Regulating glottal airflow in phonation: Application of the maximum power transfer theorem to a low dimensional phonation model

Ingo R. Titze[a)]
Department of Speech Pathology and Audiology and The National Center for Voice and Speech, 330 Wendell Johnson Speech and Hearing Center, The University of Iowa, Iowa City, Iowa 52242-1012

(Received 30 November 2000; revised 10 September 2001; accepted 18 September 2001)

Two competing views of regulating glottal airflow for maximum vocal output are investigated theoretically. The maximum power transfer theorem is used as a guide. A wide epilarynx tube (laryngeal vestibule) matches well with low glottal resistance (believed to correspond to the ''yawn-sigh'' approach in voice therapy), whereas a narrow epilarynx tube matches well with a higher glottal resistance (believed to correspond to the ''twang-belt'' approach). A simulation model is used to calculate mean flows, peak flows, and oral radiated pressure for an impedance ratio between the vocal tract (the load) and the glottis (the source). Results show that when the impedance ratio approaches 1.0, maximum power is transferred and radiated from the mouth. A full update of the equations used for simulating driving pressures, glottal flow, and vocal tract input pressures is provided as a programming guide for those interested in model development. © *2002 Acoustical Society of America.* [DOI: 10.1121/1.1417526]

PACS numbers: 43.70.Bk [AL]

I. INTRODUCTION

Speech language pathologists and singing teachers have generated two competing views (and accompanying behavioral strategies) about the management of airflow in phonation. On the one hand, there is the strategy of using a ''sigh'' to release air with the voice (Linklater, 1976; Colton and Casper, 1996; Brown, 1996), or using a *flow phonation mode* (Sundberg, 1987). This flow mode strategy helps to obtain maximum peak-to-peak glottal airflow. On the other hand, there is the opposite strategy of increasing the adduction of the vocal folds, as in *belt* (Sullivan, 1985; Bestebreurtje and Schutte, 2000) and some country-western singing (Sundberg *et al.*, 1999) to decrease both the average glottal flow and the peak flow for (perhaps) greater glottal efficiency. Even in some classical singing approaches, airflow reduction is sometimes encouraged by the mental image of ''drinking in the air'' rather than blowing out the air.

In this paper, a few data sets will be presented that simulate a ''tight adduction'' case and a ''loose adduction'' case with a computer model of phonation. One objective of the study is to show that both techniques can lead to an optimum acoustic output at the mouth, but the vocal tract configuration has to be matched to the glottal configuration. Tight adduction of the vocal folds requires a narrower supraglottal airway, whereas looser adduction requires a wider airway to maximize the output power. An underlying guiding principle is the maximum power transfer theorem in electric circuits and transmission systems, which states that the internal impedance of the source should match the impedance of the load for maximum power transfer.

A second objective of the study is to update the aerodynamic driving force equations for a low-dimensional model of vocal fold vibration in detail. Some changes have occurred since publication of the three-mass body-cover model (Story and Titze, 1995), particularly with regard to flow separation from the glottal wall and collision forces. In order to continue explorations with this low-dimensional body-cover model, it is important to provide the aerodynamic detail as a programming guide. This dual objective makes this paper somewhat of a nontraditional mixture between model development and a clinical application. But this mixture is justified by the fact that there is an unfortunate history of ''modeling for modeling sake,'' by this and other authors, with insufficient benefit to practitioners in voice and speech. This paper is an attempt to steer toward application while also maintaining a theoretical forward thrust.

II. THE MAXIMUM POWER TRANSFER THEOREM

For readers who are not familiar with the maximum power transfer theorem, a brief summary is provided. As illustrated in Fig. 1, assume a simple electric circuit with a voltage source v, an internal source resistance R_S, and a load resistance R_L. The current that flows is $i=v/(R_L+R_S)$ and the power delivered to the load is

$$p=i^2R_L=v^2R_L/(R_L+R_S)^2. \tag{1}$$

Now consider the variation of this power as the load resistance R_L is changed and the source resistance R_S and voltage v are held constant. For $R_L=0$, the power is clearly zero. For R_L approaching infinity, the power is also zero, since the denominator in Eq. (1) is of higher power in R_L than the numerator. This suggests that there is an intermediate value of R_L for which the power is maximized. By differentiating p with respect to R_L in Eq. (1) and setting the derivative to zero, it is easily shown that the power is maximum for $R_L=R_S$, with a value $p=v^2/4R_S$. Figure 2 (top trace) shows

a)Electronic mail: ingo-titze@uiowa.edu

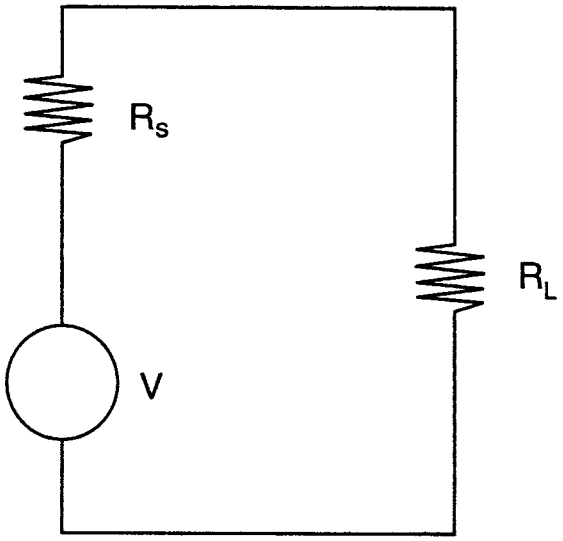

FIG. 1. Simple resistive circuit illustrating internal (source) resistance R_S and load resistance R_L.

the load power normalized to the maximum value, plotted as a function of R_L/R_S.

The circuit is not maximally *efficient*, however, when the power to the load is maximized. If we define efficiency as the ratio of load power to total power consumed in the circuit,

$$e = \frac{i^2 R_L}{i^2(R_L + R_S)} = \frac{R_L}{R_L + R_S}, \tag{2}$$

then an equal amount of power is lost to the internal resistance as is delivered to the load, resulting in an efficiency of 50%. Maximum efficiency occurs when R_L becomes infinite (Fig. 2, bottom trace). Thus, there is a trade-off between maximum power transfer and maximum efficiency of power consumption in a simple circuit of this kind.

To apply the maximum power transfer principle to voice production, we must realize that pressure and flow are not as simply related to each other as voltage and current are in a resistive circuit. Since the acoustic load is a complex impedance, the quantitative nature of the maximum power theorem is not the same, but qualitatively the principle is expected to

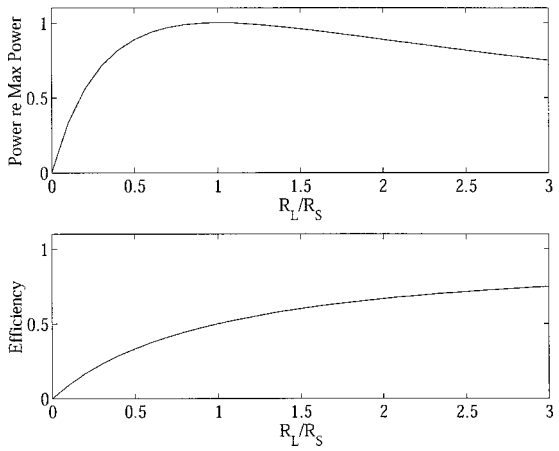

FIG. 2. (Top) power delivered to the load relative to maximum power $v^2/4R_S$, and (bottom) efficiency of power delivered to the load; both are plotted as a function of the load resistance to source resistance ratio R_L/R_S.

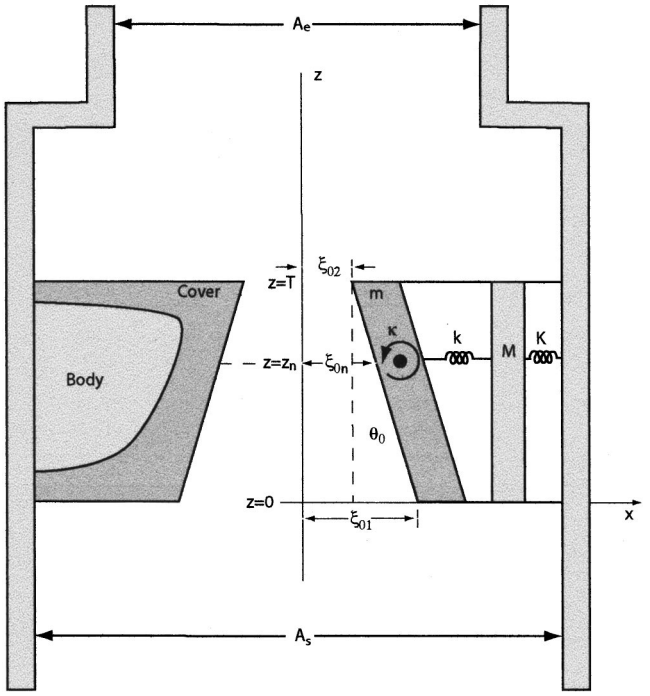

FIG. 3. Bar-plate representation of a body-cover model of the vocal folds.

be preserved: for maximum acoustic power to be delivered to the vocal tract, the internal glottal impedance should be of the same order of magnitude as the acoustic load impedance of the vocal tract. For maximum efficiency, the load impedance should be considerably higher than the glottal source impedance.

But vocal fold oscillation is a nonlinear process, for which glottal impedance calculations must be reinterpreted somewhat. The pressure source is steady (nonoscillatory), whereas the flow is nonsteady (oscillatory). For this reason, a time-domain simulation will be used to derive pressure–flow relations for various vocal tract geometries and then to determine maximum power transfer conditions numerically.

III. TIME-DOMAIN SIMULATION OF GLOTTAL AIRFLOW

A bar-plate version of the body-cover model of the vocal folds (Fig. 3) was used to vary glottal adduction of the vocal folds and the supraglottal area. The full mathematical details of the model are found in the Appendix. This low-dimensional model is patterned after Liljencrants (1991) in terms of the glottal configuration and is essentially equivalent to the three-mass model of Story and Titze (1995) in terms of its tissue characteristics. It is recognized that low-dimensional models with ''bar'' masses do not accurately represent partial glottal closure, which limits the normal variability of acoustic excitation of the vocal tract. In particular, cases for which the vocal folds do not collide and the mean glottal flow is excessive will not be representative of normal phonation. Although this may require some scrutiny of the data, the simplicity of models outweighs this acknowledged weakness. A high-dimensional finite element model (Alipour *et al.*, 2000) could have been used for this study, but not all of the mathematical details are yet available to everyone;

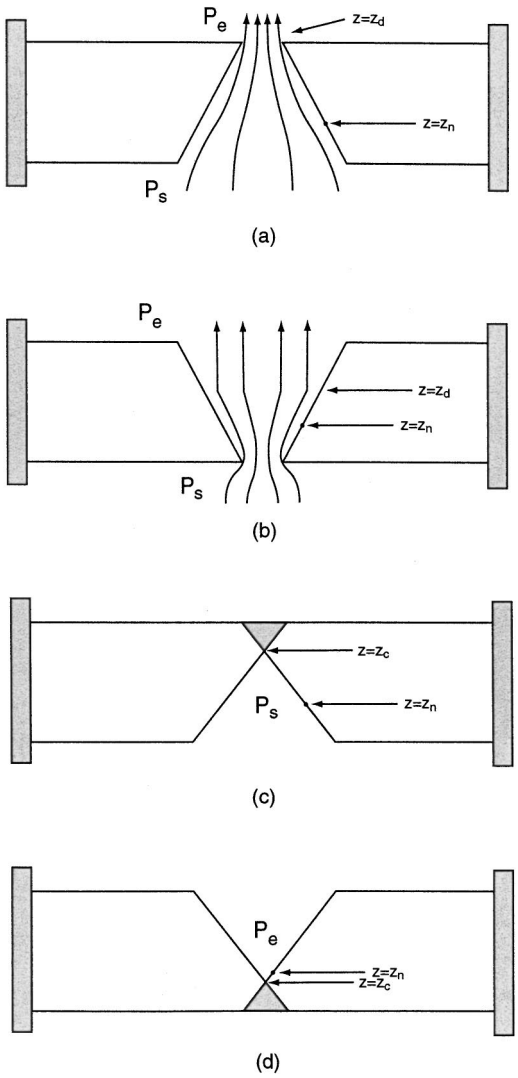

FIG. 4. Sketches used for intraglottal pressure calculations, (a) convergent glottis with no vocal fold contact, (b) divergent glottis with no vocal fold contact, (c) convergent glottis with contact on top, and (d) divergent glottis with contact at bottom.

hence, for sake of replication, the author selected this low-dimensional bar-plate model, which actually makes the study stronger if the hypotheses can be born out with this limited model. A strength of the current model is its full interactivity with the vocal tract. Acoustic pressures above and below the glottis not only define the transglottal pressure (and thereby the flow), but they also determine the driving forces on the vocal folds.

The exact equivalence between the tissue properties of the bar-plate and the three-mass versions of the body-cover model is investigated in a companion paper (Titze and Story, in press). The vocal fold configuration is controlled by ξ_{02}, the glottal half-width at the vocal processes, and θ_0, the pre-phonatory convergence angle of the glottis. The vocal tract configuration is controlled by the vowel shape and the epilaryngeal (supraglottal) tube area A_e. The subglottal area A_s is held constant for this study. Appropriate ratios of the control parameters are chosen to show under which conditions maximum power transfer occurs. Figure 4 shows various conditions of the glottis for contact and noncontact between the folds. The mathematics for the pressures and flows are given in the Appendix.

IV. PROCEDURE

A total of 120 simulations served as an initial data set for study. These were all simulations for which oscillation was either self-sustained (with limit cycles) or damped (with a point attractor). To simplify the parameter space, only two independent variables were chosen for investigation, the neutral glottal area at the top of the vocal folds ($2L\xi_{02}$ in Fig. 3), and the epilarynx tube area A_e. The convergence angle θ_0 was held constant (see below). The neutral glottal area could be negative to allow an initial vocal fold overlap. The epilarynx tube area was chosen to range between 0.2 and 2.0 cm^2 because such values were considered to be physiologically likely (Story, 1995).

The following parameters were held constant:

$P_L = 0.8$ kPa (lung pressure)
$A_s = 3.0\ \text{cm}^2$ (subglottal area)
$\theta_0 = 0.0333$ rad $= 1.91°$ (convergence angle)
$K = 200$ N/m (body stiffness)
$k = 50$ N/m (cover stiffness)
$\kappa = 5 \times 10^{-5}$ N m/rad (torsional cover stiffness)
$M = 10^{-4}$ kg (body mass)
$m = 10^{-4}$ kg (cover mass)
$I = 10^{-10}\ \text{kg m}^2$ (cover moment of inertia)
$z_n = 0.5$ T (nodal point)
$B = 0.2(KM)^{1/2}$ (0.1 damping ratio of body)
$b = 0.2(km)^{1/2}$ (0.1 damping ratio of cover in translation)
$B_c = 0.2(\kappa I_c)^{1/2}$ (0.1 damping ratio of cover in rotation)
$L = 1.5$ cm (length of vocal fold)
$T = 0.3$ cm (thickness of vocal fold)
$\rho = 1.14\ \text{kg/m}^3$ (density of air)
$c = 350$ m/s (sound velocity of warm, moist air)
Vowel = /ɑ/ or /i/, 44 section vocal tract (see Fig. 6)
Nasal coupling = 0

With these parameters, the fundamental frequency of vocal fold oscillation was around 120 Hz, an average male speaking F_0. There was some variation of F_0 with neutral glottal area and epilarynx tube area, but that is not the focus of this investigation and was of little consequence.

V. RESULTS

Figure 5 shows the parameter space for oscillation when the vocal tract was configured as an /ɑ/ vowel. Circles indicate limit cycles (sustained oscillation) and dots indicate point attractors (damped oscillation). Note that oscillations were sustainable only for neutral glottal areas ranging from -0.15 to $+0.15\ \text{cm}^2$ for the chosen viscoelastic constants of the bar-plate model. For the /i/ vowel, the oscillation space was more restricted, with most positive neutral areas yielding damped oscillation only.

Figure 6 shows a sample output for one simulation. On top is the vocal tract diameter function, showing, from left to right: trachea, epilarynx tube, pharynx, mouth, and nose. Below the diameter function are nine simulated wave forms, all 100 ms long, described in the figure caption. The neutral glottal area was $-0.06\ \text{cm}^2$ (initially overlapped on top) and

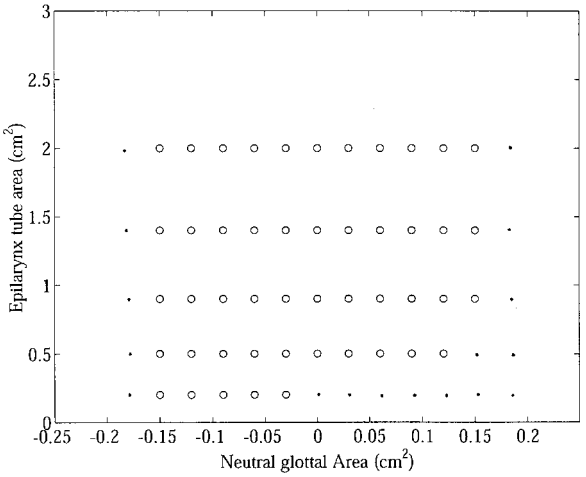

FIG. 5. Parameter space for oscillation. Circles indicate self-sustained oscillation (limit cycles) and dots indicate damped oscillation (point attractors).

the epilarynx tube area was 0.5 cm^2. This particular case was chosen as an example because it shows a period-tripling bifurcation in the glottal area (labeled AG here). More severe bifurcations were seen in a number of cases, particularly for larger negative neutral areas (where the vocal folds further overlapped prior to oscillation). Such bifurcations created somewhat uneven power calculations, as will be seen later, because the vocal tract excitations differed slightly across

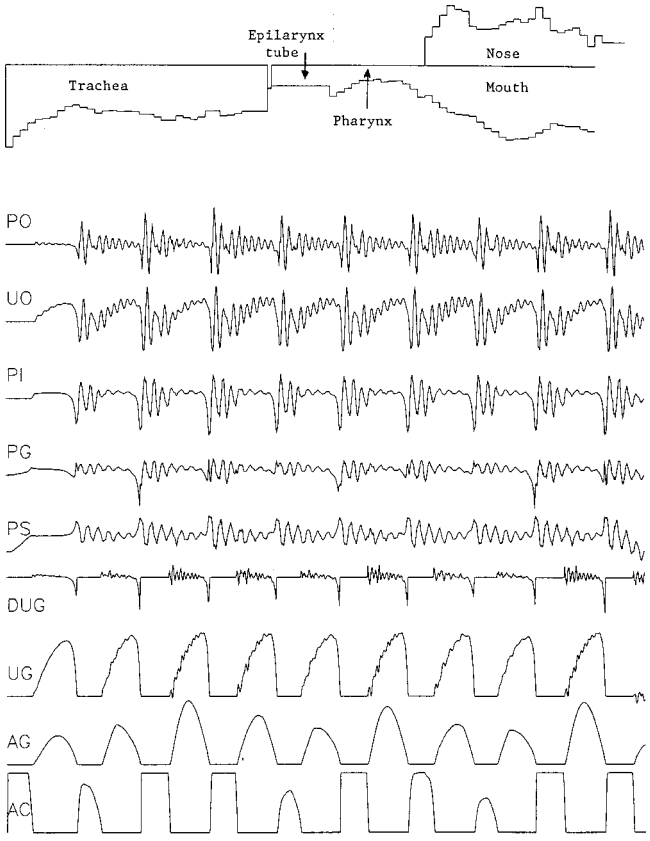

FIG. 6. Example of one of the simulations, showing from top to bottom: vocal tract diameter function, radiated mouth pressure (PO), mouth flow (UO), vocal tract input pressure (PI), intraglottal pressure (PG), subglottal pressure (PS), derivative of glottal flow (DUG), glottal flow (UG), glottal area (AG), and contact area (AC).

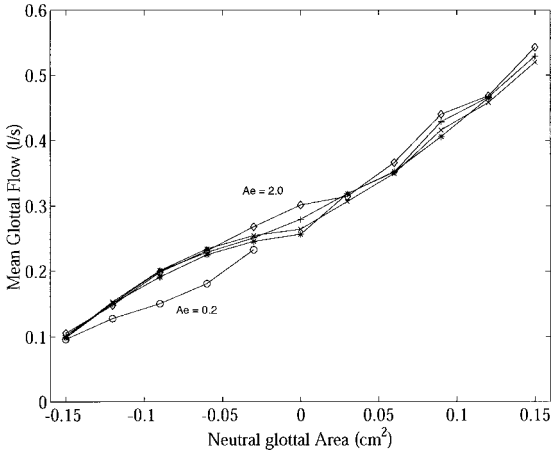

FIG. 7. Mean glottal flow vs neutral glottal area for five values of epilarynx tube area A_e. The lung pressure was kept at 0.8 kPa in all cases.

cycles; but the overall results were not compromised severely by this unevenness. Power and flow calculations were performed on the wave forms labeled PO and UG in Fig. 6.

Figure 7 shows the mean glottal flow as a function of the neutral glottal area. Epilarynx tube area is the parameter on the curves. Note that this mean flow versus neutral area relation is essentially a proportionality, with little dependence on A_e, suggesting that *for mean flow, the glottis is a high impedance source*. The mean glottal area is considerably less than 0.2 cm^2, the smallest vocal tract area chosen. Since mean glottal flow is a more measurable quantity on human subjects than neutral glottal area, we will henceforth consider mean glottal flow to be one of the two independent control parameters. This is also the parameter that has the greatest clinical relevance in the management of airflow in phonation. Readers who wish to reconvert subsequent data sets back to neutral area can simply draw a straight line through the family of curves in Fig. 7 and use this line as a nomogram for reconversion.

For nonsteady (oscillatory) flow, a different picture is obtained. Figure 8 shows peak glottal flow as a function of mean glottal flow for this model. The data set shows the first important effect of A_e. Note that small values of A_e (0.2 and 0.5 cm^2) restrict the peak flow, but values of 1.0 cm^2 and

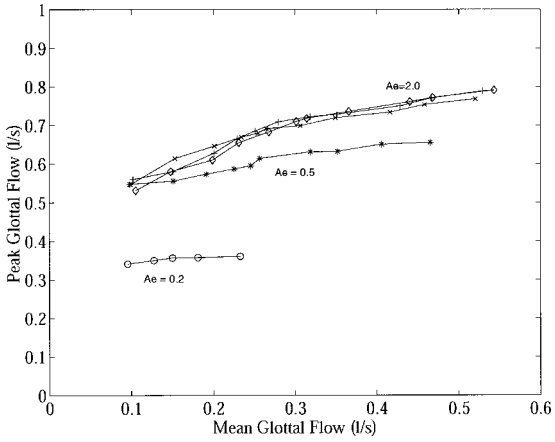

FIG. 8. Peak glottal flow vs mean glottal flow for five values of epilarynx tube area A_e. The lung pressure was 0.8 kPa in all cases.

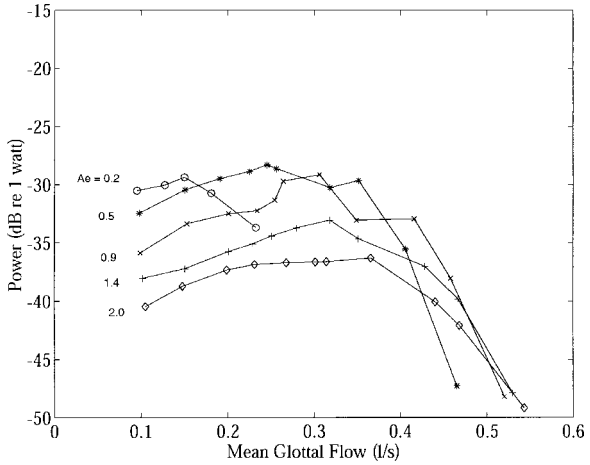

FIG. 9. Radiated oral power (in dB relative to 1.0 W) vs mean glottal flow for five values of epilarynx tube area A_e. The lung pressure was 0.8 kPa in all cases.

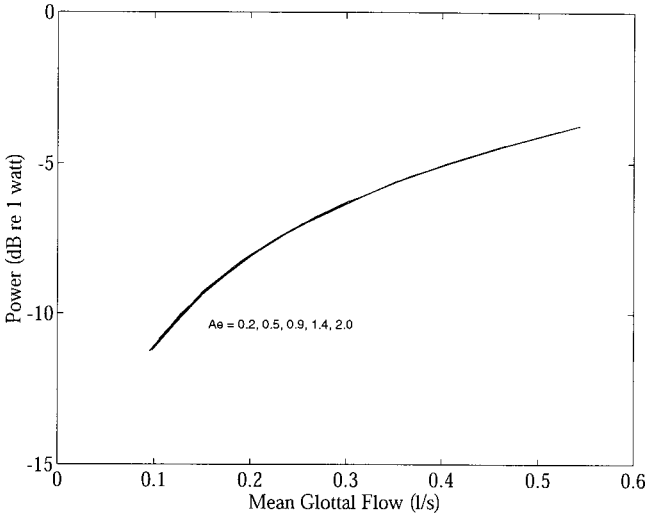

FIG. 10. Aerodynamic power (in dB relative to 1.0 W) vs mean glottal flow for five values of epilarynx tube area A_e. The lung pressure was 0.8 kPa in all cases.

higher have little effect on the peak flow. The peak flow is then governed only by the glottal resistance, not by the vocal tract impedance. We can therefore say that, acoustically, the glottis acts as a high impedance (constant flow) source only for values of $A_e > 1.0\,\mathrm{cm}^2$, with the internal glottal resistance limiting the flow. Sundberg (1987) has given glottal valving as an explanation of the ''flow mode.'' The present results suggest that the ''flow mode'' appears to be governed more by epilarynx tube area than by glottal valving. As an example, for a mean glottal flow of 0.2 l/s, the peak flow can be essentially doubled by a 5:1 widening the epilarynx tube (from 0.2 to 1.0 cm^2). On the other hand, a 5:1 increase in mean glottal flow (from 0.1 to 0.5 l/s) increases the peak flow by only 50%.

The widening of the epilarynx tube to switch to a ''flow mode'' may not always be an advantage as far as vocal output power is concerned, as seen in Fig. 9. Here we show radiated oral power (in dB relative to 1 W) as a function of mean glottal flow for a simulated /ɑ/ vowel. This radiated power was computed as P_o^2/R_r, where P_o is the oral radiated pressure and R_r is the radiation resistance $[128\rho c/(9\pi^2)$; Flanagan, 1965]. Note that as the epilarynx tube area increases (top to bottom curve), the peak output power occurs for higher mean glottal flows. This is predicted by the maximum power transfer theorem; as the load impedance decreases, the source resistance must decrease also, thus maintaining an optimum R_L/R_S ratio. But the curves have too much of a downturn when the mean flow increases much beyond 0.3 l/s. The reason for this sharp downturn was given earlier; for bar masses, the glottal flow waveforms become nearly sinusoidal for positive neutral areas (no partial glottal closure). Since high frequencies radiate much better from the mouth than low frequencies, the radiated power suffers dramatically with decreased glottal adduction (as measured here by mean glottal flow). In a higher-dimensional model (Titze and Talkin, 1979; Alipour *et al.*, 2000), this can be remedied by allowing anterior–posterior variations in vocal fold contact. For the current bar-plate model, mean glottal flows above 0.3 l/s, for which there was no vocal fold collision, will henceforth be eliminated from the data sets for maximum power transfer considerations. This will make the proof of a downturn in the output power with increasing glottal width more difficult, but still possible with appropriate impedance ratios.

Consider now the aerodynamic power $P_L u_{\mathrm{mean}}$, where u_{mean} is the mean glottal flow. This is shown in Fig. 10. Since P_L was held constant, the relation must be a proportionality, which shows up as a logarithmic curve on a semilog plot. (Data points are not shown because the curves are so tightly overlapped.) The fact that the curve is identical for all values of A_e is an internal verification of the consistancy across A_e calculations, since the actual values plotted for u_{mean} were from different A_e curves.

Figure 11 shows glottal efficiency calculations, which (in dB) are simply the difference between the radiated output power in Fig. 9 and the aerodynamic power in Fig. 10. This glottal efficiency was expected to be a monotonically increasing function with mean glottal flow, as predicted in Fig. 2 (bottom trace), because the source resistance is decreasing.

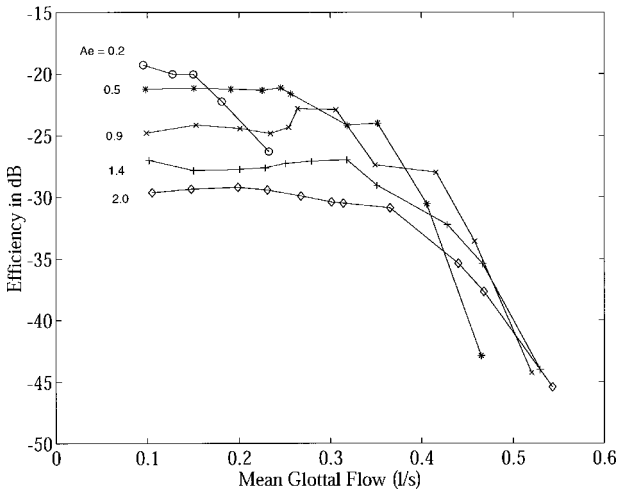

FIG. 11. Glottal efficiency (in dB relative to 1.0, or 100%) vs mean glottal flow for five values of epilarynx tube area A_e. The lung pressure was 0.8 kPa in all cases.

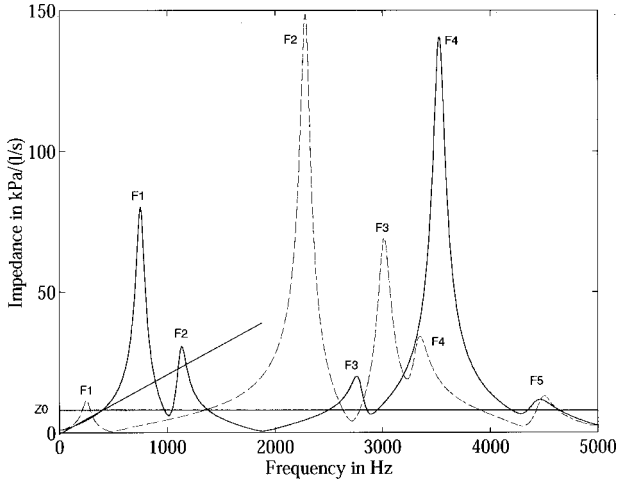

FIG. 12. Calculated input impedance of the vocal tract for the vowel /ɑ/ (solid curve) and the vowel /i/ (dashed curve). The sloping straight line is a low-frequency approximation. The horizontal straight line is the characteristic impedance of the epilarynx tube, $Z_0=\rho c/A_e$.

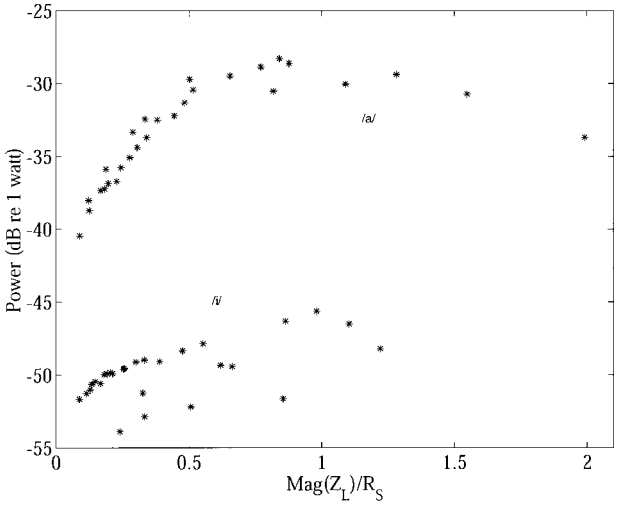

FIG. 13. Radiated oral power (in dB relative to 1.0 W) plotted against the ratio of magnitude of input impedance to glottal source resistance.

But, because of the aforementioned high frequency loss with no collision, and hence poorer radiation, the efficiency curves remained at best a constant (for low mean glottal flow) and at worst dropped off dramatically when mean flow was above 0.3 l/s.

To define a ratio similar to the R_L/R_S ratio for verification of the maximum power transfer theorem, the vocal tract input impedance was first calculated as outlined by Sondhi and Schroeter (1987) and Titze and Story (1997). For the /ɑ/ configuration shown on top of Fig. 6 (with no nasal coupling), the magnitude of the impedance is shown in Fig. 12 as a function of frequency. The solid curve is for the vowel /ɑ/ and the dashed curve is for the vowel /i/. Inspection of the real and imaginary parts of the impedance curves showed that, for low fundamental frequencies ($F_0<F_1$), the impedance is essentially a positive (inertive) reactance, with the resistance playing a minor role. A straight-line approximation to this low-frequency impedance is also shown, computed as

$$Z_L=jZ_0(F_0/350), \tag{3}$$

where $j=(-1)^{1/2}$, and

$$Z_0=\rho c/A_e, \tag{4}$$

the characteristic acoustic impedance of the epilarynx tube. Note that $Z_L=jZ_0$ occurs in Eq. (3) when $F_0=350$ Hz. At this frequency, the straight-line impedance approximation crosses the constant horizontal line marked Z_0 on the vertical axis in Fig. 12.

Ideally, a source impedance match to the load impedance Z_L would be Z_L^*, its complex conjugate. This would be a compliant source impedance. But the glottal impedance is not compliant, which will make an impedance ratio a complex number. In terms of absolute magnitudes, however, the following ratio is defined:

$$\frac{\mathrm{Mag}(Z_L)}{R_S}=\frac{(\rho c/A_e)(F_0/350)}{P_L/u_{\mathrm{mean}}}, \tag{5}$$

where the glottal resistance R_S is taken to be the lung pressure divided by the mean glottal flow. This expression is limited by the straight-line approximation to the inertive rectance, which is good to F_0 slightly greater than $F_1/2$. In the simulation data sets described previously, this approximation was easily met, since F_0 was around 120 Hz and F_1 was 750 Hz for /ɑ/ and 250 Hz for /i/.

Figure 13 (top) is a replot of the entire family of data points for oral radiated power shown in Fig. 9 for the /ɑ/ vowel, with the abscissa now being the impedance ratio in Eq. (5). A second data set for the /i/ vowel (bottom) is added to the graph. Overall, the radiated power for /i/ is much less than for /ɑ/ because of the greater mouth constriction and smaller lip opening. Note that most of the points for each data set follow a curved path, and that maximization of radiated power occurs at the value of 1.0 for this impedance ratio. (As stated before, data points for $u_{\mathrm{mean}}>0.3$ l/s were eliminated for reasons described.) The similarity between these collections of data points and the idealized maximum power transfer curve (top of Fig. 2) gives support to the maximum power transfer theorem being applicable to vocal fold vibration. In particular, the steep rise before the peak and the more shallow decline after the peak lend credence to the hypothesis. The points falling lower than expected are mainly cases for which there was vocal roughness in the form of subharmonics.

VI. CONCLUSIONS

Mean glottal airflow (or, alternatively, glottal resistance) has been a target for optimizing vocal output power in voice therapy and singing training. The current investigation suggests that the optimization process should involve both the vocal tract and the vocal folds. It appears that an impedance matching between the two might take place. In general, a wide epilarynx tube (from the ventrical to the laryngeal vestibule) requires a low glottal resistance for maximum power transfer. Conversely, a narrow epilarynx tube requires a high glottal resistance (more adduction) for maximum power transfer. What Sundberg (1987) has called the ''flow mode''

appears to be a condition where the vocal tract impedance is considerably smaller than the glottal impedance, making the glottis a flow source acoustically, as for steady-flow (aerodynamic) conditions.

Vocologists (those who habilitate voices) have some choices in guiding a speaker or singer. If the desired (or acquired) voice quality is to be bright and ''twangy,'' as in some forms of belting, gospel singing, or some regional dialects, the vocal tract can be more narrow in the epilaryngeal and pharyngeal region (Estill *et al.*, 1996; Story *et al.*, 2001). For such a vocal tract configuration, a well-adducted pair of vocal folds, with relatively high glottal resistance, would be a good match. Because of this higher glottal resistance, lung pressures would likely also be on the high side. Conversely, if the desired (or acquired) voice quality is to be ''yawny,'' as in crooning, sobbing, or a mellow speech dialect (Estill *et al.*, 1996), the epilaryngeal and pharyngeal vocal tract can be wider. For this configuration, a lesser degree of adduction, with lower glottal resistance and probably lower lung pressure, is a good match.

It is already known that ''yawn-sigh'' is a good combination for voice therapy. Sigh involves a glottal posture with low glottal impedance that matches a ''yawny'' vocal tract. Less is known about the ''twang-belt'' combination in voice training and therapy. Here the voice is sometimes initiated with a creaky production, a tighter state of vocal fold adduction. This is a match for twang, a tighter vocal tract configuration. Some vocologists shy away from a twang-belt approach to voice therapy because they fear hyperfunction and excessive vocal fold collision. But since mean glottal flow is smaller, and hence presumably also the amplitude of vibration of the vocal folds, it is not clear that one or the other of these techniques is necessarily more healthy. For the time being, one must keep an open mind about high pressure, low flow production as a viable alternative to low pressure, high flow production. The choice depends to a large degree on the natural state of the vocal tract and the voice quality to be achieved with it.

As a future investigation, it would be worthwhile to examine if the subglottal (tracheal) impedance could assume a compliant characteristic to provide a true impedance match as a complex conjugate to the supraglottal impedance. It would also be instructive to test maximum power transfer for conditions where the fundamental frequency is at or above the first formant frequency. Research is presently ongoing in this area.

ACKNOWLEDGMENT

This work was supported by the National Institution of Health, Grant No. R01 DC02532-05.

APPENDIX

1. Equations of motion

The motion of the cover (plate) can be described with two degrees of freedom if we assume a rotation θ about a nodal point z_n (see Fig. 3) and a translation ξ of the nodal point,

$$I_c\ddot{\theta}+B_c\dot{\theta}+\kappa\theta=T_a, \tag{A1}$$

$$m\ddot{\xi}+b(\dot{\xi}-\dot{\xi}_b)+k(\xi-\xi_b)=F_a, \tag{A2}$$

where T_a is the applied aerodynamic torque, I_c is the moment of inertia for rotation of the cover, B_c is the rotational damping, κ is the rotational stiffness, m is the mass of the cover, b is translational damping, k is the translational stiffness, ξ_b is the displacement of the body, and F_a is the aerodynamic force at the nodal point. Similarly, the equation of motion for the body is written as

$$M\ddot{\xi}_b+b(\dot{\xi}_b-\dot{\xi})+k(\xi_b-\xi)+K\xi_b+B\dot{\xi}_b=0, \tag{A3}$$

where M is the mass, B is the damping, and K is the stiffness of the body.

2. Glottal area simulation

The glottal area for a left–right symmetrical bar-plate model is defined as

$$a(z)=2L[\xi_{0n}-(z-z_n)\tan(\theta_0+\theta)+\xi], \tag{A4}$$

where L is the length of the glottis and ξ_{0n} is the prephonatory displacement from the midline at the nodal point z_n, measured from the bottom of the vocal folds. With this area function, the glottal entry and exit areas are defined as

$$a_1=2L\ \mathrm{Max}\{\delta,[\xi_{0n}-(0-z_n)\tan(\theta_0+\theta)+\xi]\}, \tag{A5}$$

$$a_2=2L\ \mathrm{Max}\{\delta,[\xi_{0n}-(T-z_n)\tan(\theta_0+\theta)+\xi]\}, \tag{A6}$$

where δ is the minimum allowable area for glottal flow. Note that both the dynamic variables θ and ξ enter these equations for glottal area; since glottal area determines flow and also intraglottal pressure, aerodynamic coupling of the equations of motion (A1)–(A3) is guaranteed.

3. Aerodynamic driving pressure and driving torque

The assumed linear variation of the glottal area from entry to exit allows the pressure in the glottis to be integrated over the medial surface to obtain both the net driving torque on the cover and the net driving force on the body. Using ideal Bernoulli conditions for any point z upstream of the flow detachment point z_d [Figs. 4(a) and 4(b)]

$$P_s+\tfrac{1}{2}\rho v_s^2=P(z)+\tfrac{1}{2}u^2/a^2(z), \tag{A7}$$

where P_s is the subglottal pressure, v_s is the subglottal (trachea) particle velocity, $P(z)$ is the intraglottal pressure at any point z, u is the flow, and $a(z)$ is glottal area corresponding to the intraglottal pressure $P(z)$. Now let P_{kd} be defined as the kinetic pressure at the point of flow detachment and P_{ks} the kinetic pressure in the trachea,

$$P_{kd}=\tfrac{1}{2}\rho v_d^2, \tag{A8}$$

$$P_{ks}=\tfrac{1}{2}\rho v_s^2, \tag{A9}$$

where v_d is the particle velocity at flow detachment. Defining $a_d=a(z_d)$, the intraglottal pressure in the attached region is then

$$P(z)=P_s+P_{ks}-P_{kd}a_d^2/a^2(z). \tag{A10}$$

We now argue that the kinetic pressure P_{ks} in the trachea is negligible compared to P_s. Typically, P_s is on the order of 0.5–1.0 kPa in speech (Holmberg *et al.*, 1988), but can reach 5.0 kPa in singing (Bouhuys *et al.*, 1968). For a typical mean tracheal flow of 0.2 l/s (Holmberg *et al.*, 1988) and a tracheal cross section of 3 cm^2 (Story, 1995), the particle velocity is 67 cm/s, which leads to a kinetic pressure of 0.0002 kPa according to Eq. (11). Since this is several orders of magnitude lower than P_s, we will neglect P_{ks} from this point on.

The mean aerodynamic forces over the medial surface are now obtained by integration. For example, for the lower portion of a divergent glottis for which the detachment point z_d is *below* the nodal point z_n,

$$F_l = L z_d \frac{1}{z_d} \int_0^{z_d} P(z)\,dz + L(z_n - z_d)(P_s - P_{kd}), \tag{A11}$$

where we have assumed that the jet pressure above detachment is $P_s - P_{kd}$, the value obtained by letting $z = z_d$ in Eq. (A10). Performing the integration,

$$F_l = L P_s z_d - L P_{kd} a_d^2 \left[(-a^{-1}) \left(\frac{da}{dz} \right)^{-1} \right]_0^{z_d} + L(z_n - z_d)(P_s - P_{kd}). \tag{A12}$$

In evaluating the limits of this integration, it is necessary to assume that the glottal area gradient

$$\frac{da}{dz} = (a_d - a_1)/z_d \tag{A13}$$

is independent of z, but this assumption has already been made with linearization of the glottal area.

Defining $a_1 = a(o)$, we obtain the following for a *divergent* glottis:

$$F_l = L z_n P_s - L \left(z_n - z_d + \frac{a_d}{a_1} z_d \right) P_{kd} \quad \text{for } z_d \leqslant z_n. \tag{A14}$$

The upper force requires no integration because the jet pressure does not change from $z = z_n$ to $z = T$,

$$F_u = L(T - z_n)(P_s - P_{kd}) \quad \text{for } z_d \leqslant z_n. \tag{A15}$$

If the detachment point z_d is *above* the nodal point z_n, then for a *divergent* glottis,

$$F_l = L z_n \left(P_s - \frac{a_d^2}{a_n a_1} P_{kd} \right) \quad \text{for } z_d > z_n, \tag{A16}$$

$$F_u = L(T - z_n) P_s - L \left[(T - z_d) + \frac{a_d}{a_n}(z_d - z_n) \right] P_{kd} \quad \text{for } z_d > z_n, \tag{A17}$$

where $a_n = a(z_n)$ is the glottal area at the nodal point.

For a convergent glottis, the flow remains attached over the entire glottis [i.e., $z_d = T$ and $a_d = a_2 = a(T)$]. This allows Eqs. (A14)–(A17) to be replaced by two simpler equations,

$$F_l = L z_n \left(P_s - \frac{a_2^2}{a_n a_1} P_{kd} \right), \tag{A18}$$

$$F_u = L(T - z_n) \left(P_s - \frac{a_2}{a_n} P_{kd} \right), \tag{A19}$$

and the detachment point obviously does not need to be computed.

The total force F_a in Eq. (A2) is

$$F_a = F_l + F_u. \tag{A20}$$

To obtain the exact aerodynamic torque, we would need to integrate the differential torque $P(z) z\,dz$ over the entire plate surface. But we will employ an approximation that will lead not only to a simpler expression, but will also allow for a more direct comparison of the driving forces to those of the two-mass model (Ishizaha and Flanagan, 1972), and the three-mass model (Story and Titze, 1995). The torque approximation is obtained by using the two average forces F_l and F_u with their respective moment arms, $z_n/2$ and $(T - z_n)/2$,

$$T_a = F_l \left(\frac{z_n}{2} \right) - F_u \left(\frac{T - z_n}{2} \right). \tag{A21}$$

This is the torque used in Eq. (A1).

4. Calculation of the kinetic pressure at flow detachment

The only remaining unknown in the above-mentioned force and torque equations is the kinetic pressure P_{kd}. If we make the assumption that pressure recovery and flow reattachment are well downstream from the beginning of the jet, and that the jet pressure is constant, then the glottal exit condition can be written as

$$P_2 = P_d = P_e - k_e P_{kd}, \tag{A22}$$

where P_e is the (recovered) pressure in the epilarynx tube, P_2 is the exit pressure, and k_e is the pressure recovery coefficient (Ishizaka and Flanagan, 1972), written as

$$k_e = 2 \frac{a_d}{A_e} \left(1 - \frac{a_d}{A_e} \right). \tag{A23}$$

Using once again Bernoulli's equation to the point of flow detachment,

$$P_s = P_d + P_{kd}, \tag{A24}$$

the kinetic pressure P_{kd} in all forgoing equations can be replaced by

$$P_{kd} = (P_s - P_e)/(1 - k_e). \tag{A25}$$

Now we use the Liljencrants (1991) and Pelorson *et al.* (1994) criterion for flow detachment,

$$a_d = a(z_d) = \mathrm{Min}(a_2, 1.2 a_1). \tag{A26}$$

The nodal point area in the foregoing force and torque relations is defined as

$$a_n = \mathrm{Max}[\,\delta, 2L(\xi_{0n} + \xi)\,]. \tag{A27}$$

Substituting $z = z_d$ into Eq. (A4) and equating the result to $1.2\,a_1$ from Eq. (A5) yields the height of the separation point

$$z_d = \mathrm{Min}\left\{ T, \mathrm{Max}\left[0, -\frac{0.2}{\tan(\theta_0+\theta)}[\xi_{0n}+z_n\tan(\theta_0+\theta)+\xi] \right] \right\}, \tag{A28}$$

where the limits $0<z_d<T$ have been imposed.

5. Driving pressures for glottal closure

When there is contact between portions of the vocal fold surfaces, we also need to invoke the hydrostatic pressure, defined here as the mean between the subglottal pressure and the supraglottal (epilaryngeal) pressure,

$$P_h = (P_s+P_e)/2. \tag{A29}$$

Consider three separate conditions for the driving force and torque as determined by the glottal configuration. The first condition is illustrated in Fig. 4(c) and applies to *upper contact only*:

$$a_1 > \delta, \quad a_2 \leqslant \delta, \tag{A30}$$

$$F_l = Lz_nP_s \quad \text{for } z_c \geqslant z_n, \tag{A31}$$

$$F_u = L(z_c-z_n)P_s + L(T-z_c)P_h \quad \text{for } z_c \geqslant z_n, \tag{A32}$$

$$F_l = Lz_cP_s + L(z_n-z_c)P_h \quad \text{for } z_c < z_n, \tag{A33}$$

$$F_u = L(T-z_n)P_h \quad \text{for } z_c < z_n. \tag{A34}$$

The second condition is for *lower contact only* [Fig. 4(d)]:

$$a_1 \leqslant \delta, \quad a_2 > \delta, \tag{A35}$$

$$F_l = Lz_cP_h + L(z_n-z_c)P_e \quad \text{for } z_c < z_n, \tag{A36}$$

$$F_u = L(T-z_n)P_e \quad \text{for } z_c \leqslant z_n, \tag{A37}$$

$$F_l = Lz_nP_h \quad \text{for } z_c \geqslant z_n, \tag{A38}$$

$$F_u = L(z_c-z_n)P_h + L(T-z_c)P_e \quad \text{for } z_c > z_n. \tag{A39}$$

The third condition is for *contact at both bottom and top*:

$$a_1 \leqslant \delta, \quad a_2 \leqslant \delta, \tag{A40}$$

$$F_l = Lz_nP_h, \tag{A41}$$

$$F_u = L(T-z_n)P_h. \tag{A42}$$

In all expressions, the total driving force is again (F_l+F_u) and the torque is evaluated by Eq. (A21) as before.

6. Calculation of contact point and contact area

The contact point z_c needs to be known for all the above-mentioned force and torque calculations. This point is determined by setting the medial surface displacement to zero. From Eq. (A4),

$$\xi_{0n} - (z_c-z_n)\tan(\theta_0+\theta) + \xi = 0, \tag{A43}$$

which gives the result

$$z_c = \mathrm{Min}\left\{ T, \mathrm{Max}\left[0,\ z_n + \frac{\xi_{0n}+\xi}{\tan(\theta_0+\theta)} \right] \right\}, \tag{A44}$$

where the contact point is limited to the range $0<z_c<T$. The vocal fold contact area is now easily computed as

$$a_c = L(T-z_c), \quad a_1 > 0, \quad a_2 \leqslant 0 \tag{A45}$$

$$= Lz_c, \quad a_1 \leqslant 0, \quad a_2 > 0 \tag{A46}$$

$$= LT, \quad a_1 \leqslant 0, \quad a_2 \leqslant 0. \tag{A47}$$

7. Glottal area and glottal flow

If a wave-reflection analog is used for vocal tract acoustics (Liljencrants, 1985; Story, 1995), an analytic solution for the interactive flow is obtainable as described earlier (Titze, 1984). Subglottally, the relations for pressure and flow are

$$P_s = P_s^+ + P_s^-, \tag{A48}$$

$$u = \frac{A_s}{\rho c}(P_s^+ - P_s^-), \tag{A49}$$

where P_s^+ is a forward (rostrally) traveling wave in the trachea, P_s^- is a backward (caudally) traveling wave in the trachea, A_s is the subglottal tube area, c is the velocity of sound, and u is the glottal flow.

Similarly, the pressure and flow relations in the supraglottal region are

$$P_e = P_e^+ + P_e^-, \tag{A50}$$

$$u = \frac{A_e}{\rho c}(P_e^+ - P_e^-), \tag{A51}$$

where the subscript "*e*" stands for "epilarynx tube," which is immediately above the vocal folds. The known pressures in the above-given expressions are P_s^+ and P_e^-, the waves *incident* upon the glottis. The unknown pressures are P_s^- and P_e^+, the departing waves, and the total pressures P_s and P_e. In addition, the flow u is unknown. Thus, there are five unknowns in four equations, which suggests that another equation is needed for an analytical solution. This equation is Eq. (A25), rewritten here as

$$P_{kd} = \frac{1}{2}\rho\frac{u^2}{a_d^2} = \frac{(P_s-P_e)}{1-k_e}. \tag{A52}$$

Substituting P_s and P_e from Eqs. (A48) and (A50) into Eq. (A52), and eliminating the unknown wave pressures P_s^- and P_e^+ from Eqs. (A49) and (A51) in favor of the known wave pressures P_s^+ and P_e^-, a quadratic equation for the flow u is obtained

$$u = \frac{a_dc}{(1-k_e)}\left\{ -\left(\frac{a_d}{A^*}\right) \pm \left[\left(\frac{a_d}{A^*}\right)^2 + \frac{4(1-k_e)}{\rho c^2}(P_s^+ - P_e^-) \right]^{1/2} \right\}, \tag{A53}$$

where

$$A^* = A_sA_e/(A_s+A_e) \tag{A54}$$

is an effective vocal tract area that includes both the tracheal tube and the epilarynx tube. In Eq. (A53), the plus sign is used when the second term in the square brackets is negative, and vice versa.

From an aerodynamic point of view, the glottal area is clearly a_d, the flow detachment area. But from a visual point of view, the area defined by light projected through the glottis is

$$a_g = \mathrm{Max}[0, \mathrm{Min}(a_1, a_2)]. \quad \text{(A55)}$$

This is the area that a photoglottograph would measure.

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

Bestebreurtje, M. E., and Schutte, H. K. (**2000**). ''Resonance strategies for the belting style,'' J. Voice **14**, 194–204.

Bouhuys, A., Mead, J., Proctor, D., and Stevens, K. (**1968**). ''Pressure-flow events during singing,'' Ann. N.Y. Acad. Sci. **155**, 165–176.

Brown, O. L. (**1996**). *Discover Your Voice* (Singular Publishing, San Diego).

Colton, R., and Casper, J. K. (**1996**). *Understanding Voice Problems*, 2nd ed. (Williams & Wilkins, Baltimore).

Estill, J., Fujimura, O., Sawada, M., and Beechler, K. (**1996**). ''Temporal perturbation and voice qualities,'' in *Vocal Fold Physiology: Controlling Complexity and Chaos*, edited by P. J. Davis and N. H. Fletcher (Singular Publishing, San Diego), pp. 237–252.

Flanagan, J. L. (**1965**). *Speech Analysis, Synthesis, and Perception* (Springer, New York).

Holmberg, E., Hillman, R., and Perkell, J. (**1988**). ''Glottal airflow and transglottal air pressure measurements for male and female speakers in soft, normal, and loud voice,'' J. Acoust. Soc. Am. **84**, 511–529.

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

Liljencrants, J. (**1985**). ''Speech synthesis with a reflection-type line analog,'' Ph.D. dissertation, Royal Institute of Technology, Stockholm, Sweden.

Liljencrants, J. (**1991**). ''A translating and rotating mass model of the vocal folds,'' STL Quarterly Progress and Status Report 1, Speech Transmission Laboratory, Royal Institute of Technology (KTH), Stockholm, Sweden.

Linklater, K. (**1976**). *Freeing the Natural Voice* (Drama Book, New York).

Pelorson, X., Hirschberg, A., Van Hassel, R., Wijnands, A., and Auregan, V. (**1994**). ''Theoretical and experimental study of quasi-steady flow separation within the glottis during phonation: Application to a modified two-mass model,'' J. Acoust. Soc. Am. **96**, 3416–3431.

Sondhi, M. M., and Schroeter, J. (**1987**). ''A hybrid time-frequency domain articulatory speech synthesizer,'' IEEE Trans. Acoust., Speech, Signal Process. **35**, 955–967.

Story, B. H. (**1995**). ''Physiologically-based speech simulation using an enhanced wave-reflection model of the vocal tract,'' Ph.D. dissertation, University of Iowa.

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

Story, B. H., Titze, I. R., and Hoffman, E. A. (**2001**). ''The relationship of vocal tract shape to three voice qualities,'' J. Acoust. Soc. Am. **109**, 1651–1667.

Sullivan, J. (**1985**). *The Phenomena of the Belt/pop Voice* (Sullivan, Denver).

Sundberg, J. (**1987**). *The Science of the Singing Voice* (Northern Illinois U.P., Dekalb).

Sundberg, J., Cleveland, T. F., Stone, R. E., and Iwarsson, J. (**1999**). ''Voice source characteristics of six premier country singers,'' J. Voice **13**, 168–183.

Titze, I. R. (**1984**). ''Parametrization of glottal area, glottal flow, and vocal fold contact area,'' J. Acoust. Soc. Am. **75**, 570–580.

Titze, I. R., and Story, B. H. (**in press**). ''Rulse for controlling low-dimensional vocal fold models with muscle activities,'' J. Acoust. Soc. Am.

Titze, I. R., and Story, B. H. (**1997**). ''Acoustic interaction of the voice source with the lower vocal tract,'' J. Acoust. Soc. Am. **101**, 2234–2243.

Titze, I. R., and Talkin, D. (**1979**). ''A theoretical study of the effects of various laryngeal configurations on the acoustics of phonation,'' J. Acoust. Soc. Am. **66**, 60–74.