Influence of acoustic loading on an effective single mass model of the vocal folds

Matías Zañartu[a)]
School of Electrical and Computer Engineering and Ray W. Herrick Laboratories, Purdue University, 140 S. Intramural Drive, West Lafayette, Indiana 47907

Luc Mongeau[b)]
School of Mechanical Engineering and Ray W. Herrick Laboratories, Purdue University, West Lafayette, Indiana 47907

George R. Wodicka
Weldon School of Biomedical Engineering and School of Electrical and Computer Engineering, Purdue University, West Lafayette, Indiana 47907

(Received 10 July 2006; revised 8 November 2006; accepted 20 November 2006)

Three-way interactions between sound waves in the subglottal and supraglottal tracts, the vibrations of the vocal folds, and laryngeal flow were investigated. Sound wave propagation was modeled using a wave reflection analog method. An effective single-degree-of-freedom model was designed to model vocal-fold vibrations. The effects of orifice geometry changes on the flow were considered by enforcing a time-varying discharge coefficient within a Bernoulli flow model. The resulting single-degree-of-freedom model allowed for energy transfer from flow to structural vibrations, an essential feature usually incorporated through the use of higher order models. The relative importance of acoustic loading and the time-varying flow resistance for fluid-structure energy transfer was established for various configurations. The results showed that acoustic loading contributed more significantly to the net energy transfer than the time-varying flow resistance, especially for less inertive supraglottal loads. The contribution of supraglottal loading was found to be more significant than that of subglottal loading. Subglottal loading was found to reduce the net energy transfer to the vocal-fold oscillation during phonation, balancing the effects of the supraglottal load. © *2007 Acoustical Society of America.* [DOI: 10.1121/1.2409491]

PACS number(s): 43.70.Bk, 43.70.Aj [AL] Pages: 1119–1129

I. INTRODUCTION

Mathematical models of human voice production may be used to study normal and pathologic phonation, to predict the effects of phonosurgery, or to aid voice therapy. Source-resonator coupling, produced by fluid-sound interactions at the glottis, is a key factor in the production of self-oscillations. Fluid-sound interactions affect voice efficiency, color the spectra of the voice source, and modify the stability of the dynamic voice production system, resulting in a change in oscillation frequency and onset pressure. Control of efficiency and spectral content can be used in singing (Titze, 2004). The instabilities produced by the flow-sound interaction could help to explain irregular phonation (Mergell and Herzel, 1997). The effects of subglottal tract loading may be critical, especially in studies involving excised larynx and synthetic models of the vocal folds, where the subglottal tract used to supply the flow may play a comparatively large role (Austin and Titze, 1997; Zhang *et al.*, 2006).

A. Mathematical models of voice production

Early models of vocal-fold vibrations were based on myoelastic-aerodynamic theory (Van den Berg, 1958). In the single-mass model of Flanagan and Landgraf (1968), the absence of the vocal tract made it impossible to obtain self-sustained oscillations. Earlier excised larynx experiments showed that self-sustained oscillation can be obtained in the absence of a vocal tract (Van den Berg and Tan, 1959). The effects introduced by the delay between the upper and lower portion of the vocal folds during oscillation were subsequently incorporated to create the well-known two-mass model (Ishizaka and Flanagan, 1972). This model allowed for acoustic feedback and self-sustained oscillations with or without the presence of acoustic loading. Multimass models have been developed to better account for the asymmetry between the lower and upper parts of the folds, and the presence of a mucosal wave. The "body-cover model" (Titze and Story, 1997) allowed the effects of acoustic loading on the volumetric flow rate to be considered using a wave reflection analog approach (Kelly and Lochbaum, 1962; Liljencrants, 1985; Rahim, 1994; Story, 1995). The flow model used in these investigations was shown to be stable and able to handle time-varying flow and irregular geometries (Titze, 1984). Fulcher *et al.* (2006) designed a so-called "effective one-mass model," i.e., a single-mass model of the vocal folds

[a)]Author to whom correspondence should be addressed. Electronic mail: mzanartu@purdue.edu
[b)]Present address: Department of Mechanical Engineering, McGill University, 817 Sherbrooke Street West, Montreal, Quebec, Canada, H3A 2K6.

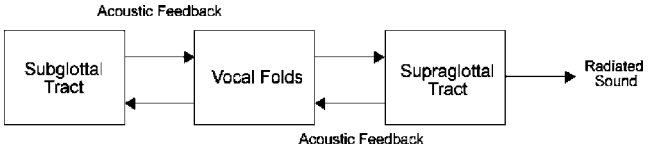

FIG. 1. Diagram of an interactive voice production system.

excited by a negative Coulomb damping force. The negative Coulomb damping force was assumed to be equivalent to the aerodynamic forces on the vocal folds, thus keeping the driving force in phase with tissue velocity. The pressure force for the inward, closing phase of the cycle was smaller by a factor of two than that for the outward opening phase. The model exhibited self-sustained oscillations without the need for acoustic loading. This model was conceptually accurate, but did not include a quantitative flow model, acoustic loading, or collision effects.

High-order models with thousands of degree of freedom (e.g., finite-element models) have provided additional insight on phonation mechanisms that are not captured by low-order models. The modal response of simplified vocal-fold models (Berry and Titze, 1996) and fluid-structure interactions (Thompson *et al.*, 2005) have been studied using the finite-elements method. Such models often use an incompressible flow formulation, and ignore acoustic loading effects. Simplified, unrealistic boundary conditions such as constant pressure or constant velocity are typically imposed at the inflow and outflow boundaries of the computational domain, ignoring the effects of sound waves in the subglottal and supraglottal spaces on transglottal pressures.

B. Previous observations of acoustic coupling in phonation

A schematic illustrating the acoustic coupling problem in voice production is shown in Fig. 1. The interactions between fluid flow through the glottis and radiated sound waves are regulated by the impedances of the glottis and the vocal tract (Rothenberg, 1981; Fant and Lin, 1987; Titze and Story, 1997). The impedance of the supraglottal tract has been argued to be primarily governed by the impedance of the lower tract, i.e., the epilarynx. Narrow entry sections in the vocal tract result in a high impedance, comparable with the impedance of the glottis, which favors flow-sound interactions, (Titze and Story, 1997). It has been shown that under these conditions, flow-sound interactions affect the volumetric flow rate waveform. Skewing, ripple or depressions in the open phase, as well as short-term variations of formant frequencies and bandwidths, were observed in cases with strong interactions (Rothenberg, 1981; Fant and Lin, 1987). Phonation is known to be facilitated by the combined effects of the time-varying geometry of the orifice and a delayed vocal-tract response. The effective damping of vocal-fold vibration decreases with vocal-tract inertance (positively reactive) and increases with vocal-tract resistance. In addition, a net compliant vocal-tract load (negatively reactive) tends to squelch oscillation (Titze, 1988).

Subglottal tract coupling effects are less well understood. Experimental studies using excised larynges (Austin and Titze, 1997) and physical rubber models (Zhang *et al.*, 2006) indicate that a subglottal tract length much longer than the physiological length is required to obtain self-sustained oscillations. Different subglottal tract configurations were found to affect the phonation threshold pressure. Strong interactions between laryngeal dynamics and subglottal acoustics were observed. The results of these studies show the importance of including the subglottal tract in laryngeal dynamics.

C. Research objectives

In the present study, the factors that control the coupling between the vibration of the vocal folds, the sound field in the subglottal and supraglottal tracts, and the airflow through the larynx were investigated. The goal was to determine how significant acoustic loading effects were relative to the time-varying flow resistance of the glottis during normal phonation. The individual contributions of the vocal tract and the subglottal tract to the overall acoustic load impedance were identified. In order to achieve these goals, a dynamic model of phonation that included the effects of the sound field in the vocal tract and the supraglottal tract for several configurations was developed.

II. EFFECTIVE SINGLE-DEGREE-OF-FREEDOM MODEL

A. Model description

An enhanced version of the "effective one-mass model" of Fulcher *et al.* (2006) was used. The driving force associated with the negative Coulomb damping was defined based on the aerodynamic forces acting on the vocal folds. Fluid-structure interactions, fluid-sound interactions, and collision effects were included. Figure 2(a) shows schematics of the vocal-fold model. The motion of the folds was assumed to be bilaterally symmetric. The equation of motion for each fold is given by

$$m\ddot{y} + b\dot{y} + k(y - y_o) = F_P + F_H, \tag{1}$$

where m is the body mass of one vocal fold, y is the displacement of the mass at the glottal end, y_o is the equilibrium position of the mass without the action of external forces, b is the viscous damping constant, and k is the spring constant. The terms F_P and F_H are fluid pressure and impact forces, as described in detail in the following sections.

B. Pressure force

The net medial-lateral fluid pressure force, F_P, applied to the walls of the vocal folds is the main driving force for the oscillator. The wall pressure is affected by fluid-sound interactions, which cause instantaneous changes in transglottal pressures, and the air flow. The flow was assumed to be inviscid, irrotational, and approximately incompressible. The "quasi-steady approximation" was made, i.e., the time-varying flow was treated as a sequence of stationary flows for similar geometries and boundary conditions at each time step. Under these conditions, Bernoulli's equation is valid along a streamline in the region from the subglottal to the

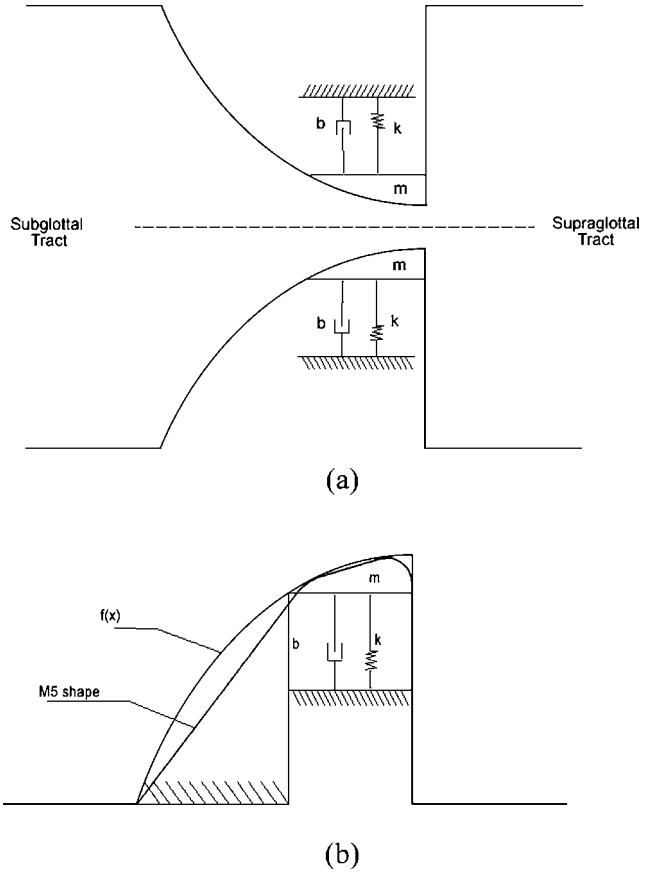

FIG. 2. Basic diagram of the single-degree-of-freedom. (a) Symmetric representation. (b) Comparison with M5 model profile (Scherer *et al.*, 2001).

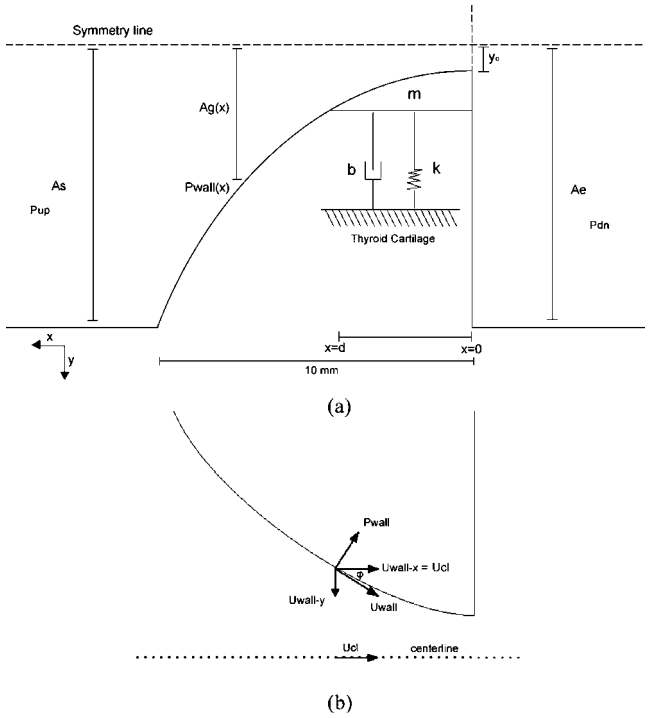

FIG. 3. Details considered to compute the pressure force. (a) Areas and pressures in the glottis. (b) Pressure and velocity along the wall.

glottal opening. The flow is assumed to be detached from the wall at the orifice discharge to form a jet. Aerodynamic effects after/at the glottal opening (pressure recovery, dynamic head loss, momentum diffusion, etc.) are included through the use of time-varying orifice discharge coefficient (herein referred to as ODC). The wave reflection analog supplies the subglottal and supraglottal acoustic pressures used in the calculations. From conservation of mass, the relation between subglottal and glottal flow velocities is

$$A_s u_{\mathrm{up}} = A_g(x) u_{\mathrm{cl}}(x), \tag{2}$$

where A_s and u_{up} are subglottal cross-sectional area and upstream (subglottal) pressure. $A_g(x)$ and $u_{\mathrm{cl}}(x)$ are functions that describe cross-sectional area and centerline velocity along the depth of the folds (or length in the flow direction). For simplicity, the function that describes the glottal area $A_g(x)$ is assumed to be quadratic. This area is obtained using a shape function $f(x)$, such that $A_g(x)$ is equal to $f(x)$ times the length of the folds. A 2D comparison between this geometry and the M5 geometry of Scherer *et al.* (2001) is shown in Fig. 2(b). To relate pressure and velocity, Bernoulli's equation was used along the centerline (or symmetry line) between the subglottal section [$x=d$ in Fig. 3(a)] and the opening of the vocal folds [$x=0$ in Fig. 3(a)]. The pressure distribution and centerline velocity along the depth of the folds, $p_{\mathrm{cl}}(x)$ and $u_{\mathrm{cl}}(x)$, are related through

$$\frac{1}{\rho} p_{\mathrm{cl}}(d) + \frac{1}{2} u_{\mathrm{cl}}^2(d) = \frac{1}{\rho} p_{\mathrm{cl}}(0) + \frac{1}{2} u_{\mathrm{cl}}^2(0), \tag{3}$$

where ρ is the air density. The total pressure at the different locations is known to be the sum of static and dynamic pressures (Mongeau *et al.*, 1997; Zhang *et al.*, 2002; Thompson *et al.*, 2005; Park and Mongeau, 2007). Thus, assuming the total head at $x=d$ is equal to the subglottal head, and that the centerline pressure at $x=0$ is equal to the supraglottal value, Eq. (3) can be rewritten as

$$\frac{1}{\rho}(p_{\mathrm{up}} + p'_{\mathrm{up}}) \frac{1}{2} u_{\mathrm{up}}^2 = \frac{1}{\rho}(p_{\mathrm{dn}} + p'_{\mathrm{dn}}) + \frac{1}{2} u_{\mathrm{cl}}^2(0), \tag{4}$$

where p_{up} and p'_{up} are the upstream steady and upstream acoustic pressures, u_{up} is the upstream flow velocity, and p_{dn} and p'_{dn} are the downstream steady and downstream acoustic pressures. Using Eq. (4) to relate centerline velocities between the orifice and the subglottal section, and defining β as the area ratio, then

$$u_{\mathrm{up}} = u_{\mathrm{cl}} \frac{A_g}{A_s} = u_{\mathrm{cl}} \beta, \tag{5}$$

where for simplicity $u_{\mathrm{cl}}=u_{\mathrm{cl}}(0)$ and $A_g=A_g(0)$. Note that all equations up to this point correspond to ideal flow. Thus, substitution of Eq. (5) in Eq. (4) yields the ideal velocity at the glottal opening,

$$u_{\mathrm{cl(ideal)}} = \sqrt{\frac{2\Delta p + (p'_{\mathrm{up}} - p'_{\mathrm{dn}})}{\rho(1-\beta^2)}}, \tag{6}$$

where Δp is the difference between the static pressure upstream and downstream. Since the volumetric flow rate Q

must be constant across the system, it can be computed at the glottal opening as

$$Q = u_{\mathrm{cl(ideal)}} A_g. \tag{7}$$

Aerodynamic effects that occur after/at the glottal opening, such as pressure recovery, dynamic head loss, and momentum diffusion may be characterized through the use of an empirical ODC (Zhang *et al.*, 2002; Park and Mongeau, 2007). The actual volumetric flow rate is

$$Q = c_d(t) u_{\mathrm{cl}} A_g, \tag{8}$$

where $c_d(t)$ is the time-varying ODC. Using Eq. (6), the flow velocity at the opening, from Eq. (7) is

$$u_{\mathrm{cl(actual)}} = c_d(t) \sqrt{\frac{2\Delta p + (p'_{\mathrm{up}} - p'_{\mathrm{dn}})}{\rho(1-\beta^2)}}. \tag{9}$$

This last expression was used with Eqs. (5) and (2) to obtain corrected expressions for the upstream flow velocity, u_{up}, and center velocity along the flow direction, $u_{\mathrm{cl}}(x)$. Bernoulli's flow equation between the subglottal section and any point in the glottis along the centerline yields the centerline pressure distribution as a function of the position in the glottis,

$$p_{\mathrm{cl}}(x) = p_{\mathrm{up}} + p'_{\mathrm{up}} + \frac{\rho}{2}(u_{\mathrm{up}}^2 - u_{\mathrm{cl}}^2(x)). \tag{10}$$

The axial flow velocity is approximately constant along x, which implies that the x component of the flow velocity near the wall is the same as along the centerline, as illustrated in Fig. 3(b). The relation between centerline velocity (u_{cl}) and the wall velocity (u_{wall}) can be obtained from the wall geometry, as shown in Fig. 3(b), yielding

$$u_{\mathrm{cl}}(x) = u_{\mathrm{wall}}(x)\cos\varphi, \tag{11}$$

$$\tan\varphi = \frac{d}{dx}(f(x)) = f'(x), \tag{12}$$

$$u_{\mathrm{wall}}(x) = \frac{u_{\mathrm{cl}}(x)}{\cos(\arctan(f'(x)))}. \tag{13}$$

Using Bernoulli's flow equation again for any pair of points in the glottis, the wall pressure over the folds as function of position, $p_{\mathrm{wall}}(x)$, is

$$p_{\mathrm{wall}}(x) = p_{\mathrm{up}} + p'_{\mathrm{up}} + \frac{\rho}{2}(u_{\mathrm{up}}^2 - u_{\mathrm{wall}}^2(x)). \tag{14}$$

The pressure force acting on the folds in Eq. (1) is the integral of the normal pressure $p_{\mathrm{wall}}(x)$ over the surface of the folds. Therefore,

$$F_p = L\int_0^d p_{\mathrm{wall}}(x)dx. \tag{15}$$

C. Orifice discharge coefficient (ODC)

Experimental data are available for the ODC of typical glottal orifices (Park and Mongeau, 2007; Zhang *et al.*, 2002; Scherer *et al.*, 2001). Values for the discharge coefficient were selected from Park and Mongeau (2007) for an opening converging case and a closing divergent case. Values were selected for Re=7000 since the quasi-steady approximation was found to be accurate for that range. Peak values of the discharge coefficient were c_d=0.85 for the opening phase and c_d=1.34 for the closing phase, i.e.,

$$c_d(t) = \begin{cases} 0.85, & v(t) > 0 \\ 1.34, & v(t) \le 0. \end{cases} \tag{16}$$

Such a step function change is physiologically unrealistic, but consistent with the Coulomb force excitation in Fulcher's original model. The following smoothing function was used in the present study to make the time evolution of the ODC more regular:

$$c_d(t) = 1.095 - 3.55\frac{v}{|v_{\max}|}, \quad v_{\max} \neq 0. \tag{17}$$

D. Collision force

A Hertz impact force (Stronge, 2000) was used to describe the collision force, F_H in Eq. (1). The impact force on each fold is given by

$$F_H = \frac{4}{3}\delta^{3/2}(1 + b_H\dot{\delta})\left(\frac{E\sqrt{r}}{1-\mu^2}\right), \tag{18}$$

where δ is the penetration of each vocal fold through the contact plane, which corresponds to the absolute value of the position of the mass from the centerline y=0. Similarly, $\dot{\delta}$ is the magnitude of the velocity of the mass during collision. The relevant vocal-fold material properties are the damping constant b_H, the Young modulus E, and the Poisson ratio μ. The parameter r is the radius of curvature of the colliding surface, which is approximately the depth of the folds, d. The values of the parameters required to compute F_H are shown in Table I. An increment in the damping ratio η_H of the oscillator was included following the guidelines suggested by Story and Titze (1995). The damping ratio, η_H, was incremented by 40% with respect to its normal value. Note that the damping ratio defines the damping constant b, as expressed during collision as

$$b_H = 2\eta_H\sqrt{mk}. \tag{19}$$

The pressure force was adjusted, since it acts onto part of the folds surface during collision. The pressure force during collision was assumed to be the product of the upstream pressure and the effective surface in contact with the fluid. From the model geometry, shown in Fig. 2, the supraglottal pressure does not affect the vibration of the vocal folds during collision. This is in contrast with models allowing convergent/divergent modes of vibration (e.g., the two-mass model or the body-cover model), in which both subglottal and supraglottal pressures can act over the vocal-fold surface during this phase of the cycle.

E. Parameter values

The values for the parameters used in the model were largely taken from previous studies (Titze, 2002; Story and

TABLE I. Parameters and constants used in the vocal-fold model for the numerical simulations.

Parameter	Symbol	Nominal value
Length of the vocal folds	L	0.015 [m]
Depth of the vocal folds	d	0.003 [m]
Equilibrium position of the mass	y_o	0.0001 [m]
Mass of one fold	m	2×10^{-4} [kg]
Stiffness (spring constant)	k	200 [N/m]
Damping ratio	η	0.1
Damping ratio (collision)	η_H	0.5
Subglottal lung pressure	p_{up}	800 [Pa]
Supraglottal pressure	p_{dn}	0 [Pa]
Density of air	ρ	1.15 [kg/m^3]
Speed of sound	c	350 [m/s]
Young modulus of the folds	E	8 [kPa]
Poisson ratio	μ	0.4

Titze, 1995; Berry and Titze, 1996). These values are shown in Table I. The values for the areas of the supraglottal tract were from Story *et al.* (1996). A simplified flow diagram of the source model is shown in Fig. 4.

III. IMPLEMENTATION OF THE WAVE REFLECTION ANALOG TECHNIQUE

The wave reflection analog technique (Kelly and Lochbaum, 1962; Liljencrants, 1985; Rahim, 1994; Story, 1995) is a time-domain description of the propagation of one-dimensional, planar acoustics waves through a collection of uniform, consecutive cylindrical tubes. Little has been done to model sound in the subglottal system using this technique, contrasting the wealth of information for the vocal tract. A subglottal area function along with an adequate subglottal attenuation factor are proposed in this section.

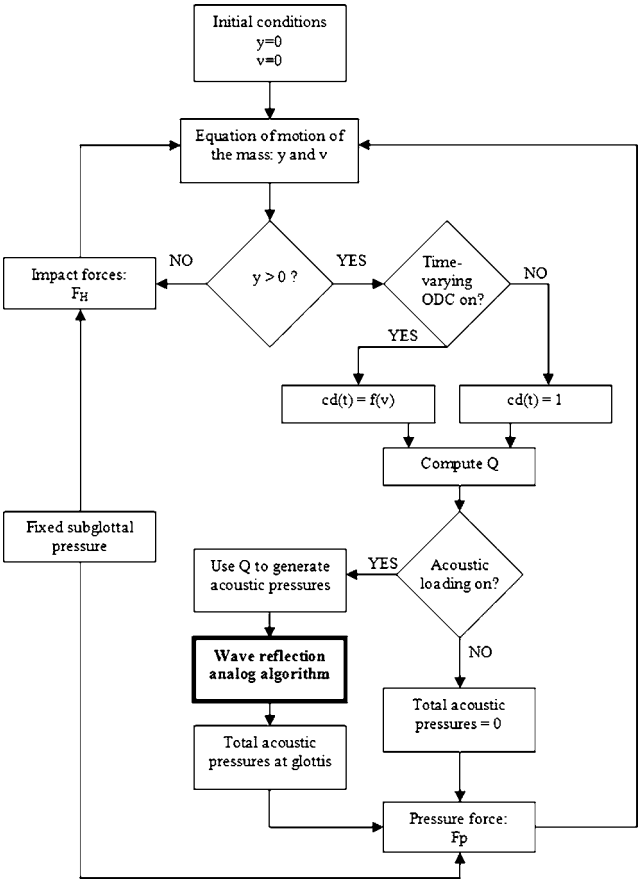

FIG. 4. Simplified flow diagram of the source model computation.

A. Subglottal area function

The area function for the subglottal tract was based on anatomical studies of the respiratory system (Weibel, 1963), in which the total cross-sectional area of the airways was measured as a function of the distance from the larynx using silicone rubber casts and excised adult lungs. Based on these data, an equivalent area function with no tube branching was denned. Although tube branching can potentially be included using a wave reflection analog approach, it was not included in this study. To compensate the absence of branching, a boundary correction (the boundary appears partially closed), first introduced by Titze (1984), was imposed at the terminal section of the equivalent subglottal model. The area denned for the terminal section produces a reflection coefficient ($r \approx 0.8$) that yielded the targeted subglottal resonances and bandwidths. Table II shows the values of the area function for 62 cylindrical sections of length 3.9683 [mm] and a sampling frequency of 44.1 [kHz].

TABLE II. Proposed subglottal area function with uniform sections of length 3.9683 [mm].

Section	Cross-sectional area in [cm]^2
1-10	2.7; 2.7; 2.7; 2.7; 2.7; 2.7; 2.7; 2.7; 2.7; 2.7;
11-20	2.7; 2.7; 2.7; 2.7; 2.7; 2.7; 2.7; 2.7; 2.7; 2.7;
21-30	2.7; 2.7; 2.7; 2.7; 2.7; 2.7; 2.7; 2.7; 2.7; 2.7;
31-40	2.2; 2.2; 2.2; 2.2; 2.2; 2.2; 2.2; 2.2; 2.2; 2.2;
41-50	2.2; 2.0; 2.0; 2.0; 2.0; 2.0; 2.0; 1.8; 1.8; 1.8;
51-60	2.7; 2.7; 2.7; 3.6; 3.6; 3.6; 4.9; 4.9; 6.7; 6.7;
61-61	9.8; 1.1;

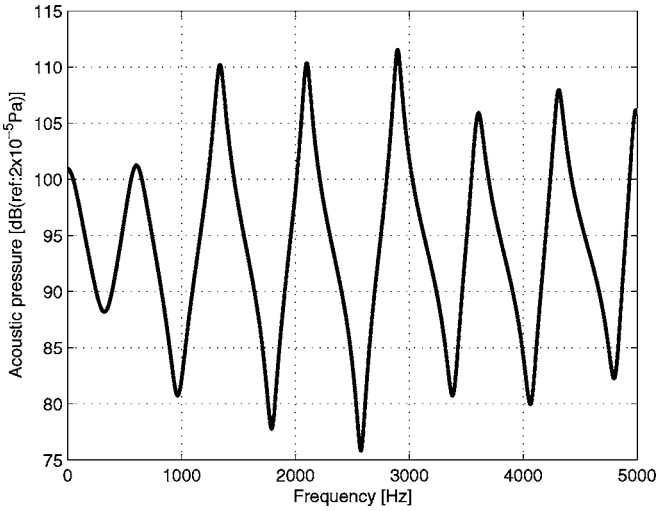

FIG. 5. Subglottal tract: Spectrum of the acoustic pressure at the glottis. Tract excited by a 1-Pa impulse.

B. Subglottal attenuation factor

Yielding walls, viscous dissipation, and heat conduction losses are the main sources of sound energy dissipation within the tracts. Following a procedure similar to that for the vocal tract (Rahim, 1994), heat conduction losses were characterized by an exponential attenuation factor. Viscous friction and heat conduction were assumed to be the same as for the vocal tract. Assuming that the main yielding walls losses in the subglottal tract are observed below 3000 Hz, the average attenuation factor was assumed to be

$$\alpha \approx \frac{0.0112}{\sqrt{A_K}}, \quad \text{and} \tag{20}$$

$$\gamma_K = e^{l\alpha_K} \approx 1 - l_{\alpha K}, \tag{21}$$

where α_K is the attenuation factor in cm^{-1}, A_K is the cross-sectional area of the kth section, γ_K is the propagation loss factor for the kth section, and l is the length of all sections.

C. Subglottal transfer function and verification studies

Figure 5 shows the transfer function of the subglottal system, constructed using the impulse response of the tract at the glottis with a sampling frequency of 44.1 KHz. It can be observed that the subglottal resonances are 613, 1341, 2100, and 2896 Hz; the half-power band widths are between 200 and 300 Hz for all formants. These values are in agreement with the findings of previous studies of subglottal acoustics (Harper *et al.* 2001, 2003; Stevens, 2000). The accuracy of the wave reflection analog scheme was evaluated through comparisons with predictions from harmonic wave propagation in tubes with simple geometries. The effects of different boundary conditions, including termination impedances, were evaluated as well as the effects of the loss factor. Acoustic coupling effects between subglottal and supraglottal tracts on vowel spectra were investigated. The details of these studies are available from Zañartu (2006).

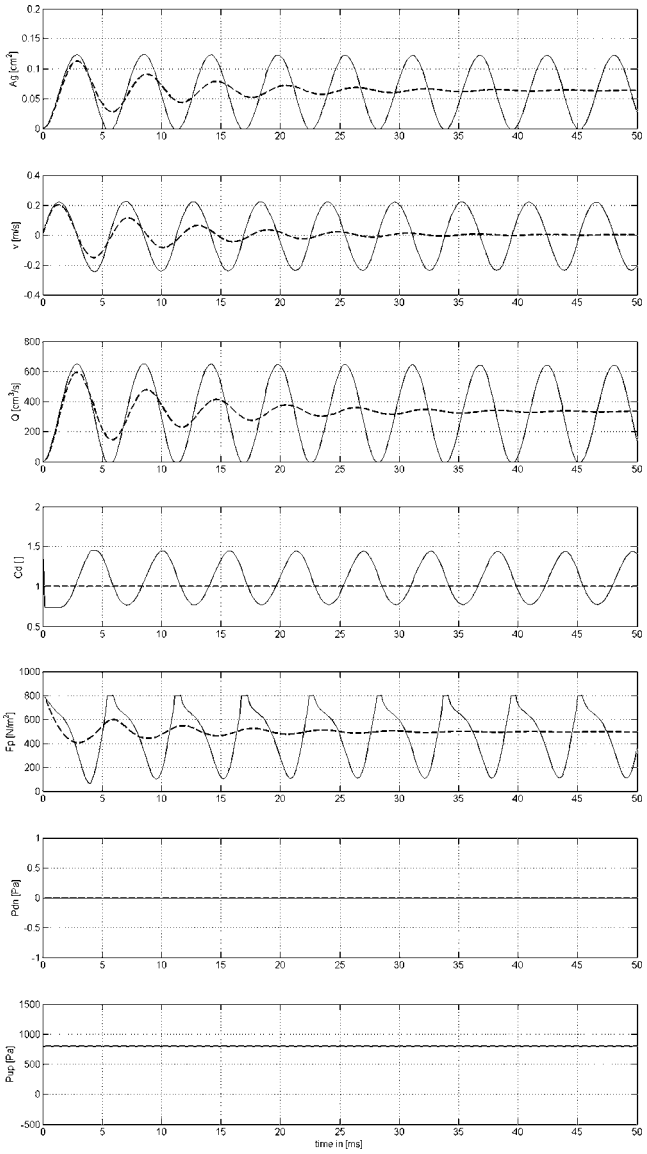

FIG. 6. Time history of key variables with no acoustic loading. A_g: glottal area [cm^2]; v: velocity of the mass [m/s]; Q: volumetric flow rate [cm^3/s]; c_d: orifice discharge coefficient; F_p: pressure force [N/m^2]; P_{dn}: downstream pressure [Pa]; P_{up}: upstream pressure [Pa]. Nomenclature: —: with ODC; - - -: without ODC.

IV. RESULTS

Section IV A describes the relative importance of aerodynamic and acoustic pressures with respect to flow-induced oscillations. Section IV B describes the contribution of each tract to the overall acoustic pressure excitation.

A. Relative influence of aerodynamic factors and acoustic coupling on stability

Three different acoustic loads were considered: (1) no acoustic loading; (2) subglottal tract and a close front unrounded vowel /i/ (IPA 301); and (3) subglottal tract and an open back unrounded vowel /a/ (IPA 305). These vowels were selected because their first formants are significantly different. The corresponding area functions were taken from Story *et al.* (1996). The area function describing the subglot-

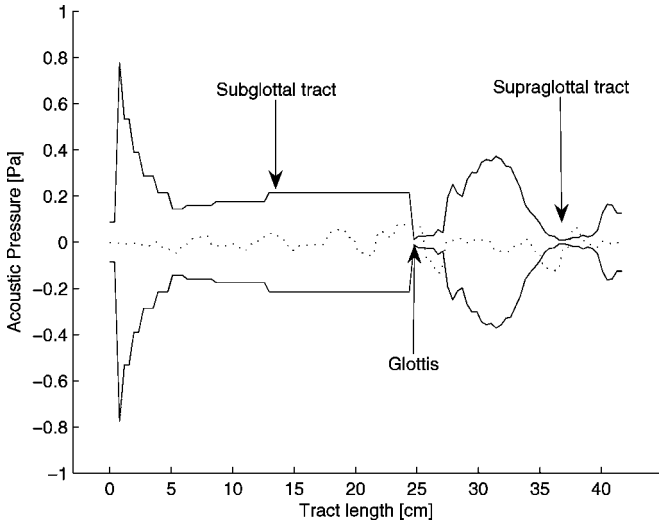

FIG. 7. Complete system geometry using the vowel /i/ with a snapshot of the acoustic pressure amplitude vs position at $t=5$ ms superimposed.

tal tract is shown in Table I. Results were obtained with and without the effects of the time-varying ODC.

1. No acoustic loading

The time history of key variables is presented in Fig. 6 for cases with and without the effect of the time-varying ODC. The subglottal static pressure was maintained at 800 Pa. In Fig. 6 (A_g) corresponds to the glottal opening obtained from the displacement of the mass. The volumetric flow rate (Q) was computed using several variables of the model as in Titze (1984, 2002). The ODC (c_d) and the pressure force (F_P) were computed using Eqs. (17) and (15). The downstream pressure (p_{dn}) is the acoustic pressure in the vocal tract at the glottis. The upstream pressure (p_{up}) was obtained as for the subglottal tract, but adding the static lung pressure. The fundamental frequency (f_0) was 180 Hz. This frequency was greater than the natural frequency of oscillation, which was expected to be 160 Hz. Self-oscillations were only obtained when the time-varying ODC was used. From the time history, it can be seen that the time-varying ODC is in phase with the velocity of the mass. Collision effects were not significant. Since there was no acoustic loading, the acoustic pressures upstream and downstream were negligible.

2. Vowel /i/ with subglottal tract

The supraglottal acoustic loading defined by the vowel /i/ had a first formant at 225 Hz and a second at 2486 Hz. The fact that the first formant is relatively close to the fundamental frequency makes this case particularly interesting. The vowel /i/ presents a loading that appears to be closer to a resistive loading than an inertive one. From Story *et al.* (1996), the section connected to the glottis is very narrow, with area 0.33 cm^2. This produces high vocal-tract impedance and strong coupling, given that the acoustic pressure within the tract is large. The subglottal tract had its first two resonances at 613 and 1341 Hz, and a cross-sectional area of 2.7 cm^2 at the glottis (see Table I). The model geometry is

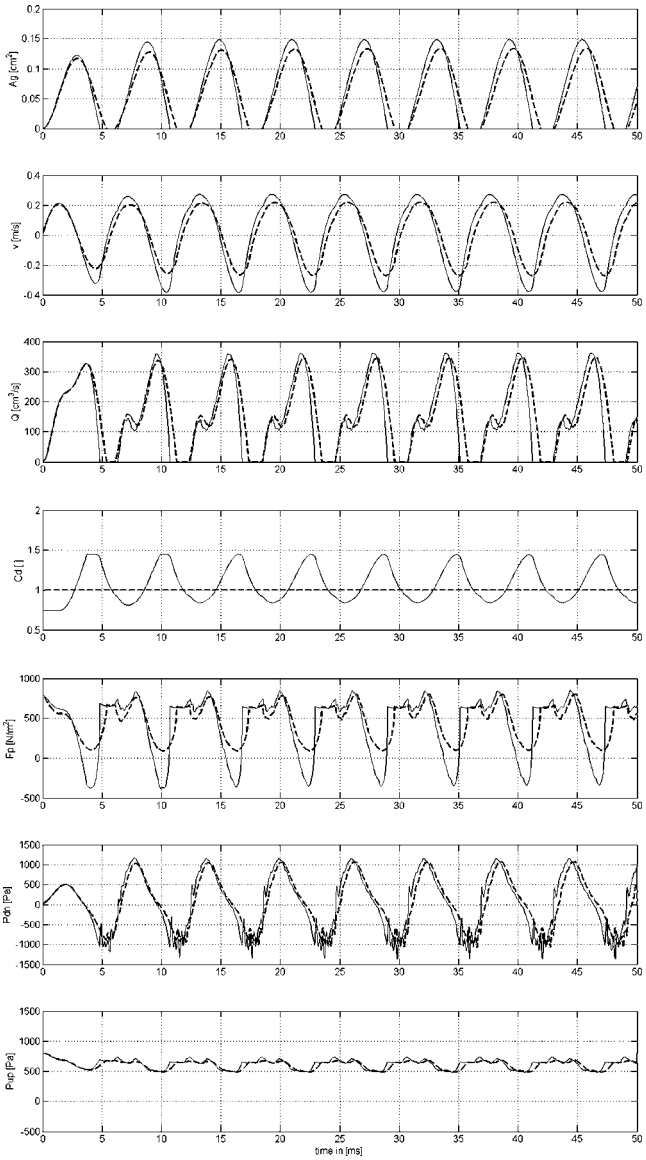

FIG. 8. Time history of key variables for vowel /i/ with subglottal tract. A_g: glottal area [cm^2]; v: velocity of the mass [m/s]; Q: volumetric flow rate [cm^3/s]; c_d: orifice discharge coefficient; F_P: pressure force [$\mathrm{N/m}^2$]; P_{dn}: downstream pressure [Pa]; P_{up}: upstream pressure [Pa]. Nomenclature: —: with ODC; - - -: without ODC.

illustrated in Fig. 7. The time history of key variables is presented in Fig. 8, using the same nomenclature as before. The oscillation reached a steady state after the second cycle, regardless of the effects of the time-varying ODC. This indicates that acoustic loading was the main factor that led to self-sustained oscillations. The collisions were larger than for the no-load case. A decrease (relative to the no-load case) in fundamental frequency, which was 170 Hz, was observed. The amplitude of the flow rate was reduced and its waveform was slightly skewed rightward in time. A depression was observed in the flow rate during the opening phase, mainly caused by the large peak in the downstream pressure. This peak is associated with the pressure oscillation associated with the low-frequency, first supraglottal formant. Although the subglottal tract was not responsible for the production of this depression, simulations without the subglottal tract suggested a significant contribution. The presence of the sub-

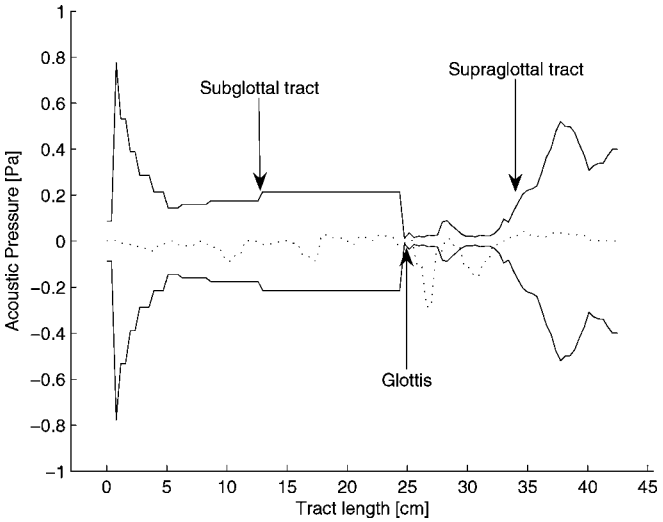

FIG. 9. Complete system geometry using the vowel /A/ with a snapshot of the acoustic pressure amplitude vs position at $t=5$ ms superimposed.

glottal system slightly reduced the depth of this depression, suggesting that it may have a damping effect on the volumetric flow rate oscillations. The contribution of each tract is studied in Sec. IV B. The coupling between source and tracts was evaluated by comparing the pressure force with the downstream and upstream pressures. During the opening phase, the downstream pressure reached a peak value and decayed almost in phase with the pressure force and the mass velocity, with a slight time lead. During collision, faster oscillations in the downstream pressure occurred, mostly associated with the high second formant frequency of the supraglottal tract. A clear coupling between source and subglottal tract was observed during collision.

The data obtained using the time-varying ODC were in close agreement with those from other interactive models of voice production (Rothenberg, 1981; Fant and Lin, 1987; Alipour *et al.*, 2000; Story and Titze, 1995; Story, 2002). The introduction of the time-varying ODC did not have a significant impact on most of the variables. This suggests that the effects of the fluid-structure interactions were less significant than the fluid-sound interaction in this simulation.

3. Vowel /a/ with subglottal tract

The supraglottal load given by the vowel /a/ has a first formant at 786 Hz and a second at 1147 Hz. From Story *et al.* (1996), the cross-sectional area at the glottis was 0.45 cm^2. The first formant is much greater than the fundamental frequency of oscillation, implying that the supraglottal load impedance is inertive. The subglottal tract was the same as in the previous case. A schematic of the system and its sound field is shown in Fig. 9. The time history of several variables is shown in Fig. 10. As for the previous case, the oscillation reached a steady state after the second cycle. The collisions were comparable, and the fundamental frequency of oscillation was 190 Hz. A ripple during the opening phase of the volumetric flow rate was evident, in place of the large depression previously observed. This phenomenon can be explained by the increased inertance of the supraglottal tract due to the greater first formant frequency. As for the vowel

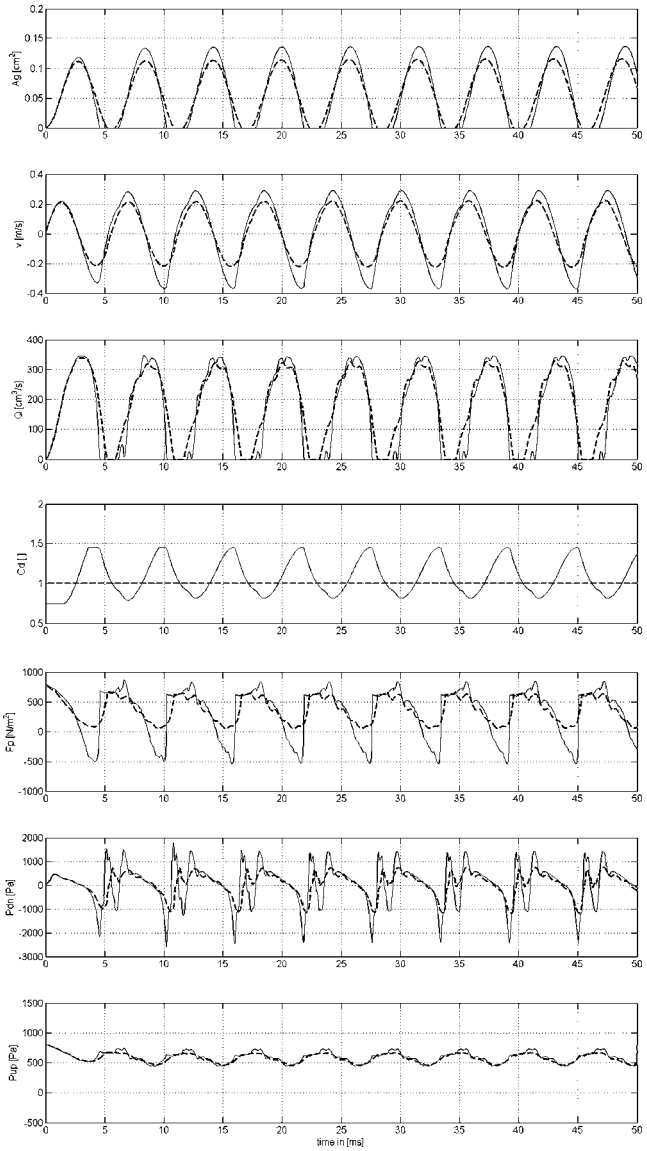

FIG. 10. Time history of key variables for vowel /a/ with subglottal tract. A_g: glottal area [cm^2]; v: velocity of the mass [m/s]; Q: volumetric flow rate [cm^3/s]; c_d: orifice discharge coefficient; F_P: pressure force [N/m^2]; P_{dn}: downstream pressure [Pa]; P_{up}: upstream pressure [Pa]. Nomenclature: —: with ODC; - - -: without ODC.

/i/, the reduced variation in the upstream pressure confirmed that the subglottal tract did not affect the production of the ripple, but it did reduce its magnitude. A detailed analysis of the contribution of the subglottal tract is presented in Sec. IV B. In comparison with the previous vowel, a higher degree of coupling with the supraglottal tract was observed since the structure of the downstream pressure closely resembled that of the pressure force.

Acoustic loading yielded self-sustained oscillations. Only minor variations were observed when the effects of fluid-structure interactions were added. The data were again in excellent agreement with those reported in other interactive models of voice production. The better coupled load (inertive vocal tract) produced less pronounced variations in the volumetric flow rate, thus introducing fewer changes in the sound source due to acoustic coupling.

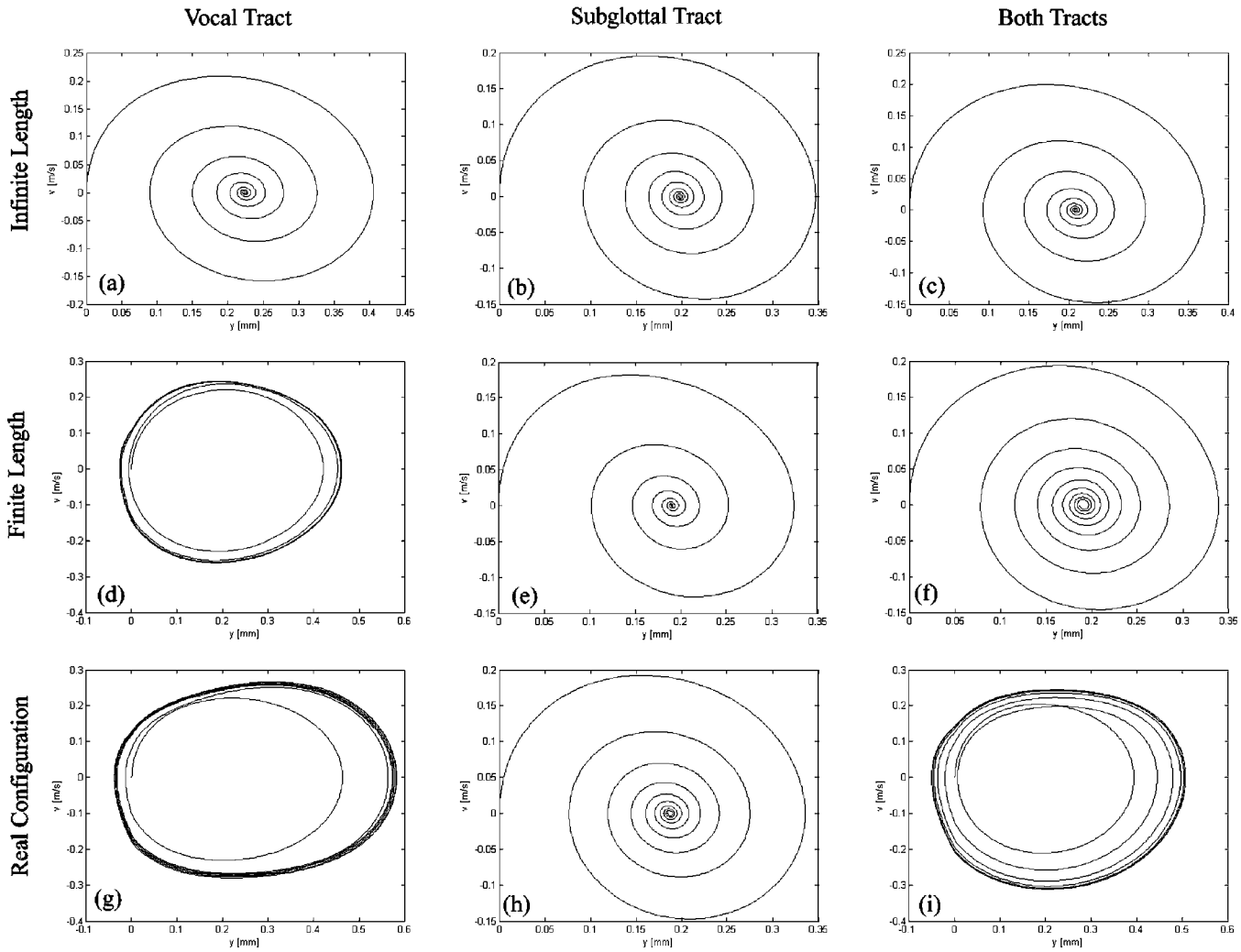

FIG. 11. Phase plots for different acoustic loadings. (a) Supraglottal: Infinitely long tube; (b) Subglottal: Infinitely long tube; (c) Both: Infinitely long tubes; (d) Supraglottal: Uniform tube, length 17 [cm] and area 1 [cm^2]; (e) Subglottal: Uniform tube, length 17 [cm] and area 1 [cm^2]; (f) Both: Uniform tube, length 17 [cm] and area 1 [cm^2]; (g) Supraglottal: MRI vowel /i/; (h) Subglottal: Proposed subglottal design; (i) Both: MRI vowel /i/ and proposed subglottal design.

B. Contributions of subglottal and supraglottal loading on stability

To evaluate the contribution of each tract (subglottal and supraglottal) to the overall dynamic stability of the system, the sound field in each tract was selectively activated and deactivated, while keeping the ODC constant. The same loads used in the previous section were included, as well as other acoustic loads, strategically selected to illustrate the contribution of the sound field in each tract.

Phase plots were used to illustrate the behavior of the system for each specific acoustic load. The phase plot or phase portrait is a geometric representation of the trajectories of a dynamical system in the phase plane (displacement vs velocity). This representation was used to identify limiting cycles and to evaluate stability. Self-sustaining oscillations were easily identified as a closed loop in the phase plot. Phase plots for the three types of loads and for three different locations are shown in Fig. 11. Every row in the group represents a loading case (infinitely long tubes, finite uniform tubes, and realistic loads). For each load, supraglottal and subglottal tracts were tested separately, and together. Every column represents the configuration of the acoustic load (supraglottal, subglottal, and both together). From the first row of Fig. 11, it can be observed that compliant loads (such as long tubes with low first formant frequencies) did not result in a steady state for any configuration. Inertive loads (such as the finite uniform tube, second row in Fig. 11) yielded self-sustained oscillations only if located supraglotally. The last row of Fig. 11 presents realistic loading profiles. The vowel /i/ led to a steady state in absence of subglottal loading. The addition of a subglottal load quenched the oscillations, indicating a damping effect or negative energy transfer.

V. DISCUSSION

Fluid-sound interactions appear to be more significant than fluid-structure interactions in maintaining self-sustained oscillations, particularly for more resistive supraglottal loads. For the subglottal tract, the acoustic field affects the aerodynamics of phonation, but not as significantly as the acoustic field of the vocal tract. The results indicate that (1) the acoustic field in the subglottal system reduces the effects intro-

duced by the vocal tract, and (2) if there is no presence of fluid-structure interactions, subglottal tract configurations (in the absence of a vocal tract) do not yield self-sustained oscillations. Subglottal loading may hamper the net energy transfer to the vocal folds oscillation during phonation, counterbalancing the effects introduced by the vocal tract. Expanding upon the behavior recently noted by Zhang *et al.* (2006), the results suggest that the inertance theory of acoustic loading (Titze, 1988) does not directly apply to the subglottal loading. The subglottal tract impedance closely resembles the behavior of a net compliant vocal-tract load, increasing the effective damping and reducing vocal-fold vibration. Product of the dipole source, the subglottal tract is driven by an acoustic pressure with an opposite polarity to that of the vocal tract. This change of sign in the subglottal impedance is consistent with the proposition that an inertive subglottal tract behaves like a compliant vocal tract, and vice versa. More work needs to be done, however, to completely understand the effects of subglottal impedance on oscillatory conditions.

Similar questions need to be investigated using more complex and extensive interactive models of phonation. More accurate numerical and experimental models of the subglottal system including several levels of tube branching, yielding walls, and other losses should be incorporated. The predictions of the current model need to be explained analytically and verified experimentally. Further research should evaluate the importance of acoustic loading under time-varying vocal-tract shapes (e.g., consonant followed by a vowel). Fast transitions observed for the vocal tract in real articulation could produce fast changes in the sound source that are likely to produce irregular patterns of vibration.

VI. CONCLUSIONS

The influence of arbitrary acoustic loading on the self-oscillation characteristics of an effective single mass model of the vocal folds was investigated. Numerical models for sound wave propagation in the subglottal tract and the vocal tract were developed based on a wave reflection analog method. A single-degree-of-freedom model with adjustable orifice discharge coefficient was designed to model vocal-fold vibrations and the airflow through the larynx, improving the model proposed by Fulcher *et al.* (2006). These models were fully coupled to study complex flow-structural-sound interactions, or how the sound field in the tracts affects source properties. The role of fluid-sound interactions was investigated by activating and deactivating some of the features of the model. The results indicated that fluid-sound interactions are generally more significant than fluid-structure interactions in maintaining self-sustained oscillation, although specific vocal-tract shapes can yield different trends. The results suggested that the acoustic field in the subglottal tract plays a less significant role than in vocal tract. However, subglottal loading tends to absorb energy when the energy transfer to the vocal-fold oscillation from hydrodynamic and vocal-tract coupling is positive, possibly damping oscillations. These findings underline the need of further research on the acoustic field in the subglottal system and acoustic interactions with the sound source.

ACKNOWLEDGMENTS

This study was supported by the research grant No. R01 DC05788 from the National Institute of Deafness and Other Communication Disorders, National Institutes of Health. The first author of this paper also acknowledges the support provided by the Fulbright Program and the Institute of International Education.

Alipour, F., Berry, D. A., and Titze, I. R. (**2000**). "A finite-element model of vocal-fold vibration," J. Acoust. Soc. Am. **108**(6), 3003–3012.

Austin, S. F., and Titze, I. R. (**1997**). "The effect of subglottal resonance upon vocal fold vibration," J. Voice **11**(4), 391402.

Berry, D. A., and Titze, I. R. (**1996**). "Normal modes in a continuum model of vocal fold tissues," J. Acoust. Soc. Am. **100**(5), 3345–3354.

Fant, G., and Lin, Q. (**1987**). "Glottal source - vocal tract acoustic interaction," Speech Transmission Laboratory Quarterly Progress and Status Report, Royal Institute of Technology, Stockholm, pp. 013–027.

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

Fulcher, L. P., Scherer, R. C., Melnykov, A., and Gateva, V. (**2006**). "Negative Coulomb damping, limiting cycles, and self-oscillation of the vocal folds," Am. J. Phys. **74**(5), 386–393.

Harper, V. P., Kraman, S. S., Pasterkamp, H., and Wodicka, G. R. (**2001**). "An acoustic model of the respiratory tract," IEEE Trans. Biomed. Eng. **48**, 543–550.

Harper, V. P., Pasterkamp, H., Kiyokawa, H., and Wodicka, G. R. (**2003**). "Modeling and measurement of flow effects on tracheal sounds," IEEE Trans. Biomed. Eng. **50**, 1–10.

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.

Kelly, J. L., and Lochbaum, C. C. (**1962**). "Speech synthesis," *Proceedings of the Fourth International Congress on Acoustics*, Copenhagen, pp. 1–4.

Liljencrants, J. (**1985**). "Speech Synthesis with a Reflection-type Line Analog," D.S. dissertation, Dept. of Speech Commun. and Music Acoust., Royal Inst. of Tech., Stockholm, Sweden.

Mergell, P., and Herzel, H. (**1997**). "Modeling biphonation - The role of the vocal tract," Speech Commun. **22**, 141154.

Mongeau, L., Franchek, N., Coker, C. H., and Kubli, R. A. (**1997**). "Characteristics of a pulsating jet through a small modulated orifice, with application to voice production," J. Acoust. Soc. Am. **102**(2), 1121–1133.

Park, J. B., and Mongeau, L. (**2007**). "Instantaneous orifice discharge coefficient of a physical, driven model of the human larynx," J. Acoust. Soc. Am. **121**(1), 442–455.

Rahim, M. G. (**1994**). *Artificial Neural Networks for Speech Analysis/Synthesis*, (Kluwer Academic, New York).

Rothenberg, M. (**1981**). "Acoustic interaction between the glottal source and the vocal tract," in *Vocal Fold Physiology*, edited by K. N. Stevens and M. Hirano (University of Tokyo Press, Tokyo).

Scherer, R. C., Shinwari, D., DeWitt, K. J., Zhang, C., Kucinschi, B. R., and Afjeh, A. A. (**2001**). "Intraglottal pressure profiles for a symmetric and oblique glottis with a divergence angle of 10 degrees," J. Acoust. Soc. Am. **109**(4), 1616–1630.

Stevens, K. N. (**2000**). *Acoustic Phonetics*, 1st ed. (The MIT Press, Boston).

Story, B. H. (**1995**). "Physiologically-Based Speech Simulation using an Enhanced Wave-Reflection Model of the Vocal Tract," Ph.D. dissertation, University of Iowa.

Story, B. H. (**2002**). "An overview of the physiology, physics and modeling of the sound source for vowels," Acoust. Sci. & Tech. **23**(4), 195–206.

Story, B. H., and Titze, I. R. (**1995**). "Voice simulation with a body-cover model of the vocal folds," J. Acoust. Soc. Am. **97**(2), 1249–1260.

Story, B. H., Titze, I. R., and Hoffman, E. A. (**1996**). "Vocal tract are functions from magnetic resonance imaging," J. Acoust. Soc. Am. **100**(1), 537–554.

Stronge, W. J. (**2000**). *Impact Mechanics*, 1st ed. (Cambridge University Press, New York).

Thompson, S. L., Mongeau, L., and Frankel, S. H. (**2005**). "Aerodynamic

transfer of energy to the vocal folds," J. Acoust. Soc. Am. **118**(3), 1689–1700.
Titze, I. R. (**1984**). "Parametrization of glottal area, glottal flow and vocal fold contact area," J. Acoust. Soc. Am. **75**(2), 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. (**2002**). "Regulating glottal airflow in phonation: Application of the maximum power transfer theorem to a low dimensional phonation model," J. Acoust. Soc. Am. **111**(1), 367–376.
Titze, I. R. (**2004**). "A theoretical study of $f0$-$f1$ interaction with application to resonant speaking and singing voice," J. Voice **18**(3), 292–298.
Titze, I., and Story, B. (**1997**). "Acoustic interactions of the voice source with the lower vocal tract," J. Acoust. Soc. Am. **101**, 2234–2243.
Van den Berg, J. W. (**1958**). "Myoelastic-aerodynamic theory of voice production," J. Speech Hear. Res. **1**, 227–244.
Van den Berg, J. W., and Tan, T. S. (**1959**). "Results of experiments with human larynxes," Pract. Otorhinolaryngol. (Basel) **21**, 425–450.
Weibel, E. R. (**1963**). *Morphometry of the Human Lung*, 1st ed. (Springer, New York).
Zañartu, M. (**2006**). "Influence of Acoustic Loading on the Flow-Induced Oscillations of Single Mass Models of the Human Larynx," M.S. thesis, School of Electrical and Computer Engineering, Purdue University.
Zhang, Z., Mongeau, L., and Frankel, S. H. (**2002**). "Experimental verification of the quasi-steady approximation for aerodynamic sound generation by pulsating jets in tubes," J. Acoust. Soc. Am. **112**(4), 1652–1663.
Zhang, Z., Neubauer, J., and Berry, D. A. (**2006**). "The influence of subglottal acoustics on laboratory models of phonation," J. Acoust. Soc. Am. **120**(3), 1558–1569.