A Reduced-Order Flow Model for Fluid–Structure Interaction Simulation of Vocal Fold Vibration

Zheng Li

Department of Mechanical Engineering,
Vanderbilt University,
2301 Vanderbilt Place,
Nashville, TN 37235-1592

Ye Chen

Department of Mechanical Engineering,
Vanderbilt University,
2301 Vanderbilt Place,
Nashville, TN 37235-1592

Siyan Chang

Department of Mechanical Engineering,
Vanderbilt University,
2301 Vanderbilt Place,
Nashville, TN 37235-1592

Haoxiang Luo¹

Department of Mechanical Engineering,
Vanderbilt University,
2301 Vanderbilt Place,
Nashville, TN 37235-1592
e-mail: haoxiang.luo@vanderbilt.edu

We present a novel reduced-order glottal airflow model that can be coupled with the three-dimensional (3D) solid mechanics model of the vocal fold tissue to simulate the fluid–structure interaction (FSI) during voice production. This type of hybrid FSI models have potential applications in the estimation of the tissue properties that are unknown due to patient variations and/or neuromuscular activities. In this work, the flow is simplified to a one-dimensional (1D) momentum equation-based model incorporating the entrance effect and energy loss in the glottis. The performance of the flow model is assessed using a simplified yet 3D vocal fold configuration. We use the immersed-boundary method-based 3D FSI simulation as a benchmark to evaluate the momentum-based model as well as the Bernoulli-based 1D flow models. The results show that the new model has significantly better performance than the Bernoulli models in terms of prediction about the vocal fold vibration frequency, amplitude, and phase delay. Furthermore, the comparison results are consistent for different medial thicknesses of the vocal fold, subglottal pressures, and tissue material behaviors, indicating that the new model has better robustness than previous reduced-order models. [DOI: 10.1115/1.4044033]

Keywords: vocal fold, fluid–structure interaction, reduced-order model, phonation

1 Introduction

Vocal fold vibration during phonation involves aerodynamic interaction of a pulsatile glottal jet and the soft vocal fold tissue stretched between the laryngeal cartilages. The unsteady airflow is responsible for activating and sustaining the vocal fold vibration, and the oscillation pattern of the vocal fold in turn modulates the airflow. This dynamic interaction determines many basic characteristics of voice [1]. An advanced computational model of the fluid–structure interaction (FSI) for the vocal fold will find useful applications in the understanding, diagnosis, and treatment of voice disorders. Many FSI models have been developed in the past with various levels of complexity. In terms of spatial setup, these models can be generally categorized depending on whether the airflow and the tissue, respectively, assume a zero-, one-, two-, or three-dimensional (3D) configuration. Within each configuration, the models can still differ significantly from one another depending on how to treat various details such as the structural tissue layers, elastic properties of the tissue, and anatomical features of the larynx. In early stages, discrete or lumped-mass systems were created to understand the onset of phonation [2–4]. In these models, the vocal fold was simplified to two or more mass blocks connected to elastic springs, and the Bernoulli equation or other simplified flow equations were used to model the airflow. Despite its simplicity, such models can capture self-induced oscillations and have been used extensively to understand the basic effects of governing parameters, e.g., the subglottal pressure and tissue stiffness, and also to investigate the characteristic behavior of normal and abnormal phonation, e.g., chaotic vibration and vocal fold polyps [3]. With the development of high-performance computing hardware and software, continuum-mechanics-based computational models have been increasingly used for vocal fold modeling. For example, both two-dimensional and 3D finite element models have been developed to simulate the vocal fold

deformation [5–9]. More recently, high-resolution simulations have been more frequently used in the FSI modeling of the vocal fold. Examples of previous works include Thomson et al. [10], Luo et al. [11,12], and Zheng et al. [13]. Using the intensive, typically parallelized computations, many of these studies have reported the unsteady vortex structures in the airflow and their interaction with the vocal fold.

As modern medical imaging technology is being advanced, internal anatomy of human bodies can be viewed with unprecedented details using noninvasive approaches such as computed tomography and magnetic resonance imaging. Such imaging modalities may provide 3D geometry of the larynx as well as the interior structure of the tissues [14–16]. The images generated by these techniques could be used to develop more sophisticated computational models that have much realistic representation of the laryngeal anatomy. Compared with the previous computational models that are based on greatly simplified geometries (even for continuum-based models), the anatomical models are a significant step closer to patient-specific and high-fidelity modeling of phonation, which is eventually needed for clinical care of voices of individual patients. Some recent works, e.g., Refs. [17] and [18], provide insights into the development toward such medical imaging-based models of the vocal fold. More details about the development and improvement of vocal fold modeling can be found in review papers of Alipour et al. [19] and Mittal et al. [20]. Only a brief summary of literature is provided here to setup the context for this study.

One issue related to patient-specific modeling is that even if a patient's anatomy could be reconstructed with high fidelity, there are still a few other modeling parameters whose values cannot be specified with accuracy, for example, the elastic properties of the tissue material that may vary from patient to patient. Even for the same patient, the effective stiffness of the tissue highly depends on neurological control of various muscle groups and consequently the adduction state of the vocal fold [21]. These uncertain parameters are important to capture the patient-specific vibration features [22]. Therefore, either ad hoc assumptions have to be made, or some parameter identification approach must be used to

¹Corresponding author.

Manuscript received December 31, 2018; final manuscript received May 30, 2019; published online October 7, 2019. Assoc. Editor: C. Alberto Figueroa.

estimate those parameters. It will be too expensive to perform parameter identification using 3D FSI models due to their high computational cost. One possible method is thus to use a reduced-order model to determine those unknown material properties, which could then be used to enhance fidelity of the 3D models.

To construct a reduced-order FSI model for vocal fold vibration, it may be appropriate to simplify the description of the flow rather than the description of the tissue mechanics, especially when accurate capture of the vibration characteristics is desirable. This is because the vocal fold deformation is three-dimensional and can be complicated, requiring at least a 3D model representation; furthermore, 3D simulation of the turbulent glottal flow is typically much more expensive than 3D simulation of the tissue deformation, and reducing flow simulation can largely lower the overall computational cost. For such a purpose, the Bernoulli equation has been most widely used in the past to describe the pressure and velocity of the glottal flow. A comparison of the Bernoulli equation with the Navier–Stokes equation was studied by Decker and Thomson [23], who used a two-dimensional setup, assuming either steady flow or FSI, to assess the accuracy of the Bernoulli principle. Their comparison showed that all Bernoulli-based models result in similar predictions of the mean intraglottal pressure, maximum orifice area, and vibration frequency; however, those predictions rely on the heuristic specification of flow separation location in the Bernoulli models, and the location is quite different from that obtained from the simulation based on the Navier–Stokes equation.

In a previous work [21], the authors coupled an anatomical vocal fold model that was based on the magnetic resonance imaging scan of the rabbit larynx with a Bernoulli-based flow model to perform fast FSI simulations. Their flow model was calibrated a priori using 3D flow simulation of the same larynx, in which the 3D flow data were used to setup the curved flow path along the airway for the one-dimensional (1D) model and also to specify the proper location of flow separation. Using a hybrid FSI model of the 1D flow and the 3D tissue and trying to match the model prediction with the experimental measurement of the vibration, they estimated the elastic constants of the vocal fold of each subject. Next, with the material properties identified for individual samples, the updated 3D FSI simulations were able to capture the specific vibration characteristics for each subject used in the study. In a later study, the same authors compared the hybrid FSI model with 3D FSI by using a simplified vocal fold geometry to more thoroughly assess the performance of such Bernoulli-based 1D flow models [24]. They found that that model prediction can be sensitive to the subjective specification of the separation location; in some case where the medial thickness of the vocal fold is small, the hybrid FSI model leads to a significantly different vibration mode of the vocal fold than the 3D FSI model.

From these previous studies, it is clear that the Bernoulli equation has serious limitation in its capability to satisfactorily compute the pressure in the flow for a given geometrical configuration of the glottis. To address this limitation, in this study we adopt a 1D momentum equation-based flow model that was originally developed to solve separated flow in the collapsible tube [25,26]. This model has been recently introduced for vocal fold modeling [27]. It includes the viscous effect as well as the pressure loss associated with flow separation that is typically encountered in a divergent channel. A recent review [28] discussed more applications of this model as well as its advantages and limitations. One particular limitation has to do with the significant viscous dissipation upstream the narrowest section of the channel that is not accounted for by the model [28,29]. In this study, we incorporate an empirical function in the 1D flow model to account for the entrance effect that takes place as the flow enters the glottal gap. By introducing a correction coefficient to the cross section area, we exclude the boundary layer at the glottal entrance and limit the momentum equation to the core of the flow. Thus, our model is less affected by the viscous dissipation upstream the narrowest section that was discussed previously [29]. The objective of this

study is to describe this new flow model and assess its effectiveness in predicting vocal fold vibration when coupled with 3D tissue mechanics in a hybrid FSI model. To do so, we use an idealized yet 3D vocal fold geometry and perform 3D FSI simulations as the benchmark results for this setup. In order to test its robustness, we vary the medial thickness of the vocal fold (and thus the shape of the glottis during vibration), the subglottal pressure, as well as the material model of the tissue, and we also include the Bernoulli-based models as supplementary references. In Sec. 2, we will first describe this 1D flow model and then introduce the validation setup. The results and discussions will be given in Sec. 3. Finally, we will provide concluding remarks in Sec. 4.

2 Method

2.1 The One-Dimensional Viscous Flow Model. Figure 1 illustrates generally how an incompressible flow behaves while going through a glottis-like gap. Along the flow, the pressure first decreases to a minimal level at the narrowest section and then increases in the divergent section. However, due to the momentum losses to viscous resistance and velocity fluctuations, especially the losses associated with flow separation, the pressure at the exit does not recover to its full level at an entrance location with the same cross-sectional area as the exit (i.e., x_1 and x_2 in Fig. 1), and the corresponding total pressure experiences significant loss after the minimum section as illustrated in Fig. 1. The amount of loss in the form of pressure decrease depends on the specific geometry as well as the Reynolds number. To describe the flow without resorting to the 3D Navier–Stokes equation, we use the 1D model developed by Cancelli and Pedley for flow in a collapsible tube [25]. In this model, the authors considered energy loss in a tube with increasing cross section, where the flow may experience significant viscous resistance as well as separation. The model consists of the unsteady continuity and momentum equations as follows:

$$\frac{\partial A}{\partial t} + \frac{\partial Au}{\partial x} = 0 \quad (1)$$

$$\rho \frac{\partial u}{\partial t} + \rho u \frac{\partial u}{\partial x} + \frac{\partial p}{\partial x} - \tau \frac{s}{A} = 0 \quad (2)$$

$$\tau = \tau_{\text{fric}} + \tau_{\chi} \quad (3)$$

where ρ , u , and p are, respectively, the density, velocity, and pressure, s is the perimeter around the cross-sectional area, A , and τ is the total stress that combines the viscous stress, τ_{fric} , and an

Fig. 1 Illustration of pressure and total pressure along the center of the flow through a glottis-like gap

additional stress, τ_χ , that causes loss of kinetic energy due to separation. The viscous term can be estimated based on fully developed flow in a tube of constant cross section, i.e., $\tau_{\text{fric}} = -2\mu(s/A)u$. The loss of kinetic energy term can be modeled according to

$$\tau_\chi = \frac{A}{s} (1 - \chi) \rho u \frac{\partial u}{\partial x} \quad (4)$$

where $0 \leq \chi \leq 1$ is a constant representing pressure recovery. Plugging Eq. (4) into Eq. (2), one obtains a term like $(1 - \chi) \rho u (\partial u / \partial x)$, which cancels part of the advection term and leaves only $\chi \rho u (\partial u / \partial x)$ in the momentum equation. Thus, $\chi = 1$ means there is no separation loss, and $\chi = 0$ means all kinetic energy is lost and there would be no pressure recovery. In Ref. [25], the sign of the production of the velocity and pressure gradient is used to determine the choice of χ . Here, we simply set χ according to converging or diverging section, i.e.,

$$\chi = \begin{cases} 1, & \text{before the minimum section} \\ \chi_{\min}, & \text{after the minimum section} \end{cases} \quad (5)$$

where $\chi_{\min}$ is the minimum value of χ . With this definition of χ , the pressure in the converging section can be fully converted to kinetic energy—that is, the Bernoulli equation is satisfied (or, the unsteady Bernoulli equation is satisfied if the unsteady term is significant); in addition, the pressure loss in the diverging section is accounted for.

Given a specific geometrical configuration of the glottis, Eqs. (1)–(3) can be solved straightforwardly using a numerical method such as the finite difference method. Similar to the 3D flow, the boundary conditions of the 1D model includes $p = P_{\text{sub}}$ at an upstream location $x = x_0$, $p = P_e$ at the glottal exit $x = x_2$, and the velocity u has zero derivative at $x = x_0$. Here, we set x_0 to be the location where the cross section of the flow domain starts to change. P_e is generally close to the domain outlet pressure, P_{out} , but it may vary a little depending on the specific geometry of the vocal fold. We will choose its constant value based on the 3D FSI results.

While Eq. (2) is a reasonable description of the flow momentum along the centerline of the converging-diverging channel, the continuity equation, Eq. (1), assumes that the velocity profile in a cross section is nearly uniform. In reality, when the flow enters the glottal gap, the sudden narrowing of the cross section at the entrance causes a significant vena contracta effect. That is, the flow under inertia is more focused to the center as shown in Fig. 2, rather than following the exact shape of the channel. The vena contracta effect is more influenced by the high curvature of the entrance and may still exist even if the channel is straight or converging after the entrance. As a result, if the actual cross section area A is used for flow continuity, significant error could be introduced to the intraglottal velocity. As it is shown later in Sec. 3, without the entrance effect, the velocity at the minimum section may be over-estimated, leading to an erroneous prediction of strong negative pressure at the location.

To account for such an entrance effect, we define the effective cross section area, A , and the actual cross section area, A_0 , and we introduce their ratio, α

$$\alpha(x) = \frac{A(x)}{A_0(x)} \quad (6)$$

when considering the mass conservation along the channel. In this study, the actual area A_0 is calculated directly from the instantaneous 3D vocal fold geometry, and $\alpha(x)$ will be determined empirically from the 3D FSI simulation as discussed later. With the function $\alpha(x)$ determined a priori, the effective area A can be calculated from Eq. (6) and is then used in the continuity equation, Eq. (1).

Fig. 2 Schematic of airflow enters the glottis where the sudden change of geometry at the inlet introduces a vena contracta effect. A_0 is the actual cross-sectional area, and A is the effective area.

2.2 Setup of the Three-Dimensional Fluid–Structure Interaction Model. We use a simple geometrical setup in the full 3D FSI simulation to provide benchmark for the reduced-order FSI model. The setup is illustrated in Fig. 3, where a rectangular box represents the airway and its walls are assumed to be rigid. The total length of the rectangular box is 12 cm and the vocal fold starts from 2 cm from the inlet. The flow is assumed to be incompressible and is governed by the viscous Navier–Stokes equation in the full 3D model. A pair of vocal fold is placed symmetrically in the box with length $L = 20$ mm, width $W = 13$ mm, and depth $D = 10$ mm. The cross section of the vocal fold is uniform and has roughly a trapezoidal shape. The initial glottal gap is at 0.4 mm. The details of the cross section were described in Refs. [11,24]. It was found previously that the medial surface thickness T significantly affects the configuration of the glottal shape during vibration and thus the intraglottal flow [24]. Therefore, two different medial surface thicknesses are used here, a large one with $T = 3.50$ mm and a small one with $T = 1.75$ mm.

For the boundary conditions, the left, right, anterior, and posterior surfaces of the vocal fold, i.e., all the sides attached to the

Fig. 3 The vocal fold model and computational domain used in the study for 3D FSI simulation

Fig. 4 Vortices in the supraglottal region obtained from 3D FSI simulations for ((a) and (b)) small medial thickness $T = 1.75$ mm and ((c) and (d)) large medial thickness $T = 3.5$ mm. The isosurface is defined using the λ -criterion with the contour level at 20 s^{-1} . ((a) and (c)) Opening phase, and ((b) and (d)) closing phase.

rectangular box, are treated as fixed surfaces, while the other surfaces in contact with airflow are free to move. During vibration, the vocal fold is allowed to have a small gap of 0.2 mm for a minimal amount of flow to go through. The airflow is driven from left (inlet) to right (outlet) by a constant pressure drop between the subglottal pressure P_{sub} at the inlet and reference supraglottal pressure $P_{\text{out}} = 0 \text{ kPa}$ at the outlet. The pressure drop is around 1 kPa.

As the current study is focused on the reduced-order modeling for the glottal flow, the tissue model of the vocal fold is not of primary concern. Thus, despite that the real tissue is anisotropic and has a multilayer structure, the vocal fold here is assumed to be isotropic and homogenous. Nevertheless, we adopt two different constitutive laws for the tissue material, the Saint-Venant model and a hyperelastic, two-parameter Mooney–Rivlin model. In both material models, nonlinear strains have been incorporated. More detail of the Saint-Venant model can be found in our group's previous work [30]. The Mooney–Rivlin model is one of popular models for representing large deformations of soft tissues. The strain energy density function for this model is given as

$$W = \alpha_{10}(\bar{I}_1 - 3) + \alpha_{01}(\bar{I}_2 - 3) + K/2(J - 1)^2 \quad (7)$$

where K represents the bulk modulus, α_{10} and α_{01} are the material constants, and $J = \det(F)$ with F stands for the deformation gradient. In addition, $\bar{I}_1$ and $\bar{I}_2$ are the invariants based on J and the principal stretches of the deformation gradient. Further detail of this model for the vocal fold can also be found in our group's previous work [30].

In both tissue models, the material density is $\rho_s = 1040 \text{ kg/m}^3$ and mass damping is 0.05 s^{-1} . In the Saint-Venant model, Young's modulus is set to be $E = 15 \text{ kPa}$, and Poisson's ratio is $\nu = 0.475$. For the Mooney–Rivlin model, $\alpha_{10} = 2.29 \text{ kPa}$ and $\alpha_{01} = 0.25 \text{ kPa}$ are used in the Mooney–Rivlin model to match the specified stiffness of the Saint-Venant model at small strain. The air density is $\rho = 1.13 \text{ kg/m}^3$. Thus, the characteristic intraglottal velocity is $V = \sqrt{2(P_{\text{sub}} - P_{\text{out}})/\rho} = 42.1 \text{ m/s}$. We define the jet Reynolds number using $\text{Re}_J = \rho V d / \mu$, where $d \sim 1 \text{ mm}$ is the characteristic glottal gap during opening phase and μ is the air viscosity. In the current study, we set $\text{Re}_J = 210$. If the channel height is used in the definition of the Reynolds number, we have $\text{Re} = 4200$.

2.3 Numerical Method and Mesh Refinement Study for Three-Dimensional Simulation. A finite element method is used to solve the tissue deformation [30]. The vocal fold is meshed

with approximately 18,000 20-nodes hexahedral elements and 80,000 vertex nodes. No-slip and no-penetration wall conditions are specified for all flow domain boundaries except the inlet and outlet, where either the inlet or outlet pressure is applied and the velocity is assumed to have a zero normal derivative. An immersed-boundary method is adopted for the flow simulation [30–32]. A nonuniform Cartesian grid with $320 \times 98 \times 72$ points is used to discretize the flow domain. The subglottal pressure is set to be $P_{\text{sub}} = 0.75, 1.0$ or 1.25 kPa , which is within the range of the onset pressure for normal human phonation. The time-step size $\Delta t = 0.0025 \text{ ms}$ is used for the FSI simulation, which leads to about 4000 steps for each vibration cycle that is approximately at 100 Hz.

A grid convergence study is done for the case with medial thickness $T = 3.5 \text{ mm}$ and $P_{\text{sub}} = 1.0 \text{ kPa}$, while the Mooney–Rivlin model is employed for the material behavior. The nonuniform Cartesian grid is doubled in the region around the vocal fold and also in the flow region immediately downstream the vocal fold, and the time-step size is reduced to 0.002 ms. From the results, a relative difference of 2.0% is observed between the baseline mesh and the fine mesh for the vibration frequency, 4.5% for the vibration amplitude, and 3.9% for the phase delay between the inferior or superior points on the medial surface. As seen later, these errors are much smaller when compared to the differences between the reduced-order model and the full 3D model. Therefore, the baseline mesh is considered to be acceptable for further investigations in this work.

3 Results and Discussions

3.1 Results From Three-Dimensional Fluid–Structure Interaction Simulations. The 3D FSI simulations provide full-flow field data, including the velocity and pressure, to benchmark the reduced-order FSI model. Figure 4 shows a snapshot of the vortex structures in the flow for $P_{\text{sub}} = 1.0 \text{ kPa}$ and the Mooney–Rivlin model. These vortex structures are unsteady and generally follow the pulsatile jet to go downstream while interacting with one another and eventually dissipating. Fortunately, even though these vortices correspond to some degree of pressure fluctuations in the flow, they do not drastically change the pressure in the supraglottal region. Thus, we could assume a constant supraglottal pressure in the reduced-order model for P_e . From the 3D simulations, we obtain that the pressure is around $P_e = 0 \text{ Pa}$ for $T = 3.5 \text{ mm}$ and $P_e = -100 \text{ Pa}$ for $T = 1.75 \text{ mm}$. In both

Fig. 5 Streamline plot and contours of the streamwise velocity within the $z = 0$ slice for $T = 1.75$ mm

$T = 1.75$ mm and 3.5 mm cases, the vocal fold exhibits a second-mode like vibration, where the oscillation resembles the second eigenmode of the current vocal fold structure and is primarily in the lateral or y -direction [12]. The amplitude of vibration is $d = 0.89$ mm for $T = 1.75$ mm and $d = 1.22$ mm for $T = 3.5$ mm. Here, the amplitude of vibration d is defined as the maximum y -displacement of the medial surface at the glottal exit (the gap width at the exit is thus $2d$). In the case of $T = 1.75$ mm, the glottis does not close completely (or reach the minimal gap) during closing phase, so flow is continuous despite being oscillatory. In the case of $T = 3.5$ mm, the glottis reaches the minimal gap and the glottal channel is significantly longer than the small thickness case; thus, the jet is nearly completely cut off during closing phase, while during opening phase the jet has shorter penetration downstream.

We use the velocity field data from the 3D simulations to calculate the area correction coefficient, $\alpha(x)$ which is defined in Eq. (6). As shown in Fig. 5, the velocity profile across the glottis is nonuniform, and as stated in Sec. 2.1, introducing such area correction would improve accuracy of the 1D mass conservation equation. This coefficient is calculated by computing the ratio between the average streamwise velocity in a cross section, u_{avg} , and the maximum streamwise velocity of the cross section, u_{max} , i.e.,

$$\alpha(x) = \frac{u_{\text{avg}}}{u_{\text{max}}} = \frac{\frac{1}{A_0} \int u \, dS}{u_{\text{max}}} \quad (8)$$

If the flow inside the glottis is symmetric, then u_{max} is also the centerline velocity. During vibration, $\alpha(x)$ varies somewhat depending on the instantaneous shape of the glottal gap. However, the coefficient is around 1 at the inlet and 0.75 at the exit from the simulation results. The inlet location x_a and the exit location x_b are marked out in Fig. 5. We use a quadratic function to approximate α and demand that it has zero derivative at x_b . The result gives the following expression for $\alpha(x)$:

$$\alpha(x) = \frac{(x - x_b)^2}{4(x_a - x_b)^2} + 0.75 \quad \text{for } x_a \leq x \leq x_b \quad (9)$$

As shown later, it turns out that the same function for α can be used for different cases of the medial thickness, subglottal pressure, and tissue behavior.

3.2 Comparison of the Reduced-Order Models. We consider two setups of the momentum-based reduced-order FSI model. In the first, the flow is described by Eqs. (1)–(3) but no entrance effect is incorporated (i.e., $\alpha(x) = 1$); this model is denoted by M1. The second model includes the entrance effect for the flow using the area correction coefficient given by Eq. (8), and this model is denoted by M2. In addition to these two models, we consider two Bernoulli-based flow models, i.e., removing the last term in Eq. (2) that involves τ . These two models are denoted by B1 and B2, respectively. For B1, we set the location of flow separation always at the minimum cross section area within the glottis, at which and further downstream the pressure is set to $p = P_e$. For B2, we set the separation location always at the glottal exit with the same pressure specification. These four models are summarized in Table 1. For all reduced-order models, the same vocal fold model as that in the 3D model is used for FSI simulation. In each case, data collection is done after sustained vibration is established.

Figure 6 shows the vibration pattern of the vocal fold within the midplane $z = 0$ that is predicted by different FSI models. In this case, $T = 1.75$ mm, $P_{\text{sub}} = 1.0$ kPa, and the Mooney–Rivlin model are used. Furthermore, for all cases with the smaller medial thickness, $\chi_{\text{min}} = 0.5$ has been adopted; and for all cases with the larger medial thickness, $\chi_{\text{min}} = 0.2$ has been adopted to account for more pressure loss in this case. The original rest shape of the vocal fold is included in the figure as a reference. The vibration pattern from

Table 1 Four reduced-order flow models used for comparison

Model	Governing equation	Entrance effect	Separation location
B1	Bernoulli equation	No	Minimal area
B2	Bernoulli equation	No	Glottal exit
M1	1D momentum equation	No	—
M2	1D momentum equation	Yes	—

Fig. 6 Vibration pattern in the midplane $z = 0$ obtained by different FSI models for the vocal fold with small medial thickness: (a) 3D FSI simulation, (b) B1, (c) B2, (d) M1, and (e) M2. The dashed lines indicate the original shape of the vocal fold.

Table 2 Reduced-order models compared with 3D FSI for vocal fold in terms of vibration frequency f , amplitude d , and phase delay ϕ

T (mm)	Model	f (Hz)	Difference	d (mm)	Difference	ϕ (deg)	Difference (deg)
1.75	3D FSI	132	—	0.89	—	−19	—
	B1	78	40.1%	1.27	42.8%	−9	10
	B2	140	7.7%	1.44	62.0%	10	29
	M1	144	10.8%	1.38	54.9%	21	40
	M2	144	7.7%	0.78	12.5%	−15	4
3.50	3D FSI	140	—	1.22	—	157	—
	B1	136	2.3%	1.31	7.4%	98	59
	B2	—	—	—	—	—	—
	M1	144	2.9%	1.25	2.5%	186	29
	M2	144	2.9%	1.20	1.4%	161	4

the 3D simulation indicates a second-mode like vibration, where the vocal fold oscillation is primarily in the lateral or y -direction. In contrast, model B1 produces a vibration pattern that resembles the first eigenmode of the vocal fold, where the vocal fold oscillates primarily in the streamwise or x -direction and the eigen frequency is at 73 Hz for the current vocal fold structure. Even though this oscillation mode also leads to opening and closing of the glottis, the vibration frequency is significantly lower than the

second eigenmode whose frequency is at 126 Hz for the current vocal fold structure. The other three reduced-order models, B2, M1, and M2 all produce a second-mode like vibration. However, the amplitude of the vibration is different among these models. For B2 and M1, the vibration amplitude is significantly greater than that of the 3D FSI model. Detailed comparisons are shown in Table 2 for the vibration frequency f , amplitude d , and phase delay ϕ between the glottal inlet and exit (i.e., the phase difference

Fig. 7 Vibration pattern in the midplane $z = 0$ obtained by different FSI models for the vocal fold with large medial thickness: (a) 3D FSI simulation, (b) B1, (c) M1, and (d) M2. The dashed lines indicate the original shape of the vocal fold.

between points 1 and 2 indicated in Fig. 5). The data in the table show that for all these quantities, M2 produces the closest result to that of the 3D FSI model.

For the large medial thickness case of $T = 3.5$ mm, the comparison of the vibration pattern is shown in Fig. 7, where $P_{\text{sub}} = 1.0$ kPa and the Mooney–Rivlin model are used. In this case, B1, M1, and M2 achieved the second-mode dominant vibration and are similar to the 3D FSI model. B2 model did not reach a steady vibration pattern, and instead the deformation becomes overly large, causing the simulation to diverge. Further examining the patterns shown in this figure, we see that in B1 the glottis switches between a divergent shape and a straight shape in a vibration cycle, while in both M1 and M2, the glottis has a convergent shape during opening, a divergent shape during closing, and a straight shape in between, i.e., a similar sequence of deformation to that predicted by the 3D FSI model. The quantitative comparison in Table 1 shows that even though M1 and M2 produce close results to the 3D FSI, M2 is better in terms of the vibration amplitude and the phase delay.

3.3 Comparison of Pressure Distribution. To further study the difference among the four reduced-order models, we compare the pressure distribution along the flow, especially in the region within the glottis, since the pressure directly causes the vocal fold displacement and provides the mechanism for sustained vibration. For the comparison, we use the vocal fold deformation obtained from the 3D FSI simulation and calculate the pressure distribution using the four 1D flow models summarized in Table 1. This way, the glottal configuration is the same across different models, and we can focus on the pressure calculation given by these flow models.

Figure 8 shows the pressure distribution for the case of $T = 1.75$ mm at both closing and opening phases. At the closing

phase, M2 has the closest distribution to the 3D FSI among the four reduced-order models. For B1, B2, and M1, the pressure drops too fast as the flow enters the glottis. This is because for M2, introducing the entrance effect moves the minimal cross-sectional area further downstream within the glottis and thus improves the pressure prediction near the glottal entrance. At the opening phase, B1 and M2 both produce reasonable pressure distribution close to the 3D FSI result, though M2 also generates a negative pressure zone in the glottis like the 3D FSI model. Both B2 and M1 have a presence of strong negative pressure, especially for B2 since its separation location is set at the glottal exit and thus over-predicts the negative pressure at the narrowest cross section. Overall, we see that including the entrance effect allows M2 to have better performance than the other 1D models.

Figure 9 shows the pressure distribution for the case of $T = 3.5$ mm at opening, maximum opening, and closing phases, where the glottal shape is convergent, straight, and divergent, respectively. For the closing phase in (a), the glottal exit has the minimal cross-sectional area. Consequently, B1 and B2 have the same separation point and thus the identical pressure prediction. M1 also produces a very close result to those of B1 and B2 since no pressure loss in the divergent section is involved and the frictional stress is small. M2 gives better pressure prediction near the entrance (between $x = -0.4$ and -0.2 cm). After $x = -0.1$ cm, all reduced models have similar results.

When the glottis is straight, Fig. 9(b) shows that M2 again has the best prediction, and the pressure drops too fast near the entrance for the other 1D models. For the closing phase with a divergent glottis, Fig. 9(c) shows that B1, M1, and M2 all have significant error with low pressure prediction near the entrance. B2 predicts an overly low pressure; thus, in the FSI simulation, this model broke down and did not reach a converging result. Similar to the 3D FSI, M2 also produces a negative pressure zone with proper magnitude, despite that the location of the zone is

Fig. 8 Comparison of pressure distribution for different models where $T = 1.75$ mm: (a) closing phase and (b) opening phase

Fig. 9 Comparison of pressure distribution for different models where $T = 3.5$ mm: (a) opening phase, (b) maximum opening, and (c) closing phase

Fig. 10 Comparison of B1, M2, and 3D FSI at different subglottal pressures for $T = 1.75$ mm: (a) vibration frequency, (b) amplitude, and (c) phase delay

slightly upstream in comparison with the 3D FSI. This negative pressure zone is instrumental for the closure of the glottis.

3.4 Effects of the Subglottal Pressure and Tissue Model. To further assess the performance of model M2, we vary the

subglottal pressure, so that $P_{\text{sub}} = 0.75, 1.0$, or 1.25 kPa, and we repeat the comparison of M2 with the 3D FSI model. In this study, only B1 is selected as a reference since B2 may not produce a converged result. The other parameters remain the same for this comparison. Figure 10 shows the comparison of the vibration

Fig. 11 Comparison of B1, M2, and 3D FSI at different subglottal pressures for $T = 3.5$ mm: (a) vibration frequency, (b) amplitude, and (c) phase delay

Table 3 Comparison of the reduced-order models with the 3D FSI using a Saint-Venant material model for the vocal fold tissue

T (mm)	Model	f (Hz)	Difference	d (mm)	Difference	ϕ (deg)	Difference (deg)
1.75	3D FSI	126	—	1.07	—	−9	—
	B1	72	42.9%	1.35	26.2%	−5	4
	M2	128	1.6%	1.05	1.9%	5	14
3.50	3D FSI	133	—	0.97	—	76	—
	B1	128	3.8%	1.25	28.9%	62	14
	M2	128	3.8%	1.06	9.3%	83	7

frequency, amplitude, and phase delay for $T = 1.75$ mm as P_{sub} is varied. The 3D FSI results show that as the subglottal pressure is raised, the vibration frequency only slightly increases; the vibration amplitude more than doubled; and the phase delay remains nearly constant. For all three subglottal pressure levels, the results from M2 agree well with the 3D FSI, while the B1 produces a different vibration mode that has a much lower frequency.

Figure 11 shows the comparison for $T = 3.5$ mm under different subglottal pressures. In this case, the Bernoulli-based model, B1, predicts the correct vibration mode whose frequency and amplitude are close to those of the 3D FSI; however, the momentum-based model, M2, has clearly better performance in that its frequency, amplitude, and phase delay all have visibly better agreement with those of the 3D FSI. These results indicate that the performance of model M2 is not significantly influenced by the subglottal pressure.

All the results presented here have been based on a Mooney–Rivlin model for the material behavior of the vocal fold tissue. To further expand the study and ensure that the new model is insensitive to the choice of the material parameters of the vocal fold, we used a Saint-Venant tissue model to repeat both the 3D and reduced-order FSI simulation. $P_{sub} = 1.0$ kPa, and the same two medial thickness values are used. The results are shown in Table 3 for B1, M2, and 3D FSI. The data show that M2 again has better agreement with the 3D FSI in terms of vibration frequency, amplitude, and phase delay for both $T = 1.75$ mm and 3.5 mm. Therefore, the choice of the material model of the vocal fold does not affect the performance of model M2 for the FSI simulation.

3.5 Further Discussions. Unlike the Bernoulli-based models, the 1D flow model introduced in this work does not need an explicit specification of the separation point in the glottis. Instead, partial loss of the kinetic energy in the divergent section is considered in the new model. This consideration allows for incomplete pressure recovery in the section, including downstream the separation point. Thus, the intraglottal pressure predicted by the model goes through a much smoother transition from the glottal entrance to the exit, a feature that the actual pressure distribution should

have as shown by the 3D simulation results in Sec. 3.4. We point out that even though the energy consideration in the present model is different from boundary layer separation consideration in the Bernoulli model, the two perspectives are still largely congruent with each other. This is because the boundary layer separation leads to significant mixing of the separated flow and is the primary reason for energy loss in the flow.

The introduction of a correction to the cross-sectional area addresses the effect of the geometry at the glottal entrance and thus further improves the pressure prediction by the new model. Although the numerical solution to this model requires an iteration process, the computational cost of the simple governing partial difference equation is minimal compared to that of solving the 3D Navier–Stokes equation or that of solving the 3D solid mechanics of the tissue deformation. Finally, the present 1D flow model can be reduced to a Bernoulli model by setting $\chi = 1$ before an assumed separation point, i.e., allowing for pressure recovery prior to flow separation, and also by setting $\alpha = 1$, i.e., ignoring the entrance effect. On the other hand, simply modifying the location of separation in the Bernoulli model does not give the same result as the current model. Therefore, the current approach offers more flexibility than the Bernoulli-based approach.

It should be pointed out that the current flow model has several limitations. Like the Bernoulli models, the current model reduces the glottal flow to 1D, thus drastically simplifying the flow behavior. Only a constant value of the pressure recovery coefficient χ is assumed in the model for the divergent section. This treatment ignores the difference before and after the separation point that may be located somewhere between the minimum cross section and the glottal exit. As a result, it should be anticipated that this model does not always provide significantly better accuracy than the Bernoulli-based models even though it generally has more advantages. The current model contains a free parameter, χ , and a function, $\alpha(x)$, which need to be set properly. The former depends on channel length of the glottis, as shown by the effect of the medial thickness in the current study, and also likely by the degree of the divergent angle if the angle is large; the latter should be more influenced by the curvature of the vocal fold surface near the glottal entrance as well as the Reynolds number. In the current

study, $\alpha(x)$ is empirically determined using 3D simulation data from a specific vocal fold model setup. In the future, it would be desirable if a more general form of this function is determined for broader situations.

4 Conclusions

We have introduced a new one-dimensional flow model for the glottal flow and for the FSI simulation of vocal fold vibration. This model is based on the reduced momentum equation with the entrance effect and the energy loss/partial pressure recovery in the divergent section included. We used a simple vocal fold configuration and performed 3D FSI simulations to assess the performance of the reduced-order FSI model. Two different medial thicknesses, three subglottal pressures, and two material behaviors were considered in the study. The results show that after incorporating the entrance effect, the momentum-based flow model provides significantly more accurate predictions of the vibration characteristics than the Bernoulli-based models as well as the momentum-based model that does not incorporate the entrance effect. Therefore, the new model offers a useful approach in the applications of the reduced-order FSI model for the vocal fold such as parameter identification.

Acknowledgment

This research was supported by an NIH grant 1R01DC016236-01A1 from the National Institute of Deafness and Other Communication Disorders (NIDCD). ZL, YC, and HL designed the reduced-order model; YC developed the 1D flow code and performed the 3D FSI simulations; ZL performed most of the study and analyses; SC contributed to the study by setting up the FSI models and performing initial simulations; ZL and HL did most of the manuscript writing, but all authors have made contributions.

Funding Data

- NIH (Grant No. 1R01DC016236-01A1, Funder ID: 10.13039/1000000009).

References

- [1] Titze, I. R., 1994, *Principles of Voice Production*, Prentice-Hall, Englewood Cliffs, NJ.
- [2] de Vries, M. P., Schutte, H. K., Veldman, A. E. P., and Verkerke, G. J., 2002, "Glottal Flow Through a Two-Mass Model: Comparison of Navier-Stokes Solutions With Simplified Models," *J. Acoust. Soc. Am.*, **111**(4), pp. 1847–1853.
- [3] Jiang, J. J., and Zhang, Y., 2002, "Chaotic Vibration Induced by Turbulent Noise in a Two-Mass Model of Vocal Folds," *J. Acoust. Soc. Am.*, **112**(5), pp. 2127–2133.
- [4] Erath, B. D., Zañartu, M., Stewart, K. C., Plesniak, M. W., Sommer, D. E., and Peterson, S. D., 2013, "A Review of Lumped-Element Models of Voiced Speech," *Speech Commun.*, **55**(5), pp. 667–690.
- [5] Alipour, F., Berry, D. A., and Titze, I. R., 2000, "A Finite-Element Model of Vocal-Fold Vibration," *J. Acoust. Soc. Am.*, **108**(6), pp. 3003–3012.
- [6] Hunter, E. J., Titze, I. R., and Alipour, F., 2004, "A Three-Dimensional Model of Vocal Fold Abduction/Adduction," *J. Acoust. Soc. Am.*, **115**(4), pp. 1747–1759.
- [7] Cook, D. D., and Mongeau, L., 2007, "Sensitivity of a Continuum Vocal Fold Model to Geometric Parameters, Constraints, and Boundary Conditions," *J. Acoust. Soc. Am.*, **121**(4), pp. 2247–2253.

- [8] Shurtz, T. E., and Thomson, S. L., 2013, "Influence of Numerical Model Decisions on the Flow-Induced Vibration of a Computational Vocal Fold Model," *Comput. Struct.*, **122**, pp. 44–54.
- [9] Zhang, Z., 2017, "Effect of Vocal Fold Stiffness on Voice Production in a Three-Dimensional Body-Cover Phonation Model," *J. Acoust. Soc. Am.*, **142**(4), pp. 2311–2321.
- [10] Thomson, S. L., Mongeau, L., and Frankel, S. H., 2005, "Aerodynamic Transfer of Energy to the Vocal Folds," *J. Acoust. Soc. Am.*, **118**(3), pp. 1689–1700.
- [11] Luo, H., Mittal, R., Zheng, X., Bielamowicz, S. A., Walsh, R. J., and Hahn, J. K., 2008, "An Immersed-Boundary Method for Flow-Structure Interaction in Biological Systems With Application to Phonation," *J. Comput. Phys.*, **227**(22), pp. 9303–9332.
- [12] Luo, H., Mittal, R., and Bielamowicz, S. A., 2009, "Analysis of Flow-Structure Interaction in the Larynx During Phonation Using an Immersed-Boundary Method," *J. Acoust. Soc. Am.*, **126**(2), pp. 816–824.
- [13] Zheng, X., Xue, Q., Mittal, R., and Bielamowicz, S., 2010, "A Coupled Sharp-Interface Immersed Boundary-Finite-Element Method for Flow-Structure Interaction With Application to Human Phonation," *ASME J. Biomech. Eng.*, **132**(11), p. 111003.
- [14] Swanson, E. R., Abdollahian, D., Ohno, T., Ge, P., Zeale, D. L., and Rousseau, B., 2009, "Characterization of Raised Phonation in an Evoked Rabbit Phonation Model," *Laryngoscope*, **119**(7), pp. 1439–1443.
- [15] Swanson, E. R., Ohno, T., Abdollahian, D., Garrett, C. G., and Rousseau, B., 2010, "Effects of Raised-Intensity Phonation on Inflammatory Mediator Gene Expression in Normal Rabbit Vocal Fold," *Otolaryngol. Head Neck Surg.*, **143**(4), pp. 567–572.
- [16] Novaleski, C. K., Kojima, T., Chang, S., Luo, H., Valenzuela, C. V., and Rousseau, B., 2016, "Nonstimulated Rabbit Phonation Model: Cricothyroid Approximation," *Laryngoscope*, **126**(7), pp. 1589–1594.
- [17] Xue, Q., Zheng, X., Mittal, R., and Bielamowicz, S., 2014, "Subject-Specific Computational Modeling of Human Phonation," *J. Acoust. Soc. Am.*, **135**(3), pp. 1445–1456.
- [18] Mittal, R., Zheng, X., Bhardwaj, R., Seo, J. H., Xue, Q., and Bielamowicz, S., 2011, "Toward a Simulation-Based Tool for the Treatment of Vocal Fold Paralysis," *Front. Physiol.*, **2**, pp. 1–15.
- [19] Alipour, F., Brucker, C., Cook, D., Gommel, A., Kaltenbacher, M., Mattheus, W., Mongeau, L., Nauman, E., Schwarze, R., Tokuda, I., and Zorner, S., 2011, "Mathematical Models and Numerical Schemes for the Simulation of Human Phonation," *Curr. Bioinf.*, **6**(3), pp. 323–343.
- [20] Mittal, R., Erath, B. D., and Plesniak, M. W., 2013, "Fluid Dynamics of Human Phonation and Speech," *Annu. Rev. Fluid Mech.*, **45**(1), pp. 437–467.
- [21] Chang, S., Novaleski, C. K., Kojima, T., Mizuta, M., Luo, H., and Rousseau, B., 2015, "Subject-Specific Computational Modeling of Evoked Rabbit Phonation," *ASME J. Biomech. Eng.*, **138**(1), p. 011005.
- [22] Mendelsohn, A. H., 2011, "Phonation Threshold Pressure and Onset Frequency in a Two-Layer Physical Model of the Vocal Folds," *J. Acoust. Soc. Am.*, **130**(5), pp. 2961–2968.
- [23] Decker, G. Z., and Thomson, S. L., 2007, "Computational Simulations of Vocal Fold Vibration: Bernoulli Versus Navier-Stokes," *J. Voice*, **21**(3), pp. 273–284.
- [24] Chang, S., 2016, "Computational Fluid-Structure Interaction for Vocal Fold Modeling," Ph.D. thesis, Vanderbilt University, Nashville, TN.
- [25] Cancelli, C., and Pedley, T., 1985, "A Separated-Flow Model for Collapsible-Tube Oscillations," *J. Fluid Mech.*, **157**, pp. 375–404.
- [26] Anderson, P., Fels, S., and Green, S., 2013, "Implementation and Validation of a 1D Fluid Model for Collapsible Channels," *ASME J. Biomech. Eng.*, **135**(11), p. 111006.
- [27] Vasudevan, A., Zappi, V., Anderson, P., and Fels, S., 2017, "A Fast Robust 1D Flow Model for a Self-Oscillating Coupled 2D FEM Vocal Fold Simulation," Annual Conference of the International Speech Communication Association (INTERSPEECH 2017), Stockholm, Sweden, Aug. 20–24, pp. 3482–3486.
- [28] Heil, M., and Hazel, A. L., 2011, "Fluid-Structure Interaction in Internal Physiological Flows," *Annu. Rev. Fluid Mech.*, **43**(1), pp. 141–162.
- [29] Luo, X., and Pedley, T., 1996, "A Numerical Simulation of Unsteady Flow in a Two-Dimensional Collapsible Channel," *J. Fluid Mech.*, **314**, pp. 191–225.
- [30] Tian, F. B., Dai, H., Luo, H., Doyle, J. F., and Rousseau, B., 2014, "Fluid-Structure Interaction Involving Large Deformations: 3D Simulations and Applications to Biological Systems," *J. Comput. Phys.*, **258**, pp. 451–469.
- [31] Luo, H., Dai, H., Ferreira de Sousa, P. J. S. A., and Yin, B., 2012, "On the Numerical Oscillation of the Direct-Forcing Immersed-Boundary Method for Moving Boundaries," *Comput. Fluids*, **56**, pp. 61–76.
- [32] Chen, Y., and Luo, H., 2018, "A Computational Study of the Three-Dimensional Fluid-Structure Interaction of Aortic Valve," *J. Fluids Struct.*, **80**, pp. 332–349.