Tongue-tip trills and vocal-tract wall compliance

Richard S. McGowan
Haskins Laboratories, 270 Crown Street, New Haven, Connecticut 06511

(Received 24 May 1991; accepted for publication 6 February 1992)

The effect of vocal-tract wall compliance on tongue-tip trills is to create a favorable pressure-flow relation at the tongue tip for sustained vibration. The governing equations are derived for a model based on this mechanism, and data on unvoiced trills are used to help set parameters for a numerical simulation of the model.

PACS numbers: 43.70.Aj, 43.70.Bk

INTRODUCTION

The majority of /r/ sounds with an alveolar-dental place of articulation are specified as interrupted: trills, taps, or flaps (Maddieson, 1984). There has been relatively little to explicate the mechanics of tongue-tip trills, and this paper is intended to provide a credible model of how vocal-tract wall compliance can help sustain these trills. Tongue-tip trills can be produced with the sides of the body of the tongue braced against the teeth and with the tongue tip free to move to and from the hard palate. Because the trill is the result of interaction between air and the solid tongue tip, the control of the tongue tip, glottal aperture, and the lungs are all important for initiating and sustaining the trill. It is not surprising, therefore, that tongue-tip trills, as well as sibilants, are the final speech gestures to be mastered by native Spanish speakers (Jimenez, 1987). There is also evidence in English, a language without trills, that tongue-tip control for sibilants has not matured by the 7th year (McGowan and Nittrouer, 1988). Here, we concentrate on the trill as a case of solid–air interaction that results in vibration in the vocal tract.

In the simplest terms, tongue-tip trills might be described in the following way. The tongue tip is viewed as a sprung trap door, with the oral cavity providing a container, or plenum, where air can have a greater pressure than atmospheric. The pressure difference between the plenum and the atmosphere causes the trap door to open as the torque provided by that pressure difference is sufficient to overcome the torsional tension of the trap door spring. The air rushes out, and the air pressure in the plenum drops so that the torsion spring closes the trap door over the mouth of the plenum. The plenum is refilled by air flowing from the lungs through the glottal resistance, and the cyclic process continues.

Energy is exchanged from the air to the tongue tip when the tip is moving away from the region of higher air pressure in the oral cavity and toward the region of lower air pressure outside the mouth, that is, during the outward movement of the tongue. The opposite exchange of energy occurs during the inward movement of the tongue tip, if the tongue tip can be supposed to be moving toward a region of higher pressure in its return toward the hard palate. These statements follow from the definition of mechanical work done on a particle of mass as the integral of the force in the direction of motion against increments of path length. The air pressure provides force distributed over a collection of masses comprising the tongue tip: The total work done on the tongue, or energy input to the tongue, is the sum of the work done by the air on each incremental mass of the tongue.

The simple description of the physics of tongue-tip trills is not specific enough to rule out the possibility that the work done by the air on the tongue in opening is equal to work done by the tongue on the air during the closing phase. The air in the lungs provides a subglottal pressure higher than atmospheric pressure, and the glottis provides a resistance to flow due to the formation of rotational air motion and viscous boundary layers. The constriction at the tongue tip also provides a resistance that is time varying. Assuming that the flow is always egressive, and the subglottal pressure is constant, the intraoral pressure depends only on the areas of the glottal opening and the tongue-tip constriction, and not the direction of the tongue-tip movement. With the glottis at a constant area or varying at such a fast rate, as in voicing, that it has a constant area considered in relation to the time scale of the tongue trill, the net pressure on the tongue tip depends only on its position, because this determines the tongue-tip constriction area, so no net work is done on the tongue in a complete cycle. Under these conditions as much energy is removed from the air in any path increment of the trap door opening as is given back to the air during the inward movement over the same path increment. This symmetry is analogous to the state of affairs with the one-mass model of the vocal folds when vocal-tract loading is not considered.

There are several ways to add details to the simple description so that the symmetry in the outward and inward movements is broken. The description can be modified to provide for more than one degree of freedom for the tongue tip. If the tongue tip can follow two different paths in the inward and outward motion, then the symmetry in energy exchange no longer has to be the case. One way for this to happen is for surface wave motion of the epithelium to occur, which, in turn, may be modeled by a two-mass system much like vocal fold vibration. There is no experimental evidence one way or another on whether the tongue tip moves with more than one degree of freedom. The vocal folds, during certain modes of phonation, possess a noticeable wave motion on their epithelia, and, until there is a decisive experiment, this motion cannot be ruled out for the tongue tip.

Possibilities for single-degree-of freedom motion for the

 0001-4966/92/052903-08$00.80

tongue tip can also be considered. Here, again, vocal fold models provide some guide (Flanagan and Landgraf, 1968). When the epithelia of the vocal folds are put under sufficient tension, as in falsetto mode, surface wave motion ceases and each fold moves with a single degree of freedom. The necessary asymmetry in the opening and closing of the vocal folds during single-degree-of-freedom motion is attributed to acoustic loads provided by the vocal tract on the glottal flow system (Titze, 1988). In particular, when the reactive part of the impedance looking into the vocal tract from the glottis is inductive, there is more pressure on the folds during lateral movement than during medial movement. This results because the inductive load provides a phase shift of the glottal area and glottal volume velocity with respect to the pressure in the glottis.

A capacitive load might be responsible for a phase shift in the tongue-tip constriction area and volume velocity with respect to oral pressure in the case of tongue-tip trills. A capacitive load in the vocal tract would not provide the correct phase shift in glottal area with respect to glottal pressure in the case of single-degree-of-freedom phonation. However, a capacitive load behind the vibrating object in the case of tongue-tip trills should provide the correct phase shift to sustain this vibration. There may be sufficient capacitance at low frequencies to create a substantial phase shift if the compliance of the vocal-tract walls is taken into account. [Using the tense cheek data (Ishizaka *et al.*, 1975; Westbury, 1983) for vocal-tract walls behind the tongue tip, the resonance frequency of the walls relevant to tongue-tip trills is around 60 Hz. The tongue-tip trills observed by the author were between 25 and 35 Hz.] Vocal-tract loading as applied to tongue-tip trills is explored in the model presented in the next section.

The model of the effect of vocal-tract compliance presented here is in the spirit of some earlier vocal fold vibration models, such as the one-mass and two-mass models (Flanagan and Landgraf, 1968; Ishizaka and Matsudaira, 1972). These models have been invaluable tools for understanding the solid–flow interaction and resulting aerodynamic forces on the folds in the production of voice, and they can be incorporated into an articulatory synthesizer. Not only would such a model be helpful in understanding tongue-tip trills, but also the solid–flow interactions and aerodynamic forces on the tongue tip during alveolar stop closure and release, and during sibilant production. Modeling of aerodynamic forces on supraglottal structures during consonant production, including taps and trills, has received very little attention (see, however, Stevens, 1991). Understanding the balance between articulator movements and forces, including the larynx, and aerodynamics is essential for articulatory modeling and synthesis.

I. DESCRIPTION OF THE MODEL

Some of the previous lumped element models of stop consonant aerodynamics, which incorporate the mechanical properties of the vocal-tract walls (Rothenberg, 1968; Müller and Brown, 1980; Westbury, 1983) can be used to model the dynamics of tongue trills, because the frequencies of interest are well below 100 Hz, thus making phase differences along the length of the vocal tract negligible. A model for the effect of the compliance of the vocal tract on tongue trills is illustrated in Figs. 1 and 2. The subglottal system is assumed to be a constant pressure source, at pressure P_S, that couples to the upper vocal tract through the glottal area A_G with resistance R_G, and glottal inductance L_G. The tongue tip forms a tight constriction of variable area A_C, flow resistance R_C, and inductance L_C. The supraglottal vocal tract walls behind the tongue-tip constriction are modeled as a lumped, second-order, mass-spring system, with inductance L_W, resistance R_W, and capacitance C_W. The compliance of the air between the tongue tip constriction and the glottis is taken into account with a capacitance C_A. The tongue tip is assumed to move as a trap door, with moment of inertia I, hinged at the connection with the tongue body, and sprung with both linear and nonlinear, cubic springs with constants κ and $\eta\kappa$, respectively, opposing the torque τ, provided by the intraoral pressure P_C. Resistance in the tongue tip is provided by the resistive constant r. The use of a damped vibrator with a cubic nonlinearity is reminiscent of the sprung masses in the two-mass model (Ishizaka and Flanagan, 1972).

Calculations were performed from Kirchoff's current and voltage laws applied to this lumped element circuit, coupled with the dynamical equation for the trap door tongue tip. With volume velocities U_G, through the glottis, U_W of the vocal tract walls, U_C through the tongue constriction, and U_A, due to air compliance, the governing equations are

$$P_C = R_C U_C + L_C \frac{dU_C}{dt} = \frac{1}{C_A} \int U_A \, dt ,$$

$$= L_W \frac{dU_W}{dt} + R_W U_W + \frac{1}{C_W} \int U_W \, dt , \qquad (1)$$

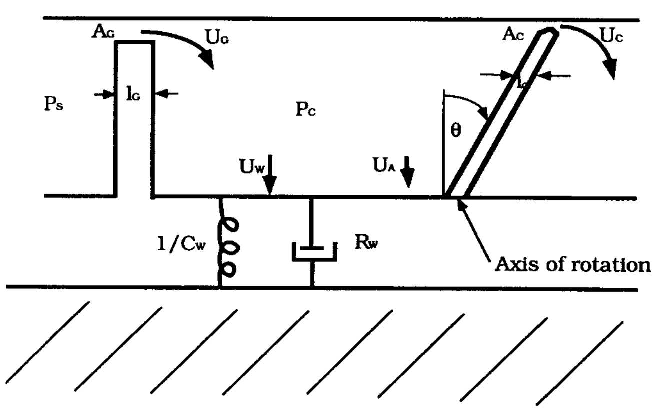

FIG. 1. Model vocal tract for tongue-tip trills. P_S = subglottal pressure, P_C = cavity or intraoral pressure, U_G = volume velocity at the glottis, U_C = volume velocity at the tongue-tip constriction, U_W = volume velocity of the vocal tract walls, U_A = volume velocity due to the compliance of the air, A_G = glottal channel area, A_C = constriction channel area, l_G = glottal channel length, l_C = thickness of the tongue tip, θ = angle of rotation of the tongue tip, C_W = wall compliance, and R_W = wall resistance. Not shown in the figure are the tongue tip length h_T and the dimension of the tongue tip into the paper, the breadth b_T. The mass of the wall is not explicitly indicated.

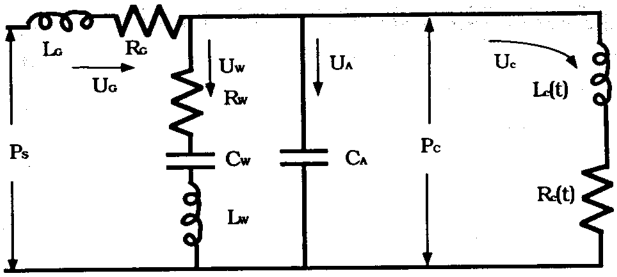

FIG. 2. Circuit diagram corresponding to Fig. 1. C_A = capacitance of the air contained in the oral cavity, L_G = glottal inductance, R_G = glottal flow resistance, L_C = time-varying constriction inductance, R_C = time-varying constriction resistance, and L_W = wall inductance.

$$P_S - P_C = R_G U_G + L_G \frac{dU_G}{dt}, \quad (2)$$

$$U_G = U_W + U_C + U_A, \quad (3)$$

$$\frac{d^2\theta}{dt^2} + \frac{r}{I}\frac{d\theta}{dt} + \frac{\kappa}{I}(1 + \eta\theta^2)\theta = \frac{\tau}{I}, \quad (4)$$

where $\theta = 0$ is the rest angle. The tongue tip is presumed to be a rectangular slab with height h_T, thickness l_C, and breadth (dimension into the paper) $= b_T$. The slab has a sharp upper edge so that it articulates with the hard palate along a line. The torque τ and moment of inertia I are then

$$\tau = \frac{1}{2} b_T h_T^2 \left[P_C - K \frac{\rho_A}{2} \left(\frac{U_C}{A_C} \right)^2 \right], \quad (5)$$

$$I = \rho_T l_C b_T (h_T^3/3), \quad (6)$$

where ρ_T is the density of the tongue tip and ρ_A = density of air. Here, K is a parameter that quantifies the reduction in pressure resulting from flow along the tongue-tip surface due to the Bernoulli effect. The relationship between the angle of rotation θ and the constriction area A_C is given by

$$A_C = b_T h_T [1 - \cos(\theta)], \quad (7)$$

where the maximum $A_C = b_T h_T$. It can be noted that the mechanics of the air and the tongue tip are coupled through Eqs. (1) and (4), because the elements R_C and L_C depend on the constriction area A_C, and the torque τ depends on P_C [Eq. (5)].

For the frequencies of interest, which are less than 100 Hz, the inductance of the air in the glottal and tongue constriction can be neglected (Rothenberg, 1968, p. 18). However, the inductances of the constrictions were included for completeness' sake, and they may be excluded for less intensive calculations. A simplifying assumption that is harder to justify is that of constant subglottal pressure. With a high impedance in the upper vocal tract and relatively open glottis, there should be important perturbations to the volume velocity at the glottis, and hence to the pressure below the glottis (Rothenberg, 1968). However, the fundamentals of this model of trills can be illustrated without the mechanical properties of the subglottal system, which, if taken into account, would have added a considerable amount of computational burden to the numerical simulation.

II. METHOD: PARAMETER ESTIMATES

To simulate the effect of the loading of the vocal tract on tongue trills it was necessary to set numerical values for the parameters involved. Experimental results were used as much as possible, but it was not possible to derive all the numerical values with reference to experimental results. The mechanical parameters for the tongue tip were especially problematical, but reasonable order-of-magnitude estimates could be made for some of these.

In one experimental procedure intraoral pressure and oral flow were recorded simultaneously in two subjects producing a series of voiced and voiceless trills. Each trill was held for approximately 2 s. The intraoral pressure was measured using a Millar MT-10 catheter pressure transducer, and the oral–nasal flow was measured using a Glottal enterprises vented pneumotacograph ("Rothenberg mask") without inverse filtering. The pressure transducer was passed through a hole in the center of the mask, which was subsequently sealed with clay, and through the nares to the oral cavity. The intraoral pressure, oral–nasal flow, and acoustic signal were simultaneously recorded on an FM tape recorder.

Because the variations in pressure and flow due to glottal vibrations were present in the voiced samples, the voiceless samples were used to extract pressure and flow. Representative maximum, sustained intraoral pressures during a voiceless trill were used for an estimate of subglottal pressure. One subject, AL, had maximum, sustained intraoral pressures of approximately 10-cm H_2O for voiceless trills. The other subject, RM, had approximately 16-cm H_2O for maximum, sustained intraoral pressures during his voiceless trills. Maximum tongue-tip constriction area can be estimated from the nonlinear relation between pressure and flow (Stevens, 1971), using the maximum volume flow, taken near minimum intraoral pressure, and the intraoral pressure for that volume flow. The measured volume velocity was assumed to be from the tongue-tip constriction alone, and the turbulent loss coefficient at the tongue-tip constriction was taken to be 0.9, which corresponds to a gradual entrance to and an abrupt exit from the constriction region. The maximum constriction opening for subject AL was found to be 6.7×10^{-5} m^2, and 2.7×10^{-5} m^2 for subject RM, based on representative maximum volume velocities of 1.6×10^{-3} m^3/s and 9.2×10^{-4} m^3/s with corresponding intraoral pressures of 540 and 1043 Pa, respectively. These figures were based on representative measurements, and, although the measurements were of typical tokens, they were not averages of many tokens. A constriction channel length of 0.002 m was used to compute the constriction inductance.

For simulation purposes, the glottis was presumed to be rectangular with a gradual entrance and an abrupt exit. The area of the abducted glottis during voiceless trills was assumed to be 5×10^{-5} m^2 or 4×10^{-5} m^2, and for an adducted glottis appropriate during voicing the average open area was taken to be between 8×10^{-6} m^2 and 5×10^{-6} m^2 (close to the figures given by Müller and Brown, 1980). A glottal length of 0.004 m was used to compute glottal inductance.

There has been data collected on vocal tract wall impedance at the neck and at the cheeks in different states of tension (Ishizaka *et al.*, 1975). These data did not completely specify the lumped capacitance, resistance, and inductance for the vocal-tract walls in the model proposed here, but they were taken to provide approximate estimates for these parameters. Because the tongue is braced against the teeth, the walls behind the tongue constriction consist of tongue, hard and soft palates, velum, pharynx, and the superior regions of the larynx. In the simulations for the tongue trill, the values used were those of the lumped parameters derived by Westbury (1983) from the data of Ishizaka *et al.* (1975) for the tense cheek condition and a vocal-tract wall area of 1.25×10^{-2} m^2. In MKS units, the wall capacitance C_W was 2.5×10^{-9}, the wall inductance L_W was 2.0×10^{3}, and the wall resistance R_W was 1.9×10^{6}. The capacitance of the air between the glottis and the tongue tip C_A was estimated to be 8.1×10^{-10} in MKS units.

The mechanical parameters of the trap door tongue tip, τ/I, κ/I, r/I, and η were the most difficult parameters to estimate. Using Eqs. (5) and (6), and the fact that torque is equal to force times distance to the pivot:

$$\frac{\tau}{I}=\beta\left[P_C-K\frac{\rho_A}{2}\left(\frac{U_C}{A_C}\right)^2\right], \tag{8}$$

with $\beta=(3/2)/(\rho_T h_T l_C)$. The parameters used to calculate β were a tongue tip thickness l_C of 0.005 m, the distance from the "hinge" to the constriction h_T of 0.01 m, and the density ρ_T of 10^3 kg/m^3. (Note, that with maximum A_C and h_T set, that b_T is determined.) With these estimates $\beta=30$ in MKS units. However, because β was only given with order-of-magnitude estimates of the dimensions of the tongue tip, this parameter was varied in conjunction with the other tongue-tip parameters, so that sustained oscillations with a frequency of about 30 Hz were obtained. The coefficient that quantifies the Bernoulli effect on the back of the tongue tip K in Eq. (8) was estimated to be 0.25, because, on the average, the particle velocity in this region should be about 1/2 of that through the tongue-tip constriction.

While keeping β within a factor of two of the order-of-magnitude estimate given above, the remaining tongue-tip parameters were varied to obtain sustained oscillation with a frequency between 25 and 35 Hz. That the oscillations were sustained for a given set of parameters was checked by simulating a 2-s trill. It was found that with the nonlinearity parameter $\eta=400$, and with damping ratios ($=r/2\sqrt{\kappa I}$) somewhere between 0.5 and 0.6 sustained trilling would result with κ/I between 30 and 240, a range of 3 in tongue-tip natural frequency.

The numerical simulations were done using a fourth-order Runge–Kutta extrapolation with the time step set to 10^{-6} s. The small time step was necessary because the system of equations behaved as a stiff system and was unstable at larger time steps. This could be attributed to the inclusion of the inductive elements at the glottis and the constriction.

III. RESULTS AND DISCUSSION

Four simulations were run to test the model. Two of these were based on the parameters derived from data taken on AL, and two were based on parameters derived from data taken on RM, one each for an abducted glottis and one each for a relatively constricted glottis. The wall impedances, constriction rest areas, and subglottal pressures were all determined as described in the previous section. The latter two parameters were subject dependent, and the wall impedance was not. The abducted glottis condition was with $A_G=5.0\times10^{-5}$ m^2 for simulation based on AL's data, and $A_G=4.0\times10^{-5}$ m^2 for simulations based on RM's data. The relatively more adducted glottis condition was run with $A_G=8.0\times10^{-6}$ m^2 for simulations based on AL's data, and with $A_G=5.0\times10^{-6}$ m^2 for simulations based on RM's data. The larger glottal area for AL for the relatively closed glottis was used because tongue-tip parameters could not be adjusted to obtain an approximate 30-Hz oscillation at smaller glottal openings. The subglottal pressure for AL was not sufficient to sustain oscillations at higher glottal resistances. Because this subglottal pressure was inferred from measurements taken during voiceless trills, it was possible that AL elevated his subglottal pressure to sustain voiced trills, when the average glottal opening could have been somewhat smaller than 8.0×10^{-6} m^2.

To obtain oscillations between 25 and 35 Hz in the open glottis condition for simulations based on AL's data, β was set equal to 24.0 and κ/I was set to 30.0 with a damping ratio of 0.58, and for simulations based on RM's data β was set equal to 25.0 and κ/I was set to 35.0 with a damping ratio of 0.55. (All parameter values are reported in MKS units.) In the relatively adducted glottis condition for AL, β was set equal to 14.0 and κ/I was set to 240.0, with a damping ratio of 0.58, and for RM β was set equal to 16.0 and κ/I was set to 240.0, with a damping ratio of 0.55.

Figures 3 and 4 show the simulation results for the open glottis condition for AL and RM, respectively. These figures show the traces of the area of the constriction, intraoral pressure, and constriction, glottal, and wall volume velocities in the final 200 ms of a 210-ms simulation. A sample of measured intraoral pressure and oral volume velocity traces ($=$ constriction volume velocity, U_C) during one of AL's unvoiced trills is shown in Fig. 5. The simulated intraoral pressure traces shown in Fig. 3 have minima less than the measured intraoral pressure traces, indicating that the resistance to flow has been underestimated for maximum constriction area. Further, the minimum measured constriction volume velocity is nonzero, indicating that the tongue-tip constriction may not close completely. In general, a major difference between the measured and simulated intraoral pressures and constriction volume velocity U_C is the smoothness of the simulation traces as compared to the more abrupt changes at the releases and closures of the tongue tip in the measured traces. This could be caused by a couple of factors. In the simulation of AL's voiceless trills, the area of the constriction A_C did not reach the estimated maximum, so that there was not the saturation of A_C and U_C that was seen in the simulation of RM's voiceless trills shown in Fig. 4, which resulted in the "squaring off" of some of the pulses. The other factor is that the detailed features of the waves depend on the aerodynamics in the region of the tongue tip, which has only been modeled crudely here.

The relative phases of the various simulated quantities

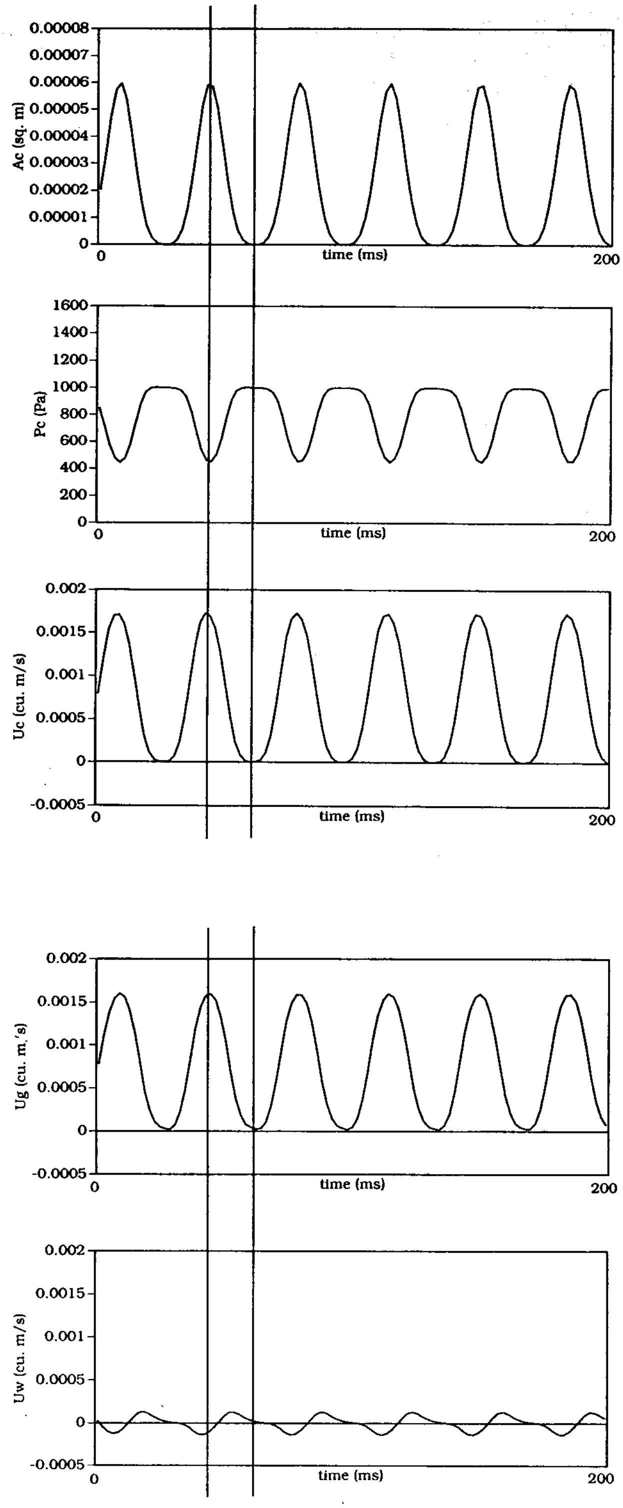

FIG. 3. Time traces of simulation results based on parameters for subject AL in the abducted glottis condition. Vertical lines indicate the relative phases between traces. The left line corresponds to the maximum constriction area and the right line to the minimum constriction area.

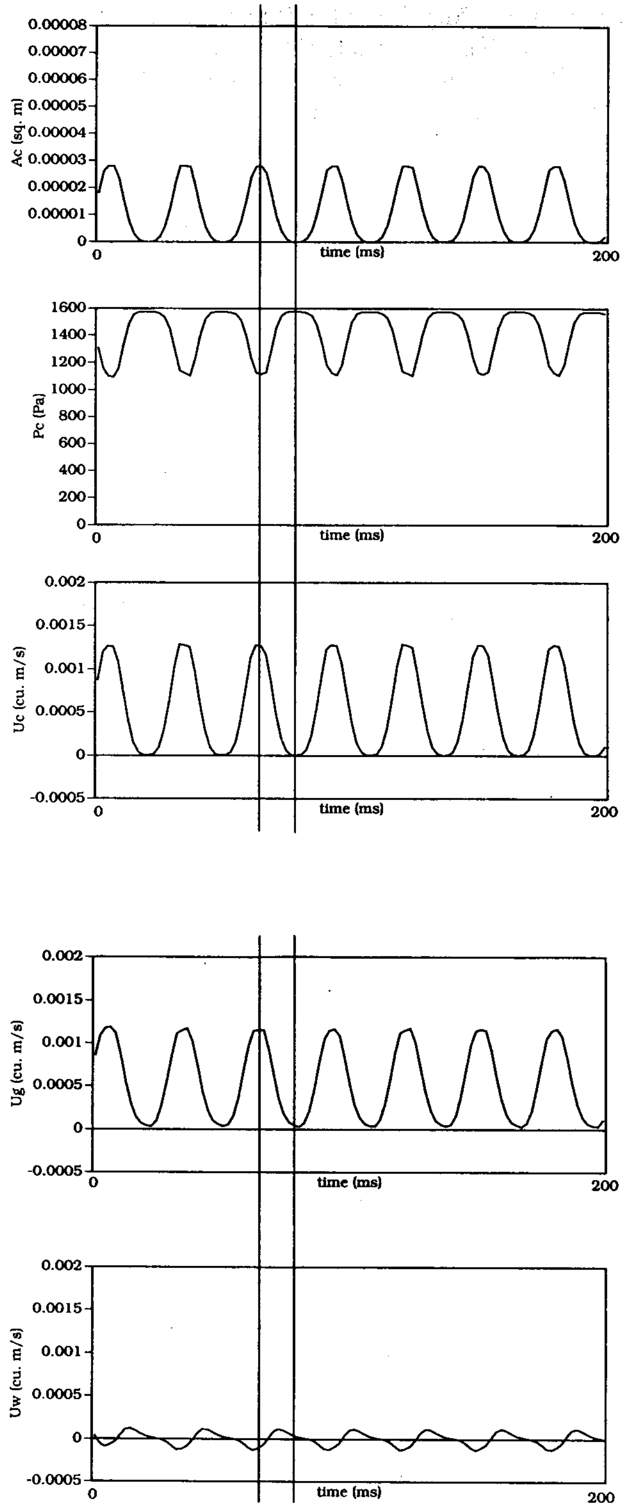

FIG. 4. Time traces of simulation results based on parameters for subject RM in the abducted glottis condition. Vertical lines indicate the relative phases between traces. The left line corresponds to the maximum constriction area and the right line to the minimum constriction area.

are of interest because they show the effects of the compliant vocal tract behind the constriction. In Figs. 3 and 4, two vertical lines have been drawn so that relative phases between the traces can be observed. The vertical line on the left corresponds to maximum constriction area and the vertical line on the right corresponds to minimum constriction area. The tongue-tip constriction volume velocity (U_C) maxima is either simultaneous with, or leads the tongue-tip constric-

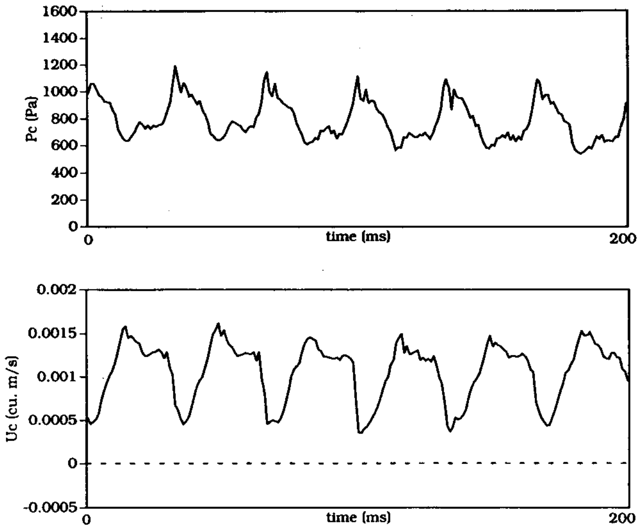

FIG. 5. Time trace of recorded intraoral pressure and constriction volume velocity for subject AL.

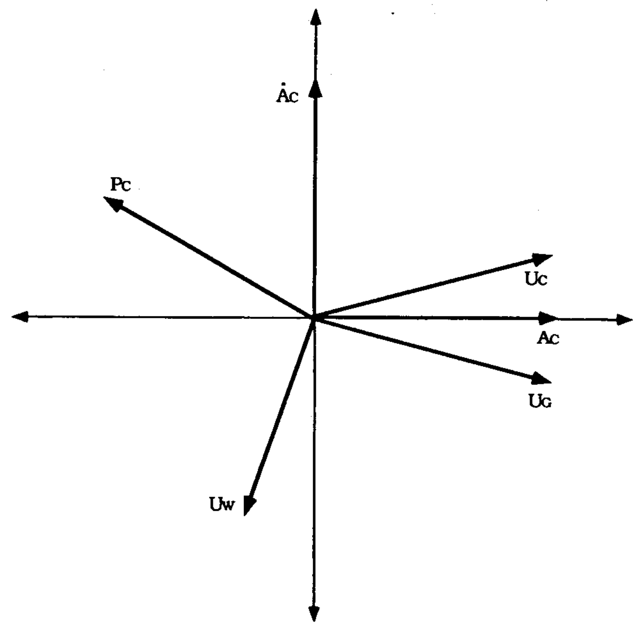

FIG. 6. Vector diagram for the results shown in Fig. 3.

tion area (A_C) maxima by up to 20°, and the maxima of A_C, in turn, lead the intraoral pressure (P_C) minima by about 30°. Also, the glottal volume velocity (U_G) maxima follow the tongue-tip constriction volume velocity (U_C) maxima by about 30° to 35°. While maximum negative (= minimum) wall volume velocity (U_W) is attained 20° to 30° before the maximum constriction volume velocity (U_C), the maximum positive wall volume velocity is attained at about 60° before the minimum constriction volume velocity, and the wall volume velocity reaches zero near the minimum of the constriction volume velocity. (Positive wall volume velocity corresponds to outward movement). The reason that the wall volume velocity extrema are not all in the same phase relation with the constriction volume velocity extrema appears to be the highly nonsinusoidal character of the wall volume velocity. However, the relation between U_W and U_C means that the walls have a maximum inward velocity as the tongue-tip constriction volume velocity goes from the minimum to the maximum, that the maximum outward wall volume velocity occurs as the tongue tip is closing, and that the walls are maximally inflated near the minimum constriction volume velocity (i.e., during closure). Thus it is seen that the tongue-tip constriction volume velocity maxima lead the maxima of the glottal volume velocity because the deflating vocal-tract walls add to the volume velocity through the tongue-tip constriction. The relative phases of the maxima of the various physical quantities from Fig. 3 are summarized in a vector diagram of Fig. 6. While the time derivative of A_C, $\dot{A}_C$, is assumed to be 90° out of phase with A_C, the nonsinusoidal character of the traces limits the applicabilty of the vector diagram.

In general, the pressure must be greater during the opening phase than it is during the closing phase in order for energy to be exchanged from the air to the tongue over a complete cycle. The compliance of the vocal-tract walls can cause asymmetry between the opening and closing of the tongue tip. During the closing phase of the tongue-tip constriction, the walls are expanding because of their compliance, thus storing some of the potential energy that would otherwise be stored in the compression of the air alone. During the opening phase, with the constriction volume velocity growing, the vocal-tract walls begin to discharge air through the constriction. Because there is a time constant associated with the charging and discharging of the compliant vocal-tract walls, there is a lag in intraoral pressure buildup during closure and a lag in intraoral pressure loss during opening, with the net result being higher intraoral pressure during opening than during closing for a given constriction area, or tongue-tip position. It can be noted that the capacitance of the enclosed air serves some of the same function, but the necessary phase shift should be greatly enhanced by the walls because of their larger capacitance, which is in parallel to that of the air.

The motional impedance illustrates the exchange of energy, as well as the reactive part of the impedance due to the air acting on the tongue tip. The ratio of $-P_C$ to the time derivative of θ, is proportional to the motional impedance of each element of the trap door tongue tip. Because θ and A_C are monotonically related for θ between 0 and 180° [see Eq. (7)], the ratio of $-P_C$ to $\dot{A}_C$ will have the same sign as the motional impedance. It can be seen from Fig. 6 that the ratio of these two quantities falls within the third quadrant, and, therefore, both the real and imaginary parts of the motional impedance are negative. The negative motional resistance indicates that energy is being fed into the tongue tip, while the negative reactive part indicates a capacitive load. This latter fact is the result of the low natural frequency of the tongue tip compared with other natural frequencies. The range of natural frequencies for the trap door is from 0.8 to 2.5 Hz, while the natural frequency for the walls is about 60 Hz, and higher for other systems (e.g., the glottal inductance coupled with the wall capacitance, with $A_G = 5\times10^{-5}$ m^2, has a natural frequency of about 320 Hz).

Figures 7 and 8 show time traces of the constriction area, the intraoral pressure, and the wall, the constriction, and the glottal volume velocities for simulations based on the two subjects in a relatively adducted glottis condition. Because of the higher glottal resistance, the glottal volume velocity was less variable and the wall volume velocity tracks the tongue-tip constriction volume velocity more closely than in the open glottis condition. As a result, the vocal-tract wall volume velocity maxima occurred closer to the minima of the tongue-tip constriction volume velocity, than for the

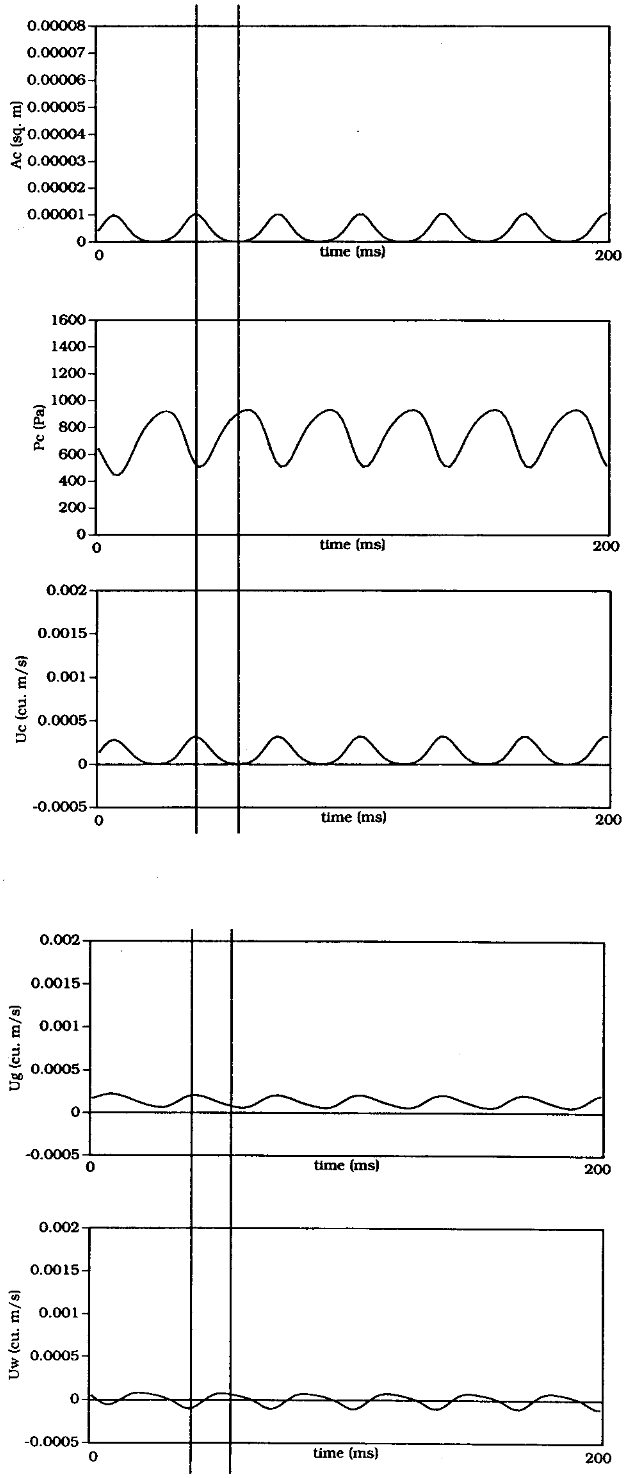

FIG. 7. Time traces of simulation results based on parameters for subject AL in the adducted glottis condition. Vertical lines indicate the relative phases between traces. The left line corresponds to the maximum constriction area and the right line to the minimum constriction area.

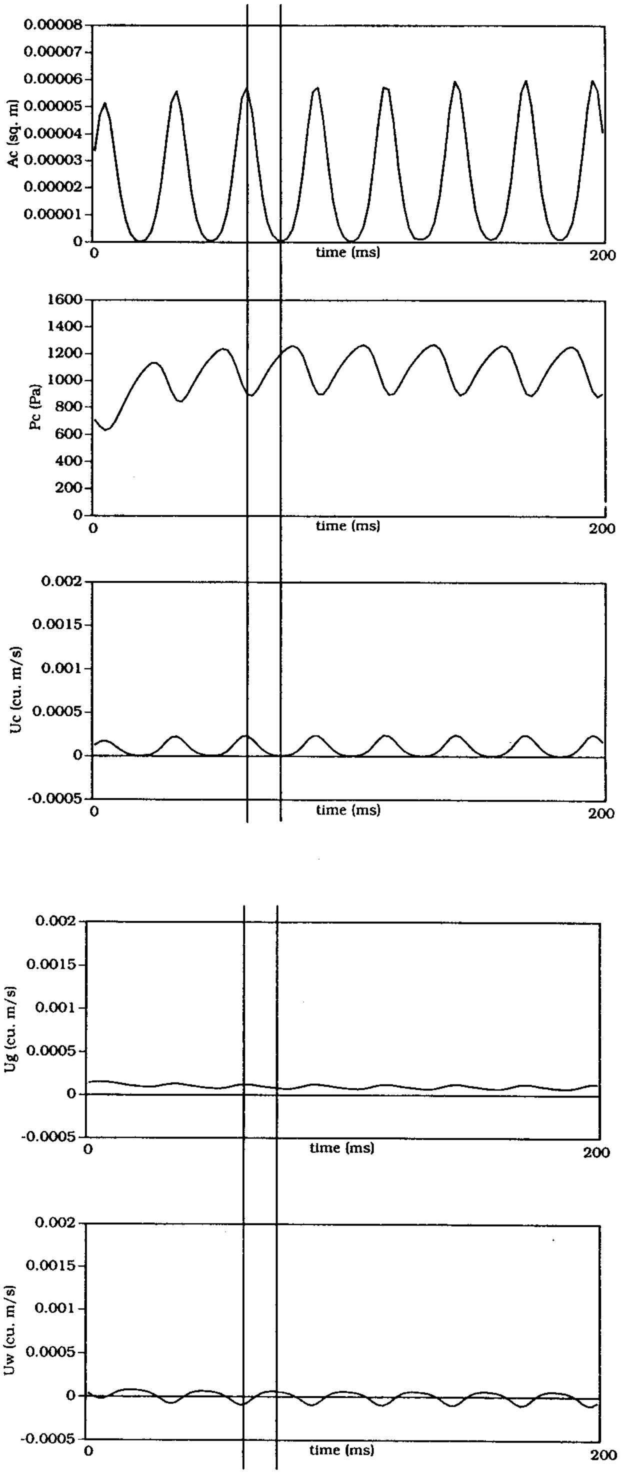

FIG. 8. Time traces of simulation results based on parameters for subject RM in the adducted glottis condition. Vertical lines indicate the relative phases between traces. The left line corresponds to the maximum constriction area and the right line to the minimum constriction area.

abducted glottis condition. Also, the deflation of the vocal-tract walls was a proportionately greater contributor to the tongue-tip constriction volume velocity than the glottal volume velocity in the adducted glottis condition. This is in contrast to the case of the open glottis, where the variation in glottal volume velocity was the major contributor to tongue-tip constriction volume velocity, with the wall vibration providing a perturbation. While the glottal volume velocity remains egressive to fill the vocal-tract walls, even as the tongue-tip constriction is closed, the maximum outward wall volume velocities were reduced more than the maximum inward wall volume velocities in going from an abducted to an adducted glottis condition. Thus glottal flow is less effective in filling the walls at closure in the adducted glottis condition.

IV. CONCLUSION

The simulations have shown that the vocal tract wall compliance provides a viable mechanism for sustained oscillation during tongue-tip trills. Use was made of parameters derived from previous experiments, and of pressure and flow measurements taken by the author for two subjects producing sustained voiceless trills. Unknown parameters were adjusted to produce trills with frequencies between 25 and 35 Hz, as were observed for the two subjects. These unknown parameters involved the linear dimensions of the tongue tip and the ratio of torsion spring constant to the moment of inertia of the tongue tip, as well as the amount of nonlinearity and the damping ratio. The simulations were run in both an open glottis condition and a relatively constricted glottis condition with parameters derived from both subjects. The former condition was intended to simulate voiceless trilling, and the later condition the average open area of the glottis for voiced trills.

Much of the detailed mechanics has not been included in this model, such as the nonlinearities involving collisions of the tongue and the hard palate. The details of the flow near the tongue tip can also be important. The model here is limited to the compliance of the vocal tract as a mechanism for the phase shift between intraoral pressure and tongue-tip position necessary for sustained oscillation. This situation is analogous to the one-mass model of vocal fold oscillation, where the inductive loading of the supraglottal tract is invoked to explain sustained oscillation. The two-mass model provides another mechanism for converting flow energy into vibratory energy with a more detailed picture of the fluid mechanics in the glottis involving flow separation, while an inductive load in the supraglottal tract can still provide favorable phase shifts for sustained oscillation. It is possible that flow separation also plays a role in tongue-tip trills, along with the compliance of the vocal tract. It is known that flow separation is not a quasisteady phenomenon, even at low frequencies (e.g., Sobey, 1983; Pedley and Stephanoff, 1985). Thus it is entirely possible that favorable energy exchange from the air to the tongue tip can be obtained with separation depending not only on tongue-tip position, but also on tongue-tip velocity, without involving more than one degree of freedom for the tongue tip. Separation phenomena that depend on velocity have been known to cause objects to vibrate, as with the galloping vibrations of transmission lines in wind (den Hartog, 1985). [Also, den Hartog (1985) describes mechanical valves where the compliance of the air can have an effect on stability.] Research into these unsteady separation mechanisms for trills and other vibrations in the vocal tract needs to be pursued.

This paper illustrates the importance of simultaneously including articulatory and aerodynamic components in speech production modeling. Some work in modeling and articulatory synthesis that couples aerodynamics and articulation has been done (e.g., Scully, 1990). Further progress requires that this coupling be pursued.

ACKNOWLEDGMENTS

Part of this research grew out of a conversation the author had with Dr. Ronald Scherer. This research was supported by NIH Grants DC 00121 and DC 00865 to Haskins Laboratories. The author thanks anonymous reviewers and Dr. Philip Rubin for their comments.

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

den Hartog, J. P. (1985). *Mechanical Vibrations* (Dover, New York).

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

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., and Matsudaira, M. (1972). "Fluid mechanical considerations for vocal-fold vibration (Monograph 8)," Speech Commun. Res. Lab Inc., Santa Barbara, CA.

Jimenez, B. C. (1987). "Acquisition of Spanish consonants in children aged 3–5, 7 months," Lang. Speech Hear. Serv. Schools, **18**, 357–363.

Maddieson, I. (1984). *Patterns of Sound* (Cambridge U.P., Cambridge, England).

McGowan, R. S., and Nittrouer, S. (1988). "Differences in fricative production between children and adults," J. Acoust. Soc. Am. **83**, 229–236.

Müller, E. M., and Brown, W. S. (1980). "Variations in the supraglottal air pressure waveform and their articulatory consequences," in *Speech and Language: Advances in Basic Research and Practice, Vol. 4*, edited by N. Lass (Academic, New York).

Pedley, T. J., and Stephanoff, K. D. (1985). "Flow along a channel with a time dependent indentation in one wall: the generation of vorticity waves," J. Fluid Mech. **160**, 337–367.

Rothenberg, M. (1968). *The breath-stream dynamics of simple-released-plosive production* (Krager, Basel, Switzerland).

Scully, C. (1990). "Articulatory Synthesis," in *Speech Production and Speech Modeling*, edited by W. J. Hardcastle and A. Marchal (Kluwer, Dordrecht).

Sobey, I. J. (1983). "The occurrence of separation in unsteady flow," J. Fluid Mech. **134**, 247–257.

Stevens, K. N. (1971). "Airflow and turbulence noise for fricative and stop consonants: Static considerations," J. Acoust. Soc. Am. **50**, 1180–1192.

Stevens, K. N. (1991). "Models for voicing and aspiration stop consonants," J. Acoust. Soc. Am. Suppl. 1 **89**, S1978.

Titze, I. R. (1988). "On the physics of small-amplitude oscillation of the vocal folds," J. Acoust. Soc. Am. **83**, 1536–1553.

Westbury, J. R. (1983). "Enlargement of the supraglottal cavity and its relation to stop consonant voicing," J. Acoust. Soc. Am. **73**, 1322–1336.