Voice simulation with a body-cover model of the vocal folds

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

(Received 16 February 1994; accepted for publication 21 October 1994)

A simple, low-dimensional model of the body-cover vocal-fold structure is proposed as a research tool to study both normal and pathological vocal-fold vibration. It maintains the simplicity of a two-mass model but allows for physiologically relevant adjustments and separate vibration of the body and the cover. The classic two-mass model of the vocal folds [K. Ishizaka and J. L. Flanagan, Bell Syst. Tech. J. **51**, 1233–1268 (1972)] has been extended to a three-mass model in order to more realistically represent the body-cover vocal-fold structure [M. Hirano, Folia Phoniar. **26**, 89–94 (1974)]. The model consists of two "cover" masses coupled laterally to a "body" mass by nonlinear springs and viscous damping elements. The body mass, which represents muscle tissue, is further coupled laterally to a rigid wall (assumed to represent the thyroid cartilage) by a nonlinear spring and a damping element. The two cover springs are intended to represent the elastic properties of the epithelium and the lamina propria while the body spring simulates the tension produced by contraction of the thyroarytenoid muscle. Thus contractions of the cricothyroid and thyroarytenoid muscles are incorporated in the values used for the stiffness parameters of the body and cover springs. Additionally, the two cover masses are coupled to each other through a linear spring which can represent vertical mucosal wave propagation. Simulations show reasonable similarity to observed vocal-fold motion, measured vertical phase difference, and mucosal wave velocity, as well as experimentally obtained intraglottal pressure.

PACS numbers: 43.70.Aj, 43.70.Gr

INTRODUCTION

Simulation of vocal-fold vibration typically has one of two purposes. Either the model is used as the voice source for a speech synthesis system, or it is used to study specific aspects of the mechanics of the vibration. Hirano (1974) states that if the purpose of a vocal-fold model is for physiological or clinical purposes, then it must have parameters that can be varied to simulate different conditions produced by various laryngeal adjustments or by pathological variations. This paper describes a new model of self-sustained vocal-fold vibration based on the body-cover concept of vocal-fold structure (Hirano, 1974). It is simple and low dimensional but allows for physiologically relevant adjustments. It is intended to be used primarily as a research tool to study both normal and pathological vocal-fold vibration rather than as a voice source for a speech synthesis system, although the latter is not necessarily ruled out.

The body-cover concept (Hirano, 1974) is generally used to describe the vocal-fold structure (Fig. 1). It suggests that the vocal fold can be divided into two tissue layers with different mechanical properties. The body layer consists of muscle fibers and some tightly connected collagen fibers of the vocal ligament. The cover layer consists of pliable, noncontractile tissue (the epithelium, the superficial layer, and the intermediate layers of the lamina propria) that acts as a flexible sheath around the body layer. The cover typically is loosely connected to the body during vibration. The motion of the cover layer is usually observed as a surface wave that propagates from the bottom of the vocal fold to the top (in Fig. 1) thus experiencing movement in both the lateral and vertical directions. Self-sustained vocal-fold oscillation is highly dependent on this surface-wave behavior (typically referred to as the vertical phase difference) and is the primary mechanism for transferring energy from the glottal flow stream to the tissue to fuel the vibration. The body layer is primarily involved in lateral motion. Based on his findings, Hirano (1974) suggests that the vocal fold should be treated as a double structured vibrator whose stiffness parameters should be based on the relative activations of the thyroarytenoid (TA) and cricothyroid (CT) muscles. Thus the resultant vibration of the vocal folds is composed of the coupled oscillations of the body and cover layers.

Based on the body-cover concept, a parsimonious model of vocal-fold vibration would include mass and stiffness elements that represent the properties of the different tissue layers as well as allowing for the effect of vertical phase difference in the cover layer. Several lumped-element models have been proposed throughout the years that successfully simulate the vertical phase difference but do not include the effect of separate tissue layers. These models will now be briefly reviewed.

A. Review of previous models

Flanagan and Landgraf (1968) modeled vocal-fold vibration with a single mass-spring oscillator driven by airflow from the lungs. The model produced reasonable self-sustained oscillations only with an inertive vocal tract load. Because of its single degree of freedom in the tissue, it could not produce the vertical phase difference needed for flow-induced oscillation. The two-mass model of Ishizaka and Flanagan (1972), subsequently referred to as the IF72 model in this paper, was able to sustain oscillation with or without a

 0001-4966/95/97(2)/1249/12/$6.00

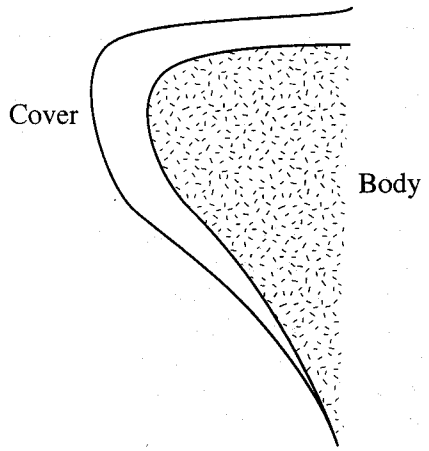

FIG. 1. Body-cover structure of the vocal folds (coronal plane).

vocal tract and provided the degrees of freedom necessary to produce the vertical phase difference. This model has been widely used as a simple, low-dimensional model of the vocal folds. Both the Flanagan–Landgraf (1968) and the IF72 models were initially created to serve as a voice source for a speech synthesis system, but the IF72 model was later used to study various pathologies of the vocal folds (Ishizaka and Isshiki, 1976). However, there is no direct physiological correlation between the spring stiffnesses and the effects of muscle contractions in either of these models.

Titze (1973, 1974) represented both the vertical and longitudinal modes of vocal-fold vibration with a 16-mass model. This model, subsequently called T73, consisted of eight coupled longitudinal sections, each with two masses in the coronal plane. In contrast to the IF72 model, the upper mass was coupled to only the lower mass and not to a rigid lateral boundary. The two masses (in each longitudinal section) were also allowed to have a vertical degree of freedom which simulated two-dimensional trajectories of the vocal-fold tissue. The T73 model was specifically designed to study the mechanics of vocal-fold vibration with the hope that it would lead to a better understanding of the workings of the voice source.

Koizumi *et al.* (1987) has described several variations of the simple two-mass model that incorporate elements of both the IF72 and T73 models. Their modifications were intended to produce a more natural sounding artificial voice than the IF72 model. The intended use was primarily as a glottal source for speech synthesis. However, these models also have been used as a tool to study vocal pathologies (Smith *et al.*, 1992).

In addition, Wong *et al.* (1991) have combined the two-mass approach of the IF72 model with the longitudinal discretization of the T73 model to create a ten-mass model which was used specifically to study vocal pathologies.

The "two-mass"-type models discussed above may work well as voice source models for speech synthesis but they all are deficient in the sense that they do not capture the layered structure of the vocal folds. In the IF72 model, the lower mass is made thicker (vertical dimension in the coronal plane) and more massive than the upper element as an attempt to include the effects of the body layer. But, because the provision does not exist for coupled oscillation of *both* layers, the two-mass model is essentially a "cover" model rather than a "body-cover" model. This may be a valid model for laryngeal conditions in which only the cover is in vibration but is not adequate if conditions are such that both body and cover are in vibration.

Other more complex models that simulate the layered structure of the vocal folds provide a more precise physiological representation of human tissue but are not well suited for use in a speech synthesis system because of their computational complexity. The continuum mechanics model of Titze and Talkin (1979) and a more recent finite element implementation (Alipour-Haghighi and Titze, 1983) are examples. The continuum models obviously have as their purpose the detailed study of tissue movement throughout the vocal fold during vibration. Because these models have a very large number of degrees of freedom, they are capable of producing a rather complex vibratory pattern composed of many different modes. Thus studies of vocal pathologies as well as normal phonation are natural areas of pursuit with these models but the complexity that results from using many degrees of freedom makes these models computationally intensive. However, recent modal analyses using the finite element model have shown that the vocal-fold vibration is largely dominated by the first two to three modes of vibration (Berry *et al.*, 1994). The contributions of the higher modes have been found to be much less significant than the lower ones. This finding implies that a simple, low-dimensional model may capture enough of the vibratory characteristics to serve as a useful research tool in cases where fine detail is not required.

I. THE MODEL

A three-mass model of the vocal folds is proposed here as a lumped-element approximation of the body-cover structure. The model is essentially the classic IF72 two-mass model with a third mass added to simulate the effect of the body component. It will be used to simulate a section of the vocal fold 0.3 cm in thickness (vertical dimension in coronal view), 0.23 cm in depth, and 1.0 cm in length. The cover portion of this 0.3-cm-thick vocal fold is divided into two equally thick elements. A single larger mass is used to simulate the body layer. This discretization of the body-cover structure into a lumped-element system is shown in Fig. 2.

Both the upper and lower cover masses are coupled to the body mass through nonlinear springs and damping elements. The two cover masses are also coupled to each other through a coupling spring. This coupling spring accounts for the shear forces in the cover. The body mass is coupled to a rigid boundary, i.e., the thyroid cartilage. For the purposes of this paper, the right and left vocal folds are considered to be symmetric; that is, the same movement is considered to occur on the left as on the right. The asymmetric case will be the subject of a future study.

The springs k_u and k_l are considered to represent the stiffness of the cover tissue as well as the effective coupling stiffnesses between the body and cover, which vary primarily with the contraction of the CT muscle. Spring k_b represents

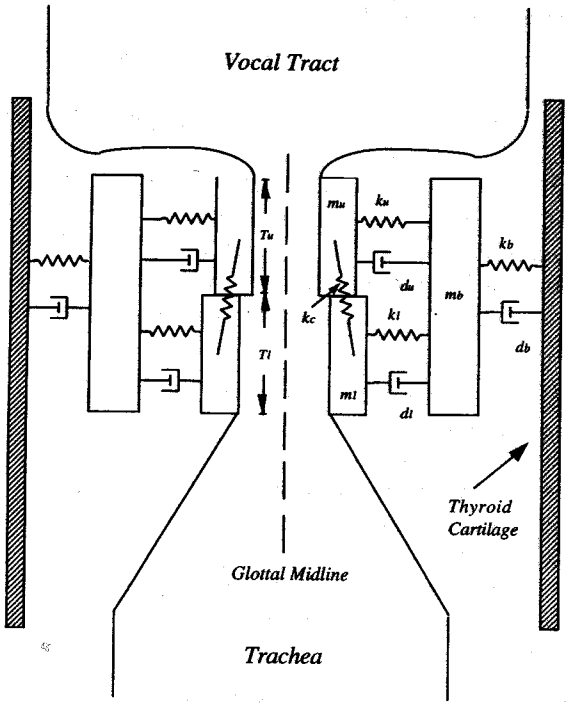

FIG. 2. Lumped-element representation of the body-cover structure of the vocal folds (coronal plane).

the effective stiffness of the body, which varies with both the TA contraction and CT contraction. The values used for the masses and stiffnesses (discussed later) were obtained from considerations of the physical properties of body and cover tissues.

A wave-reflection analog vocal tract (Kelly and Lochbaum, 1962; Liljencrants, 1985) was attached to the three-mass model both supraglottally and subglottally. This allowed for interaction of the vocal folds with the acoustic pressures generated above and below the vocal folds.

A. Equations of motion

The equations of motion for the three masses are written in terms of the coupling forces and the external driving forces that are exerted on each mass,

$$F_u = m_u\ddot{x}_u = F_{ku} + F_{du} - F_{kc} + F_{eu} + F_{u\mathrm{Col}}, \tag{1a}$$

$$F_l = m_l\ddot{x}_l = F_{kl} + F_{dl} + F_{kc} + F_{el} + F_{l\mathrm{Col}}, \tag{1b}$$

$$F_b = m_b\ddot{x}_b = F_{kb} + F_{db} - [F_{ku} + F_{dl} + F_{kl} + F_{dl}], \tag{1c}$$

where the following forces are identified:

F_{du}, F_{dl}, and F_{db}—forces due to damping,
F_{ku}, F_{kl}, and F_{kb}—lateral spring forces,
F_{kc}—spring force due to the coupling of m_u and m_l,
$F_{u\,\mathrm{Col}}$ and $F_{l\,\mathrm{Col}}$—forces generated only during collision with the opposite vocal fold,
F_{eu} and F_{el}—external forces generated by the glottal flow,
F_u, F_l, and F_b—forces of the accelerating masses.

If x_u, x_l, and x_b are defined to be the displacements of each mass and x_{uo}, x_{lo}, and x_{bo} are the initial positions of each mass, then the displacements from equilibrium will be $(x_u - x_{uo})$, $(x_l - x_{lo})$, and $(x_b - x_{bo})$. With these definitions, the equations for the lateral spring forces can be written. For the upper mass,

$$F_{ku} = -k_u[\{(x_u - x_{u0}) - (x_b - x_{b0})\} + \eta_u\{(x_u - x_{u0}) - (x_b - x_{b0})\}^3], \tag{2}$$

where η_u is the coefficient of nonlinearity and $[(x_u - x_{uo}) - (x_b - x_{bo})]$ gives the net compression or elongation of the spring. Similarly, for the lower and body masses the spring forces are

$$F_{kl} = -k_l[\{(x_l - x_{l0}) - (x_b - x_{b0})\} + \eta_l\{(x_l - x_{l0}) - (x_b - x_{b0})\}^3] \tag{3}$$

and

$$F_{kb} = -k_b[(x_b - x_{b0}) + \eta_b(x_b - x_{b0})^3]. \tag{4}$$

All of the coefficients of nonlinearity (η's) were set to a value of 100, as was done in the IF72 model. The coupling force is determined by the relative displacements of the upper and lower masses and a linear spring (k_c). This force is given by

$$F_{kc} = -k_c[(x_l - x_{lo}) - (x_u - x_{uo})]. \tag{5}$$

When the left and right vocal folds are in collision an extra nonlinear spring is switched on to simulate the effect of impact. The extra spring forces have the form

$$F_{u\,\mathrm{Col}} = -h_{u\,\mathrm{Col}}[(x_u - x_{u\,\mathrm{Col}}) + \eta_u(x_u - x_{u\,\mathrm{Col}})^3], \tag{6a}$$

$$F_{l\,\mathrm{Col}} = -h_{l\,\mathrm{Col}}[(x_l - x_{l\,\mathrm{Col}}) + \eta_l(x_l - x_{l\,\mathrm{Col}})^3], \tag{6b}$$

where the h's are the linear spring coefficients, with values of $3k_u$ and $3k_l$. The η's are nonlinear coefficients, set to 500 as in the IF72 model; the $x_{u\,\mathrm{Col}}$ and $x_{l\,\mathrm{Col}}$ are the displacements where collision occurs for medial motion.

The damping forces are given by

$$F_{dl} = -d_l(\dot{x}_l - \dot{x}_b), \tag{7a}$$

$$F_{du} = -d_u(\dot{x}_u - \dot{x}_b), \tag{7b}$$

$$F_{db} = -d_b\dot{x}_b, \tag{7c}$$

where the d's are the damping coefficients. The damping coefficients are computed with the following equations during an open-glottis condition:

$$d_l = 2\zeta_l(m_l k_l)^{1/2}, \tag{8a}$$

$$d_u = 2\zeta_u(m_u k_u)^{1/2}, \tag{8b}$$

$$d_b = 2\zeta_b(m_b k_b)^{1/2}. \tag{8c}$$

The collision of the right and left vocal folds and the subsequent "closed-glottis" interval presents somewhat of a problem for a lumped-element model. In a continuum model the vocal folds are regarded as a distributed system that is composed of a large number of small masslike elements. During a collision the surface elements of the left vocal fold collide with the surface elements of the right vocal fold and will consequently experience no more displacement during the closed-glottis interval. However, all of the remaining elements are still free to move and will continue to do so as

momentum is redistributed into other directions and the internal damping forces dissipate much of the energy in the system. An example of a collision of a distributed system is that of an automobile crash test. When the automobile collides with the stationary wall the front bumper first comes to a stop, but all of the other elements of the car, including the test dummies, have the momentum to continue their motion until the energy in the system has been dissipated or redirected. With a lumped-element model, all of the mass in the system is concentrated into a few point masses. In the present model we have three point (or bar) masses. Thus if a collision were simulated by simply halting the motion of the point mass it would have the effect of bringing an entire massive system to an instantaneous stop, which is a highly unnatural event. Ishizaka and Flanagan (1972) proposed that a collision of point masses could be simulated by switching on a nonlinear spring at the instant of collision which opposes continued movement [see Eqs. 6(a) and 6(b)] and also by increasing the damping ratio (ζ) for each mass element in a stepwise fashion (i.e., ζ is increased to be $\zeta+1.0$). In the IF72 model, the $\zeta+1.0$ (closed-glottis) damping ratio is nearly 10× greater than that for the open glottis. Titze (1976) disputes that glottal closure increases the energy losses in the tissue by such a large amount. However, the same approach of the IF72 model has been adopted here except that only 0.4 is added to the open-glottis ratio (i.e., $\zeta=\zeta+0.4$) during collision. It will be shown later that this value doubles the open-glottis damping ratio for the cover masses and brings it close to critical damping. Thus the tissue losses during collision will be increased, but not to the high levels as used in the IF72 model. It is realized that this approximation for collision damping is a weakness in the argument for a physiologically realistic model but if the simplicity of the lumped-element model is to be maintained, approximations such as this must be accepted.

During the closed-glottis condition the damping ratio (ζ) is increased in a stepwise fashion for the upper and lower masses while the body mass damping is not changed. Leaving the body damping unchanged is justified by assuming that only the elements in collision suffer increased energy losses. Since the left and right vocal folds are considered to be symmetric and zero displacement represents the glottal midline, any negative displacement of the cover masses would imply that the left and right folds are in collision. The collision damping is shown in Eqs. (9) and (10):

$$\zeta_u=\begin{cases}\zeta_u, & x_u\geq 0.0,\\ \zeta_u+0.4, & x_u<0.0,\end{cases} \tag{9}$$

$$\zeta_l=\begin{cases}\zeta_l, & x_l\geq 0.0,\\ \zeta_l+0.4, & x_l<0.0.\end{cases} \tag{10}$$

The glottal area, assuming symmetric vocal folds, is calculated for the upper and lower mass regions by

$$a_u=\begin{cases}2x_uL_g, & x_u\geq 0,\\ 0, & x_u<0,\end{cases} \tag{11a}$$

$$a_l=\begin{cases}2x_lL_g, & x_l\geq 0,\\ 0, & x_l<0,\end{cases} \tag{11b}$$

where L_g is the length of the glottis.

B. Pressure equations

The equations of motion for the three-mass system are coupled to the aerodynamic driving forces via the glottal area. The intraglottal pressure exerts a force on the upper and lower cover masses which is the driving force that produces oscillation. This pressure will depend on the open area of the glottis in the upper and lower mass regions. Titze (personal communication) has outlined a general method of calculating the intraglottal pressure which will be adapted here for the three-mass model. The method is based on the following assumptions:

(1) The flow detaches at the minimum glottal diameter.
(2) Bernoulli-type flow exists from the subglottal region to the minimum glottal diameter.
(3) A constant diameter jet exists from the minimum diameter to the glottal exit. Pressure is considered to be constant in this region.
(4) Pressure recovery after glottal exit (expansion and reattachment) follows the equations derived by Ishizaka and Matsudaira (1972).

The derivation of simplified equations for the pressure in the glottis is now presented within the framework of a lumped-element model. In the region where Bernoulli flow is applicable, the pressure can be computed at any point upstream of the minimum area as

$$P(a)=P_s-\frac{1}{2}\rho u^2\left(\frac{1}{a^2}-\frac{1}{a_s^2}\right), \tag{12}$$

where P_s is the subglottal pressure, ρ is the density of air, u is the flow, a is the cross-sectional area of the channel, and a_s is the corresponding subglottal duct area. Now let P_m represent the pressure at the minimum cross-sectional area in the glottis. Then from continuity of pressure at the boundary of the Bernoulli and jet regimes,

$$P_m=P_s-\frac{1}{2}\rho u^2\left(\frac{1}{a_m^2}-\frac{1}{a_s^2}\right), \tag{13}$$

and from pressure recovery as derived by Ishizaka and Matsudaira (1972),

$$P_m=P_i-\frac{1}{2}\rho k_e u^2\left(\frac{1}{a_m^2}\right). \tag{14}$$

In the last equation, P_i is the supraglottal pressure (vocal tract input pressure), a_m is the minimum cross-sectional area within the glottis, and k_e is the exit pressure coefficient (defined below). Equating (13) and (14), solving for u^2, and substituting the result into (12) gives an equation for the pressure at any point upstream of the minimum cross-sectional area:

$$P=P_s-(P_s-P_i)\times\frac{a^{-2}-a_s^{-2}}{a_m^{-2}(1-k_e)-a_s^{-2}}\quad\text{Bernoulli regime.} \tag{15}$$

At any point downstream of the minimum diameter,

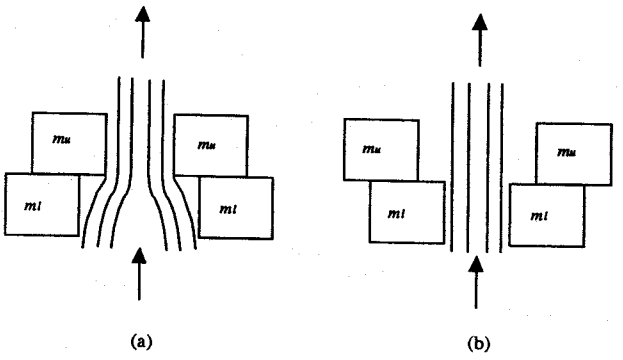

FIG. 3. Assumed flow patterns in the glottis, (a) Bernoulli region in lower section, jet region in upper; (b) jet region in both upper and lower sections.

$$P=P_i-(P_s-P_i)\frac{k_e a_m^{-2}}{a_m^{-2}(1-k_e)-a_s^{-2}} \quad \text{Jet regime.} \tag{16}$$

If collision occurs at the location of the upper or lower mass, the flow is temporarily blocked and the pressures are computed as

$$P=P_s \quad \text{below the collision,} \tag{17a}$$

$$P=0 \quad \text{within the collision,} \tag{17b}$$

$$P=P_i \quad \text{above the collision.} \tag{17c}$$

When the duct area a is small, Eqs. (15) and (16) can be simplified considerably. In comparison to a^{-2} and a_m^{-2}, the term a_s^{-2} is negligible. Furthermore, the pressure recovery coefficient, derived by Ishizaka and Matsudaira (1972) as

$$k_e=\frac{2a_m}{a_i}\left(1-\frac{a_m}{a_i}\right), \tag{18}$$

is also negligible because the supraglottal duct area a_i is much greater than a_m. Equations (15) and (16) then reduce to

$$P=P_s-(P_s-P_i)(a_m/a)^2 \quad \text{Bernoulli regime,} \tag{19}$$

$$P=P_i \quad \text{jet regime.} \tag{20}$$

These equations are computationally preferred because the ratio (a_m/a) is guaranteed to be greater than 0 and less than 1. Division by zero is not a concern.

In the case of the three-mass model, there are only two points in the glottis at which the pressure must be computed; at the location of the upper mass and the lower mass. Thus there are at most two different cases to deal with; either the upper or the lower region is the minimum area (Fig. 3). When the glottis is in the configuration shown in Fig. 3(a), the pressure in the region of the lower mass is computed by

$$P_l=P_s-(P_s-P_i)(a_u/a_l)^2, \tag{21}$$

where a_u and a_l are the open areas of the upper and lower glottal sections, respectively. Since the glottis is assumed to be symmetric, these areas are computed using Eqs. 11(a) and 11(b). For the glottal configuration of Fig. 3(b), the upper and lower pressures are both set to equal the input pressure to the vocal tract,

$$P_u=P_l=P_i. \tag{22}$$

The forces exerted on each cover mass by the pressures in the glottis (P_u and P_l) are

$$F_{eu}=P_u L_g T_u, \tag{23a}$$

$$F_{el}=P_l L_g T_l, \tag{23b}$$

where L_g is the length of the vocal fold that is effectively in vibration and T_u and T_l are the upper and lower vertical thicknesses (in the coronal plane).

C. Flow equations

The three-mass model is coupled to a vocal tract in which subglottal, pharyngeal, oral, and nasal sections are included. Since a wave-reflection algorithm is being used we can derive the flow equation based on the incident and reflected pressures at the glottis. The formulation of the flow equations are given in Titze (1984); only the results will be presented here. It should be noted that the glottal flow resistances and glottal flow inertances proposed by Ishizaka and Flanagan (1972) are purposely not used in this formulation. There is no proven validity for these circuit elements for time-varying flow and irregular geometry. They are based on steady flow between long, parallel plates (Poiseuille flow) and were proposed by van den Berg *et al.* (1957) for experimentation on static models. Computationally, they introduce inaccuracy and instability because they become infinite when glottal area goes to zero. The transglottal pressure coefficient k_t serves as a correction factor for overidealized flow calculations. The flow through the glottis is

$$u=\left(\frac{a_m c}{k_t}\right)\left\{\frac{-a_m}{A^*}\pm\left[\left(\frac{a_m}{A^*}\right)^2+\left(\frac{4k_t}{c^2\rho}\right)(P_s^+-P_i^-)\right]^{1/2}\right\}, \tag{24}$$

where k_t is a transglottal pressure coefficient (Scherer and Titze, 1983), ρ is the density of air, a_m is the minimum glottal area, and c is the speed of sound. A^* is defined as an effective vocal tract area for acoustic loading of the glottis:

$$\frac{1}{A^*}=\frac{1}{A_s}+\frac{1}{A_i}, \tag{25}$$

where A_s and A_i are the areas of the first sections of the subglottal and supraglottal ducts, respectively. P_s^+ and P_i^- are the incident pressures above and below the glottis and are known from previous calculations. According to Titze (1984), the plus sign in Eq. (24) is appropriate when $P_s^+>P_i^-$ and the minus sign when $P_s^+<P_i^-$. Once the flow has been computed, the reflected pressures P_s^- and P_i^+ can be found using the relations

$$P_s^-=P_s^+-(\rho c/A_s)u, \tag{26a}$$

$$P_i^+=P_i^-+(\rho c/A_i)u. \tag{26b}$$

D. Connection to the vocal tract

P_i^+ is the forward traveling (toward the lips) partial pressure component in the first supraglottal section of a wave-reflection type of vocal tract. Similarly, P_s^- is the backward traveling wave in the subglottal section immediately

below the glottis. From this point, the partial pressures are propagated in the sub- and supraglottal tracts as described by Liljencrants (1985). Appropriate attenuation factors are used to account for tract losses. The radiation impedance at the lips is assumed be represented by an inductance and resistance in parallel (Flanagan, 1972) and the subglottal system is terminated by a constant lung pressure and a resistance. The vocal tract shape used for the simulations in this paper is simply a uniform tube with a cross-sectional area of 5 cm^2; nasal sections were not used.

E. Parameter values

The numerical values used in this model should have some physiological relevance. This section demonstrates how the initial numbers were obtained and then fine tuned to produce acceptable phonation.

The masses were obtained by computing the approximate volume that is consumed by the particular element and then multiplying by the tissue density. For the cover masses, the depth is made up of the epithelial layer, the superficial layer, and the intermediate layer of the lamina propria. According to Hirano (1977) and Hirano *et al.* (1981) the epithelial layer has an approximate depth of 0.005 cm while the superficial layer and intermediate layer of the lamina propria are each about 0.03 cm deep. The depth of the cover is then

$$D_{\text{cover}}=0.005\ \text{cm}+0.03\ \text{cm}+0.03\ \text{cm} =0.065\ \text{cm}. \tag{27}$$

At this point we must assume an effective length of the vibrating vocal fold and define a vertical thickness for each mass. A length of 1 cm was chosen for the vocal-fold length and the thickness of each mass will be 0.15 cm for a total thickness of 0.3 cm. Using the thickness, depth, and length we can compute the volume consumed by the portion of the cover that we are intending to simulate. The cover mass is then the product of the volume and the tissue density. The density of the mucosa is approximately 1.02 g/cm^3 (Perlman, 1985). Thus the mass of the cover is

$$M_{\text{cover}}=(0.065\ \text{cm})(0.3\ \text{cm})(1.0\ \text{cm}) \times(1.02\ \text{g/cm}^3)=0.0199\ \text{g}. \tag{28}$$

In this model each cover element will be assumed to have the same mass so that

$$m_l=m_u=M_{\text{cover}}/2\approx 0.01\ \text{g}. \tag{29}$$

For the body mass we assume the same length and thickness as the cover. The depth of the body is determined by the combination of the depth of the deep layer of the lamina propria and the depth of the portion of the muscle that is effectively in vibration. Hirano *et al.* (1981) reported that the depth of the deep layer is approximately 0.05 cm. However, the amount of the muscle tissue that is involved in vibration depends on the amplitude of vocal-fold vibration and on the relative activation levels of the CT and TA muscles. Titze *et al.* (1989) define a ratio of the cross section of the TA muscle in vibration to the total cross-sectional area in vibration (including the cover and deep layer of the lamina propria). They suggest that this ratio may have a value of about 0.3 for a "soft" loudness condition and 0.6 for a "loud" loudness condition. Choosing a value of 0.5 for this ratio to simulate a condition between soft and loud, an effective muscle depth can be computed. Previously it was shown that the depth of the cover is ~0.065 cm and the depth of the deep layer (of the ligament) is ~0.05 cm which combine to produce a total depth of 0.115 cm. If the ratio of TA cross section to total cross-sectional area is 0.5, then the depth of the muscle will also need to be 0.115 cm, assuming that the body and cover have the same thickness. Now the depth of the body is the combination of the deep layer and the portion of vibrating muscle:

$$D_{\text{body}}=0.05\ \text{cm}+0.115\ \text{cm}=0.165\ \text{cm}. \tag{30}$$

The body mass is then the product of the volume and the tissue density. The density of the muscle has been measured to be 1.04 g/cm^3 (Perlman, 1985) and we will assume for simplicity that the density of the deep layer of the lamina propria (ligament) has the same density as the muscle. So the mass of the body element is

$$m_b=(0.165\ \text{cm})(0.3\ \text{cm})(1.0\ \text{cm})(1.04\ \text{g/cm}^3) =0.05148\ \text{g (approx. 0.05 g)}. \tag{31}$$

The values of the effective stiffness coefficients are determined mainly by the longitudinal stress in the tissue fibers. This stress can be converted to an equivalent coupling stiffness by equating the expressions for fundamental frequency of a vibrating string and a vibrating mass:

$$F_0=\left(\frac{1}{2L_g}\right)\left(\frac{\sigma}{\rho}\right)^{1/2}=\left(\frac{1}{2\pi}\right)\left(\frac{k}{m}\right)^{1/2}, \tag{32}$$

where σ is the longitudinal stress in the tissue fibers and L_g is the vocal-fold length. Now solving for k gives

$$k=\pi^2\sigma m/\rho L_g^2. \tag{33}$$

The values for mass have already been computed, the density of the tissue is known, and the length has been chosen to be 1 cm. Thus only a value for stress (σ) is needed to calculate k.

Alipour-Haghighi and Titze (1991) give passive stress–strain curves for both body and cover tissues taken from an excised canine larynx. This study shows that a strain of 10% in the cover will produce a stress of about 4.0 kPa while the same amount of strain in the body tissue produces about 3.5 kPa of stress. A strain of 10% was chosen a crude estimate for normal phonation. These values of stress can now be used in Eq. (33) to compute stiffness coefficients. With $\rho_{\text{cover}}=1.02$ g/cm^3 and $\rho_{\text{body}}=1.04$ g/cm^3, we get the following estimates of k for the cover and body:

$$k_{\text{cover}}=5.0\ \text{N/m}, \tag{34a}$$

$$k_{\text{body}}=50.0\ \text{N/m}. \tag{34b}$$

But these are only passive stiffnesses, so if there is any TA contraction the stiffness of the body will be higher. Alipour-Haghighi *et al.* (1989) obtained a stress–strain relationship for tetanically stimulated vocal-fold muscle which showed that a stress of approximately 75 kPa was produced for a strain of 10%. Using this value for the stress (σ) in Eq. (33)

gives a stiffness coefficient of ~850.0 N/m. This value represents the stiffness of the vocal-fold muscle that could be expected during supramaximal stimulation. However, during normal phonation the level of activation will be much lower than supramaximal and hence the stiffness will also be much lower. Simulations with the three-mass model showed that reasonable phonation was achieved if the body stiffness was set to 100.0 N/m which is higher than the passive stiffness but much lower than the supramaximal condition. It was also determined that the well documented vertical phase difference was best modeled if the stiffness of the upper cover element was slightly smaller than the lower. Thus the following stiffness values were used:

$$k_u = 3.5 \text{ N/m}, \tag{35a}$$

$$k_l = 5.0 \text{ N/m}, \tag{35b}$$

$$k_b = 100.0 \text{ N/m}. \tag{35c}$$

The spring constant k_c accounts for the shear forces between the masses. In a distributed system, this shear determines the mucosal wave velocity. It is this parameter over which the human appears to have the least control, its properties being primarily dependent upon vocal health. The mucosal wave velocity has been measured in the range from 1.0 to 2.0 m/s (Baer, 1975; Titze *et al.*, 1993; Sloan *et al.*, 1993). In this model, values of 0.5–3.0 N/m were used for k_c, which produced wave velocities of approximately in the range of 1.0–2.0 m/s.

Kaneko *et al.* (1972) estimated the vocal-fold damping ratio to be approximately 0.1–0.2 at a fundamental frequency of 30–40 Hz while Isshiki (1977) reported damping ratios of 0.2–0.4 at a fundamental frequency of 130 Hz. Simulations using the three-mass model have been performed using a range of damping ratios that cover both the Kaneko *et al.* (1972) study and the Isshiki (1977) study. These simulations showed that phonation is more acceptable using damping ratios closer to Isshiki's data rather than those of Kaneko. The following damping ratios have been assigned to the three elements:

$$\zeta_u = 0.4, \tag{36a}$$

$$\zeta_l = 0.4, \tag{36b}$$

$$\zeta_b = 0.2. \tag{36c}$$

The parameter values defined in this section constitute a system that simulates essentially normal phonation. This means that a steady vibration is achieved with the amplitudes of vibration, vertical phase difference, fundamental frequency, and mucosal wave velocity all in the range of the published information on these quantities.

The simulation is performed by integrating the equations of motion with a fourth-order Runge–Kutta algorithm. The sampling frequency was 22 050 Hz and all pressure, flow, and vocal tract computations are executed at every time sample. The model runs with a compute time to real-time ratio of about 100:1 on a Decstation 5000. However, no effort has been made to optimize the computational performance.

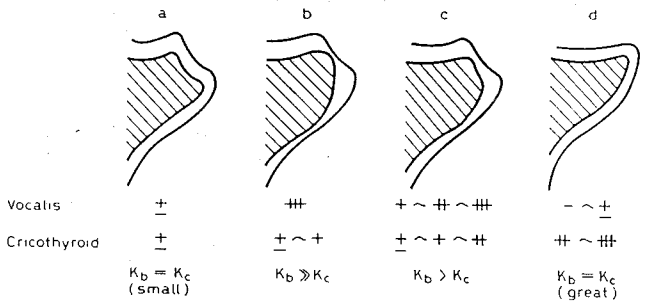

FIG. 4. Sketches of four laryngeal configurations based on different relative stiffnesses of body and cover elements. The plus and minus symbols shown below the sketches denote the activation levels of the vocalis and cricothyroid muscles; plus is associated with activation and a plus and minus together imply a lax state. The relative stiffnesses between the cover and body that result from the various activation combinations are shown on the bottom line (from Hirano, 1974).

II. SIMULATION RESULTS

A. Four cases of Hirano (1974)

It was desired to test this model on realistic laryngeal configurations. Hirano (1974) describes four typical laryngeal adjustments that are reflected in terms of the relative stiffnesses of the body and cover layer (Fig. 4). The first three configurations are primarily concerned with the amount of stiffness present in the body layer while the fourth case implies a stiffness increase in both the body and the cover.

A set of simulations was performed that attempted to cover the range of oscillations predicted by the first three cases. This was done by leaving all parameters constant except for the coupling stiffness (k_c) and the body stiffness (k_b). The parameters that are held constant are given in Table I. In these simulations, the body stiffness is intended to represent the activation level of the TA muscle. The coupling stiffness, over which humans have little control, was selected as a variable parameter so that its effect could be clearly observed. Each of these simulations are shown as small plots within the $k_b - k_c$ plane in Fig. 5. In each plot the lower two traces are the upper (solid) and lower (dashed) masses. The upper trace is the displacement of the body mass. Clearly, a more lax body layer produces a large amplitude displacement in the body and lower frequencies of vibration. Increasing

TABLE I. Parameter values that were held constant for the cases in Fig. 5.

Parameters	Values
m_u	0.01
m_l	0.01
m_b	0.05
T_u	0.15
T_l	0.15
k_u	3.5
k_l	5.0
k_b	variable
k_c	variable
x_{uo}	0.0179
x_{lo}	0.018
x_{bo}	0.30
P_l	0.80

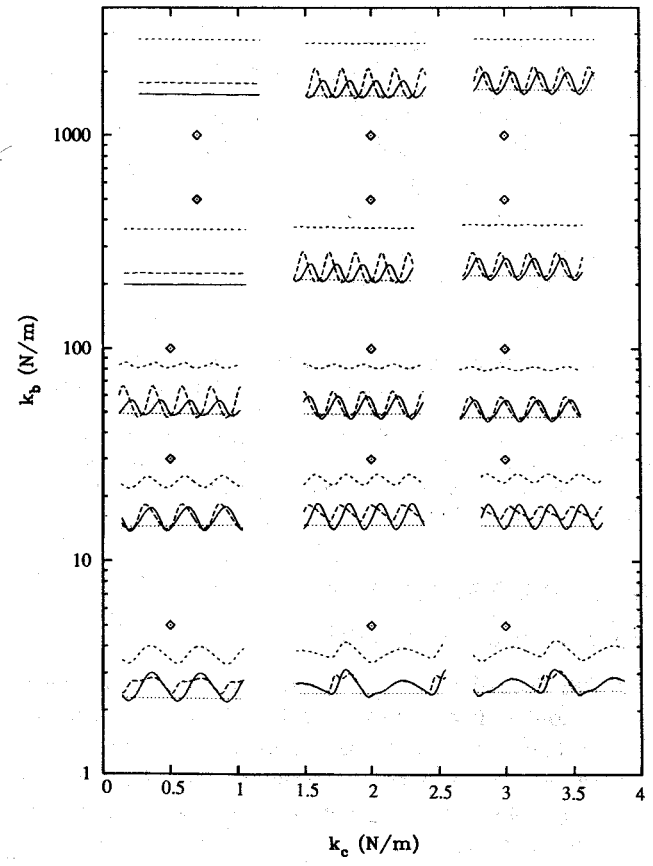

FIG. 5. Simulations are shown as small plots within the k_b-k_c plane. In each plot the lower two traces are the upper (solid) and lower (dashed) masses. The upper trace (dotted) is the displacement of the body mass. Body and coupling stiffness values range from 5 to 1000 and 0.5 to 3.0 N/m, respectively.

the body stiffness (i.e., contracting the TA) leads to lower body layer amplitudes and higher pitches. The coupling stiffness tends to control the ratio of the displacement amplitudes of the upper and lower masses and also the phase difference (time delay) between them.

Three cases that fall in the range of the parameter values used in Fig. 5 which are believed to be representative of Hirano cases a, b, and c (Fig. 4) are now examined in more detail. An additional simulation representing case d was also performed. In this case all of the spring constants were varied from their "normal" configuration as well as the mass of the body layer. The model parameters for each case are shown in Table II. For each case, the laryngeal configuration and the characteristics of the vocal-fold vibration which Hirano describes will be repeated and then followed with a comparison of the simulation results.

Figure 4(a) shows the case in which the stiffness of both the body and cover is low due to a small amount of contraction in the TA and CT muscles. Hirano (1974) suggests that both body and cover are quite lax and will be equally involved in the movement. This, he claims, is typical of soft phonation at low pitch. This case was simulated by setting the stiffness coefficients to be those given in Table II for case A. The simulation is shown in Fig. 6. In Fig. 6(a), the displacements of each of the three masses during vibration are

TABLE II. Parameter values for the four cases of Hirano (1974). m's=mass (g), k's=stiffness (N/m), T's=thickness (cm), x_o's=prephonatory displacement (cm), and P_l=kPa.

Parameters	Case A	Case B	Case C	Case D
m_u	0.01	0.01	0.01	0.01
m_l	0.01	0.01	0.01	0.01
m_b	0.05	0.05	0.05	0.015
T_u	0.15	0.15	0.15	0.15
T_l	0.15	0.15	0.15	0.15
k_u	3.5	3.5	3.5	79.0
k_l	5.0	5.0	5.0	80.0
k_b	20.0	700.0	100.0	200.0
k_c	0.5	2.0	2.0	2.0
x_{uo}	0.0179	0.0179	0.0179	0.0179
x_{lo}	0.018	0.018	0.018	0.018
x_{bo}	0.30	0.30	0.30	0.30
P_l	0.80	0.80	0.80	0.80

shown on the vertical axis while time is on the horizontal. Zero displacement is considered to be the glottal midline and any negative displacement implies collision with the opposite vocal fold. It can be seen in this figure that the cover and the body move together throughout a large portion of a cycle. This case produced a fundamental frequency of 95 Hz, a lower to upper amplitude ratio of 1.05, and a vertical phase difference of 36 deg/mm, which produces a mucosal wave velocity of 0.94 m/s. The phase difference in degrees is calculated according to Titze *et al.* (1993) by

$$\phi = 360\tau/T, \tag{37}$$

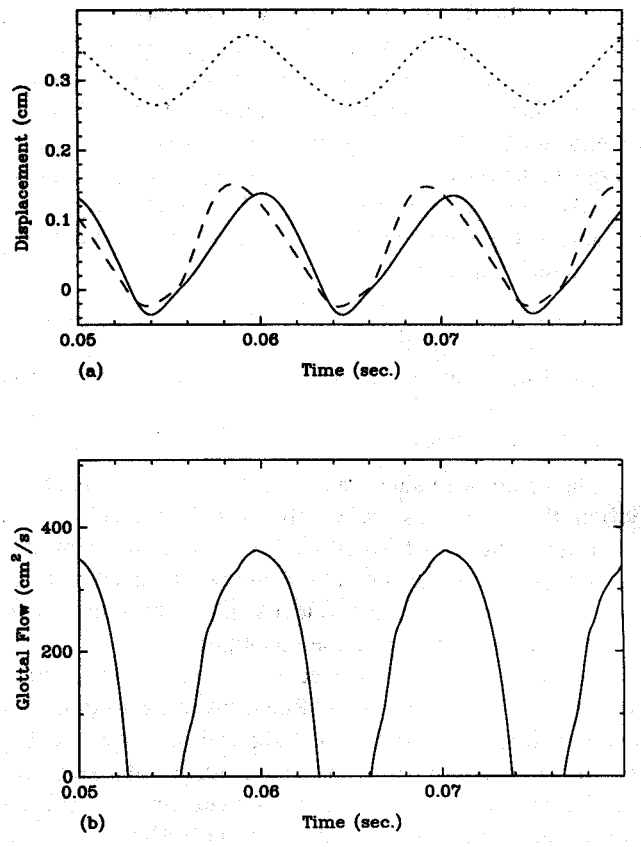

FIG. 6. Displacements of the three masses and glottal flow for case A in Table II, (a) displacement, (b) glottal flow.

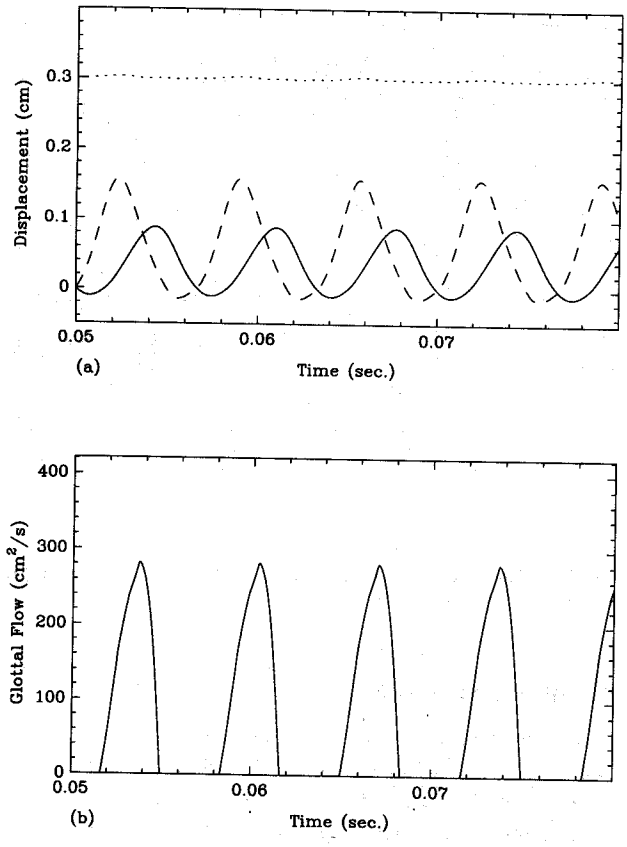

FIG. 7. Displacements of the three masses and glottal flow for case B in Table II, (a) displacement, (b) glottal flow.

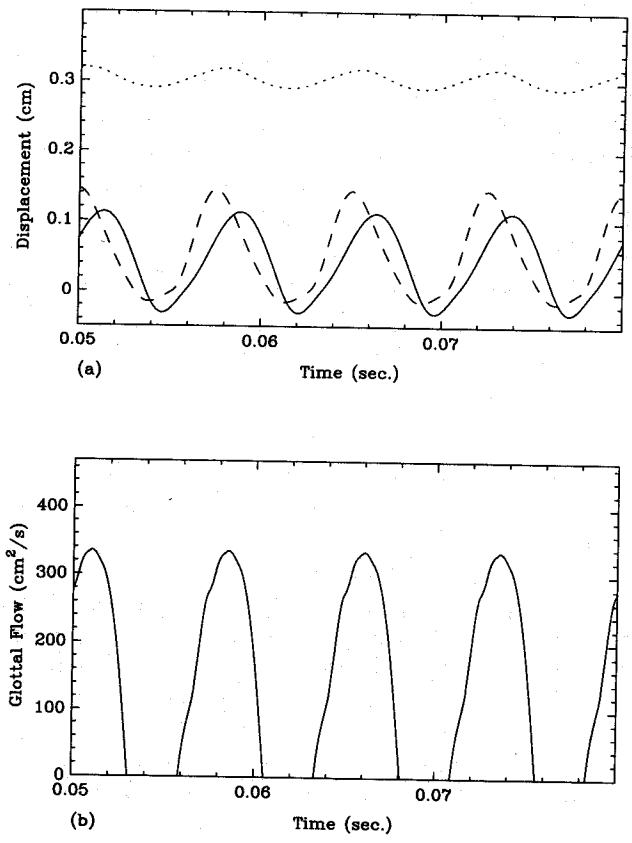

FIG. 8. Displacements of the three masses and glottal flow for case C in Table II, (a) displacement, (b) glottal flow.

where τ is the time delay between the two cover masses and T is the fundamental period of vibration. The mucosal wave velocity (c) is related to the phase difference by the following equation:

$$c = 360z/\phi T, \tag{38}$$

where z is the distance between the centers of the cover masses. Figure 6(b) is the glottal flow waveform for this case.

When the TA is contracted to a much greater degree than the CT, the stiffness of the body is much higher than that of the cover. In this case, which Hirano (1974) claims to be representative of loud heavy voices at medium pitch levels, the vibration presumably takes place mainly in the cover as sketched in Fig. 4(b). The simulation of this case is shown in Fig. 7. It is observed that there is very little movement of the body mass while considerable movement takes place in the cover masses. F_0 was 150 Hz with a vertical phase difference of 96 deg/mm. The amplitude ratio (lower to upper mass) was found to be 1.85 while the mucosal wave velocity was a low 0.32 m/s. In this case, with the high stiffness of the body, the three-mass model is effectively reduced to the structure of the IF72 model but with different parameter values based on the properties of the cover layer. Thus the "body-cover" model has been reduced to a "cover" model.

In case C the TA contraction is slightly more dominant than that of the CT. Vocal-fold movement will involve both the body and cover but may involve the cover slightly more. Hirano (1974) suggests that this case is considered to be "normal" phonation. The simulation results are shown in Fig. 8. It is observed that both the body and cover are involved in the motion with the cover having a somewhat greater amplitude. This case produced an F_0 of 134 Hz, a vertical phase difference of 43 deg/mm, and a mucosal wave velocity of 1.1 m/s. Titze *et al.* (1993) have recently reported vertical phase differences of 27–61 deg/mm with most of the data clustering in the 40–55-deg/mm range. They also measured mucosal wave velocities of 0.5–2.2 m/s. The three mass simulation for normal conditions agrees well with these numbers.

Figure 4(d) shows the conditions that Hirano (1974) suggests are used for falsetto voice. A large CT contraction imposes a large amount of passive stiffness on both the cover and the body. This means that k_u, k_l, and k_b, are all very large, but k_b is large because the deep layer of the lamina propria (part of the vocal ligament) now assumes the longitudinal tension rather than the muscle. The active tension in the muscle is nearly zero, the amplitude of vibration is very small, and there is little or no vertical phase difference. A recent study of the stress–strain properties of the human vocal ligament has shown that a stress of approximately 130 kPa is developed by a strain of 40%–50% (Min *et al.*, 1994). The calculation of the stiffness k_b for the falsetto case was based on this information. Also, the effective vibrating mass of the body is primarily the ligament and so m_b has been reduced.

The simulation in Fig. 9 shows very little movement in both the body and cover. The displacements of both of the

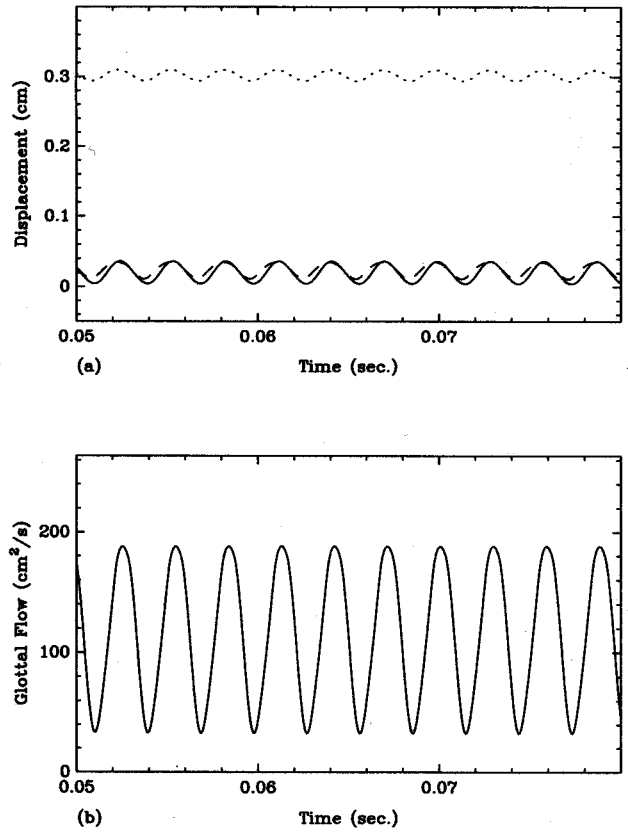

FIG. 9. Displacements of the three masses and glottal flow for case D in Table II, (a) displacement, (b) glottal flow.

cover masses are shown in this graph but they are nearly "locked" together in vibration; hence they have almost identical displacement waveforms. This means that there is very little vertical phase difference and a correspondingly high mucosal wave velocity. It should also be noted that the cover masses do not collide with the opposite vocal fold which results in a nearly sinusoidal glottal flow waveform with a significant dc offset [Fig. 9(b)]. This is quite typical of human falsetto production. The fundamental frequency has increased to 339 Hz which is an increase in frequency of more than a factor of 2.5 over the previous cases. The glottal flow in Fig. 9(b) is a nearly sinusoidal waveform with a significant dc offset resulting from the lack of collision.

Table III summarizes the results of these four cases in terms of resulting F_0, vertical phase difference (VPD), lower to upper amplitude ratio (AR), and mucosal wave velocity (MWV).

TABLE III. Simulation results of the four cases of Hirano (1974) in terms of resulting F_o, vertical phase difference, lower to upper amplitude ratio, and mucosal wave velocity.

Case	F_0 (Hz)	VPD (deg/mm)	AR	MWV (m/s)
A	95	36	1.05	0.9
B	150	96	1.85	0.3
C	134	43	1.29	1.1
D	339	17	0.84	6.9

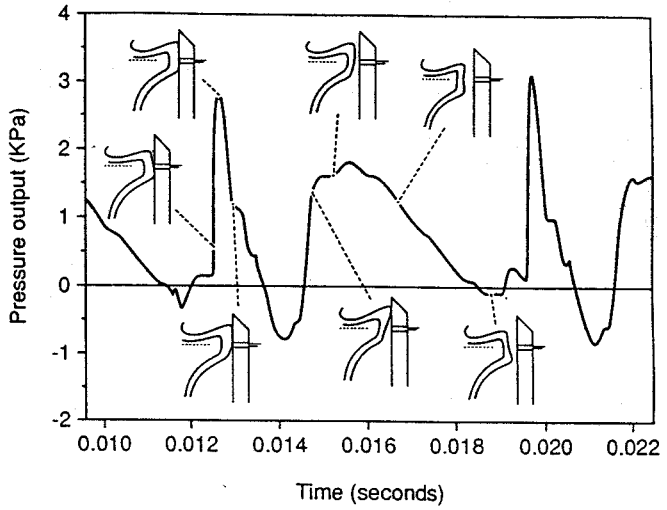

FIG. 10. Intraglottal pressure measure with excised canine hemilarynx setup (from Jiang and Titze, 1994).

B. Intraglottal pressure

As a further test of how the three-mass model can simulate characteristics of real vocal-fold vibration, intraglottal pressure calculated from the model was compared to recent measurements of this quantity. Jiang and Titze (1994) measured intraglottal pressure in an excise canine hemilarynx configuration. They obtained pressure curves like that shown in Fig. 10. This figure clearly shows the pressure peaks that result from the both the contact of the vocal folds and the aerodynamic processes. The sketches of vocal-fold position were derived from video recordings of both axial (top) and a sagittal views of vocal-fold vibration. The sagittal view provided estimates of contact area.

In the three-mass model the intraglottal pressure was defined to be the average pressure acting on the two cover masses. Using the parameters of case "c" (normal phonation) the intraglottal pressure pattern shown in Fig. 11(a) is produced. The peak at the beginning of the intraglottal pressure cycle (point A) is due to the successive collisions of the lower and upper masses. The collision forces then begin to subside as the folds start their opening phase. When the lower masses (left and right) break apart (point B), the aerodynamic pressure dramatically increases. The aerodynamic pressure then declines as the folds once again move toward closure (point C). When the right and left folds are very close to collision the pressure sinks to negative values (point D). This is the so-called Bernoulli effect. Comparing Figs. 10 and 11(a), there are obvious differences in the pressure waveforms. Some of these differences come from the fact that the experiment was performed with a real, continuous system and the model is simply a discretized representation of this system. Specifically, the "peakiness" seen in the modeled intraglottal pressure comes from discretizing the vertical aspect of the vocal folds into two sections. However, a difference that is not explained by the discretization is the presence of the large negative Bernoulli pressure in the model and its virtual absence in the experiment. It was suspected that this might be due to the fact that the model simulation

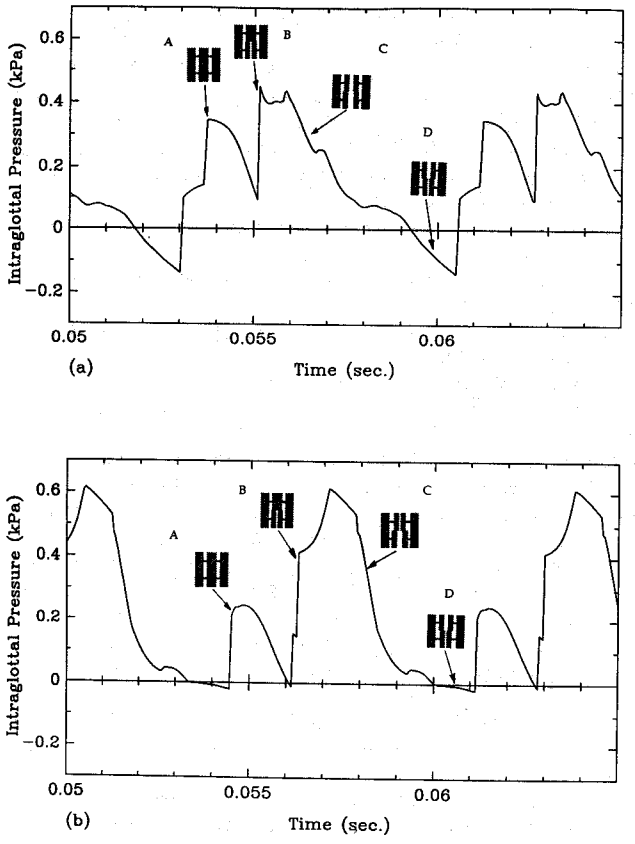

FIG. 11. Intraglottal pressure computed by the three mass model for case C in Table I. (a) includes a simulated human sized supraglottal and subglottal vocal tract, (b) Includes a simulation of the supraglottal and subglottal systems used in the experiment shown in Fig. 9.

was performed with a supra- and subglottal vocal tract that simulated the human system while the experiment was performed with an excised canine larynx without a vocal tract load. More specifically, the presence of the supraglottal vocal tract imposes an inertive load on the vibrating glottis at frequencies below the first formant. During the beginning of the open phase of the glottis, the glottal flow starts to move the mass of air in the vocal tract. As soon as the inertia of the vocal tract air mass is overcome, it will be put into forward motion. Now as the vocal folds begin to come together, the glottal area decreases, the glottal air velocity increases, and consequently the intraglottal pressure drops. Because of the inertia of the vocal tract air mass, it continues to move forward and tends to pull the air through the glottis, accentuating the increase in particle velocity and the decrease in intraglottal pressure.

The model was run again using a simulation of the supraglottal and subglottal setup used in the experiment. The supraglottal system was almost completely removed except for one section (0.79 cm) because a typical larynx dissection leaves cartilage about 0.5–1 cm higher than the vocal folds. The subglottal system consisted of a brass "tracheal" tube that was connected to a "pseudo-lung" (chamber of volume comparable to that of lungs). This was simulated by lengthening the subglottal system in the model and creating the chamber volume. The resulting simulation is shown in Fig. 11(b).

Now there is just a slight dip below zero in the Bernoulli phase (point D) which is quite compatible with the experimental curve shown in Fig. 10. It should also be noted that the fundamental frequency increased to 150 Hz in the simulation of the experimental setup. The fact that the F_0 will be higher without a vocal tract load is discussed in Ishizaka and Flanagan (1972) and in Titze (1988). It is also interesting to note that the decline in pressure slightly below zero following point A is also seen in Fig. 10 during the same phase of the cycle. This was not the case when the full supraglottal vocal tract was coupled to the three-mass model.

III. GENERAL DISCUSSION

A lumped-element model of the body-cover structure of the vocal folds has been presented. It builds on the widely used two-mass model of Ishizaka and Flanagan (1972) by adding a third mass to simulate the muscle tissue. Additionally, the values of the mass and stiffness parameters are based on measurements of vocal-fold tissue properties. This three-mass model is particularly useful as a research tool because it allows for physiologically relevant laryngeal adjustments and also simulates the interaction between the body and cover tissue layers. Contractions of the CT and TA muscles are incorporated in the values used for the stiffness parameters of the body and cover springs. Simulations have shown reasonable similarity to observed vocal-fold motion, measured vertical phase difference, and mucosal wave velocity, as well as experimentally obtained intraglottal pressure.

The three-mass model will reduce to a two-mass system when the stiffness of the body mass is set to be extremely large. Experience with this model has also shown that if the parameter values of the cover masses are set equal to those associated with the Ishizaka–Flanagan model, then their results are replicated.

ACKNOWLEDGMENT

This research was partially funded by Grant No. P60 DC00976 from the National Institutes on Deafness and Other Communication Disorders.

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

Alipour-Haghighi, F., Titze, I. R., and Perlman, A. L. (**1989**). "Tetanic contraction in vocal fold muscle," J. Speech Hear. Res. **32**, 226–231.

Alipour-Haghighi, F., and Titze, I. R. (**1983**). "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), pp. 183–190.

Baer, T. (**1975**). "Investigation of phonation using excised larynges," Doctoral dissertation, Massachusetts Institute of Technology, Cambridge, MA.

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.

Flanagan, J. L. (**1972**). *Speech Analysis, Synthesis, and Perception* (Springer-Verlag, New York).

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

Hirano, M., Kurita, S., and Nakashima, T. (**1981**). "The structure of the vocal folds," in *Vocal Fold Physiology*, edited by K. Stevens and M. Hirano (Univ. of Tokyo, Tokyo), pp. 33–41.

Hirano, M. (**1977**). "Structure and vibratory behavior of the vocal folds," in *Dynamic Aspects of Speech Production*, edited by M. Sawashima and F. S. Cooper (Univ. of Tokyo, Tokyo, Japan), pp. 13–27.

Hirano, M. (**1974**). "Morphological structure of the vocal cord as a vibrator and its variations," Folia Phoniatr. **26**, 89–94.

Ishizaka, K., and Isshiki, N. (**1976**). "Computer simulation of pathological vocal-cord vibration," J. Acoust. Soc. Am. **60**, 1193–1198.

Ishizaka, K., and Flanagan, J. L. (**1972**). "Synthesis of voiced sounds from a two-mass model of the vocal cords," Bell Syst. Tech. J. **51**, 1233–1268.

Ishizaka, K., and Matsudaira, M. (**1972**). "Fluid mechanical considerations of vocal cord vibration," Monogr. **8**, Speech Commun. Res. Lab., Santa Barbara, CA.

Isshiki, N. (**1977**). *Functional Surgery of the Larynx* (Kyoto University, Kyoto, Japan), pp. 62–67.

Jiang, J. J., and Titze, I. R. (**1994**). "Measurement of vocal fold intraglottal pressure and impact stress," J. Voice **8**, 132–144.

Kaneko, T., Asano, H., Naito, J., Kobayashi, N., Hayashi, K., and Kitamura, T. (**1972**). "Biomechanics of the vocal cords—On damping ratio," J. Jpn. Bronchoesophagol. Soc. **25**, 133–138.

Kelly, J., and Lochbaum C. (**1962**). "Speech synthesis," Prac. Fourth International Congress On Acoustics, Paper G42 (unpublished), pp. 4.

Koizumi, T., Taniguchi, S., and Hiromitsu, S. (**1987**). "Two-mass models of the vocal cords for natural sounding voice synthesis," J. Acoust. Soc. Am. **82**, 1179–1192.

Liljencrants, J. (**1985**). "Speech synthesis with a reflection-type line analog," D.S. dissertation, Dept. of Speech Comm. and Music Acous., Royal Inst. of Tech., Stockholm, Sweden.

Min, Y., Titze, I. R., Alipour-Haghighi, F., and Hoffman, H. (**1994**). (unpublished).

Perlman, A. L. (**1985**). "A technique for measuring the elastic properties of vocal fold tissue," Doctoral disseration, The University of Iowa, Iowa City, Iowa.

Scherer, R., and Titze, I. R. (**1983**). "Pressure flow relationships in a model of the laryngeal airway with a diverging glottis," in *Vocal Fold Physiology: Current Research and Clinical Issues*, edited by D. Bless and J. Abbs (College Hill, San Diego), pp. 179–193.

Sloan, S. H., Berke, G. S., Gerratt, B. R., Kreiman, J., and Ye, M. (**1993**). "Determination of vocal fold mucosal wave velocity in an in vivo canine model," Laryngoscope **103**, 947–953.

Smith, M. E., Berke, G. S., Gerratt, B. R., and Kreiman, J. (**1992**). "Laryngeal paralyses: Theoretical considerations and effects on laryngeal vibration," J. Speech Hear. Res. **35**, 545–554.

Titze, I. R. (**1973**). "The human vocal cords: A mathematical model, part I," Phonetica **28**, 129–170.

Titze, I. R. (**1974**). "The human vocal cords: A mathematical model, part II," Phonetica **29**, 1–21.

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. (**1984**). "Parameterization of the glottal area, glottal flow, and vocal fold contact area," J. Acoust. Soc. Am. **75**, 570–580.

Titze, I. R. (**1988**). "The physics of small-amplitude oscillation of the vocal folds," J. Acoust. Soc. Am. **83**, 1536–1552.

Titze, I. R., Luschei, E. S., and Hirano, M. (**1989**). "Role of the thyroarytenoid muscle in regulation of fundamental frequency," J. Voice **3**, 213–224.

Titze, I. R., Jiang, J. J., and Hsiao, T. Y. (**1993**). "Measurement of mucosal wave propagation and vertical phase difference in vocal fold vibration," Ann. Otol. Rhinol. Laryngol. **102**, 58–63.

van den Berg, J. W., Zantema, J. T., and Doornenbal, P., Jr. (**1957**). "On the air resistance and the Bernoulli effect of the human larynx," J. Acoust. Soc. Am. **29**, 626–631.

Wong, D., Ito, M. R., Cox, N. B, and Titze, I. R. (**1991**). "Observation of perturbations in a lumped-element model of the vocal folds with application to some pathological cases," J. Acoust. Soc. Am. **89**, 383–394.