Validation of theoretical models of phonation threshold pressure with data from a vocal fold mechanical replica (L)

Jorge C. Lucero[a)]
Department of Mathematics, University of Brasilia, Brasilia DF 70910-900, Brazil

Annemie Van Hirtum,[b)] Nicolas Ruty,[c)] Julien Cisonni,[d)] and Xavier Pelorson[e)]
GIPSA-lab, UMR CNRS 5216, Grenoble Universities, 961 rue de la Houille Blanche, BP 46, 38402 Saint-Martin d'Heres, France

(Received 20 May 2008; revised 26 November 2008; accepted 1 December 2008)

This paper analyzes the capability of a mucosal wave model of the vocal fold to predict values of phonation threshold lung pressure. Equations derived from the model are fitted to pressure data collected from a mechanical replica of the vocal folds. The results show that a recent extension of the model to include an arbitrary delay of the mucosal wave in its travel along the glottal channel provides a better approximation to the data than the original version of the model, which assumed a small delay. They also show that modeling the vocal tract as a simple inertive load, as has been proposed in recent analytical studies of phonation, fails to capture the effect of the vocal tract on the phonation threshold pressure with reasonable accuracy.
 [DOI: 10.1121/1.3056468]

PACS number(s): 43.70.Aj, 43.70.Bk, 43.70.Jt [CHS] Pages: 632–635

I. INTRODUCTION

The phonation threshold of lung pressure is defined as the minimum value required to initiate vocal fold oscillation. It is an important factor for building empirical laws of laryngeal aerodynamics (Titze, 1992) and represents the pressure level at which the energy transferred from the airflow to the vocal folds is large enough to overcome the energy dissipated in the tissues, so that an oscillatory movement of growing amplitude may take place (Lucero, 1999). The phonation threshold pressure value has also been interpreted as a measure of ease of phonation and proposed as a diagnostic tool for vocal health (Titze *et al.*, 1995).

Two decades ago, Titze (1988) derived an equation for the phonation threshold pressure by modeling the vocal fold oscillatory movement as a superficial mucosal wave propagating in the direction of the airflow. The equation related the threshold pressure to biomechanical parameters, namely, glottal geometry, tissue damping coefficient, and mucosal wave velocity. However, it lacked the oscillation frequency as an explicit parameter. It is well known that phonation threshold pressure increases with frequency, as demonstrated by experimental measures (e.g., Titze, 1992). In his works, Titze (1988, 1992) pointed out the missing parameter and offered a possible solution by relating the vocal fold thickness and mucosal wave velocity to the oscillation frequency.

In a recent paper (Lucero and Koenig, 2007), it was shown that the lack of the frequency factor is a consequence of one of the simplifications made in the vocal fold model: the assumption of a small time delay for the mucosal wave to travel along the vertical dimension of the vocal folds. A more general analysis for an arbitrary time delay results in an extended equation for the phonation threshold pressure, which includes the oscillation frequency explicitly.

Because a direct validation, using *in vivo* measurements on human speakers, of these theoretical predictions cannot be achieved easily, we propose to test them against *in vitro* experiments using a mechanical replica of the vocal folds. Mechanical replicas of the voice production system, such as the one introduced by Ruty *et al.* (2007), allow us to test theoretical models against experimental data quantitatively and to extract conclusions about the range of validity of those models.

II. EXTENSION OF THE MUCOSAL WAVE MODEL

Figure 1 shows a schematic of the mucosal wave model. Complete right-left symmetry of the folds is assumed, and motion of tissues is allowed only in the horizontal direction. A surface wave propagates through the superficial tissues, in the direction of the airflow (upward).

The equation of motion of the vocal fold tissues is obtained by lumping their biomechanical properties at the midpoint of the glottis and assuming that they are forced by the mean glottal pressure P_g, which yields

$$M\ddot{\xi} + B\dot{\xi} + K\xi = P_g, \tag{1}$$

where ξ is the tissue displacement at the midpoint, and M, B, K, are the mass, damping, and stiffness, respectively, per unit area of the medial surface of the vocal folds.

The glottal aerodynamics is modeled by assuming that the flow is frictionless, stationary, and incompressible. Further, we assume that the subglottal pressure is equal to a constant lung pressure P_L, the vocal tract input area is much

[a)]Electronic mail: lucero@unb.br
[b)]Electronic mail: annemie.vanhirtum@gipsa-lab.inpg.fr
[c)]Electronic mail: nicolas.ruty@gipsa-lab.inpg.fr
[d)]Electronic mail: julien.cisonni@gipsa-lab.inpg.fr
[e)]Electronic mail: pelorson@icp.inpg.fr

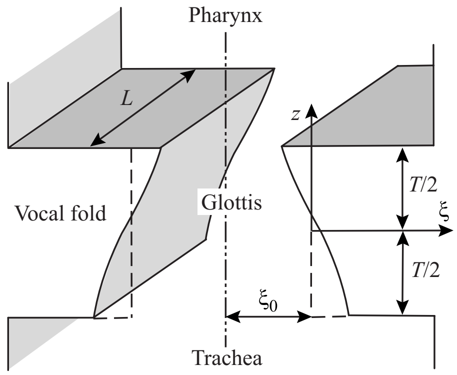

FIG. 1. Vocal fold model (after Titze, 1988).

larger than the glottal area, and the prephonatory glottal channel is rectangular. Under such conditions, the mean glottal air pressure P_g may be expressed as

$$P_g = P_i + (P_L - P_i)(1 - a_2/a_1)/k_t, \tag{2}$$

where P_i is the supraglottal pressure (at the entry of the vocal tract), k_t is a transglottal pressure coefficient, and a_1, a_2 are the glottal areas at the lower and upper edges of the glottal channel, respectively, given by

$$a_1(t) = 2L[\xi_0 + \xi(t+\tau)], \tag{3}$$

$$a_2(t) = 2L[\xi_0 + \xi(t-\tau)], \tag{4}$$

where ξ_0 is the prephonatory glottal half-width, τ is the time delay for the mucosal wave to travel half the glottal height ($T/2$ in Fig. 1), and L is the vocal fold length.

Following Chan and Titze (2006), the input pressure to the vocal tract is modeled as $P_i \approx I\dot{u}$, where I is the vocal tract inertance and $\dot{u}$ is the time derivative of the airflow. This approximation is valid when the oscillation frequency of the vocal folds (F_0) is below the first formant (F_1) of the vocal tract (Titze, 1988). For a quasisteady flow condition and small amplitude oscillations around an abducted (open) glottis, the flow derivative may be approximated by $\dot{u} \approx v_2\dot{a}_2$, where $v_2 = \sqrt{2P_L/(k_t\rho)}$ is the air particle velocity at the glottal exit, ρ is the air density, and a_2 is the glottal area at the upper edge of the vocal folds, given by Eq. (4). Therefore, we have the approximation

$$P_i = 2LIv_2\xi'(t-\tau), \tag{5}$$

where $\xi'(t-\tau) = d\xi/d(t-\tau)$.

With the above assumptions, the mean glottal air pressure is then

$$P_g = 2LIv_2\xi'(t-\tau) + \left[\frac{P_L - 2LIv_2\xi'(t-\tau)}{k_t}\right] \times \left[\frac{\xi(t+\tau) - \xi(t-\tau)}{\xi_0 + \xi(t+\tau)}\right]. \tag{6}$$

The equation of motion for the vocal fold oscillation is then given by Eqs. (1) and (6). More details on the assumptions of the model and the derivation of the equations may be easily found in the cited references.

III. OSCILLATION THRESHOLD PRESSURE

Equations (1) and (6) constitute a functional differential equation with advance and delay arguments ($t+\tau$ and $t-\tau$, respectively). It has a unique fixed point at $\xi=0$, which corresponds to the prephonatory position.

Linearization around that position produces

$$M\ddot{\xi} + B\dot{\xi} + K\xi = 2LIv_2\xi'(t-\tau) + \frac{P_L}{\xi_0 k_t}[\xi(t+\tau) - \xi(t-\tau)]. \tag{7}$$

Proposing a solution of the form $\xi(t) = Ce^{\lambda t}$, where C and λ are complex constants, and seeking nonzero solutions produces the associated characteristic equation

$$M\lambda^2 + B\lambda + K - 2LIv_2\lambda e^{-\lambda\tau} - \frac{2P_L}{k_t\xi_0}\sinh(\lambda\tau) = 0. \tag{8}$$

Let P_{th} denote the phonation threshold value of the lung pressure P_L, at which the vocal fold oscillation starts. At the threshold, a pair of complex roots of the above equation cross the imaginary axis from left to right. Next, letting $\lambda = i\omega$, $P_L = P_{th}$, and separating real and imaginary parts, we obtain the conditions

$$-\omega^2 M + K - 2LIv_2\omega\sin(\omega\tau) = 0, \tag{9}$$

$$\omega B - 2LIv_2\omega\cos(\omega\tau) - \frac{2P_{th}}{k_t\xi_0}\sin(\omega\tau) = 0, \tag{10}$$

and, from Eq. (10), we obtain

$$P_{th} = \frac{k_t\xi_0 B\omega}{2\sin(\omega\tau)} - k_t\xi_0 lIv_2\omega\cot(\omega\tau), \tag{11}$$

where $0 < (\omega\tau) < \pi$, and v_2 is computed at the threshold condition, i.e., $v_2 = \sqrt{2P_{th}/(k_t\rho)}$.

If we ignore the effect of the vocal tract by setting $I = 0$ (no vocal tract load), we obtain

$$P_{th} = \frac{k_t\xi_0 B\omega}{2\sin(\omega\tau)}, \tag{12}$$

which is the equation found by Lucero and Koenig (2007). For $\tau \to 0$, $\sin(\omega\tau) \to \omega\tau$. Eq. (12) simplifies further to Titze's (1988) result

$$P_{th} = \frac{k_t\xi_0 B}{2\tau}. \tag{13}$$

Note also that a Taylor expansion of Eq. (12) around $\omega=0$ produces

$$P_{th} = \frac{k_t\xi_0 B}{2}\left(\frac{1}{\tau} + \frac{\tau}{6}\omega^2 + \mathcal{O}(\omega^4)\right). \tag{14}$$

Keeping only the first two terms, we obtain a quadratic approximation to P_{th} in terms of ω, as proposed by Titze (1992).

Considering now $\tau \to 0$ in Eq. (11), and therefore $\sin(\omega\tau) \to \omega\tau$ and $\cot(\omega\tau) \to 1/(\omega\tau)$, we obtain

$$P_{\mathrm{th}}=\frac{k_t\xi_0 B}{2\tau}-\frac{k_t\xi_0 L I v_2}{\tau}, \tag{15}$$

which is the equation found by Chan and Titze (2006).

IV. DATA

To test the above results, we used data collected from a mechanical replica for a previous study by Ruty *et al.* (2007, Figs. 8, 9 and 10 of their paper). The replica consists of two metal half-cylinders covered with latex, which mimic the vocal fold structure in a 3:1 scale, with a similar aspect ratio. Geometrical dimensions and other parameters of the replica were chosen in order to match as closely as possible the glottal aerodynamics (see Table I of Ruty *et al.*, 2007). The cylinders are filled with water, at a controlled internal pressure P_c. The initial separation between the latex tubes decreases when P_c is increased, and becomes zero for $P_c > 5000$ Pa. The vocal tract is simulated with a downstream cylindrical resonator. Two different tubes were used, with a diameter of 25 mm, and lengths of 250 mm and 500 mm, respectively. Their dimensions were chosen in order to present a weak and a strong acoustical coupling. The first acoustical resonances of the tubes are 340 Hz, for the 250 mm tube, and 170 Hz, for the 500 mm tube. Those resonance frequencies are, respectively, higher than and comparable to the oscillation frequency of the latex structure, which is in the range of 110–170 Hz.

Measures of oscillation threshold pressure were obtained by increasing the air pressure upstream of the vocal fold replica until an oscillation of the latex structures was detected. The oscillation frequency at the oscillation onset was then computed by spectral analysis on the acoustic output signal. A threshold pressure for the oscillation offset was also measured, by decreasing the upstream pressure until the oscillation was interrupted, but those values are not used here. This procedure was repeated for various values of the water pressure P_c, and for the two cylindrical resonators.

For our analysis, we ignored all data for $P_c > 5000$ Pa, because in that range the latex tubes are in contact ($\xi_0 = 0$), and consequently the above equations produce $P_{\mathrm{th}} = 0$. Let us also recall that the mucosal wave model assumes an open prephonatory glottis, wide enough so that the effect of air viscosity may be neglected (Titze, 1988).

V. NUMERICAL RESULTS

We fitted the above theoretical equations to Ruty *et al.*'s (2007) data by a standard least squares procedure implemented in Matlab, with the oscillation threshold pressure P_{th} as the target.

In a first numerical experiment, we fitted Eq. (12) to each resonator's data, with $(k_t B)$ and τ as parameters; the results are shown in Fig. 2. The computed optimal values were $(k_t B) = 350.81$ Pa s/m, $\tau = 2.66$ ms, and $(k_t B) = 1864.0$ Pa s/m, $\tau = 2.90$ ms for the 250 and 500 mm resonators, respectively. For comparison, we also fitted Eq. (13), obtaining $(k_t B) = 3436.4$ Pa s/m, $\tau = 7.12$ ms, and $(k_t B) = 248.68$ Pa s/m, $\tau = 0.0962$ ms for the 250 and 500 mm resonators, respectively. As shown by the plots, our extended

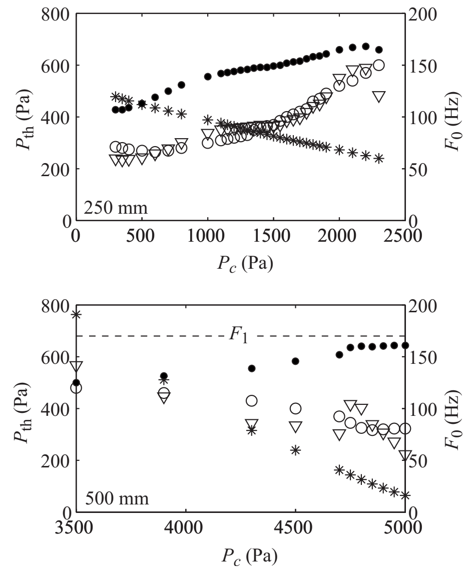

FIG. 2. Oscillation threshold pressure P_{th} and frequency F_0 vs internal pressure P_c for a 250 mm resonator (upper panel) and 500 mm resonator (lower panel). Circles: measured pressure values; triangles: theoretical pressure values given by Eq. (12), stars: theoretical pressure values given by Eq. (13); filled circles: measured oscillation frequency. The broken line in the lower panel indicates the first acoustical resonance (F_1) of the resonator.

equation (12) provides a reasonably good approximation for both resonators, better than Eq. (13). In case of the 250 mm resonator, Eq. (13) produces a decreasing P_{th} pattern, instead of the measured increasing pattern, because ξ_0 decreases when P_c increases (Ruty *et al.*, 2007, Fig. 8). The extended Eq. (12), on the other hand, is able to compensate for the decrease in ξ_0 by the increase of oscillation frequency F_0 at larger values of P_c.

In the case of the 500 mm resonator, the plot also shows the location of the first acoustical resonance F_1, at 170 Hz (for the 250 mm resonator, $F_1 = 340$ Hz falls outside the frequency range of the plot). Note that the oscillation frequency F_0 is close to F_1, particularly at large values of P_c, and therefore the pure inertance approximation for the vocal tract load does not hold.

In a second numerical experiment, we fitted Eqs. (11) and (15) to both 250 and 500 mm resonator data sets simultaneously, with k_t, B, and τ as parameters, to see how well they capture the vocal tract effect (Fig. 3). We set $L = 45$ mm (from Ruty *et al.*, 2007) and $\rho = 1.14$ kg/m^3 (from Chan and Titze, 2006). Also, the range of possible values for the transglottal coefficient k_t was limited to [1.0, 1.4] (Titze, 1988). The vocal tract inertance was computed as $I = \rho l/A$, where l is the length and A is the cross sectional area. For the 250 and 500 mm resonators, we have $I = 580.60$ kg/m^4 and $I = 1161.2$ kg/m^4, respectively. The computed optimal parameters were $k_t = 1.40$, $B = 783.95$ Pa s/m, $\tau = 1.37$ ms, for Eq. (11), and $k_t = 1.04$, $B = 1363.3$ Pa s/m, $\tau = 4.83 \times 10^{-7}$ ms, for Eq. (15).

In this experiment, the results for the 250 mm resonator are similar to those in Fig. 2: the extended Eq. (11) provides a good approximation, better than Eq. (15). Equation (15) does not predict the observed increase of P_{th} with P_c. The best approximation it can produce is by setting a very small

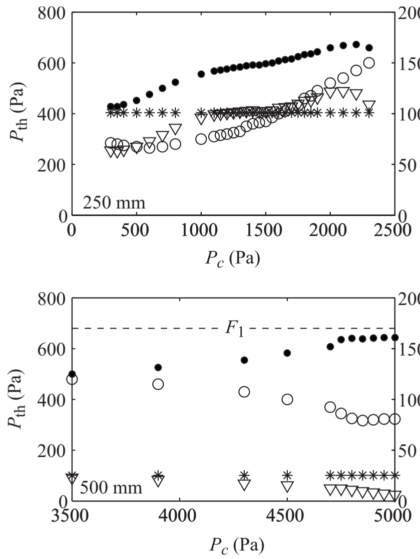

FIG. 3. Oscillation threshold pressure $P_{\rm th}$ and frequency F_0 vs internal pressure P_c for a 250 mm resonator (upper panel) and 500 mm resonator (lower panel). Circles: measured pressure values; triangles: theoretical pressure values given by Eq. (11), stars: theoretical pressure values given by Eq. (15); filled circles: measured oscillation frequency. The broken line in the lower panel indicates the first acoustical resonance (F_1) of the resonator.

value of τ, which results in an almost constant P_{th}. The results for the 500 mm resonator, on the other hand, are much poorer than those in Fig. 2: both Eqs. (11) and (15) produce values of threshold pressure much lower than the measured values.

VI. CONCLUSIONS

The above results show that the extended equation for phonation threshold pressure, given by Eq. (12), provides a better theoretical characterization than Eq. (13) previously derived by Titze (1988). In particular, the extended model contains the oscillation frequency as an explicit parameter, which was missing in the previous model, and therefore is able to capture phonation threshold versus frequency relations.

The results also show that modeling the vocal tract input pressure with the simple inertive load of Eq. (5) seems a crude approximation, which fails to model the effect of the vocal tract on the phonation threshold pressure with reasonable accuracy. However, two issues must be considered here: First, the inertive model is based on the assumption of an oscillation frequency much lower than the first vocal tract formant. This assumption does not hold well for the 500 mm resonator, for which the theoretical results are poor compared to the data. Second, a lumped impedance representation for the vocal tract may still be too simple to fit the experimental data, and a more sophisticated frequency-dependent model might be required.

Finally, note that the theoretical flow model relies on many simplifying, and thus questionable, assumptions by considering that the glottal flow is frictionless, quasi-steady, and incompressible. Of all these assumptions, the work of Ruty *et al.* (2007) tends to show that viscous effects are the most critical.

All of the above issues are currently being considered for extensions of this work.

ACKNOWLEDGMENTS

This work was supported by CAPES—Brazil under Program STIC-AmSud, MCT/CNPq, and by a Ph.D. grant from the French Ministry of Education and Research.

Chan, R. W., and Titze, I. R. (**2006**). "Dependence of phonation threshold pressure on vocal tract acoustics and vocal fold tissue mechanics," J. Acoust. Soc. Am. **119**, 2351–2362.

Lucero, J. C. (**1999**). "Theoretical study of the hysteresis phenomenon at vocal fold oscillation onset-offset," J. Acoust. Soc. Am. **105**, 423–431.

Lucero, J. C., and Koenig, L. L. (**2007**). "On the relation between the phonation threshold lung pressure and the oscillation frequency of the vocal folds," J. Acoust. Soc. Am. **121**, 3280–3283.

Ruty, N., Pelorson, X., Hirtum, A. V., Lopez-Arteaga, I., and Hirschberg, A. (**2007**). "An in vitro setup to test the relevance and the accuracy of low-order vocal folds models," J. Acoust. Soc. Am. **121**, 479–490.

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

Titze, I. R. (**1992**). "Phonation threshold pressure: a missing link in glottal aerodynamics," J. Acoust. Soc. Am. **91**, 2926–2935.

Titze, I. R., Schmidt, S. S., and Titze, M. R. (**1995**). "Phonation threshold pressure in a physical model of the vocal fold mucosa," J. Acoust. Soc. Am. **97**, 3080–3084, part 1.