Mehrdad H. Farahani and Zhaoyan Zhang: JASA Express Letters [http://dx.doi.org/10.1121/1.4959965] Published Online 2 August 2016

# Experimental validation of a three-dimensional reduced-order continuum model of phonation

**Mehrdad H. Farahani and Zhaoyan Zhang**[a)]
*Department of Head and Neck Surgery, University of California, Los Angeles, 31-24 Rehab Center, 1000 Veteran Avenue, Los Angeles, California 90095, USA*
*mhosnieh@ucla.edu, zyzhang@ucla.edu*

**Abstract:** Due to the complex nature of the phonation process, a one-dimensional (1D) glottal flow description is often used in current phonation models. Although widely used in voice research, these 1D flow-based phonation models have not been rigorously validated against experiments. In this study, a 1D glottal flow model is coupled with a three-dimensional nonlinear continuum model of the vocal fold and its predictions are compared to physical model experiments. The results show that the 1D flow-based model is able to predict the phonation threshold pressure and onset frequency within reasonable accuracy and to reproduce major vibratory features observed in the experiments.


[AL]
**Date Received:** May 9, 2016 **Date Accepted:** July 13, 2016

## 1. Introduction

Human phonation involves a complex nonlinear fluid-structure interaction within the glottis. Fully resolving this complex interaction in three dimensions is computationally challenging and time-consuming. Often, simulation of a few cycles of vocal fold vibration would take days or months, depending on the degree of complexity being modeled, which limits the use of such models in parametric studies of voice production or practical voice applications. Alternatively, reduced-order models of phonation with simplified physics have been developed to reduce the computational costs. In particular, the glottal flow is often simplified as a one-dimensional (1D) flow up to the point where the glottal flow separates from the vocal fold surface, with the flow separation point predicted based on a simple geometric measure (Ishizaka and Flanagan, 1972; Liljencrants, 1991). The 1D flow models neglect complex flow phenomena such as intraglottal and supraglottal vortices and turbulence, thus significantly reducing computational expense. While these simplifications are necessary for large-scale parametric studies (e.g., Titze and Talkin, 1979; Zhang, 2016) and able to qualitatively reproduce experimental observations (e.g., Ruty *et al.*, 2007; Zhang and Luu, 2012), they have yet to be rigorously validated against experiments. One attempt at experimental validation was by Ruty *et al.* (2007), in which a 1D flow model was coupled to two lumped-element vocal fold models and the model prediction was compared to experimental observation. They showed that the simplified models were able to qualitatively match experiments but quantitative discrepancies remained due to inaccurate estimations of model parameters and possibly the simplified flow description.

The goal of this study is to quantitatively evaluate the accuracy of a 1D flow-based phonation model in modeling the glottal fluid-structure interaction during phonation, by comparing model predictions to the experimental observation in a silicone vocal fold model as reported in Mendelsohn and Zhang (2011). Unlike the Ruty *et al.* (2007) study which used lumped-element vocal fold models, a continuum vocal fold model is used in the present study with experimentally measured vocal fold geometry and material properties, thus eliminating the need for model parameter estimation or tuning. As a first step toward a more complete validation, this study compares the phonation threshold pressure and fundamental frequency as well as the vibration pattern at onset between simulation and experiment.

## 2. Method

### *2.1 Experiment*

Details of the experiment can be found in Mendelsohn and Zhang (2011), in which phonation onset characteristics were investigated using a two-layer body-cover silicone vocal fold model of varying stiffness and geometry. Four experimental conditions from

[a)]Author to whom correspondence should be addressed.

 

Table 1. The linear Young's modulus ($E_b$), loss factor, and Rayleigh damping coefficients ($\alpha$, $\beta$) of the vocal fold body layer in the four simulation conditions. For all four conditions, the cover Young's modulus ($E_c$) is 6.13 (5.51) kPa, and the loss factor is 0.09 with the Rayleigh damping coefficients of $\alpha = 26.024941$ and $\beta = 6.913 \times 10^{-5}$. The values in parentheses are the original data reported in Mendelsohn and Zhang (2011).

| Condition | $E_b$ (kPa) | Loss factor | $\alpha$ | $\beta$ |
|---|---|---|---|---|
| 1 | 6.13 (5.51) | 0.09 | 26.024941 | $6.913 \times 10^{-5}$ |
| 2 | 11.87 (9.33) | 0.08 | 23.133281 | $6.145 \times 10^{-5}$ |
| 3 | 23.28 (21.56) | 0.06 | 17.349961 | $4.609 \times 10^{-5}$ |
| 4 | 66.25 (63.33) | 0.06 | 17.349961 | $4.609 \times 10^{-5}$ |

the Mendelsohn and Zhang (2011) study are computationally simulated in the present study (Table 1). These four conditions had identical vocal fold geometry (geometry G1 in Mendelsohn and Zhang, 2011) and cover layer Young's modulus of 5.51 kPa. The four conditions differed in the body layer Young's modulus, which was 5.51, 9.33, 21.56, and 63.33 kPa for conditions 1 to 4, respectively. The first condition had the same stiffness in the body and cover layers, thus was essentially a one-layer vocal fold model.

### *2.2 Computational method*

A three-dimensional two-layer finite element model (FEM) of the vocal fold is developed using the commercial finite element software ADINA. For simplicity, left–right symmetry in vocal fold properties and vibration about the mid-sagittal plane is assumed, and only one vocal fold is modeled in this study. Vocal fold collisions are modeled using a contact surface (Fig. 1), which prevents the vocal fold from crossing the mid-sagittal plane. To avoid singularities in the flow model during the contact, the contact surface is located at 0.002 mm away from the mid-sagittal plane. No vocal tract is considered as in the experiment. Although the experimental setup had a subglottal system, the tracheal tube was 11 cm and its effect is considered small. Thus, no subglottal system is included in this study and a constant subglottal pressure $P_{sub}$ is imposed at the entrance to the glottis. The vocal fold geometry in the computational model is the same as the geometry G1 in the experiment (Mendelsohn and Zhang, 2011), except for a small fillet added to the lateral edge of the superior surface of the vocal fold. Specifically, the vocal fold cross section in the coronal plane has a depth (along the medial-lateral direction) of 6 and 1.5 mm for the body and cover layers, respectively, and a medial surface thickness of 3 mm in the superior–inferior direction. The vocal fold length is 15 mm along the anterior–posterior (AP) direction.

A large-displacement, large-strain formulation is used for the vocal fold, which takes into account both geometric and material nonlinearities. The stress-strain curves of the four different vocal fold stiffness conditions were obtained from tensile tests at 1 Hz (higher frequency is currently not possible due to experimental limitations), from which the material loss factors are also estimated. The Young's moduli at the origin (condition of zero stain) and the corresponding loss factors as estimated from the stress-strain curves for the four stiffness conditions are listed in Table 1. Because the tensile tests were conducted using silicone models fabricated at a different time from those used in the Mendelsohn and Zhang (2011) study, the estimated Young's moduli are slightly different from those reported in Mendelsohn and Zhang (2011). A

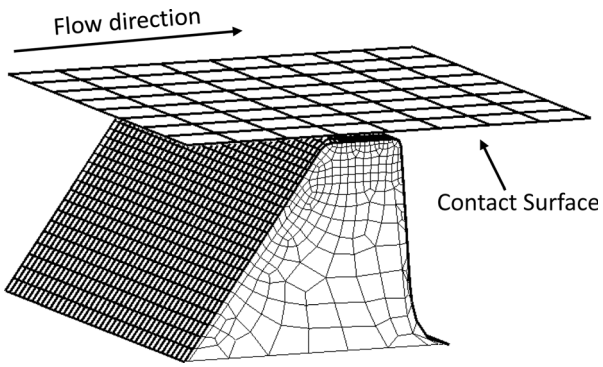


Fig. 1. The finite element model of the vocal fold and the contact surface.

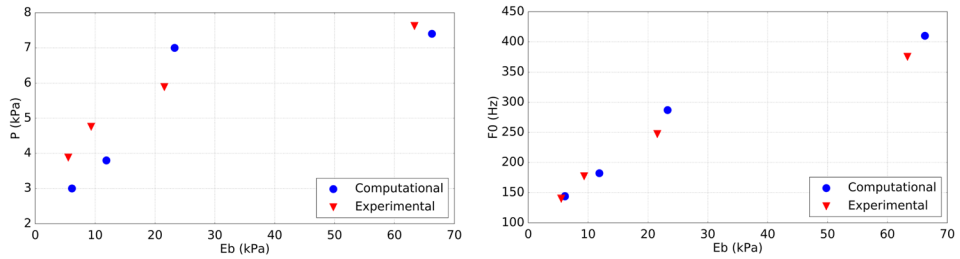


Fig. 2. (Color online) The phonation threshold pressure (left) and onset frequency (right) predicted from the simulation and measured in the experiment as a function of the body-layer Young's modulus.

Rayleigh damping scheme is used in the simulation, with its coefficients determined so that the loss factor equals the experimentally measured loss factor at 69 and 132 Hz, which roughly correspond to the frequencies of the first two *in vacuo* eigenmodes which exhibit dominant medial-lateral motion in the coronal plane. The Rayleigh damping coefficients are also listed in Table 1. Note that these data were obtained at a frequency much lower than the actual phonation frequency. The material properties during phonation may be slightly different, particularly for the loss factor which may increase noticeably with increasing frequency.

The glottal flow is modeled as a 1D quasi-steady flow up to the point of flow separation. For a convergent glottal configuration, the flow is assumed to separate from the vocal fold surface at the superior edge of the vocal fold. For a divergent glottis, the flow is assumed to separate at a location downstream of the minimum glottal constriction with a cross-sectional area 1.2 times the minimum glottal area. The air pressure downstream of the flow separation point is assumed to be the atmospheric pressure. The intraglottal pressure is determined by the time-varying glottal cross-sectional area function, which is evaluated for each time step at $N = 40$ equally spaced cross sections ($A_i$, and $i = 1, \ldots, N$), with $A_1$ at the glottal entrance and $A_N$ at the flow separation point. The intraglottal pressures at two consecutive cross sections are related by (Pelorson *et al.*, 1994)

$$P_{i+1} = P_i + \frac{1}{2}\rho Q^2\left(\frac{1}{{A_i}^2} - \frac{1}{{A_{i+1}}^2}\right) - \frac{12\mu dx}{L_0 {h_{\text{eq}}}^3}, \tag{1}$$

where $\rho$ is the air density, $Q$ is the volume flow rate, $\mu$ is the air viscosity, $L_0$ is the vocal fold length, $h_{\text{eq}} = (A_i + A_{i+1})/2L0$, and $dx$ is the axial distance between the two consecutive cross sections. The volume flow rate $Q$ can be obtained using Eq. (1) and considering the boundary conditions $P_1 = P_{\text{sub}}$ and $P_N = 0$, and the intraglottal pressure distribution can be then calculated using Eq. (1). The intraglottal pressures at these cross sections are then used to calculate by interpolation the air pressures acting on each node on the surface of the vocal fold in the FEM model, which are then coupled to the vocal fold model through a user-defined subroutine in ADINA.

For each of the four stiffness conditions in Table 1, the subglottal pressure is increased with an increment of 100 Pa in order to capture phonation onset. For each subglottal pressure, phonation is simulated for 0.4 s, using an adaptive time marching scheme with an initial time step size of 0.01 ms. Phonation onset is considered to occur at the minimum subglottal pressure at which the vibration amplitude at the end of the 0.4 s-long simulation either reaches steady-state or continues to grow. In general, each simulation takes about 1 day for the one-layer condition and about 4 days for the two-layer conditions which require a finer mesh.

## 3. Results

Figure 2 compares the phonation threshold pressure and fundamental frequency predicted from the model and those measured from the experiment. The agreement is generally good, especially for phonation onset frequency. The largest discrepancy in phonation threshold pressure (about 22.7%) occurs for condition 1, whereas the discrepancy in phonation onset frequency is the largest for condition 4 with the stiffest body layer (about 9.3%).

Figure 3 compares the predicted and experimentally observed vocal fold vibration pattern within one oscillation cycle from a superior view for conditions 1 and 4, at a subglottal pressure slightly above onset. For condition 1, which corresponds to a one-layer vocal fold, both experiment and simulation show large vocal fold deformation along the medial-lateral direction, with the maximum vibration amplitude as large as

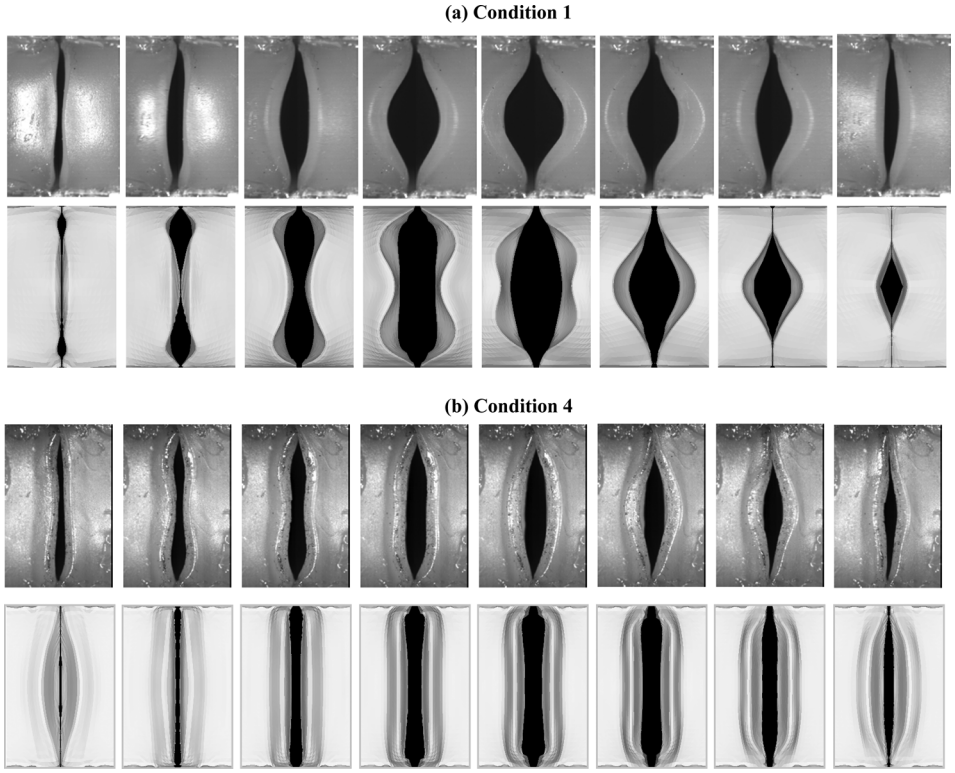


Fig. 3. Vocal fold vibration from a superior view for eight equally-spaced instants during one oscillation cycle for conditions 1 (top two rows) and 4 (bottom two rows). For each condition, the images in the first panel are from the experiment, and the second panel is from the simulation.

half of the vocal fold depth along the medial-lateral direction. Large vertical motion over the entire superior surface is also observed as the vocal fold is pushed upwards. In the experiment the vocal fold also exhibited AP out-of-phase motion, i.e., the middle portion vibrated slightly out-of-phase from the anterior or posterior portions of the vocal fold, resulting in a number eight-shaped glottis in the closing phase. This AP out-of-phase motion is also observed in the simulation, although it appears to be more strongly excited than in the experiment, resulting in a larger vibration amplitude at the anterior and posterior quarter locations and a number eight-shaped glottis even during the opening phase in the simulation. In addition, the vocal fold in the simulation is able to almost close the middle glottis (frames 1 and 2), which was not observed in the experiment.

Condition 4 corresponds to a two-layer vocal fold condition with the largest body-cover stiffness ratio. Compared to condition 1, the experiment showed that the overall glottal opening area was reduced and large vocal fold motion was confined to the medial edge of the vocal fold (as illustrated by the medial region of darker color). Both features are reproduced in the simulation. Similar to condition 1, the AP out-of-phase motion is excited more strongly in the simulation, resulting in a larger glottal opening at the anterior and posterior edges in the simulation than the experiment. The simulation also predicts a much smaller minimum glottal area and better glottal closure than the experiment.

Experiments using silicone vocal fold models also showed that the entire medial surface of the one-layer vocal fold model tends to vibrate in phase as a whole (Zhang *et al.*, 2006; Murray and Thomson, 2012). Figure 4 shows the time traces of the predicted medial-lateral positions of the upper and lower margins of the medial surface in the coronal plane. The upper and lower margins move almost in phase with each other for the one-layer vocal fold in condition 1, whereas a phase difference is observed in condition 4. The vibration amplitude in both the medial-lateral direction and the superior–inferior direction (not shown in Fig. 4) is much reduced in condition 4 compared to condition 1, consistent with the observation in Mendelsohn and Zhang (2011). Figure 4 also shows that a phase difference in the medial-lateral motion exists between the middle portion and the off-center portions of the vocal fold, which is consistent with the observation in Xuan and Zhang (2014).

## 4. Discussion

Previously, similar 1D flow models are shown to be able to qualitatively reproduce experimental observations (Ruty *et al.*, 2007; Zhang and Luu, 2012). The results of

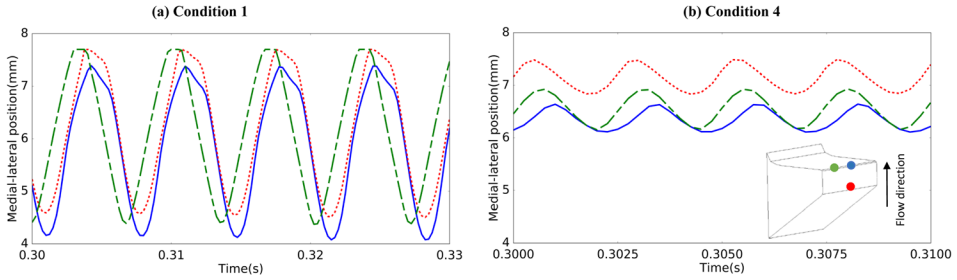


Fig. 4. (Color online) The time histories of the medial-lateral position of three points on the vocal fold medial surface for conditions 1 (left) and 4 (right). The three points are the upper (solid line) and lower (dotted line) margins of the medial surface in the coronal plane, and the upper margin (dashed-dotted line) in an off-coronal plane.

this study show that the 1D flow model is able to predict the phonation threshold pressure and onset frequency within reasonable accuracy, and reproduce major vibratory features observed in experiments. These studies support the validity of the use of 1D flow models in simulations of phonation. The good agreement with experiment also indicates a small effect of supraglottal flow structures (e.g., vortices and jet flapping) that are neglected in the 1D flow model, which is consistent with our previous experimental study (Zhang and Neubauer, 2010).

Two major differences are observed between simulation and experiment in the present study. One is the strong excitation of the AP out-of-phase motion in the simulation, particularly the large vibration amplitude at the anterior and posterior quarter locations. This larger than experiment excitation could be due to the inaccurate loss factor data which were obtained at 1 Hz. The loss factor is expected to increase considerably with increasing frequency. Because strong AP out-of-phase motion is more likely to occur in higher-order eigenmodes with higher eigenfrequencies, it is possible that with more realistic loss factor data the AP out-of-phase motion may be less excited than observed in this study. A second difference between experiment and simulation is that the vocal fold is able to close the glottis better in the simulation than in the experiment. It is currently unclear what might have caused this difference. The inaccurate loss factor data may play a role here. It is also possible that the flow features neglected in the 1D flow model may have a negative impact on the glottal closure pattern. Note that Zhang and Luu (2012), which used a similar 1D flow model but without accounting for the glottal flow viscous loss, also reported a stronger fluid-structure interaction in their numerical model than in their experiment. These issues need to be further explored in future studies, using more accurate vocal fold material properties data, a better representation of the experimental setup, and gradually refined flow models [e.g., including the glottal channel curvature effect, or using more accurate models of the glottal flow viscous loss (Fulcher and Scherer, 2011)]. Such studies would help to reveal the effects on phonation of the various flow features that are neglected in 1D flow models.

Finally, the comparison between experiment and simulation in this study is limited to phonation threshold pressure and the vocal fold vibration pattern. Because the primary end interest of voice production is the produced acoustics, future studies will aim toward validating the 1D flow models with respect to their capability in accurately predicting the output acoustics.

## Acknowledgments

This study was supported by NIH Grant Nos. R01DC011299 and R01DC009229. The authors thank Dr. Himadri Samajder for assistance in the tensile tests of the silicone models.

## References and links

Fulcher, L., and Scherer, R. (**2011**). "Phonation threshold pressure: Comparison of calculations and measurements taken with physical models of the vocal fold mucosa," J. Acoust. Soc. Am. **130**, 1597–1605.

Ishizaka, K., and Flanagan, J. L. (**1972**). "Synthesis of voiced sounds from a two-mass model of the vocal cords," Bell Sys. Tech. J. **51**(6), 1233–1268.

Liljencrants, J. (**1991**). "A translating and rotating mass model of the vocal folds," STL/QPSR **1**, 1–18. http://www.speech.kth.se/prod/publications/files/qpsr/1991/1991_32_1_001-018.pdf (Last viewed July 12, 2016).

Mendelsohn, A. H., and Zhang, Z. (**2011**). "Phonation threshold pressure and onset frequency in a two-layer physical model of the vocal folds," J. Acoust. Soc. Am. **130**(5), 2961–2968.

Murray, P. R., and Thomson, S. L. (**2012**). "Vibratory responses of synthetic, self-oscillating vocal fold models," J. Acoust. Soc. Am. **132**, 3428–3438.

Pelorson, X., Hirschberg, A., Van Hassel, R. R., Wijnands, A. P. J., and Aurégan, Y. (**1994**). "Theoretical and experimental study of quasisteady-flow separation within the glottis during phonation. Application to a modified two-mass model," J. Acoust. Soc. Am. **96**(6), 3416–3431.

Ruty, N., Pelorson, X., Van Hirtum, A., 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., 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.

Xuan, Y., and Zhang, Z. (**2014**). "Influence of embedded fibers and an epithelium layer on glottal closure pattern in a physical vocal fold model," J. Speech Lang. Hear. Res. **57**, 416–425.

Zhang, Z. (**2016**). "Cause-effect relationship between vocal fold physiology and voice production in a three-dimensional phonation model," J. Acoust. Soc. Am. **139**(4), 1493–1507.

Zhang, Z., and Luu, T. H. (**2012**). "Asymmetric vibration in a two-layer vocal fold model with left-right stiffness asymmetry: Experiment and simulation," J. Acoust. Soc. Am. **132**, 1626–1635.

Zhang, Z., and Neubauer, J. (**2010**). "On the acoustical relevance of supraglottal flow structures to low-frequency voice production," J. Acoust. Soc. Am. **128**(6), EL378–EL383.

Zhang, Z., Neubauer, J., and Berry, D. A. (**2006**). "Aerodynamically and acoustically driven modes of vibration in a physical model of the vocal folds," J. Acoust. Soc. Am. **120**, 2841–2849.