanding of the role of the vocal membranes requires detailed information on the anatomy and geometry of these structures. Because the importance of geometry has not been recognized previously, most anatomical descriptions of vocal membranes have been based on serial sections of the larynx, which distort the angle of attachment of the membranes to the rest of the vocal folds. Our simulations suggest that there is an angle of attachment for any given membrane length which is optimal for lowering threshold pressure; this prediction could easily be tested by careful measurements of this angle in intact larynges.

Our final finding was that the addition of the vocal membranes made our model more susceptible to the effects of vocal-fold asymmetry, and thus more likely to enter regimes of subharmonic or chaotic vibration. Many examples of similar vocal instabilities have been found in humans, e.g., in newborn cries (Sirviö and Michelsson, 1976; Mende *et al.*, 1990), noncry vocalizations of infants (Robb and Saxman, 1988), Russian lament (Mazo *et al.*, 1995), and also in conversational speech (Dolansky and Tjerlund, 1968; Kohler, 1996). Pathological voices in particular exhibit a wide variety of vocal instabilities (Herzel and Wendler, 1991; Herzel, 1993). The concepts of nonlinear dynamics provide a framework suitable for understanding complex vocal-fold oscillations and for classifying this menagerie of irregularities. Left–right asymmetries of the larynx, manifested in glottal geometry, tissue properties, or in nervous function, induce subharmonic, toroidal, and chaotic oscillations of the vocal folds. Sub- and supraglottal resonances and turbulent airflow can also lead to nonlinear effects (Herzel, 1996). These findings can be generalized to phonation in nonhuman mammals as well (see Wilden *et al.*, 1998). The current results suggest that vocal membranes can heighten the sensitivity of the vocal folds to asymmetries, thus increasing instability, and thus add support for the notion that the larynges of nonhuman primates may be prone to instabilities (Lieberman, 1968;

Pressman and Kelemen, 1963). Moreover, the experimentation with simple models with geometric modifications like the one explored here may provide potential and clinical valuable insights into the functioning of diseased or damaged vocal folds. In particular, the phenomenon of sulcus vocalis offers an interesting parallel to the geometric addition of the vocal membrane considered here. The term sulcus vocalis is applied to a vast array of organic changes ranging from minor vocal-fold indentations to severe scarred lesions. Like the vocal membranes, sulcus vocalis is often bilateral and extends down the whole length of the vocal fold (Hirano *et al.*, 1990). Thus, while the flaps of membrane seen in a sulcus vocalis patient differ in several ways from vocal membranes (they may be thicker, may be located on the inside of the fold instead of the top edge, and may point up or down), they may have qualitative effects on phonation dynamics similar to those described here, and could be investigated using similar modeling strategies.

V. FUTURE DIRECTIONS

The model of the vocal membranes described here only scratches the surface of the problem. Two obvious extensions of the current model are addition of a vocal tract (exploring possible effects of source-tract coupling on phonation) and allowing θ to vary dynamically (thus adding a third independent oscillator). Source-tract coupling is based on the possibility of a novel mode of phonation resulting from interaction between vibrations of the laryngeal source and vocal-tract resonances. In wind instruments, strong coupling between the source (e.g., the reed of a clarinet, or the lips of a trumpet player) and column of air in the instrument body leads to a situation in which the length of the air column effectively dictates the possible vibration frequencies of the source; there is a strong source-tract interaction (Fletcher and Rossing, 1991; Sundberg, 1991). In contrast, source and tract in human speech appear in general to be almost completely independent: a singer can produce any arbitrary pitch with any configuration of the vocal tract (Fant, 1960; Lieberman and Blumstein, 1988; Sundberg, 1987; Fletcher, 1992). Although most data gathered to date have supported the idea that source and filter are also independent in animal vocalization (as in the human voice), there are some conflicting reports, it would probably be premature to conclude that no animal vocalizations involve any source-tract coupling (for a review of the data for nonhuman primates, see Fitch and Hauser, 1995; for birds, see Nowicki and Marler, 1988).

The coupling of a resonator tube to a simplified two-mass model was studied by Mergell and Herzel (1997), who showed that the combination of small asymmetries, moderate subglottal pressures, and a strong vocal-tract coupling can induce various complex vocal-fold oscillations. For the simulation of Fig. 7 we used a relatively large asymmetry, but the results of Mergell and Herzel (1997) indicate that resonance of the fundamental frequency with sub- or supra-glottal formants and a strong source-tract coupling based on a narrow epilarynx and incomplete glottal closure can lead to instabilities at smaller laryngeal asymmetries. Because the vocal membranes provide a larger surface area upon which formant-generated pressure variations could act on the vocal

folds, the vocal membranes could lead to an increase in this destabilization. Alternatively, an increase in coupling could allow the animal to stabilize call fundamental frequency via vocal-tract resonances, as is the case in wind instruments. Studies of a model including both a vocal tract and vocal membranes are currently in progress.

Finally, although no video endoscopy or excised larynx observations of phonating animals with vocal membranes are available to support this conjecture, it seems reasonable to assume that the membranes can vary their angle of attachment to the rest of the vocal fold during phonation, acting as an independent oscillator, and thus adding another degree of freedom to the current model. Unfortunately, the data necessary to model such behavior are not currently available, particularly data on stress-strain relationships within the membrane and tensions applied across the membrane. Even details of the geometry and mass of the membrane are lacking for most species. We are currently gathering data of this sort from nonhuman primates, and hope to implement more complex models in the future. For now, the present study demonstrates that even a very simple change in vocal-fold morphology can have nontrivial effects on phonation dynamics, and demonstrates the usefulness of low-order models in exploratory analyses of nonhuman phonation.

- Alipour-Haghighi, F., and Titze, I. R. (1985). "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).
- 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.
- Brown, C. H., and Cannito, M. P. (1995). "Modes of vocal variation in Syke's monkey (*Cercopithecus albogularis*) squeals," *J. Comp. Psych.* **109**, 398–415.
- Dolansky, L., and Tjernlund, P. (1968). "On certain irregularities of voiced-speech wave-forms," *IEEE Trans. Audio Electroacoust.* **AU-16**, 51–56.
- Fant, G. (1960). *Acoustic Theory of Speech Production* (Mouton, The Hague).
- Fitch, W. T. (1997). "Vocal tract length and formant frequency dispersion correlate with body size in rhesus macaques," *J. Acoust. Soc. Am.* **102**, 1213–1222.
- Fitch, W. T., and Hauser, M. D. (1995). "Vocal production in nonhuman primates: Acoustics, physiology, and functional constraints on 'honest' advertisement," *Am. J. Prima.* **37**, 191–219.
- Fletcher, N. H. (1992). *Acoustic Systems in Biology* (Oxford U.P., Oxford, UK).
- Fletcher, N. H., and Rossing, T. D. (1991). *The Physics of Musical Instruments* (Springer-Verlag, New York).
- Gautier, J. P. (1971). "Etude morphologique et fonctionnelle des annexes extra-larynges des cercopithecinae; liaison avec les cris d'espacement," *Biol. Gabonica* **7**, 230–267.
- Griffin, D. R. (1958). *Listening in the Dark* (Yale U.P., New Haven, CT).
- Griffiths, T. A. (1978). "Modification of *M. cricothyroideus* and the larynx in the *Mormoopidae*, with reference to amplification of high-frequency pulses," *J. Mammal.* **59**(4), 724–730.
- Harrison, D. F. N. (1995). *The Anatomy and Physiology of the Mammalian Larynx* (Cambridge U.P., New York).
- Hartley, D. J., and Suthers, R. A. (1988). "The acoustics of the vocal tract in the horseshoe bat, *Rhinolophus hildebrandti*," *J. Acoust. Soc. Am.* **84**, 1201–1213.
- Hast, M. (1989). "The larynx of roaring and non-roaring cats," *J. Anat.* **163**, 117–21.
- Herzel, H., and Wendler, J. (1991). "Evidence of chaos in phonatory samples," in Proceedings of ROSPEECH, Genova, 1991 (ESCA), pp. 263–266.

- Herzel, H. (1993). "Bifurcations and chaos in voice signals," *Appl. Mech. Rev.* **46**, 399–413.
- Herzel, H. (1996). "Possible mechanisms of vocal instabilities," in *Vocal Fold Physiology: Controlling Complexity & Chaos*, edited by P. J. Davis and N. H. Fletcher (Singular, San Diego), pp. 63–75.
- Hill, W. C. O. (1957). *Primates: Comparative Anatomy and Taxonomy: 3 Pithecoidea, Platyrrhini* (Interscience, New York).
- Hill, W. C. O., and Booth, A. H. (1957). "Voice and larynx in African and Asiatic Colobidae," *J. Bombay Nat. Hist. Soc.* **54**, 309–321.
- Hirano, M. (1974). "Morphological structure of the vocal cord as a vibrator and its variations," *Folia Phoniatr.* **26**, 89–94.
- Hirano, M., Yoshida, T., Tanaka, S., and Hibi, S. (1990). "Sulcus vocalis: functional aspects," *Ann. Otol. Rhinol. Laryngol.* **99**, 679–683.
- Hirano, M. (1991). "Phonosurgical anatomy of the larynx," in *Phonosurgery: Assessment and Surgical Management of Voice Disorders*, edited by C. N. Ford and D. M. Bless (Raven, New York).
- 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.
- Kelemen, G. (1949). "Structure and performance in animal language," *Arch. Otolaryngol.* **50**, 74–744.
- Kelemen, G. (1969). "Anatomy of the larynx and the anatomical basis of vocal performance," in *The Chimpanzee, Vol. 1*, edited by G. Bourne (Karger, Basel, Germany), pp. 165–187.
- Kohler, K. J. (1996). "Articulatory reduction in German spontaneous speech," in *Proceeding of the 4th Speech Prod. Seminar, Autrans*, edited by P. Perrier (European Speech Communication Association, Grenoble), pp. 1–4.
- Kurita, S., Nagata, K., and Hirano, M. (1983). "A comparative study of the layer structure of the vocal fold," in *Vocal Fold Physiology: Contemporary Research and Clinical Issues*, edited by D. M. Bless and J. H. Abbs (College-Hill, San Diego).
- Lieberman, P. (1968). "Primate vocalization and human linguistic ability," *J. Acoust. Soc. Am.* **44**, 1574–1584.
- Lieberman, P., and Blumstein, S. E. (1988). *Speech Physiology, Speech Perception, and Acoustic Phonetics* (Cambridge U.P., Cambridge, UK).
- Lucero, J. C. (1993). "Dynamics of the two-mass model of the vocal folds: Equilibria, bifurcations, and oscillation region," *J. Acoust. Soc. Am.* **94**, 3104–3111.
- Lucero, J. C. (1995). "The minimum lung pressure to sustain vocal fold oscillation," *J. Acoust. Soc. Am.* **98**, 779–784.
- Lucero, J. C. (1996). "Nonlinear dynamics of the vocal fold oscillation," First ESCA Tutorial and Research Workshop on Speech Production Modeling—4th Speech Production Seminar, pp. 185–188.
- Mazo, M., Erickson, D., and Harvey, T. (1995). "Emotion and expression: temporal data on voice quality in Russian lament," in *Vocal Fold Physiology*, edited by O. Fujimura and M. Hirano (Singular, San Diego), pp. 173–178.
- Mende, W., Herzel, H., and Wermke, K. (1990). "Bifurcations and chaos in newborn cries," *Phys. Lett. A* **145**, 418–424.
- Mergell, P., and Herzel, H. (1997). "Modelling biphonation—The role of the vocal tract," *Speech Commun.* **22**, 141–154.
- Mergell, P., Herzel, H., Wittenberg, Th., Tigges, M., and Eysholdt, U. (1998). "Phonation onset: Modeling and high speed clottography," *J. Acoust. Soc. Am.* **104**, 467–470.
- Morton, E. S. (1977). "On the occurrence and significance of motivation-structural rules in some birds and mammal sounds," *Am. Nat.* **111**, 855–869.
- Negus, V. E. (1940). *The Comparative Anatomy and Physiology of the Larynx* (Hafner, New York).
- Nowicki, S., and Marler, P. (1988). "How do birds sing?," *Mus. Percep.* **5**, 391–426.
- Pressman, J., and Kelemen, G. (1963). "Physiology of the larynx," *Physiol. Rev.* **35**, 506–554.
- Robb, J. B., and Saxman, J. (1988). "Acoustic observations in young children's vocalizations," *J. Acoust. Soc. Am.* **83**, 1876–1882.
- Sirviö, P., and Michelsson, K. (1976). "Sound-spectrographic cry analysis of normal and abnormal newborn infants," *Folia Phoniatr.* **28**, 161–173.
- Schön-Ybarra, M. (1995). "A comparative approach to the nonhuman primate vocal tract: Implications for sound production," in *Frontiers in Primate Vocal Communication*, edited by E. Zimmerman and J. D. Newman (Plenum, New York).
- Starck, D., and Schneider, R. (1960). "Respirationsorgane," in *Primatologia III/2*, edited by H. Hofer, A. H. Schultz, and D. Starck (Karger, Basel).
- Steinecke, I., and Herzel, H. (1995). "Bifurcations in an asymmetric vocal-fold model," *J. Acoust. Soc. Am.* **97**, 1874–1884.
- Stevens, K. N. (1977). "Physics of laryngeal behavior and larynx modes," *Phonetica* **34**, 264–279.
- Story, B. H., and Titze, I. R. (1995). "Voice simulation with a body-cover model of the vocal folds," *J. Acoust. Soc. Am.* **97**, 1249–1260.
- Sundberg, J. (1987). *The Science of the Singing Voice* (Northern Illinois U.P., DeKalb, IL).
- Sundberg, J. (1991). *The Science of Musical Sounds* (Academic, New York).
- Suthers, R. A., and Fattu, J. M. (1973). "Mechanisms of sound production in echolocating bats," *Am. Zool.* **13**, 1215–1226.
- Suthers, R. A. (1988). "The production of echolocation signals by bats and birds," in *Animal Sonar: Processes and Performance*, edited by P. E. Nachtigall and P. W. B. Moore (Plenum, New York), pp. 23–45.
- Thomas, C. L. (1989). *Taber's Cyclopedic Medical Dictionary* (Davis, Philadelphia).
- 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**, 1536–1522.
- Wendler, J., Seidner, W., Kittel, G., and Eysholdt, U. (1996). *Lehrbuch der Phoniatrie und Pädaudiologie* (Georg Thieme, Stuttgart, New York).
- Wilden, I., Herzel, H., Peters, G., and Tembrock, G. (1998). "Subharmonics, biphonation, and deterministic chaos in mammal vocalization," *Bioacoustics* **9**, 171–196.