

An anatomically based, time-domain acoustic model of the subglottal system for speech production

Julio C. Ho^{a)}

Weldon School of Biomedical Engineering, Purdue University, 206 South Martin Jischke Drive,
West Lafayette, Indiana 47907

Matías Zañartu

School of Electrical and Computer Engineering, Purdue University, 140 South Martin Jischke Drive,
West Lafayette, Indiana 47907

George R. Wodicka^{b)}

Weldon School of Biomedical Engineering, Purdue University, 206 South Martin Jischke Drive,
West Lafayette, Indiana 47907

(Received 20 December 2010; accepted 22 December 2010)

A time-domain model of sound wave propagation in the branching airways of the subglottal system is presented. The model is formulated as an extension to an acoustic transmission-line modeling scheme originally developed for simulating the supraglottal system in the time-domain during speech production [Maeda (1982). *Speech Commun.* **1**, 199–229; Mokhtari *et al.* (2008). *Speech Commun.* **50**, 179–190]. The approach allows for predictions of time-varying acoustic pressure and volume velocity at any point along the various generations of subglottal airways from trachea to alveoli. In addition, the model can be configured so that its overall structure simulates different geometric forms, including airways that branch in a symmetric or asymmetric pattern. Three subglottal configurations, two symmetric and one asymmetric, were represented based on reported anatomical dimensions of the subglottal airways. Estimates of the acoustic input impedances of these subglottal configurations revealed resonant characteristics similar to those found in the previous studies. Simulations of voiced sound propagation into the subglottal airways, achieved by coupling the subglottal model to a two-mass vocal fold model and a supraglottal tract configured for different vowels, yielded predictions of time-domain sound pressure waveforms below the vocal folds that compare favorably to previous measurements in human subjects.

© 2011 Acoustical Society of America. [DOI: 10.1121/1.3543971]

PACS number(s): 43.70.Bk, 43.70.Aj [ADP]

Pages: 1531–1547

I. INTRODUCTION

The subglottal system comprises the respiratory airways that extend from the superior end of the trachea to the alveoli. During phonation, sound waves generated at the glottis (the source) propagate back and forward along these subglottal airways. Such phenomena lead to subglottal resonances that can, in turn, influence the behavior of the source and alter the characteristics of the radiated speech (Klatt and Klatt, 1990; Austin and Titze, 1997; Stevens, 1998; Zhang *et al.*, 2006; Chi and Sonderegger, 2007; Zañartu *et al.*, 2007; Titze, 2008). In computational models of human speech production, the subglottal system is typically treated as an ideal static pressure source (Flanagan and Landgraf, 1968; Ishizaka and Flanagan, 1972; Maeda, 1982; Sondhi and Schroeter, 1987; Lucero, 1993; Mokhtari *et al.*, 2008), as an acoustic circuit with predefined resonances (Koizumi *et al.*, 1985; Cranen and Boves, 1987; Cranen and Schroeter, 1995), or as a tubular “horn-like” structure that

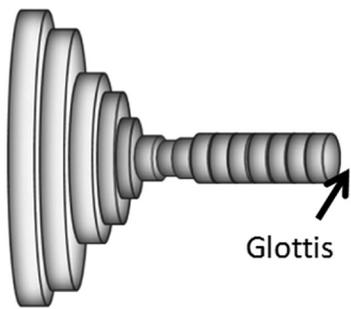
more or less approximates the total cross-sectional area changes along the subglottal airways [Fig. 1(a)] (Titze, 1984; Titze and Story, 1997; Birkholz *et al.*, 2007; Zañartu *et al.*, 2007). The lack of physiological realism in these representations, however, limits the amount of relevant information that can be extracted from these models. In order to obtain quantitative insights into the transmission of speech sounds within the subglottal airways and to correlate these mathematical predictions with the anatomy of the airways and recordings overlaying the subglottal airways (e.g., on the neck at the suprasternal notch), it would be valuable, thus, to construct a numerical model that incorporates the morphology of the subglottal airways.

In this paper, we present a mathematical framework for simulating the subglottal system as an inverted “tree-like” structure that follows the branching pattern of the tracheo-bronchial airways. Figure 1(b) shows a graphical conceptualization of the subglottal model configured for a symmetric branching case. The model is conceived as a concatenated series of short soft-walled cylindrical tube segments interconnected as a bifurcating network that stems from the trachea and terminates at the alveoli. The model can be configured so that its overall geometry follows simple symmetric branching patterns, where branches at the same

^{a)}Author to whom correspondence should be addressed. Electronic mail: hoj@purdue.edu

^{b)}Also at: the School of Electrical and Computer Engineering, Purdue University, West Lafayette, IN 47907.

a) “horn-like” Subglottal Model



b) “tree-like” Subglottal Model

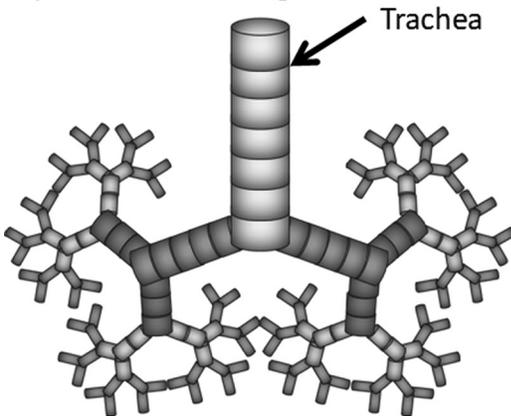


FIG. 1. Subglottal system modeled as (a) a “horn-like” structure and as (b) an inverted “tree-like” network of cylindrical tubes. For simplicity, only the first seven airway generations are shown in (b).

generation are anatomically identical (also known as a regular dichotomy), or more complex asymmetric branching patterns, where branches at the same generation may not be equal (also known as irregular dichotomy).

Many studies have shown that the subglottal system exhibits a number of acoustic resonances within the frequency range of 20 to 3000 Hz. Van den Berg (1958) measured the acoustic input impedance on cadavers of dogs and humans and found resonances at 300, 870, and 1427 Hz with peak impedances of approximately 10 dyne-s/cm⁵ (cgs acoustic ohms). These values, however, differ significantly from those found in later studies and, to our knowledge, have never been replicated. More conventionally accepted estimates of the subglottal resonances are those reported by Ishizaka *et al.* (1976), who measured the acoustic input impedance directly from the stoma of five Japanese patients with laryngectomies. They found that on average the subglottal system exhibited resonance peaks at 640, 1400, and 2100 Hz with peak impedances of roughly 40, 50, and 35 dyne-s/cm⁵, respectively, and bandwidths ranging between 140 and 250 Hz (cf. Figs. 3 and 4 in Ishizaka *et al.*, 1976). Habib *et al.* (1994a) measured the acoustic input impedance on nine normal human subjects using a technique that involved a loudspeaker connected to a rigid tube over the frequency range of 16 to 2048 Hz. On seven of the subjects, all of them males, they found resonances at approximately 606 and 1461 Hz with peak impedances of 49 dyne-s/cm⁵ each. The other two subjects, both females with lower body height and weight, exhibited only a single

resonance at 920 Hz. Boves (1984) made direct measurements of subglottal pressure using a miniature pressure transducer inserted through the posterior end of the vocal folds. By means of spectral analysis, the author estimated the subglottal resonances at 475, 1175, 1945, and 2645 Hz. In a subsequent study, Cranen and Boves (1987) used an linear predictive coding (LPC) analysis technique that considered the subglottal pressure only during the closed phase of the glottal cycle and estimated the subglottal resonances at 510, 1355, and 2290 Hz. Klatt and Klatt (1990) analyzed the spectra of aspirated sounds and associated the presence of additional poles (peaks) and zeros (dips), not associated with the supraglottal formants, with the resonance frequencies of the subglottal system. With this approach, the authors found that in females, the first four poles were at approximately 750, 1650, 2350, and 3150 Hz, while in males, the second, third, and fourth poles were at 1550, 2200, and 3275 Hz, respectively.

Various acoustic models of sound wave propagation in the subglottal system have been previously developed (Ishizaka *et al.*, 1976; Fredberg and Hoenig, 1978; Hudde and Slatky, 1989; Habib *et al.*, 1994a; Harper *et al.*, 2001). These models were generally formulated using frequency-domain techniques, which efficiently allows for the inclusion of frequency-dependent parameters, such as heat conduction losses. For the specific purpose of modeling voice production, however, the utility of frequency-domain-based subglottal models is limited by the intricacy of trying to couple them with time-domain-based vocal fold models. Although some efforts have been made toward combining time-domain vocal fold models with frequency-domain tract models (Sondhi and Schroeter, 1987), these approaches are challenging. Time-domain models of the subglottal system, on the other hand, offer a potentially more intuitive approach for linking with vocal fold models and thus simulating the processes of voice production.

The subglottal model proposed in the present study extends from a time-domain modeling scheme initially conceived by Maeda (1982). In his work, Maeda describes how the equations of sound wave propagation along a cascaded transmission-line (TL) network representation of the supraglottal system can be discretized and transformed into a matrix equation (i.e., a system of equations). Maeda’s original scheme, however, was limited in that it could only include a single side branch cavity [the nasal tract (NT)]. Recently, Mokhtari *et al.* (2008) improved on Maeda’s approach by allowing the scheme to include multiple side branch cavities of the supraglottal tract, representing the piriform fossae and the NT, while conveniently keeping the overall numerical computations into a single matrix equation.

The Sec. II begins with a brief summary of the improved time-domain scheme for simulating the supraglottal system. Following this, a detailed description of how the scheme can be further extended to model the subglottal system as a branching TL network is given. Three different subglottal configurations (two symmetric and one asymmetric) are then represented and their acoustic properties are examined. Finally, example scenarios of speech sound propagation into the subglottal airways are presented through the novel

construction of a time-domain model of the interconnected subglottal, glottal, and supraglottal systems.

II. MODEL

A. Supraglottal model

This section summarizes the enhanced time-domain modeling scheme proposed by Mokhtari *et al.* (2008). The scheme simulates the complete supraglottal system, including the NT and piriform fossae, by employing a single matrix equation. This formulation will serve as the basis for constructing the subglottal model presented in Sec. II B. For more details on the original scheme we refer the reader to Maeda (1982) and Mokhtari *et al.* (2008).

1. TL representation

One-dimensional sound wave propagation in a cylindrical tube can be modeled equivalently as electrical wave propagation in a lumped-element TL circuit [Fig. 2(a)], where the acoustic pressure and airflow (volume velocity) are analogous to the electric voltage and current, respectively (Fant, 1960; Flanagan, 1972). Using the TL circuit in Fig. 2(a) as an elementary building block, a cascaded TL network representation of the vocal tract, NT, and piriform fossae can be constructed as shown Fig. 2(b). Here, the supraglottal airways are represented as a concatenated series of short cylindrical tube segments approximating the area functions of the supraglottal tract from the glottis to the lips, nostrils, and piriform ends. At any given TL block, n in the cascaded network, the lumped elements L_n , R_n , and C_n , representing the tube's acoustic inductance, resistance, and capacitance, respectively, are defined as

$$L_n = \frac{\rho l_n}{2A_n}, \quad (1a)$$

$$R_n = \frac{4\pi\mu l_n}{A_n^2}, \quad (1b)$$

$$C_n = \frac{l_n A_n}{\rho c^2}, \quad (1c)$$

where ρ , μ , and c are the air density, air viscosity, and speed of sound in air, respectively. Typical values for these air properties are $\rho = 1.14 \times 10^{-3}$ g/cm³, $\mu = 1.86 \times 10^{-4}$ dyne/s/cm², and $c = 35\,400$ cm/s (Flanagan, 1972). The parameters A_n and l_n are the n th tube's cross-sectional area and length, respectively. The series elements L_{wn} , R_{wn} , and C_{wn} represent the locally reacting yielding walls. These are defined as

$$L_{wn} = \frac{k_{wm}}{l_n 2\sqrt{\pi A_n}}, \quad (2a)$$

$$R_{wn} = \frac{k_{wb}}{l_n 2\sqrt{\pi A_n}}, \quad (2b)$$

$$C_{wn} = \frac{l_n 2\sqrt{\pi A_n}}{k_{wk}}, \quad (2c)$$

where k_{wm} , k_{wb} , and k_{wk} are the wall's mass, resistance, and stiffness per unit area, respectively. For the supraglottal walls, these values have been estimated to be approximately $k_{wm} = 1.5$ g/cm², $k_{wb} = 1400$ g/s/cm², and $k_{wk} = 3 \times 10^5$ g/cm² (Ishizaka *et al.*, 1975). The current source U_{dn} accounts for the flow generated due to the "articulated"

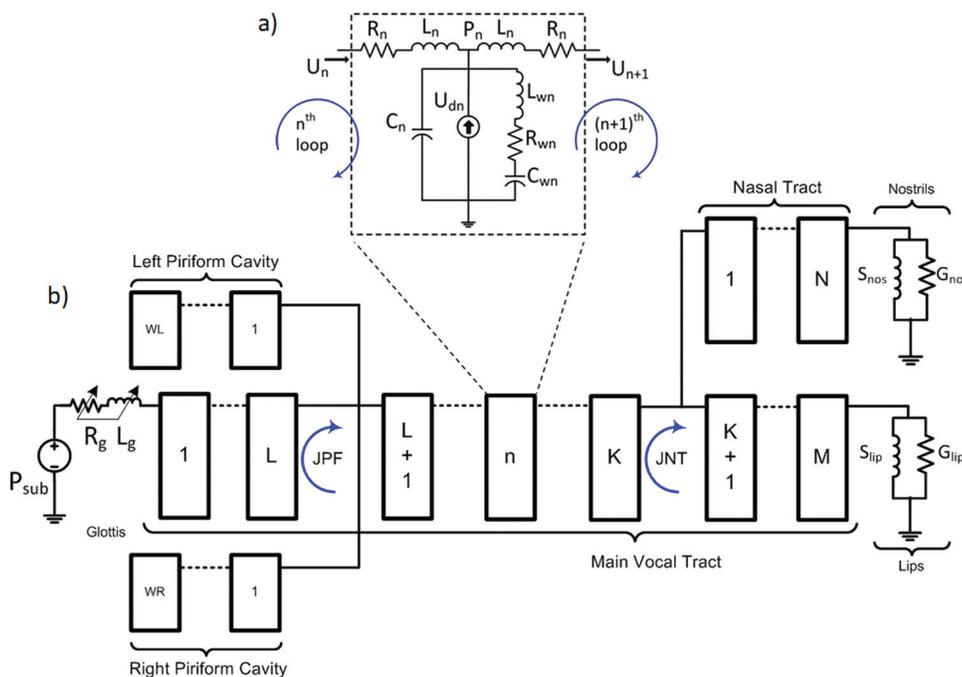


FIG. 2. (Color online) (a) TL circuit representation of a single cylindrical tube segment (or "tubelet"). (b) Cascaded TL network representation of the supraglottal system with side branches representing the piriform fossae and NT. Letters M, N, WL, and WR denote the number of tubelets used to represent the main vocal tract, NT, LP cavity, and RP cavity, respectively. The NT and piriform fossae connect to the main vocal tract between the L th and $(L+1)$ th blocks and between the K th and $(K+1)$ th blocks, respectively.

$$\Lambda_n[t-1] = 4f_s C_n[t-1] P_n[t-1] - \Lambda_n[t-2], \quad (10c)$$

$$\Gamma_n[t-1] = 4f_s L_{wn}[t-1] U_{wn}[t-1] - \Gamma_n[t-2], \quad (10d)$$

$$\psi_n[t-1] = \frac{U_{wn}[t-1]}{f_s C_{wn}[t-1]} + \psi_n[t-2], \quad (10e)$$

$$Q_n[t-1] = 4f_s U_n[t-1] (L_{n-1}[t-1] + L_n[t-1]) - Q_n[t-2], \quad (10f)$$

and where f_s is the simulation sampling frequency in hertz.

At the terminating ends (glottis, lips, and nostrils) and at the junction points (where the piriform fossae and NT attach to the main vocal tract), Eqs. (7), (8a), (9), and (10f) are computed as special cases. For the elements in \mathbf{H} , these are

$$H_1^V = -2f_s(L_g + L_1^V) - (R_g + R_1^V) - b_1^V, \quad (11a)$$

$$H_J^V = -2f_s(L_\alpha^V) - (R_\alpha^V) - (b_\alpha^V), \quad (11b)$$

$$H_I^V = -2f_s(L_I^V) - (R_I^V) - (b_I^V), \quad (11c)$$

$$H_1^T = -2f_s(L_1^T) - (R_1^T) - (b_1^T), \quad (11d)$$

$$H_{D+1}^B = -2f_s(L_D^B) - (R_D^B) - (b_D^B + b_{D+1}^B), \quad (11e)$$

where subscripts J and α stand for either JPF (junction with piriform fossae) and L or JNT (junction with nasal tract) and K (see Fig. 2), corresponding to the half loops preceding the junction points, I stands for either $L+1$ or $K+1$, corresponding to the half loops after the junction points, T stands for either RP, LP, or NT corresponding to the first half loop in the RP, LP, or NT (respectively), B and D stand for V and M or NT and N , corresponding to the terminating loops at the radiating ends, and where elements b_{D+1}^B in Eq. (11e) are given by

$$b_{D+1}^B = \frac{1}{G_{\text{rad}} + \frac{S_{\text{rad}}}{2f_s}}. \quad (12)$$

For the elements in \mathbf{F} , the corresponding special cases are calculated with the following:

$$F_1^V[t] = -P_{\text{sub}}[t] - b_1^V[t](U_{d1}^V[t] - V_1^V[t-1]) - Q_1^V[t-1], \quad (13a)$$

$$F_J^V[t] = b_\alpha^V[t](U_{d\alpha}^V[t] - V_\alpha^V[t-1]) - Q_J^V[t-1], \quad (13b)$$

$$F_I^V[t] = -b_I^V[t](U_{dI}^V[t] - V_I^V[t-1]) - Q_I^V[t-1], \quad (13c)$$

$$F_1^T[t] = -b_1^T[t](U_{d1}^T[t] - V_1^T[t-1]) - Q_1^T[t-1], \quad (13d)$$

$$F_{D+1}^B[t] = b_{D+1}^B[t](U_{dD+1}^B[t] - V_{D+1}^B[t-1]) - b_{D+1}^B[t](-V_{D+1}^B[t-1]) - Q_{D+1}^B[t-1], \quad (13e)$$

where the corresponding Q elements are given by

$$Q_1^V[t-1] = 4f_s U_1^V[t-1](L_g[t-1] + L_1^V[t-1]) - Q_1^V[t-2], \quad (14a)$$

$$Q_{\text{JPF}}^V[t-1] = 4f_s(U_1^{\text{RP}}[t-1] + U_1^{\text{LP}}[t-1] + U_{L+1}^V[t-1]) \times L_L^V[t-1] - Q_{\text{JPF}}^V[t-2], \quad (14b)$$

$$Q_{\text{JNT}}^V[t-1] = 4f_s(U_1^{\text{NT}}[t-1] + U_{K+1}^V[t-1]) \times L_K^V[t-1] - Q_{\text{JNT}}^V[t-2], \quad (14c)$$

$$Q_1^V[t-1] = 4f_s U_1^V[t-1] L_1^V[t-1] - Q_1^V[t-2], \quad (14d)$$

$$Q_1^T[t-1] = 4f_s U_1^T[t-1] L_1^T[t-1] - Q_1^T[t-2], \quad (14e)$$

$$Q_{D+1}^B[t-1] = 4f_s U_{D+1}^B[t-1] L_D^B[t-1] - Q_{D+1}^B[t-2]. \quad (14f)$$

After the matrix equation $\mathbf{F} = \mathbf{H}\mathbf{U}$ is solved for the airflows in \mathbf{U} , the acoustic pressures P_n along the main vocal tract, NT, and piriform fossae, and the “wall airflows” U_{wn} are obtained with

$$P_n[t] = b_n[t](U_n[t] - U_{n+1}[t] - U_{dn}[t] + V_n[t-1]), \quad (15a)$$

$$U_{wn}[t] = Y_{wn}[t](P_n[t] + \Gamma_n[t-1] - \psi_n[t-1]). \quad (15b)$$

At the radiating ends (lips and nostrils), P is computed as

$$P_{D+1}^B[t] = b_{D+1}^B[t](U_{D+1}^B[t] + V_{D+1}^B[t-1]), \quad (16)$$

where

$$V_{D+1}^B[t-1] = -\frac{S_{\text{rad}}}{f_s} P_{D+1}^B[t-1] + V_{D+1}^B[t-2]. \quad (17)$$

Equations (1)–(17) constitute the routine that is iteratively computed at every time step throughout the algorithm to simulate the supraglottal system in Fig. 2(b).

B. Subglottal model: mathematical derivations

The subglottal system bifurcates into progressively smaller and smaller airways dividing into approximately 24

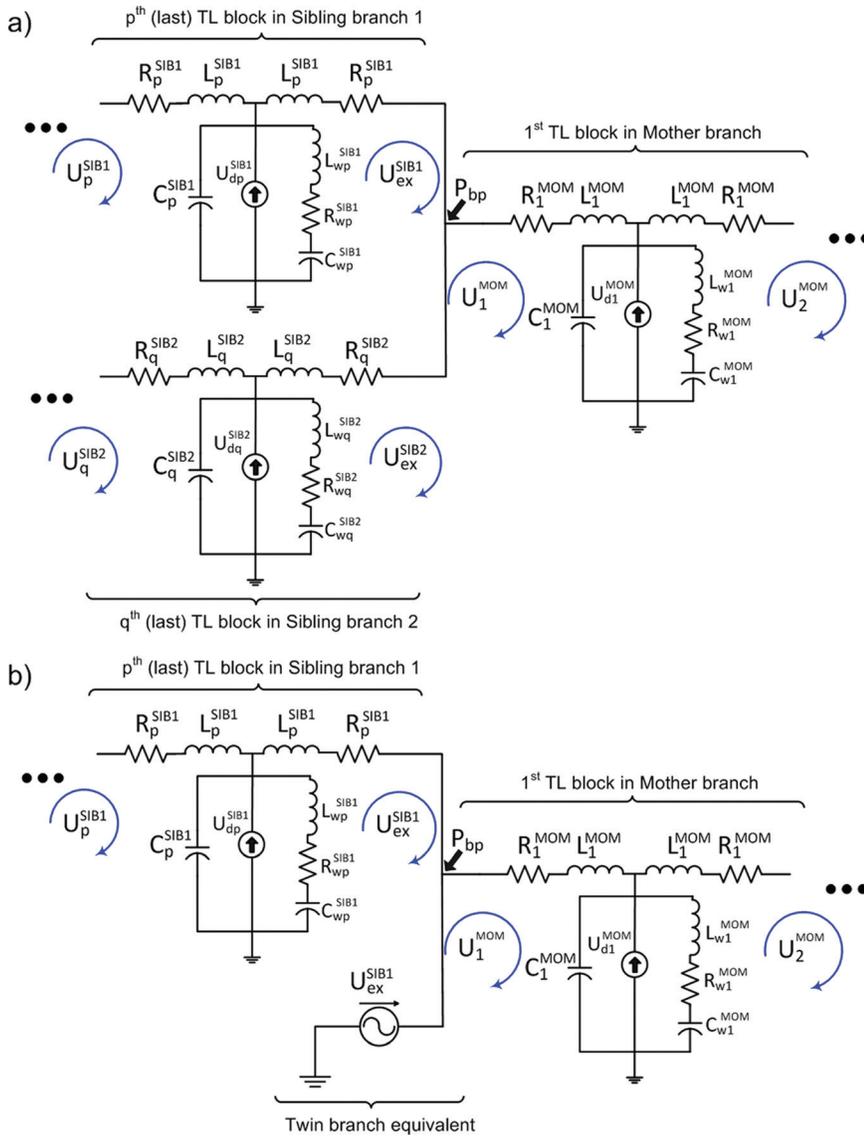


FIG. 4. (Color online) (a) TL circuit representation of a branching point in the subglottal network in Fig. 3. (b) Equivalent TL circuit representation of (a) under the assumption that both sibling sub-networks are symmetric (i.e., identical to each other).

$$X = \left\{ \left(\frac{-2f_s L_{n-1}^{\text{SUB}} - R_{n-1}^{\text{SUB}} - b_{n-1}^{\text{SUB}}}{2} \right) + (-2f_s L_n^{\text{SUB}} - R_n^{\text{SUB}} - b_n^{\text{SUB}}) \right\},$$

$$Y = b_n^{\text{SUB}}, \quad (22b)$$

$$X = \left\{ \left(\frac{-2f_s L_{n-1}^{\text{SUB}} - R_{n-1}^{\text{SUB}} - b_{n-1}^{\text{SUB}}}{2} \right) + (-2f_s L_n^{\text{SUB}} - R_n^{\text{SUB}} - b_n^{\text{SUB}}) \right\},$$

$$Y = \frac{b_n^{\text{SUB}}}{2}, \quad (22c)$$

$$X = H_n^{\text{SUB}}, \quad Y = \frac{b_n^{\text{SUB}}}{2}, \quad (22d)$$

Note that case (22a) results in Eqs. (7) and (8a) corresponding to a regular loop with no branching in neither the n th nor the $(n+1)$ th loop. For cases (22a) and (22d), the elements in \mathbf{F} are computed using Eqs. (9) and (10) as in a regular loop.

For cases (22b) and (22c), Eq. (10f) is replaced with the following:

$$Q_n[t-1] = 4f_s U_n[t-1] \left(\frac{1}{2} L_{n-1}[t-1] + L_n[t-1] \right) - Q_n[t-2]. \quad (23)$$

The special elements F_1^{SUB} and H_1^{SUB} corresponding to the alveolar boundary cases are computed as follows:

$$H_1^{\text{SUB}} = -2f_s (L_1^{\text{SUB}}) - (R_1^{\text{SUB}}) - (b_1^{\text{SUB}}), \quad (24a)$$

$$F_1^{\text{SUB}}[t] = -P_{\text{alv}}[t] - b_1^{\text{SUB}}[t] (U_{d1}^{\text{SUB}}[t] - V_1^{\text{SUB}}[t-1]) - Q_1^{\text{SUB}}[t-1], \quad (24b)$$

where

$$Q_1^{\text{SUB}}[t-1] = 4f_s U_1^{\text{SUB}}[t-1] L_1^{\text{SUB}}[t-1] - Q_1^{\text{SUB}}[t-2]. \quad (25)$$

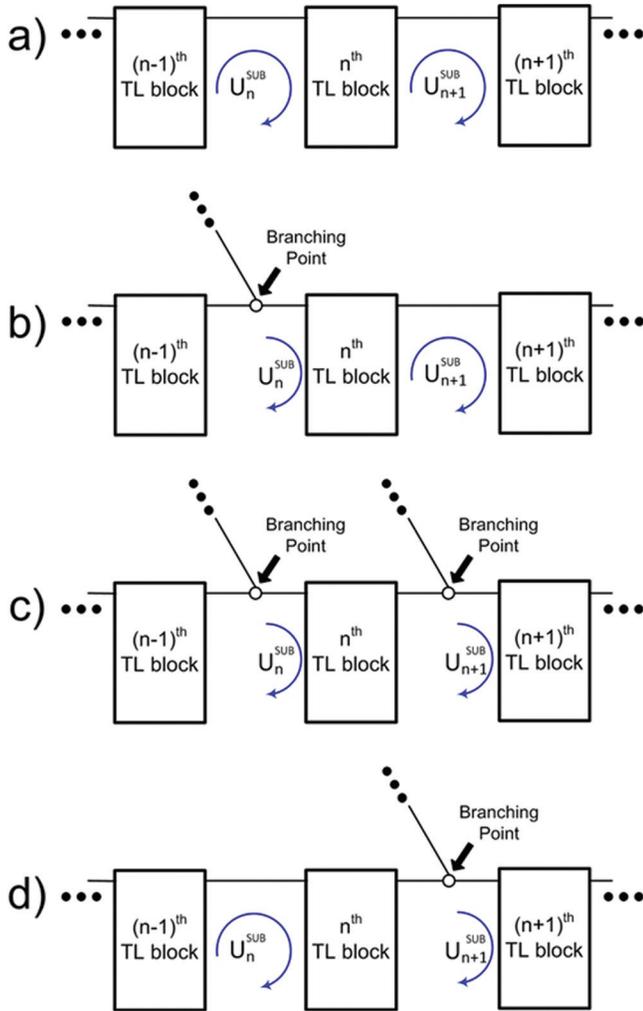


FIG. 5. (Color online) Four cases pertaining to elements X and Y in the n th row in Eq. (22).

The boundary condition at the alveolar end is approximated as an ideal pressure source, P_{alv} , simulating the static pulmonary pressure.

2. Asymmetric branching

Anatomically, the subglottal airways do not always branch in a completely symmetrical fashion. There are clear anatomical differences, for instance, between the left main bronchus (LMB) and right main bronchus (RMB), where the former is normally longer and narrower than the latter. Hence, it has been previously suggested that it is necessary to account for asymmetries into subglottal models in order to reproduce its acoustic properties accurately.

From Fig. 4(a), the generalized loop equations at the exit of the SIB1 and SIB2 can be written as

$$F_{\text{ex}}^{\text{SIB1}} = b_p^{\text{SIB1}} U_p^{\text{SIB1}} + H_{\text{ex}}^{\text{SIB1}} U_{\text{ex}}^{\text{SIB1}} - P_{bp}, \quad (26a)$$

$$F_{\text{ex}}^{\text{SIB2}} = b_q^{\text{SIB2}} U_q^{\text{SIB2}} + H_{\text{ex}}^{\text{SIB2}} U_{\text{ex}}^{\text{SIB2}} - P_{bp}, \quad (26b)$$

where P_{bp} is the pressure at the branching point. Similarly, the loop equation at the first entry section of mother branch can be written as

$$F_1^{\text{MOM}} = P_{bp} + H_1^{\text{MOM}} U_1^{\text{MOM}} + b_1^{\text{MOM}} U_2^{\text{MOM}}. \quad (27)$$

Combining Eqs. (26) and (27) and using the equality in Eq. (18) yields the following:

$$F_{\text{ex}}^{\text{SIB1}} + F_1^{\text{MOM}} = b_p^{\text{SIB1}} U_p^{\text{SIB1}} + U_{\text{ex}}^{\text{SIB1}} \{H_{\text{ex}}^{\text{SIB1}} + H_1^{\text{MOM}}\} + U_{\text{ex}}^{\text{SIB2}} H_1^{\text{MOM}} + b_1^{\text{MOM}} U_2^{\text{MOM}}, \quad (28a)$$

$$F_{\text{ex}}^{\text{SIB1}} + F_1^{\text{MOM}} = b_q^{\text{SIB2}} U_q^{\text{SIB2}} + U_{\text{ex}}^{\text{SIB2}} \{H_{\text{ex}}^{\text{SIB2}} + H_1^{\text{MOM}}\} + U_{\text{ex}}^{\text{SIB1}} H_1^{\text{MOM}} + b_1^{\text{MOM}} U_2^{\text{MOM}}, \quad (28b)$$

where

$$H_{\text{ex}}^{\text{SIB1}} = -2f_s L_p^{\text{SIB1}} - R_p^{\text{SIB1}} - b_p^{\text{SIB1}}, \quad (29a)$$

$$H_{\text{ex}}^{\text{SIB2}} = -2f_s L_q^{\text{SIB2}} - R_q^{\text{SIB2}} - b_q^{\text{SIB2}}, \quad (29b)$$

$$H_1^{\text{MOM}} = -2f_s L_1^{\text{MOM}} - R_1^{\text{MOM}} - b_1^{\text{MOM}}. \quad (29c)$$

In \mathbf{F} , elements $F_{\text{ex}}^{\text{SIB1}}$, $F_{\text{ex}}^{\text{SIB2}}$, and F_1^{MOM} are given by

$$F_{\text{ex}}^{\text{SIB1}}[t] = b_p^{\text{SIB1}} [t] (U_{dp}^{\text{SIB1}} [t] - V_p^{\text{SIB1}} [t-1]) - Q_{\text{ex}}^{\text{SIB1}} [t-1], \quad (30a)$$

$$F_{\text{ex}}^{\text{SIB2}}[t] = b_q^{\text{SIB2}} [t] (U_{dq}^{\text{SIB2}} [t] - V_q^{\text{SIB2}} [t-1]) - Q_{\text{ex}}^{\text{SIB2}} [t-1], \quad (30b)$$

$$F_1^{\text{MOM}}[t] = -b_1^{\text{MOM}} [t] (U_{d1}^{\text{MOM}} [t] - V_1^{\text{MOM}} [t-1]) - Q_1^{\text{MOM}} [t-1], \quad (30c)$$

where

$$Q_{\text{ex}}^{\text{SIB1}} [t-1] = 4f_s U_{\text{ex}}^{\text{SIB1}} [t-1] L_p^{\text{SIB1}} [t-1] - Q_{\text{ex}}^{\text{SIB1}} [t-2], \quad (31a)$$

$$Q_{\text{ex}}^{\text{SIB2}} [t-1] = 4f_s U_{\text{ex}}^{\text{SIB2}} [t-1] L_q^{\text{SIB2}} [t-1] - Q_{\text{ex}}^{\text{SIB2}} [t-2], \quad (31b)$$

$$Q_1^{\text{MOM}} [t-1] = 4f_s U_1^{\text{MOM}} [t-1] L_1^{\text{MOM}} [t-1] - Q_1^{\text{MOM}} [t-2]. \quad (31c)$$

Ideally, one could imagine that by repeating this procedure at every single branching point along the entire subglottal network, a completely asymmetrical network can be simulated. However, this approach is difficult to implement since it requires considerable computational power and memory. Here, we undertook a more operable approach primarily inspired by the anatomical data measured by Yeh and Schum (1980), who reported the dimensions of each major bronchus and averaged dimensions for the minor airways within each lobe (as detailed in Sec. II C 1). This approach incorporates only the largest asymmetries (occurring at lower airway generations) while keeping the smaller airways (at higher generations) as symmetric sub-networks. Specifically, the trachea, RMB, LMB, and intermediate bronchus (IMB) are each

modeled as separate branches that branch following Eqs. (28)–(31), whereas the right upper lobe (RUL), right middle lobe (RML), right lower lobe (RLL), left upper lobe (LUL), and left lower lobe (LLL) are each represented as smaller separate symmetric networks. Figure 6 shows a graphical conceptualization of this asymmetric model. The resulting \mathbf{F} and \mathbf{U} column matrices and the Σ and β diagonal components of \mathbf{H} for this subglottal configuration have the following forms:

$$\mathbf{F} = [F_1^{\text{LUL}}, F_2^{\text{LUL}}, \dots, (F_{\text{ex}}^{\text{LUL}} + F_1^{\text{LMB}}), F_1^{\text{LLL}}, \dots, (F_{\text{ex}}^{\text{LLL}} + F_1^{\text{LMB}}), F_2^{\text{LMB}}, \dots, (F_{\text{ex}}^{\text{LMB}} + F_1^{\text{TRA}}), F_1^{\text{RUL}}, \dots, (F_{\text{ex}}^{\text{RUL}} + F_1^{\text{RMB}}), F_1^{\text{RML}}, \dots, (F_{\text{ex}}^{\text{RML}} + F_1^{\text{IMB}}), F_1^{\text{RLL}}, \dots, (F_{\text{ex}}^{\text{RLL}} + F_1^{\text{IMB}}), F_2^{\text{IMB}}, \dots, (F_{\text{ex}}^{\text{IMB}} + F_1^{\text{RMB}}), F_2^{\text{RMB}}, \dots, (F_{\text{ex}}^{\text{RMB}} + F_1^{\text{TRA}}), F_2^{\text{TRA}}, \dots, F_{\text{ex}}^{\text{TRA}}]^T, \quad (32a)$$

$$\mathbf{U} = [U_1^{\text{LUL}}, U_2^{\text{LUL}}, \dots, U_{\text{ex}}^{\text{LUL}}, U_1^{\text{LLL}}, \dots, U_{\text{ex}}^{\text{LLL}}, U_2^{\text{LMB}}, \dots, U_{\text{ex}}^{\text{LMB}}, U_1^{\text{RUL}}, \dots, U_{\text{ex}}^{\text{RUL}}, U_1^{\text{RML}}, \dots, U_{\text{ex}}^{\text{RML}}, U_1^{\text{RLL}}, \dots, U_{\text{ex}}^{\text{RLL}}, U_1^{\text{IMB}}, \dots, U_{\text{ex}}^{\text{IMB}}, U_2^{\text{RMB}}, \dots, U_{\text{ex}}^{\text{RMB}}, U_2^{\text{TRA}}, \dots, U_{\text{ex}}^{\text{TRA}}]^T, \quad (32b)$$

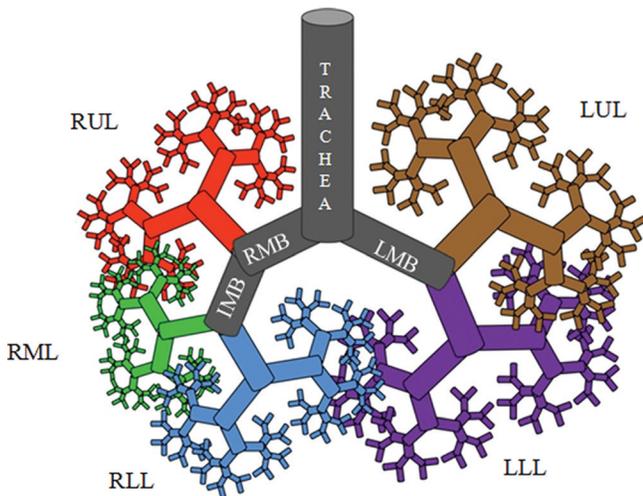


FIG. 6. (Color online) Graphical conceptualization of the Y&S_{asym} asymmetric subglottal model. The trachea, RMB, LMB, and IMB are each geometrically different from one another. The LUL, LLL, RUL, RML, and RLL are incorporated as symmetric sub-networks. For simplicity, not all generations are shown.

$$\Sigma = \{H_1^{\text{LUL}}, \dots, (H_{\text{ex}}^{\text{LUL}} + H_1^{\text{LMB}}), H_1^{\text{LLL}}, \dots, (H_{\text{ex}}^{\text{LLL}} + H_1^{\text{LMB}}), H_2^{\text{LMB}}, \dots, (H_{\text{ex}}^{\text{LMB}} + H_1^{\text{TRA}}), H_1^{\text{RUL}}, \dots, (H_{\text{ex}}^{\text{RUL}} + H_1^{\text{RMB}}), H_1^{\text{RML}}, \dots, (H_{\text{ex}}^{\text{RML}} + H_1^{\text{IMB}}), H_1^{\text{RLL}}, \dots, (H_{\text{ex}}^{\text{RLL}} + H_1^{\text{IMB}}), H_2^{\text{IMB}}, \dots, (H_{\text{ex}}^{\text{IMB}} + H_1^{\text{RMB}}), H_2^{\text{RMB}}, \dots, (H_{\text{ex}}^{\text{RMB}} + H_1^{\text{TRA}}), H_2^{\text{TRA}}, \dots, H_{\text{ex}}^{\text{TRA}}\}, \quad (32c)$$

$$\beta = \{b_1^{\text{LUL}}, \dots, b_{N_a}^{\text{LUL}}, 0, b_1^{\text{LLL}}, \dots, b_{N_b}^{\text{LLL}}, 0, b_1^{\text{LMB}}, \dots, b_{N_c}^{\text{LMB}}, 0, b_1^{\text{RUL}}, \dots, b_{N_d}^{\text{RUL}}, 0, b_1^{\text{RML}}, \dots, b_{N_e}^{\text{RML}}, 0, b_1^{\text{RLL}}, \dots, b_{N_f}^{\text{RLL}}, 0, b_1^{\text{IMB}}, \dots, b_{N_g}^{\text{IMB}}, 0, b_1^{\text{RMB}}, \dots, b_{N_h}^{\text{RMB}}, 0, b_1^{\text{TRA}}, \dots, b_{N_i}^{\text{TRA}}\}. \quad (32d)$$

where additionally the non-diagonal elements in \mathbf{H} at entries $(Na + 1, Nb + 1)$ and $(Nb + 1, Na + 1)$ are set to H_1^{LMB} , at $(Na + 1, Nb + 2)$ and $(Nb + 2, Na + 1)$ are set to b_1^{LMB} , at $(Nc + 1, Nh + 1)$ and $(Nh + 1, Nc + 1)$ are set to H_1^{TRA} , at $(Nc + 1, Nh + 2)$ and $(Nh + 2, Nc + 1)$ are set to b_1^{TRA} , at $(Nd + 1, Ng + 1)$ and $(Ng + 1, Nd + 1)$ are set to H_1^{RMB} , at $(Nd + 1, Ng + 2)$ and $(Ng + 2, Nd + 1)$ are set to b_1^{RMB} , at $(Ne + 1, Nf + 1)$ and $(Nf + 1, Ne + 1)$ are set to H_1^{IMB} , and at $(Ne + 1, Nf + 2)$ and $(Nf + 2, Ne + 1)$ are set to b_1^{IMB} .

3. Coupling with the supraglottal system

In the two previous studies by Maeda (1982) and Mokhtari *et al.* (2008), the subglottal system was modeled as an ideal static pressure source. The use of this subglottal representation led to Eqs. (11a), (13a), and (14a), where P_{sub} represents the static subglottal pressure source. With a subglottal structure coupled behind the glottis, however, it is necessary to modify those equations in order to account for the additional impedance and force components below the glottis. Thus, the following relationships are employed

$$H_1^V = -2f_s(L_s^{\text{SUB}} + L_g + L_1^V) - (R_s^{\text{SUB}} + R_g + R_1^V) - (b_s^{\text{SUB}} + b_1^V), \quad (33a)$$

$$F_1^V[t] = b_s^{\text{SUB}}[t](U_{ds}^{\text{SUB}}[t] - V_s^{\text{SUB}}[t - 1]) - b_1^V[t](U_{d1}^V[t] - V_1^V[t - 1]) - Q_1^V[t - 1], \quad (33b)$$

$$Q_1^V[t - 1] = 4f_s U_1^V[t - 1](L_s^{\text{SUB}}[t - 1] + L_g[t - 1] + L_1^V[t - 1]) - Q_1^V[t - 2], \quad (33c)$$

where elements L_g and R_g represent the total glottal inductance and resistance, respectively.

The generalized \mathbf{F} , \mathbf{U} , and \mathbf{H} matrices representing the entire respiratory tract, including both subglottal and supraglottal systems, are

$$\mathbf{F} = [F_1^{\text{SUB}}, \dots, F_s^{\text{SUB}}, F_1^V, \dots, F_{N+1}^{\text{NT}}]^T, \quad (34a)$$

$$\mathbf{U} = [U_1^{\text{SUB}}, \dots, U_s^{\text{SUB}}, U_1^V, \dots, U_{N+1}^{\text{NT}}]^T, \quad (34b)$$

TABLE II. Subglottal airway dimensions used in the Y&S_{asym} asymmetric model as a function of generation level z . Values are those reported in Yeh and Schum (1980) as the typical path lung models for the RUL, RML, RLL, LUL, and LLL, as well as the separate dimensions of the trachea, RMB, LMB, and the IMB.

Y&S _{asym}										
z	Length (cm) Diameter (cm) Trachea ($z = 0$)		Length (cm) Diameter (cm) RMB ($z = 1$)		Length (cm) Diameter (cm) LMB ($z = 1$)		Length (cm) Diameter (cm) IMB ($z = 2$)			
	10.0	2.01	3.09	1.75	5.63	1.38	3.02	1.33		
	RUL							LUL	LLL	
2	1.22	1.02	RML		RLL		1.45	1.03	1.42	1.15
3	0.800	0.760	2.27	0.720	0.88	1.01	1.08	0.835	1.33	0.905
4	1.270	0.650	1.34	0.620	1.09	0.800	1.02	0.640	1.13	0.680
5	1.250	0.579	1.63	0.528	1.33	0.650	1.09	0.535	0.891	0.559
6	0.827	0.454	1.04	0.376	1.22	0.583	1.02	0.426	1.02	0.454
7	0.988	0.355	1.04	0.317	0.796	0.471	0.751	0.341	0.836	0.365
8	0.798	0.278	0.691	0.268	0.803	0.367	0.832	0.307	0.778	0.316
9	0.557	0.216	0.527	0.199	0.880	0.347	0.555	0.234	0.771	0.298
10	0.401	0.158	0.394	0.147	0.900	0.317	0.482	0.178	0.611	0.286
11	0.350	0.118	0.266	0.106	0.591	0.249	0.388	0.135	0.544	0.211
12	0.250	0.088	0.225	0.083	0.449	0.181	0.343	0.100	0.431	0.146
13	0.194	0.070	0.172	0.064	0.337	0.134	0.267	0.078	0.302	0.102
14	0.143	0.058	0.118	0.051	0.257	0.101	0.215	0.061	0.224	0.076
15	0.119	0.053	0.105	0.048	0.222	0.077	0.175	0.055	0.188	0.061
16	0.102	0.049	0.091	0.046	0.158	0.066	0.144	0.051	0.152	0.055
17	0.089	0.047	0.080	0.044	0.131	0.058	0.118	0.048	0.124	0.051
18	0.078	0.045	0.072	0.044	0.110	0.054	0.098	0.047	0.103	0.049
19	0.070	0.044	0.066	0.043	0.094	0.051	0.082	0.045	0.087	0.047
20	0.063	0.044	0.061	0.043	0.082	0.048	0.071	0.044	0.076	0.046
21	0.057	0.043	0.056	0.043	0.072	0.046	0.060	0.044	0.066	0.045
22	0.053	0.043	0.053	0.043	0.064	0.045	0.053	0.043	0.059	0.044
23	0.025	0.030	0.025	0.030	0.058	0.044	0.025	0.030	0.053	0.043
24	–	–	–	–	0.053	0.043	–	–	0.025	0.030
25	–	–	–	–	0.025	0.030	–	–	–	–

3. Spatial discretization and simulation sampling frequency

In order to model the subglottal airways as a TL circuit, each tube segment should be no longer than $\sim 1/8$ the wave-

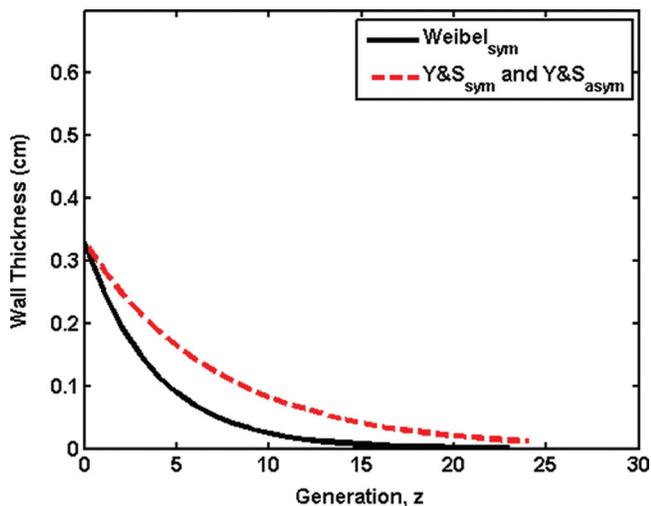


FIG. 7. (Color online) Wall thickness approximations used for the Weibel_{sym} (solid line) and Y&S_{sym} and Y&S_{asym} (dashed line) models.

length of the highest frequency of interest (Fant, 1960). Thus, for frequencies of interest up to 3 kHz, each subglottal airway should be partitioned into a number of shorter tube segments no longer than ~ 1.5 cm. In these models, we therefore partitioned them into segments no longer than 0.5 cm. This yielded impedance matrices \mathbf{H} of size 68×68 for Weibel_{sym}, 65×65 for Y&S_{sym} and 227×227 for Y&S_{asym}. With the additional elements representing the vocal folds, main vocal tract, and piriform fossae (corresponding to a vowel /a/) coupled together, the complete \mathbf{H} matrices were of size 153×153 with the Weibel_{sym} model, 150×150 with the Y&S_{sym} model, and 312×312 with the Y&S_{asym} model. To assume only planar wave propagation in the model, it is necessary to operate below the cut-on frequency of the first nonplanar propagation mode. This cut-on frequency is given by $f_{\text{cuton}} = \frac{0.5861c}{2r}$ (Eriksson, 1980), where c and r are the speed of sound and airway radius, respectively. Using a conservative estimate of maximum subglottal airway radius of 2 cm, this yields a prediction of plane wave propagation valid for frequencies up to ~ 5 kHz. As in Mokhtari *et al.* (2008), the closed ends of the piriform fossae were modeled by reducing the area of the last tubular section to 0.00001 cm^2 , thus reducing the airflow to effectively zero without causing any numerical overflow. As discussed in Maeda (1982), the

simulation sampling frequency, f_s , needs to be sufficiently high in order to avoid frequency warping effects. Here, all the simulations were carried out at $f_s = 70$ kHz. The scheme was implemented using a personal computer running MATLAB.

D. Simulations

The following sections describe two simulations that were undertaken to study the subglottal models. In the first, a simple method for computing the acoustic input impedance of each subglottal system is put forth. Estimates of the acoustic input impedance are intended for allowing comparisons of the subglottal models with the subglottal properties reported in the literature. In the second set of simulations, the subglottal models are coupled to the supraglottal models (as described in Sec. II B 3) with the classical self-oscillating, two-mass model of *Ishizaka and Flanagan (1972)*. The simulations were carried out with different supraglottal configurations representing different vowels. The purpose of this set of simulations was to explore possible scenarios of speech sound propagation into the subglottal airways.

1. Subglottal input impedance

The subglottal system has been most commonly characterized by computing (in models) and/or measuring (in physical models or humans) the acoustic input impedance from the tracheal exit. In each of the three subglottal models (W_{sym} , $Y\&S_{\text{sym}}$, and $Y\&S_{\text{asym}}$), we computed the acoustic input impedance by forcing a sinusoidal pressure source, $P_{\text{drv}} = \sin(2\pi f_{\text{drv}} t)$, at the exit of the last tracheal segment. The relation of the subglottal input impedance, $\tilde{Z}_{\text{in}}^{\text{SUB}}$, with the pressure driver and the input current, \tilde{U}_{drv} , is then,

$$\tilde{Z}_{\text{in}}^{\text{SUB}} = \frac{\tilde{P}_{\text{drv}}}{\tilde{U}_{\text{drv}}} = R_{\text{in}}^{\text{SUB}} + jX_{\text{in}}^{\text{SUB}}. \quad (37)$$

TABLE III. Parameters values used in the two-mass vocal fold model.

Parameter	Value	Units
Glottal length, l_g	1.2	cm
Lower mass, m_1	0.125	g
Upper mass, m_2	0.025	g
Upper mass thickness, d_{m1}	0.25	cm
Lower mass thickness, d_{m2}	0.05	cm
Initial upper area, A_{g01}	0.05	cm ²
Initial lower area, A_{g02}	0.05	cm ²
Lower spring stiffness, k_1	80000	dyne/cm
Upper spring stiffness, k_2	8000	dyne/cm
Coupling spring stiffness, k_c	25000	dyne/cm
Lower mass damping ratio, $\zeta_{1(\text{open})}$	0.1	g/s
Upper mass damping ratio, $\zeta_{2(\text{open})}$	0.6	g/s
Lower mass damping ratio, $\zeta_{1(\text{closed})}$	1.1	g/s
Upper mass damping ratio, $\zeta_{2(\text{closed})}$	1.6	g/s
Nonlinear spring coefficients, $\eta_{k1,2}$	100	—
Nonlinear spring coefficients, $\eta_{h1,2}$	500	—
Nonlinear spring coefficients, $h_{1,2}$	$3k_{1,2}$	dyne/cm

The magnitude of the input impedance, $|\tilde{Z}_{\text{in}}^{\text{SUB}}|$, and the resistance $R_{\text{in}}^{\text{SUB}}$ and reactance $X_{\text{in}}^{\text{SUB}}$ components are given by

$$|\tilde{Z}_{\text{in}}^{\text{SUB}}| = \frac{|\tilde{P}_{\text{drv}}|}{|\tilde{U}_{\text{drv}}|} = \frac{1}{|\tilde{U}_{\text{drv}}|}, \quad (38a)$$

$$R_{\text{in}}^{\text{SUB}} = |\tilde{Z}_{\text{in}}^{\text{SUB}}| \cos \theta, \quad (38b)$$

$$X_{\text{in}}^{\text{SUB}} = |\tilde{Z}_{\text{in}}^{\text{SUB}}| \sin \theta, \quad (38c)$$

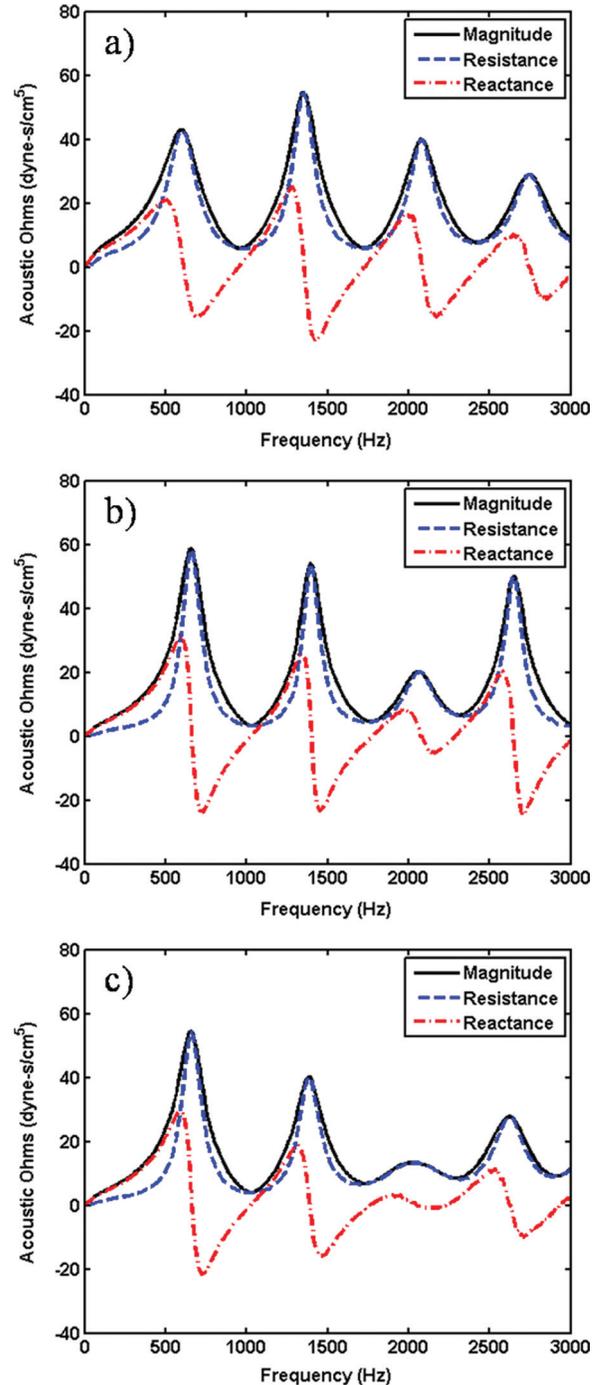


FIG. 8. (Color online) Subglottal input impedances, including their respective magnitude (solid line), resistance (dashed line), and reactance (dash-dotted line) components, computed from models (a) W_{sym} , (b) $Y\&S_{\text{sym}}$, and (c) $Y\&S_{\text{asym}}$.

TABLE IV. Summary of the subglottal resonance frequencies (SubR1-3), peak impedances (PZ1-3), and bandwidths (BW1-3) found in models $Weibel_{sym}$, $Y\&S_{sym}$, and $Y\&S_{asym}$.

	SubR1 (Hz)	SubR2 (Hz)	SubR3 (Hz)	PZ1 dyne-s/cm ⁵	PZ2 dyne-s/cm ⁵	PZ3 dyne-s/cm ⁵	BW1 (Hz)	BW2 (Hz)	BW3 (Hz)
$Weibel_{sym}$	604	1354	2077	42	52	39	196	145	172
$Y\&S_{sym}$	654	1402	2054	58	53	20	131	117	187
$Y\&S_{asym}$	654	1379	2022	54	40	14	131	145	401

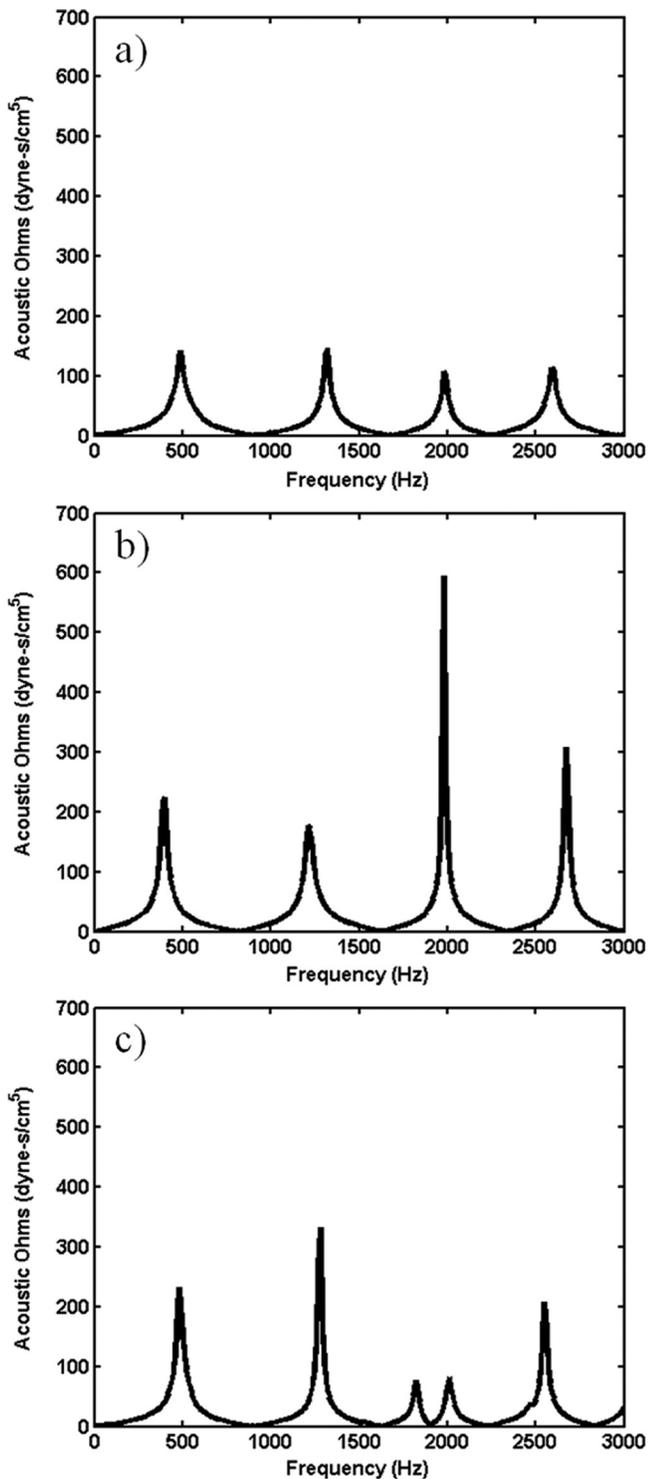


FIG. 9. Subglottal input impedances (magnitude only) estimated for the models (a) $Weibel_{sym}$, (b) $Y\&S_{sym}$, and (c) $Y\&S_{asym}$ with hard walls.

where θ is the phase difference between P_{drv} and U_{drv} (omitting the complex signs \sim here since the generated signals are purely real). Equations (37) and (38) were evaluated over the frequency range of f_{drv} from 10 to 3000 Hz.

2. Speech sound propagation into the subglottal airways

It was also desired to explore the subglottal models under conditions of voice production in order to examine the possible scenarios of speech sound propagation into the subglottal airways. For this purpose, we used the complete system equation (including both subglottal and supraglottal systems) as described in Sec. II B 3. For all cases, the subglottal and supraglottal tracts were maintained static throughout the entire simulations. The alveolar pressure, P_{alv} , was slowly increased to a nominal value of 7 cmH₂O (6864 dyne/cm) over the first 2.5 ms of the simulation. Three supraglottal tract configurations corresponding to the vowels /a/, /e/, and /i/ were considered. Here, we make use of the vocal tract and piriform fossae area functions reported in Takemoto *et al.* (2006), which were obtained using a cine-magnetic resonance imaging (MRI) technique from the normal Japanese patients. For simplicity, the NT was not included in this set of simulations. In the previous studies by Maeda (1982) and Mokhtari *et al.* (2008), the vocal folds were simulated with a forced-oscillation, single-mass model. In the present study, we opted for the self-oscillating, two-mass model of Ishizaka and Flanagan (1972) as it provides a more realistic approximation of the vocal fold vibrations. Thus, the circuit elements of the glottis L_g and R_g [in Eq. (33a)] were replaced by those of the two masses as indicated in Ishizaka and Flanagan (1972). The vocal fold parameters used in the two-mass model are summarized in Table III.

III. RESULTS AND DISCUSSION

A. Subglottal input impedance

Figure 8 shows the magnitude of the subglottal input impedances (solid lines), including their resistance (dashed lines) and reactance (dashed-dotted lines) components, estimated for each of the models (a) $Weibel_{sym}$, (b) $Y\&S_{sym}$, and (c) $Y\&S_{asym}$. Table IV summarizes the subglottal resonance frequencies, peak impedances, and bandwidths of each model. In general, these results are in good agreement with those estimated from *in vivo* measurements in laryngectomees by Ishizaka *et al.* (1976) and in normal male human subjects by Habib *et al.* (1994a), and to some extent with those estimated from spectral analysis of aspirated sounds in

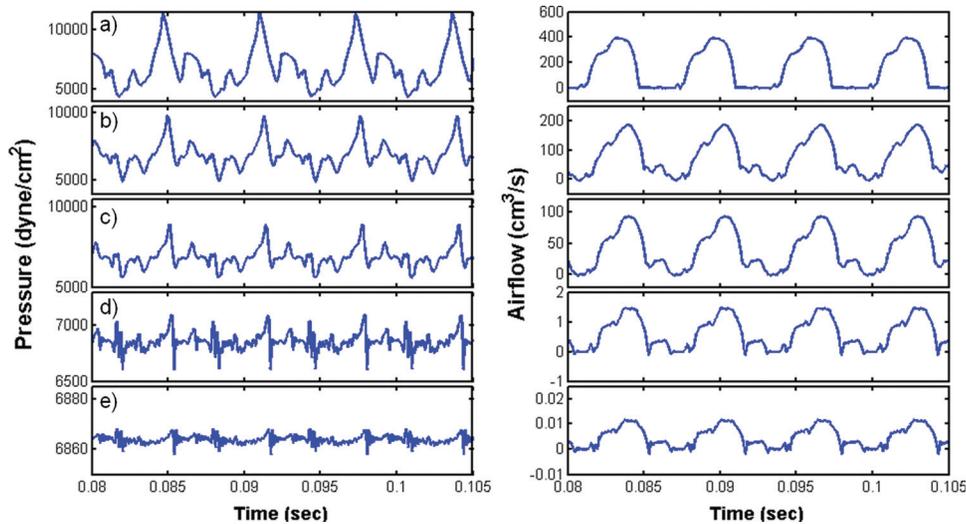


FIG. 10. (Color online) Acoustic pressure (left column) and airflow (right column) waveforms at the exit of the (a) 0th generation (i.e., directly below the vocal folds), (b) 1st generation (exit of a primary bronchus), (c) 2nd generation (exit of a secondary bronchus), (d) 8th generation, and (e) 15th generation on the Weibel_{sym} model. Source excitations were generated using a two-mass vocal fold model with a supraglottal tract configured for a vowel /a/.

male subjects by Klatt and Klatt (1990). The Weibel_{sym} model, in particular, seems to be able to replicate not only the frequency locations of each subglottal resonance but also the relative amplitudes of each peak when compared to the measurements made by Ishizaka *et al.* (cf. Fig. 3 in Ishizaka *et al.*, 1976).

A fourth subglottal resonance was also found in all models in the vicinity of 2700–2800 Hz (Fig. 8). This resonance has also been observed in other theoretical models (Fant *et al.*, 1972; Ishizaka *et al.*, 1976); however, results from measurements on human subjects have shown large inter-subject variability in this frequency range.

For comparison, the input impedances of the models were also computed using a hard-walled configuration. Perfectly rigid airways can be simulated by decoupling the yielding wall elements (L_{wn} , R_{wn} , C_{wn}) in all the TL blocks along the subglottal network. The resulting subglottal input impedances with this wall configuration are shown in Fig. 9. In general, a hard-walled configuration caused a shift of the subglottal resonances toward lower frequencies

accompanied with an increase in their impedance magnitudes. This change is in agreement with the predictions in the models of Ishizaka *et al.* (1976) and Habib *et al.* (1994a). It is interesting to notice that with this configuration, the Y&S_{asym} model exhibits an additional resonance in the vicinity of 1800 Hz. We suspect that this extra resonance results from the dissimilar lobes resonating at slightly different frequencies. When using a soft-walled configuration, these two resonances are damped and form a single observable resonance peak.

B. Speech sound propagation into the subglottal airways

Figures 10 and 11 show the results of the simulations involving speech sound propagation into the subglottal airways obtained in the Weibel_{sym} and Y&S_{sym} models, respectively, for a supraglottal system configured as a vowel /a/. Each figure shows the acoustic pressure (left column) and airflow (right column) at the exit of the (a) 0th generation

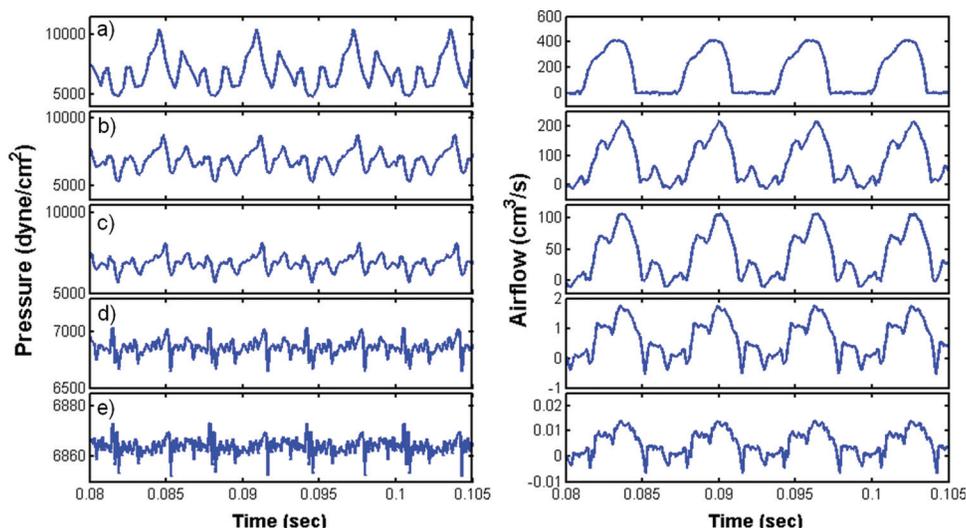


FIG. 11. (Color online) Acoustic pressure (left column) and airflow (right column) waveforms at the exit of the (a) 0th generation (i.e., directly below the vocal folds), (b) 1st generation (exit of the primary bronchus), (c) 2nd generation (exit of the secondary bronchus), (d) 8th generation, and (e) 15th generation on the Y&S_{sym} model. Source excitations were generated using a two-mass vocal fold model with a supraglottal tract configured for a vowel /a/.

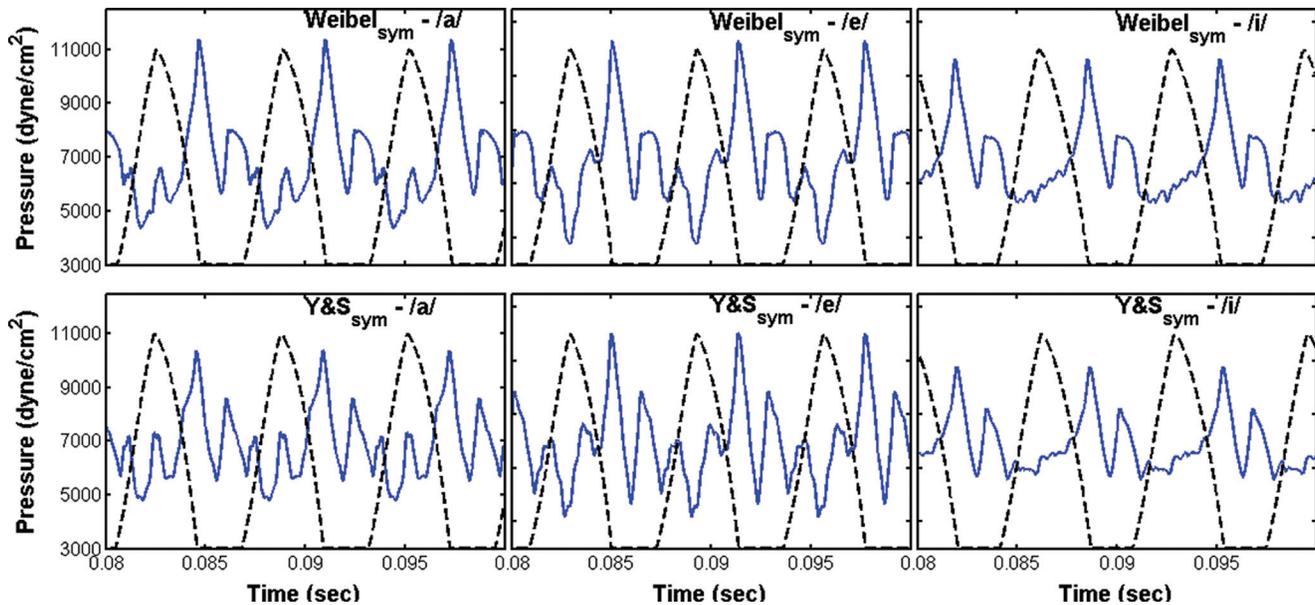


FIG. 12. (Color online) Subglottal pressure waveforms (solid lines) predicted in the Weibel_{sym} and Y&S_{sym} models for supraglottal tract configurations corresponding to vowels /a/, /e/, and /i/. Superimposed dashed lines depict the resulting glottal area in the self-oscillating, two-mass vocal fold model.

(i.e., immediately below the glottis), (b) 1st generation, (c) 2nd generation, (d) 8th generation, and (e) 15th generation. As expected, the amplitudes of the pressure and airflow waveforms decrease rapidly as a function of the generation level since the sound energy is distributed into an effectively increasing number of airways. A particularly interesting observation that can be made from the airflow waveforms in cases (b)–(e) is the consistent presence of a secondary peak during the beginning of the closed phase. The presence of a similar peak has been commonly observed in estimates of the glottal airflow obtained from inverse filtering (Gauffin and Sundberg, 1989; Hertegard and Gauffin, 1995). Our results seem to suggest, thus, that a plausible explanation for this secondary peak is that it is caused by the resonances of the subglottal system and is eventually “leaked” into the vocal tract through a glottal chink and/or an incomplete closure of the vocal folds. Further investigations with a vocal fold model that does not achieve complete glottal closure are needed to verify this hypothesis.

When comparing the results predicted in the Y&S_{sym} and Y&S_{asym} models (not shown), we observed that for most of the lower generation levels (0th, 1st, and 2nd generations), the waveforms were almost indistinguishable. Somewhat larger differences were observed at higher airway generations, however.

Figure 12 depicts the subglottal pressure waveforms directly below the vocal folds (solid lines) obtained in the Weibel_{sym} (top row) and Y&S_{sym} (bottom row) models along with their corresponding glottal areas (dashed lines) for vowels /a/, /e/, and /i/. Here, it is noticeable that the large pressure peaks occur at the instants of glottal closure. This sudden increase in the subglottal pressure is considered to be analogous to the “water hammer” effect experienced in liquid-filled pipelines (Ishizaka *et al.*, 1976; Boves, 1984; Sciamarella and Artana, 2009). During the closed phase, a damped transient response is observed. This transient pattern is consistent from vowel to vowel since the subglottal

resonances are predominant during this phase. During the open phase, the pressure tends to drop as air flows into the vocal tract. This portion is highly vowel dependent due to acoustic interactions arising between the subglottal and supraglottal tracts. Similar subglottal pressure patterns have been observed from the *in vivo* recordings made by Koike and Hirano (1973), Cranen and Boves (1985, 1987), Hertegard *et al.* (1995), Neumann *et al.* (2003), and Miller and Schutte (1991) (presented in Austin and Titze, 1997).

Based on tracheal puncture measurements on a male subject while performing repetitions of the syllable /pa/, Hertegard *et al.* (1995) estimated that the pressure peaks at the instant of glottal closure were approximately 40%–45% above the mean subglottal pressure, whereas the pressure drops occurring during the open phase were approximately 30%–35% below the mean subglottal pressure. In our simulations corresponding to a vowel /a/ (shown in Fig. 12), the pressure peaks and pressure drops, calculated over 16 glottal cycles, were found to be 66% above and 37% below the mean subglottal pressure in the Weibel_{sym} model, and 51% above and 30% below the mean subglottal pressure in the Y&S_{sym} model, respectively. Differences between our results and those obtained by Hertegard *et al.* (1995) for the maximum pressure peaks may be attributed to the fact that the two-mass vocal fold model does not account for a possible glottal chink which may act as a pressure leak during the glottal closure that decreases the amplitude of the sudden peak and that in our simulations the mean subglottal pressure was lower.

IV. CONCLUSIONS

In this paper, we have presented a time-domain computational model of sound wave propagation in the subglottal system. The model recreates the branching structure of the tracheobronchial airways and extends an acoustic TL modeling scheme initially conceived by Maeda and recently

improved by Mokhtari *et al.* for representing the supraglottal system (Maeda, 1982; Mokhtari *et al.*, 2008). The current model can be configured so that it simulates a subglottal system that branches in a symmetric or asymmetric pattern. Two symmetric models, based on the anatomical data from Weibel's model A (Weibel, 1963) and Yeh and Schum's typical path lung model for the whole lung (Yeh and Schum, 1980), and one asymmetric model, based on Yeh and Schum's typical path lung models for each lobe and major bronchi (Yeh and Schum, 1980), were constructed, and their acoustic properties were examined by computing their respective subglottal input impedances. The resonance frequencies, peak impedances, and bandwidths estimated from these three models were in good agreement with those obtained by Ishizaka *et al.* (1976) from Japanese patients with laryngectomies and by Habib *et al.* (1994a) from normal male subjects. Comparing the estimated subglottal input impedances from the symmetric and asymmetric models, the symmetric models predict reasonably realistic subglottal resonances while maintaining a reduced computational load. In particular, we found that the symmetric model Weibel_{sym} was capable of closely replicating both the frequency locations of the subglottal resonances and the relative amplitudes between each peak observed by Ishizaka *et al.* Simulations of speech sound propagation into the subglottal airways, associated with the voicing of sustained vowels, were achieved by coupling the subglottal and supraglottal systems with the classical self-oscillating, two-mass vocal fold model of Ishizaka and Flanagan (1972). The predicted subglottal pressure waveforms directly below the vocal folds were in qualitative agreement with those previously observed from *in vivo* measurements.

To our knowledge, this is the first anatomically based model of the subglottal system for speech production that allows for quantitative predictions of sound wave propagation at any level of the branching airways in the time-domain. Overall, this modeling approach provides a comprehensive framework to study the impact of the subglottal system on various voiced sounds in both health and disease. Future studies that incorporate and compare the importance of branching and depth, as well as vocal fold model complexity, on various measurable speech parameters are planned.

- Austin, S. F., and Titze, I. R. (1997). "The effect of subglottal resonance upon vocal fold vibration," *J. Voice* **11**, 391–402.
- Birkholz, P., Jackel, D., and Kroger, B. J. (2007). "Simulation of losses due to turbulence in the time-varying vocal system," *IEEE Trans. Audio, Speech, Lang. Process.* **15**, 1218–1226.
- Boves, L. (1984). *The Phonetic Basis of Perceptual Ratings of Running Speech* (Foris Publications, Dordrecht, Holland), pp. 68–88.
- Chi, X., and Sonderegger, M. (2007). "Subglottal coupling and its influence on vowel formants," *J. Acoust. Soc. Am.* **122**, 1735–1745.
- Cranen, B., and Boves, L. (1985). "Pressure measurements during speech production using semiconductor miniature pressure transducers – Impact on models for speech production," *J. Acoust. Soc. Am.* **77**, 1543–1551.
- Cranen, B., and Boves, L. (1987). "On subglottal formant analysis," *J. Acoust. Soc. Am.* **81**, 734–746.
- Cranen, B., and Schroeter, J. (1995). "Modeling a leaky glottis," *J. Phonetics* **23**, 165–177.
- Eriksson, L. J. (1980). "Higher-order mode effects in circular ducts and expansion chambers," *J. Acoust. Soc. Am.* **68**, 545–550.
- Fant, G. (1960). *Acoustic Theory of Speech Production, with Calculations Based on X-ray Studies of Russian Articulations* (Mouton, s' Gravenhage), pp. 27–28.
- Fant, G., Ishizaka, K., Lindqvist-Gauffin, J., and Sundberg, J. (1972). "Subglottal formants," *KTH Dept. Speech, Music Hear. Q. Prog. Status Rep.* **13**, 1–12.
- Flanagan, J. L. (1972). *Speech Analysis Synthesis and Perception* (Springer-Verlag, Berlin), pp. 23–53.
- Flanagan, J. L., and Landgraf, L. L. (1968). "Self-oscillating source for vocal-tract synthesizers," *IEEE Trans. Audio. Electroacoust.* **16**, 57–64.
- Fredberg, J., and Hoenig, A. (1978). "Mechanical response of the lungs at high frequencies," *J. Biomech. Eng.* **100**, 57–66.
- Gauffin, J., and Sundberg, J. (1989). "Spectral correlates of glottal voice source waveform characteristics," *J. Speech Hear. Res.* **32**, 556–565.
- Habib, R. H., Chalker, R. B., Suki, B., and Jackson, A. C. (1994a). "Airway geometry and wall mechanical-properties estimated from subglottal input Impedance in humans," *J. Appl. Physiol.* **77**, 441–451.
- Habib, R. H., Suki, S., Bates, J. H. T., and Jackson, A. C. (1994b). "Serial distribution of airway mechanical-properties in dogs—Effects of histamine," *J. Appl. Physiol.* **77**, 554–566.
- Harper, 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.
- Hertegard, S., and Gauffin, J. (1995). "Glottal area and vibratory patterns studied with simultaneous stroboscopy, flow glottography, and electroglottography," *J. Speech Hear. Res.* **38**, 85–100.
- Hertegard, S., Gauffin, J., and Lindestad, P. A. (1995). "A comparison of subglottal and intraoral pressure measurements during phonation," *J. Voice* **9**, 149–155.
- Hudde, H., and Slatky, H. (1989). "The acoustical input impedance of excised human lungs – Measurements and model-matching," *J. Acoust. Soc. Am.* **86**, 475–492.
- Ishizaka, K., and Flanagan, J. L. (1972). "Synthesis of voiced sounds from a two-mass model of vocal cords," *Bell Syst. Tech. J.* **51**, 1233–1268.
- Ishizaka, K., French, J. C., and Flanagan, J. L. (1975). "Direct determination of vocal-tract wall impedance," *IEEE Trans. Acoust., Speech, Signal Process.* **23**, 370–373.
- Ishizaka, K., Matsudaira, M., and Kaneko, T. (1976). "Input acoustic-impedance measurement of subglottal system," *J. Acoust. Soc. Am.* **60**, 190–197.
- Klatt, D. H., and Klatt, L. C. (1990). "Analysis, synthesis, and perception of voice quality variations among female and male talkers," *J. Acoust. Soc. Am.* **87**, 820–857.
- Koike, Y., and Hirano, M. (1973). "Glottal-area time function and subglottal-pressure variation," *J. Acoust. Soc. Am.* **54**, 1618–1627.
- Koizumi, T., Taniguchi, S., and Hiromitsu, S. (1985). "Glottal source vocal-tract interaction," *J. Acoust. Soc. Am.* **78**, 1541–1547.
- Lucero, J. C. (1993). "Dynamics of the two-mass model of the vocal folds: Equilibria, bifurcations, and oscillation region," *J. Acoust. Soc. Am.* **94**, 3104–3111.
- Maeda, S. (1982). "A digital simulation method of the vocal-tract system," *Speech Commun.* **1**, 199–229.
- Miller, D. G., and Schutte, H. K. (1991). *Effects of Downstream Occlusions on Pressure near the Glottis in Singing* (Singular Publishing Group, San Diego, CA), pp. 91–98.
- Mokhtari, P., Takemoto, H., and Kitamura, T. (2008). "Single-matrix formulation of a time domain acoustic model of the vocal tract with side branches," *Speech Commun.* **50**, 179–190.
- Neumann, K., Gall, V., Schutte, H. K., and Miller, D. G. (2003). "A new method to record subglottal pressure waves: Potential applications," *J. Voice* **17**, 140–159.
- Sciamarella, D., and Artana, G. (2009). "A water hammer analysis of pressure and flow in the voice production system," *Speech Commun.* **51**, 344–351.
- Sondhi, M., and Schroeter, J. (1987). "A hybrid time-frequency domain articulatory speech synthesizer," *IEEE Trans. Acoust., Speech, Signal Process.* **35**, 955–967.
- Stevens, K. N. (1998). *Acoustic Phonetics* (MIT Press, Cambridge, MA), pp. 196–198.
- Suki, B., Habib, R. H., and Jackson, A. C. (1993). "Wave propagation, input impedance, and wall mechanics of the calf trachea from 16 to 1,600 Hz," *J. Appl. Physiol.* **75**, 2755–2766.
- Takemoto, H., Honda, K., Masaki, S., Shimada, Y., and Fujimoto, I. (2006). "Measurement of temporal changes in vocal tract area function from 3D cine-MRI data," *J. Acoust. Soc. Am.* **119**, 1037–1049.

- 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. (2008). "Nonlinear source-filter coupling in phonation: Theory," *J. Acoust. Soc. Am.* **123**, 2733–2749.
- Titze, I. R., and Story, B. H. (1997). "Acoustic interactions of the voice source with the lower vocal tract," *J. Acoust. Soc. Am.* **101**, 2234–2243.
- Titze, I. R., and Sundberg, J. (1992). "Vocal intensity in speakers and singers," *J. Acoust. Soc. Am.* **91**, 2936–2946.
- van den Berg, J. (1958). "Myoelastic-aerodynamic theory of voice production," *J. Speech Lang. Hear. Res.* **1**, 227–244.
- Weibel, E. R. (1963). *Morphometry of the Human Lung* (Springer, Berlin), pp. 136–140.
- Yeh, H. C., and Schum, G. M. (1980). "Models of human-lung airways and their application to inhaled particle deposition," *Bull. Math. Biol.* **42**, 461–480.
- Zañartu, M., Mongeau, L., and Wodicka, G. R. (2007). "Influence of acoustic loading on an effective single mass model of the vocal folds," *J. Acoust. Soc. Am.* **121**, 1119–1129.
- Zhang, Z. Y., Neubauer, J., and Berry, D. A. (2006). "The influence of subglottal acoustics on laboratory models of phonation," *J. Acoust. Soc. Am.* **120**, 1558–1569.