Fall 12-16-2022

Computational Investigations of the Fluid-Structure Interaction During Phonation: The Role of Vocal Fold Elasticity and Glottal Flow Unsteadiness

Xiaojian Wang
xiaojian.wang@maine.edu

COMPUTATIONAL INVESTIGATIONS OF THE FLUID-STRUCTURE INTERACTION DURING PHONATION: THE ROLE OF VOCAL FOLD ELASTICITY AND GLOTTAL FLOW UNSTEADINESS

By

Xiaojian Wang

B.S. Shandong University, 2013

M.S. Hohai University, 2017

A DISSERTATION

Submitted in Partial Fulfillment of the

Requirements for the Degree of

Doctor of Philosophy

(in Mechanical Engineering)

The Graduate School

The University of Maine

December 2022

Advisory Committee:

Qian Xue, Associate Professor of Mechanical Engineering, Advisor

Xudong Zheng, Associate Professor of Mechanical Engineering, Co-Advisor

Lauren Ross, Associate Professor of Civil and Environmental Engineering

Andrew J. Goupee, Associate Professor of Mechanical Engineering

Jung-Hee Seo, Associate Research Professor of Mechanical Engineering, JHU

COMPUTATIONAL INVESTIGATIONS OF THE FLUID-STRUCTURE INTERACTION DURING PHONATION: THE ROLE OF VOCAL FOLD ELASTICITY AND GLOTTAL FLOW UNSTEADINESS

By Xiaojian Wang

Dissertation Advisor: Dr. Qian Xue

An Abstract of the Dissertation Presented
in Partial Fulfillment of the Requirements for the
Degree of Doctor of Philosophy
(in Mechanical Engineering)
December 2022

Human voice production arises from the biomechanical interaction between vocal fold vibrations and airflow dynamics. Changes in vocal fold stiffness can lead to changes in vocal fold vibration patterns and further changes in voice outcomes. A good knowledge of the cause-and-effect relationship between vocal fold stiffness and voice production can not only deepen the understanding of voice production mechanisms but also benefit the treatment of voice disorders associated with vocal fold stiffness changes. This constitutes the first objective of this dissertation. The second objective of this dissertation is to further examine the range of validity of the quasi-steady assumption of glottal flow during phonation. The assumption is of vital importance for phonation modeling since it enables to eliminate the unsteady aspects of glottal flow, which greatly simplifies the flow modeling.

A three-dimensional flow-structure interaction model of voice production is employed to investigate the effects of vocal fold stiffness parameters on voice production. The vocal fold is modeled as the cover-ligament-body structure with a transversely isotropic constitutive relation. Stiffness parameters in both the transverse plane and the longitudinal direction of each layer of the vocal fold are systematically varied. The results show that varying the stiffness parameters has obvious monotonic effects on the fundamental frequency, glottal flow rate and glottal opening, but has non-monotonic

effects on the glottal divergent angle, open quotient and closing velocity. Compared to the transverse stiffness parameters, the longitudinal stiffness parameters generally have more significant impacts on glottal flows and vocal fold vibrations. Additionally, the sensitivity analysis reveals that the stiffness parameters of the ligament layer have the largest effect on most output measures.

Next, flow-structure interaction simulations are carried out to study the effect of fiber orientation in the conus elasticus on voice production. Two continuum vocal fold models with different fiber orientations in the conus elasticus are constructed. The more realistic fiber orientation (caudal-cranial) in the conus elasticus is found to yield smaller structural stiffness and larger deflection at the junction of the conus elasticus and ligament than the anterior-posterior fiber orientation, which facilitates vocal fold vibrations and eventually causes a larger peak flow rate and higher speed quotient. The generated voice is also found to have a lower fundamental frequency and smaller spectral slope.

Finally, the validity of the quasi-steady assumption for glottal flow is systematically examined by considering the voice frequency range, complexity of glottal shapes and air inertia in the vocal tract. The results show that at the normal speech frequency (~ 100 Hz), the dynamics of the quasi-steady flow greatly resembles that of a dynamic flow, and the glottal flow and glottal pressure predicted by the quasi-steady approximation have very small errors. However, the assumption produces huge errors at high frequencies (~ 500 Hz). In addition, air inertia in the vocal tract can undermine the validity of the assumption via the nonlinear interaction with the unsteady glottal flow. The role of glottal shapes in the validation is found to be insignificant.

ACKNOWLEDGEMENTS

First and foremost, my deepest gratitude goes to my advisor Dr. Qian Xue and co-advisor Dr. Xudong Zheng, for their invaluable supervision and priceless criticisms at every stage of my PhD study. It would be impossible for me to finish this dissertation without their guidance and encouragement. I would also like to thank Dr. Ingo Titze at the University of Utah for the inspiring discussion on the research of quasi-steady flow assumption. I likewise thank my doctoral committee members, Dr. Jung-Hee Seo, Dr. Lauren Ross and Dr. Andrew J. Goupee for their interest in my work.

I would also like to thank all the members in the Complex Flow Modeling & Simulation Lab, Dr. Geng Liu, Dr. Ngoc Pham, Dr. Yang Zhang, Dr. Weili Jiang, Dr. Biao Geng, Dariush Bodaghi, Mohammadreza Movahhedi and Yasaman Asiaee for a cherished time spent together in the lab. Special thanks to Dr. Weili Jiang and Dr. Biao Geng for the discussion on the CFD code and simulation setup.

I would also like to thank my friends and family for their support and love all through my studies.

Finally, I would like to thank the financial support provided by the National Institutes of Health (National Institute on Deafness and Other Communication Disorders Grant No. 1R03DC014562), the Department of Mechanical Engineering and the Department of Physics and Astronomy. I would also like to thank the Advanced Computing Group (ACG) at the University of Maine and the Extreme Science and Engineering Discovery Environment (XSEDE) for providing the computing resources.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1 INTRODUCTION

1.1. Overview and motivation

The human voice is a highly versatile signal which not only can be used to express emotions, but it can also convey paralinguistic information such as age, gender, personality and social status. The voice production is a complex aerodynamic-structure-acoustic interaction process. The systems that function together to produce the voice include the subglottal system (lungs, bronchi and trachea), the vocal folds and the resonating system (pharynx, oral and nasal cavities). During voice production, the air is driven from the lungs through the larynx, producing the self-sustained oscillation of the vocal folds. The vocal folds in turn modulate the airflow, turning the airflow into pulsatile jet exiting the glottis. The sound produced by the oscillating vocal folds (source of voice) is fed into and filtered by the resonating system, where it is either enhanced or attenuated to make for speech sounds.

Situated in the larynx, the two band-like vocal folds play a critical role in voice production. Prior to phonation, the vocal folds are brought together by the laryngeal adductor muscles to close the glottis. The air pressure builds up below the closed glottis, and the vocal folds vibrations are initiated when the pressure exceeds a certain threshold pressure. The voice control is primarily achieved through the laryngeal muscle activation which governs the position, shape and mechanical properties of vocal fold and further determines the vibratory patterns of vocal fold. Some of the cause-effect relationships between the biomechanical properties of vocal fold and the produced voice were already established. The most well-known principle is that vocal fold elongation increases vocal fold stiffness, which leads to the increase of fundamental frequency, or voice pitch. The stiffness conditions of different vocal fold layers can vary dramatically, depending on the activation levels of different muscles. Four typical voice registers (from modal to falsetto) can be produced through proper muscle activations which adjust the relative stiffness conditions of the cover and body layer of the vocal fold (Hirano, 1974). However, in addition to the stiffness change, muscle activation also causes changes in vocal fold geometry, position, effective mass in

vibration, etc. It's not clear whether the produced voices mainly result from the stiffness change or the combined effect of all the factors. To deepen the understanding of the realistic voice control mechanism, a study on the separate effect of vocal fold stiffness on voice production is needed. In addition, under vocal fold pathologies (edema, lesion paralysis, etc.), the biomechanical properties of the vocal fold can be altered, which usually results in voice disorders. A better understanding of the roles of vocal fold stiffness in phonation would benefit the medical diagnosis and treatment of voice disorders. For example, an implant is usually placed into the paralyzed vocal fold during surgical procedures to restore its function, such an understanding will help determine the materials used for the implant and the optimal placement of the implant.

The pressure drop across the larynx drives air through the glottis, forming the glottal flow. During the opening phase of the glottal cycle, the glottis has a shape of converging nozzle, and the airflow is attached to the entire medial surface of the vocal folds. During the closing phase, the glottis takes on a shape of diverging nozzle, and the airflow detaches from the vocal fold surface and presents many complex features such as asymmetric jet motions, shear-layer instabilities, vortex shedding and turbulence. Direct simulation of these three-dimensional details requires to numerically solve the time-dependent partial differential equations (Navier-Stokes equations), which is computationally very expensive. Considering the fact that the dimensions of the glottis are small and the vibration speed of vocal fold is much slower than the airflow velocity, the quasi-steady assumption can be applied to the modeling of glottal flow. The assumption assumes that the unsteady aspects of glottal flow induced by flow acceleration and vocal fold vibration are negligible, and the flow can be modeled as a sequency of steady flows through static glottises. Although the quasi-steady assumption based glottal flow models have been widely used in voice production modeling, the accuracy and range of validity of the assumption have not been rigorously verified by considering the complexity of glottal geometry, voice frequency range and effect of vocal tract inertance.

The purpose of this dissertation is threefold.

1) The first is to use a 3D continuum vocal fold model coupled with a Navier-Stokes equation based unsteady incompressible glottal flow model to better understand the effect of vocal fold stiffness on voice production. Stiffness parameters in both the transverse plane and the longitudinal direction of each layer of the vocal fold are systematically varied, and a parametric analysis is conducted to evaluate the sensitivity of the acoustic and vibration output measures to each stiffness parameter.
2) The second is to investigate the influence of fiber orientation of conus elasticus on its stiffness and voice production with the aim of improving the accuracy of voice simulation models. Two vocal fold models with different fiber orientations in the conus elasticus are built, and the produced vocal fold vibration, aerodynamic and acoustic parameters are compared and analyzed.
3) The third is to further examine the validity of the quasi-steady assumption for glottal flow by considering the voice frequency range, complexity of glottal shapes and air inertia in the vocal tract. The effects of flow unsteadiness and vocal fold vibration are separately examined by performing the dynamic, quasi-steady and pseudo static simulations. The validity of the assumption is assessed through quantifying the errors of glottal flow and glottal wall pressure between dynamic and quasi/pseudo simulations. A momentum budget analysis is also carried out to evaluate the magnitude of unsteady acceleration in the airflow and its relationship to the quantified errors.

1.2. Laryngeal cartilages and muscles

The laryngeal skeleton is mainly made up of four cartilages: the cricoid cartilage, the thyroid cartilage and a pair of arytenoid cartilages (see Figure 1.1). The cricoid cartilage is a ring-shaped structure that locates at the most inferior part of the laryngeal skeleton. The structure has a narrow and thin arch anteriorly but becomes broader and taller posteriorly. There are a pair of cricoarytenoid articular facets near the top for

the attachment of the arytenoid cartilages and a pair of cricothyroid articular facets on the lateral sides for the attachment of the thyroid cartilage. The main functions of the cricoid cartilage are to (1) maintain the patency of the airway and (2) provide attachments for laryngeal muscles, ligaments and cartilages. Above the cricoid cartilage lies the largest laryngeal cartilage: the thyroid cartilage. The thyroid cartilage consists of two laminae which meet at the midline anteriorly. The thyroid angle formed by the two laminae varies from 90 (for males) to 120 (for females) degrees, and the sharper thyroid angle in adult males results in the prominent Adam's apple. Above the prominence in the midline is the V-shaped thyroid notch which separates the two laminae. The thyroid cartilage is not enclosed on the posterior aspect, where two sets of cornua characterize the posterior borders of the cartilage. The superior cornua go upward and connect with the hyoid bone, while the inferior cornua go downward and join with the cricoid cartilage. The functions of the thyroid cartilage include protecting the vocal folds, adjusting the pitch of voice (through changing the angle relative to the cricoid cartilage) and providing attachments for ligaments and muscles. The anterior portion of the vocal fold is tightly connected to the thyroid cartilage through the anterior macula flava and commissure tendon. The paired arytenoid cartilages are pyramid-shaped structures and sit on the top posterolateral surface of the cricoid cartilage. Each of the arytenoid cartilage has two processes: the vocal process which provides attachment for the vocal ligament and the muscular process which serves as an attachment for muscles. The inferior surface of the arytenoid cartilage articulates with the upper articular facet of the cricoid cartilage, forming the cricoarytenoid joint. The cricoarytenoid joint enables rocking and gliding motion of the arytenoid cartilage, which is necessary for the control of abduction and adduction of the vocal folds.

Voice control is achieved through the activation of intrinsic laryngeal muscles, which changes the geometry and mechanical properties of the vocal folds. The main intrinsic laryngeal muscles include the cricothyroid (CT) muscle, the thyroarytenoid (TA) muscle, the lateral cricoarytenoid (LCA) muscle, the posterior cricoarytenoid (PCA) muscle and the interarytenoid (IA) muscle (see Figure 1.2). The CT muscle

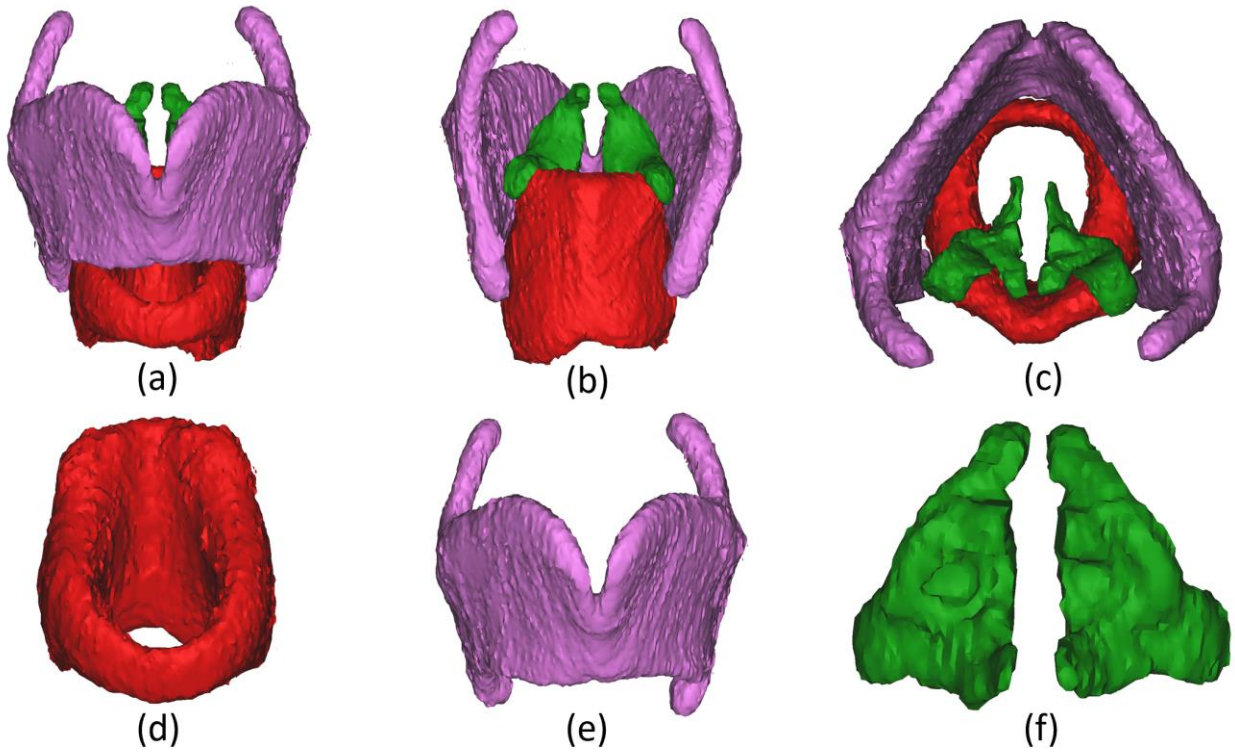

Figure 1.1 Skeleton of the larynx: (a) front view (b) back view (c) top view (d) cricoid cartilage (e) thyroid cartilage (f) arytenoid cartilage.

is composed of two bundles: the pars recta and the pars oblique. Contraction of the CT muscle tends to draw the anterior part of the cricoid and thyroid cartilage closer to each other and make the arytenoid cartilage farther from the thyroid cartilage. Both of these two effects result in the elongation and the increase in the tension and stiffness of the vocal folds, which makes the CT muscle the primary muscle for controlling the voice pitch. The TA muscle forms the body layer of the vocal folds and is generally considered to be composed of the vocalis muscle and the muscularis muscle. Different from the CT muscle, activation of the TA muscle tends to draw the arytenoid cartilage forward, thus shortening the vocal folds. During this process, the TA muscle becomes tight while the cover layer of the vocal fold becomes slack, which leads to the increase in the stiffness of the body layer while the decrease in the stiffness of the cover layer of the vocal folds. The effect of TA muscle activation on pitch-control thus depends on the vibration amplitude of the vocal folds. If the vibration amplitude is low and the vibration is limited to the cover layer, the result of TA muscle activation is more likely to decrease the vibration frequency. On the

contrary, if the vibration amplitude is high and the vibration involves a large part of the body layer, TA muscle activation tends to increase the vibration frequency (Titze et al., 1989). In addition, TA muscle activation can increase the vertical thickness of the vocal folds by bulging the inferior part of the medial surface. The LCA muscle arises from the superior lateral surface of the cricoid cartilage and connects to the muscular process of the arytenoid cartilage. Contraction of the LCA muscle draws the muscular process forward, and due to the rocking and rotation of the arytenoid, the vocal process will move medially to close the glottis. Therefore, the LCA muscle mainly functions to adduct the vocal folds. The PCA muscle originates from the posterior surface of the cricoid and inserts into the muscular process of the arytenoid cartilage. Contraction of this muscle draws the muscular process posteriorly and rotates the vocal process laterally to open the glottis. Contrary to the LCA muscle, the PCA muscle functions to abduct the vocal folds. The IA muscle can be divided into two parts: the transverse part and the oblique part. Both of the two parts run between the posterior surface of the two arytenoid cartilages. The function of IA muscle is to assist the LCA muscle in adducting the vocal folds through bringing the two arytenoids closer together.

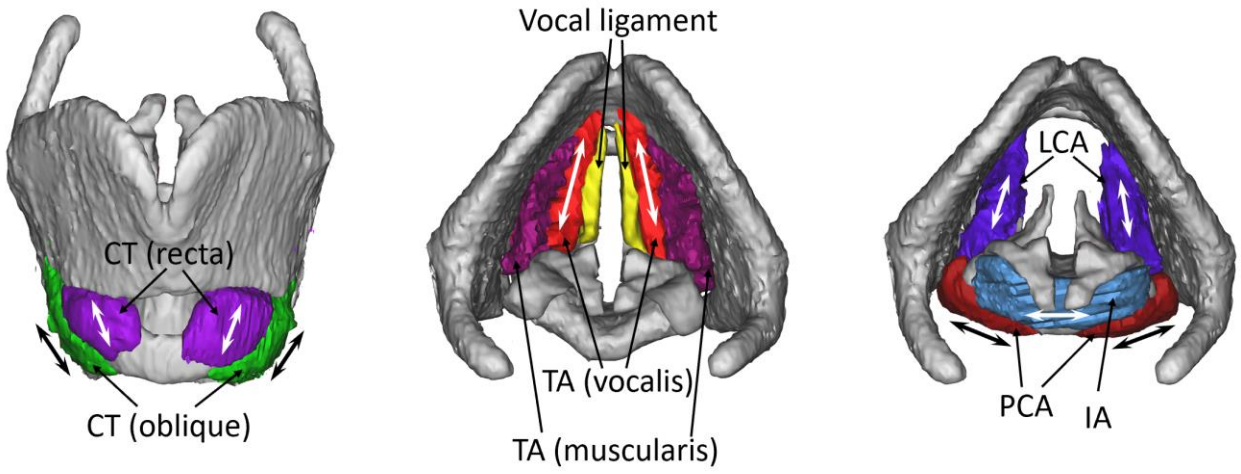

Figure 1.2 Main intrinsic laryngeal muscles. Double line arrows indicate the direction of muscle activation.

1.3. Vocal fold structure and self-sustained vibration

1.3.1. Layered structure of vocal fold

The human vocal folds are a paired structure with a length on the order of 1 cm, a width and a height of approximately 0.5-1 cm. They are composed of five tissue layers: the epithelium, the superficial, intermediate and deep layer of the lamina propria, and the thyroarytenoid muscle (see Figure 1.3). The epithelium layer is the outermost layer of approximately 0.1 mm thickness, which functions to protect the inner delicate layers and helps maintain the shape of the vocal fold (Sataloff, 2017). With microvilli on the surface, the epithelium layer can also help minimize slippage when the two vocal folds contact (Sato, 2018). Beneath the epithelium lies the superficial lamina propria (Reinke's space), which can be histologically distinguished by the loose connective tissue with sparse elastic or collagen fibers. The thickness of this layer is approximately 0.17 mm (Prades et al., 2010). The superficial layer of the lamina propria is connected to the epithelium layer through a supporting structure called the basement membrane (Sato, 2018), and the two layers constitute the mucosa of the vocal fold, in which the propagating mucosal wave is essential for normal vibration and phonation. Abnormalities in the mucosal wave can cause irregular vibrations of the vocal fold and further cause voice disorders. Deeper to the superficial lamina propria are the intermediate and deep layer of the lamina propria. The thickness of these two layers is approximately 0.63 mm and 0.45 mm, respectively (Prades et al., 2010). The intermediate and deep layer are distinguished from the superficial layer by the dense and bundled elastic or collagen fibers within them, which provides elasticity or resilience to the tissue when subjected to a force. These two layers constitute what is known as the vocal ligament. The thyroarytenoid muscle (the deepest layer) makes up the main body of the vocal folds, and the muscle activation in this layer regulates the geometry and stiffness of the vocal folds, which plays an essential role in vocal fold posturing and voice control.

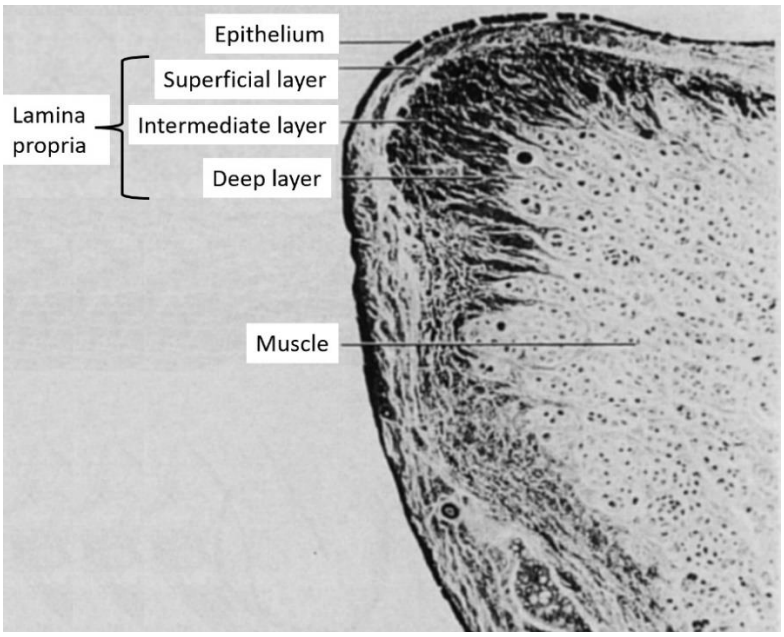

Figure 1.3 A coronal section of vocal fold showing its inner layers (from Hirano et al., 1981).

Based on the histological observations and mechanical properties, the five tissue layers are often regrouped into two or three mechanical layers. In the classic two-layer scheme, Hirano (1974) observed that the fibers in the deep layer of the lamina propria is firmly connected with the thyroarytenoid muscle and believed the vibratory movements of the two structures should be the same. Therefore, the deep lamina propria and the thyroarytenoid muscle were grouped into a body layer. The rest of the tissue layers are loosely connected to and can move differently from the body layer and were grouped into a cover layer. From a mechanical point of view, the five layers can withstand different levels of mechanical stress and thus function differently during voice production, they can also be reclassified into three layers: the cover, composed of the epithelium and the superficial lamina propria; the ligament, composed of the intermediate and deep layers of the lamina propria; and the body, which is the thyroarytenoid muscle (Hirano, 1974; Titze, 2000). The two-layer or three-layer scheme is of vital importance for studying the mechanisms of voice production, especially in physical modeling and computational simulation.

1.3.2 Two mechanisms of sustained vibration

A normal vibratory cycle of vocal fold in the coronal section is shown in Figure 1.4 in the series of time-frame diagrams. In the first diagram, the two vocal folds are in full contact, and the glottis is closed. The diagrams 2 to 4 indicate the opening phase of the glottis, during which the vocal folds are gradually pushed apart by the airflow with the inferior portion leading the superior portion and form a convergent glottal shape. As the inferior portion of the vocal folds arrives at the maximum lateral displacement, it starts to move medially due to the elastic restoring force of vocal fold tissues. Since the superior portion continues its lateral motion, there will be a moment at which the inferior and superior "meet" in the same position in the medial-lateral direction. This forms the maximum glottal opening and is shown in diagram 5. The diagrams 6 to 8 represent the closing phase of the glottis, during which the inferior takes the lead in moving medially, and a divergent glottal shape is formed. The glottis will be closed again when the inferior parts of the two vocal folds collide with each other (diagram 9). In the end, the superior parts also come into collision with each other, and the glottis is completely closed. This pattern cyclically repeats at the fundamental frequency of vocal fold vibration.

The fact that the vocal folds can maintain the periodic vibration means that there is positive net energy being continuously transferred to them to counteract the energy dissipation due to damping, inelastic collision, etc. This energy transfer is achieved through the asymmetric aerodynamic driving forces generated by the alternating convergent-divergent glottal geometry and the inertance of the vocal tract, which will be introduced below.

If neglecting the viscous effect in the boundary layer and the unsteady effects, the glottal flow during normal phonation can be well described by the Bernoulli's equation. The intraglottal pressure P_i of a cross section with an area A_i can be written as

$$P_i = P_0\left(1 - \frac{{A_s}^2}{{A_i}^2}\right) \tag{1.1}$$

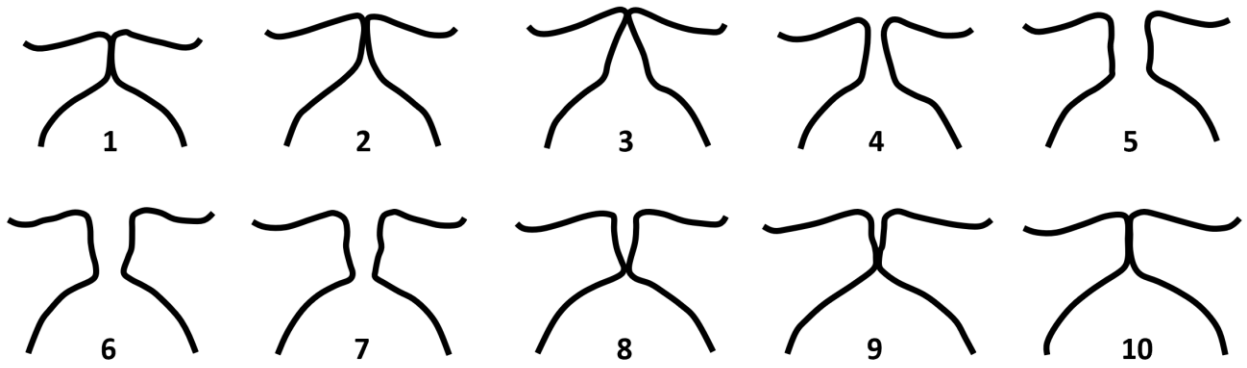

Figure 1.4 Diagram showing the coronal profiles of vocal fold in a normal vibratory cycle.

where P_0 is the subglottal pressure, and A_s is the cross-sectional area of the glottis at the position of flow separation. Here the flow pressure downstream of the separation point is assumed to be the same as the supraglottal pressure which is commonly simplified to zero gauge pressure. During the opening phase (diagrams 2-4), the glottis presents a convergent shape with the cross-sectional area gradually increasing from the superior to the inferior, and the airflow often separates at the exit of the glottis. Therefore, A_i is larger than A_s, which means the glottal pressure will be positive. As the glottis changes to the divergent shape during the closing phase, the flow separation location generally moves downward. The flow pressure will be zero for the cross sections above the separation location and be negative for those below the separation location since $A_s > A_i$. As a result, the glottal pressure is greater in the opening phase than the closing phase, and during each vibratory cycle energy is transferred to the vocal folds from the airflow to sustain the vibration.

The inertance of the air column in the vocal tract can also generate asymmetric glottal pressure and facilitate the vibration of vocal folds. To demonstrate this mechanism, the supraglottal pressure (P_1) is included in the Bernoulli's equation which is rewritten as

$$P_i = P_1 + (P_0 - P_1)(1 - \frac{{A_s}^2}{{A_i}^2}) \tag{1.2}$$

The glottis is assumed to be straight from the inferior to the superior to eliminate the effect of alternating convergent-divergent glottal shape. Since $A_s = A_i$, the intraglottal pressure P_i becomes the same as the

supraglottal pressure P_1. The supraglottal pressure is the input pressure of the vocal tract which can be calculated as

$$P_1 = I\frac{dU}{dt} \tag{1.3}$$

where I is the inertance of the air column in the vocal tract, U is the volume velocity of the air column, and $\frac{dU}{dt}$ is the rate of change of the airflow. During the opening phase of the glottis, the glottal flow increases, and the rate of change of the flow is positive, which causes the supraglottal and intraglottal pressure to be positive. During the closing phase, the supraglottal and intraglottal pressure becomes negative since the flowrate decreases. Therefore, a positive pressure drives the vocal folds outward in the opening phase while a negative pressure assists the vocal folds in moving medially in the closing phase. This results in a positive net energy transfer to the vocal folds, which means the sustained vibration is possible even under the single effect of vocal tract inertance.

In real phonation, the alternating convergent-divergent glottal geometry and the inertance of the vocal tract work simultaneously to generate the asymmetric pattern of intraglottal pressure, which is the key factor for initiating and sustaining the vibrations of vocal fold.

1.4 Literature review

1.4.1 Computational models for voice production

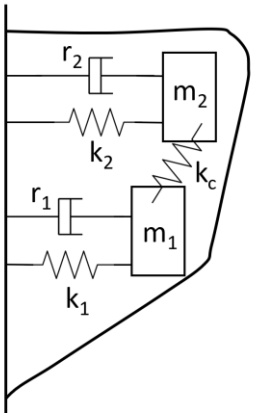

Figure 1.5 A typical configuration of two-mass model for the vocal fold.

The advent of the digital computer enables voice production to be simulated through numerically solving the mathematical equations that govern the dynamics of fluids and solids. In the early computational models, the vocal folds were represented by lumped masses with springs and dampers being used to characterize the viscoelastic property of the vocal fold structure. A typical model configuration, in which two masses are used to approximate the vocal fold structure, is shown in Figure 1.5. The governing equations of motion for the two masses were derived by Ishizaka and Flanagan (1972), which are

$$\begin{aligned} m_1\frac{d^2x_1}{dt^2} + r_1\frac{dx_1}{dt} + k_1(x_1) + k_c(x_1 - x_2) &= F_1 \\ m_2\frac{d^2x_2}{dt^2} + r_2\frac{dx_2}{dt} + k_2(x_2) + k_c(x_2 - x_1) &= F_2 \end{aligned} \tag{1.4}$$

where the second and third terms account for the viscous loss and nonlinear elasticity of the vocal fold, respectively, and the fourth term accounts for the interaction between the two masses. The forces F_1 and F_2 are the aerodynamic forces exerted on the vocal fold and are obtained from the glottal pressure. Ishizaka and Flanagan (1972) used a modified one-dimensional Bernoulli equation to describe the glottal flow, in which the viscous effect, local losses, inertance and pressure recovery of the airflow were considered. The coupling between the solid and flow equations was achieved through the relationship between the displacement of the masses and glottal area. Although this two-mass model can produce physiologically more realistic behaviors (e.g., vertical phase difference) and improved intraglottal pressure distribution over the initially proposed one-mass model (Flanagan and Landgraf, 1968), some other details of vocal fold behavior still cannot be captured due to the small number of degree of freedoms. Later studies tried to improve the performance of the model through incorporating longitudinal phase difference (Titze, 1973,1974; Wong et al., 1991; Schwarz et al., 2008), body-cover structure (Story and Titze, 1995), mucosal wave of the upper medial surface (Tokuda et al., 2007, 2010) and three-dimensional dynamics of the vocal folds (Yang et al, 2010). In addition to these studies, efforts were also made to optimize the correlation between model parameters and tissue properties. For example, Titze and Story

(2002) proposed a set of empirical rules for relating the geometrical and viscoelastic parameters of the model to muscle activities. Yang et al. (2011, 2012) mathematically optimized the biomechanical parameters of the model based on the vocal fold dynamics extracted from experiments. Nevertheless, an inherent drawback of the lumped mass model is the weak relevance between its parameters and the biomechanical properties of vocal fold tissues. Additionally, due to the greatly simplified geometrical representation of vocal fold and the small number of degree of freedoms, the lumped mass model cannot well approximate the highly three-dimensional features of vocal fold motion.

A milestone in the development of vocal fold model is the introduction of continuum models which overcome the shortcomings of lumped mass models. In early attempts to develop the continuum model (Titze and Strong, 1975; Titze, 1976; Titze and Talkin, 1979), vocal fold tissue was modeled as a linear elastic, orthotropic material, and a finite-element or finite-difference method was employed. The focus of these studies was to explore how the selected parameters (e.g., elastic moduli, incompressibility, boundary conditions) affect model performance and provide directions for optimizing the initial continuum model. Based on the findings of these studies, Alipour et al. (2000) developed a finite element-based continuum model which incorporates more realistic geometries and layered structure of the vocal folds. More importantly, the material parameters used in the model are from experimental data on the biomechanical properties of vocal fold tissues, which makes it more physiologically relevant. Although the model developed by Alipour et al. (2000) can generate reasonable glottal waveforms and vocal fold vibrations, the linear materials make it inappropriate for modeling the stiffness changes induced by muscle activation. To further improve the vocal fold modeling, subsequent studies employed nonlinear constitutive models such as modified Kelvin fiber model (Hunter et al., 2004), Ogden model (Zhang et al., 2009; Smith and Hunter, 2014), Mooney-Rivlin model (Yin and Zhang, 2013, 2014), Yeoh model (Yin and Zhang, 2016) and Holzapfel-Gasser-Ogden model (Ngoc et al., 2018; Geng et al., 2020). The application of these nonlinear models greatly advances the understanding of the relationship between muscle activity

and voice control. Moreover, considering that the geometry and structure of the vocal folds also have a significant impact on voice production, recent studies used images from computed tomography (CT) scan or magnetic resonance imaging (MRI) to create the realistic larynx model (Wu and Zhang, 2019; Geng et al., 2020; Jiang et al., 2022), which largely increases the fidelity of computational models.

The fact that intraglottal pressure provides the driving force of vocal fold vibration highlights the importance of glottal flow modeling. Traditionally, the glottal flow is assumed to be quasi-steady, inviscid and incompressible, and the 1D Bernoulli's equation is used to approximate the intraglottal pressure-flow relationship. Flow properties, such as viscosity and unsteadiness, can be considered with correction terms in the equation. For example, Ruty et al. (2007) added a Poiseuille term in the Bernoulli flow model. They found that for very small glottal apertures, the threshold pressure and fundamental frequency predicted by the model were improved and become closer to experimental data. The biggest advantage of the Bernoulli flow model is that it is computationally efficient and easily implemented. However, it cannot duplicate many significant flow behaviors such as flow separation and vortex shedding. Alipour and Titze (1995) coupled a 2D unsteady incompressible Navier-Stokes flow model with a 2D finite-element vocal fold model to simulate the glottal airflow and vocal fold vibration. Their model captured vortices around the glottis and produced the jet velocity waveform and pressure profiles that are similar to the experimental observations. Luo et al. (2009) also employed the 2D incompressible Navier-Stokes equations to describe the glottal airflow, but they applied an immersed-boundary method to model the fluid-structure interaction. Glottal flow features, including the time-varying jet asymmetry and flow separation, were observed in their simulations, which is consistent with the experimental results. Later, Zheng et al. (2010, 2011) expanded the immersed-boundary method based numerical model of phonation for 3D simulation. In their studies, the glottal airflow was governed by the 3D, incompressible and fully unsteady Navier-Stokes equations. In addition to the asymmetric jet deflection captured by the 2D model, the experimentally observed three-dimensionality of the glottal flow, such as the longitudinal vena

contracta and development of 3D vortex structures, were also reproduced by their models. In summary, the development of the computational models of voice production has undergone a process from low fidelity to high fidelity, for both the solid and fluid models.

1.4.2 Vocal fold biomechanics

1.4.2.1 Biomechanical properties of vocal fold

Just as the thickness, tension and length of a guitar string can be varied to change the pitch of the sound it produces, the geometry and stiffness of the vocal fold tissues are also regulated to generate different voices. However, unlike the guitar string which can be easily adjusted through following some specific rules, the vocal fold tissues are quite irregular, and the results of the regulation are less predictable and highly dependent on the biomechanical properties of vocal fold tissues.

The fibers in the lamina propria and the thyroarytenoid muscle run roughly along the length direction of the vocal fold, which suggests the elastic property of the vocal fold tissue is anisotropic. Generally, the tissue is more difficult to stretch in the fiber direction than in the transverse plane, reflecting a larger strength or stiffness in the fiber direction.

Stress-strain curves from the experiments were usually used to estimate the Young's modulus of vocal fold tissues. The stress and strain of the tissues exhibit an approximate exponential relationship, with a relatively shallow slope at small strains (< 15%) and a steep slope at large strains (> 25%) (Ishizaka and Kaneko, 1968; Tran et al., 1993; Min et al., 1995; Chan et al., 2007; Kelleher et al., 2013a, b; Oren et al., 2014). This kind of stress-strain curve indicates the nonlinear elastic property of vocal fold tissues, but it also indicates that a linear elastic material can be used to approximate the vocal fold tissues at low strains. The longitudinal Young's modulus of vocal fold layers decrease in the order of the epithelium, the lamina propria and the thyroarytenoid muscle (Hirano, 1985). At low strains (< 15%), Alipour and Titze (1991) reported the mean longitudinal Young's moduli for the cover and body tissues of canine vocal folds as 41.9 kPa ± 7.1 kPa and 20.7 kPa ± 2.4 kPa, respectively. Min et al. (1995) conducted longitudinal stretch

tests on the human vocal ligament and reported the mean Young's modulus as 33.1 kPa ± 10.4 kPa at low strains. The longitudinal Young's moduli for the cover and ligament of human vocal folds were also estimated by Kelleher et al. (2013a) through conducting uniaxial tensile tests. At small stretches, the average Young's moduli of all subjects for the cover and ligament are 300.0 kPa and 26.8 kPa, respectively. Compared to their longitudinal counterparts, the transverse Young's moduli of vocal fold tissues are much smaller. Under a transverse tension test, Hirano (1985) found that the transverse modulus of the lamina propria is about a half order of magnitude (elongations < 10%) and one order of magnitude smaller (elongations about 40%) than the longitudinal modulus. Using the indentation method, Chhetri et al. (2011) estimated the mean Young's moduli of the cover of human vocal folds are 2.9 kPa at the superior location, 4.8 kPa at the middle location and 7.5 kPa at the inferior location. For the thyroarytenoid muscle, the mean Young's modulus is 2.0 kPa. These moduli should be regarded as the transverse modulus since the direction of the force applied by the indenter is in the transverse plane. In the transverse indentation performed by Kelleher et al. (2013a), the mean transverse moduli of the cover and ligament are 8.6 kPa and 2.5 kPa, respectively. The different Young's modulus for the same tissue could be due to the different devices and methods used in the experiments. Conducting experiments on different subjects or specimens could also result in the difference. For example, the elastic modulus of the human vocal fold cover was reported to be larger than that of the ligament for females, while the opposite was observed for males (Chan et al., 2007).

Due to the fact that human tissues contain quite a high percentage of water, vocal fold tissues are also considered as incompressible (Berry and Titze, 1996). The incompressibility assumption helps reduce the material parameters needed in the computational models of vocal fold (Cook et al., 2008). In addition, like other soft human tissues, there is also viscosity in vocal fold tissues. The changes in viscosity can affect the threshold pressure and contact stress of vocal fold oscillation (Finkelhor et al., 1988; Erath et al., 2017).

The stiffness of vocal fold as a whole not only depends on its material properties, but also depends on the geometry, boundary condition, fiber orientation, etc. For example, a vocal fold separated from the laryngeal cartilages is less stiff than those remain attached to the cartilages (Chhetri, 2011).

1.4.2.2 Vocal fold dynamics

Vibration of the vocal folds is an integral part of voice production. Different vocal fold vibration patterns can lead to different voice types or qualities. The dynamics of vocal fold vibration, including vibration amplitudes, velocities, accelerations and phase delays, are highly sensitive to the biomechanical and geometric properties of the vocal folds.

Previous studies have shown that vocal fold dynamics are strongly dependent on its stiffness. Using the two-mass lumped model, Ishizaka and Flanagan (1972) were the first to investigate the effects of vocal fold stiffness on its vibration. Their results showed that increasing the vocal fold stiffness generally reduces the vibration amplitude. In addition, decreasing the flexural stiffness (coupling stiffness of the two masses) increases the vertical phase difference. In their body-cover three-mass model, Story and Ttize (1995) found that different stiffness conditions of the cover and body layer can produce quite different vibration patterns. More specifically, a large stiffness in the cover or body layer decreases their respective vibration amplitude, and a large body-cover stiffness ratio results in a big vertical phase difference. Tokuda et al. (2007) used their vertically arranged three-mass model to simulate the register transitions. They observed that gradually increasing the stiffness of the middle mass can lead to the transitions of vocal fold vibration from chest-like to falsetto-like. Moreover, investigations of vocal fold paralysis with lumped mass models demonstrated that tension imbalance between the left and right vocal fold can induce irregular vocal fold vibrations characterized by toroidal oscillations (Steinecke and Herzel, 1995), period-doubling (Herzel and Knudsen, 1995; Xue et al., 2010) and chaotic motion (Herzel et al., 1995; Erath et al., 2011). More recently, continuum and synthetic vocal fold models were adopted to examine the relation between vocal fold stiffness and its dynamics. Using a body-cover continuum model, Zhang (2009) showed

that increasing the body-cover stiffness ratio, which corresponds to the condition of high TA contraction level or low CT contraction level, reduces the vibration amplitude of the body layer and restricts the vocal fold motion to the medial surface, resulting in increased efficiency of flow modulation and sound production. Similar findings were also reported by Mendelsohn and Zhang (2011). Zhang (2010) used a physical model to investigate the effect of left-right asymmetric body-layer stiffness on the vibratory patterns of the vocal folds. The study showed that the whole vibration is dominated by the rigid-body model at low subglottal pressure and by the soft-body model at high subglottal pressure, and the other vocal fold follows the vibration at the same frequency. A following study conducted by Zhang and Luu (2012) used both physical and numerical vocal fold models to further investigate this problem. They reported that the vocal fold vibratory behavior is dominated by the soft fold under the condition of large left-right stiffness difference. When the left-right stiffness difference is small, both folds vibrate with similar amplitude, but the stiff fold is ahead of the soft fold in vibration. A more comprehensive study of the effects of vocal fold stiffness on voice production was conducted by Zhang (2017) with a three-dimensional, body-cover, transversely isotropic continuum model. In the study, changes in both the transverse and longitudinal stiffness of the vocal fold were found to significantly affect the fundamental frequency, phonation threshold pressure, the glottal area and flow rate. However, the effects of varying vocal fold stiffness on the closed quotient, vertical phase difference and spectral slope were small and inconsistent. The study concluded that there is no direct cause-effect relationship between the vocal fold stiffness and voice types.

The dynamics of vocal folds are also highly affected by their geometry. Since the fluid-structure interaction and mucosal wave propagation primarily occur on the medial surface of the vocal folds, the medial surface shape is of great importance in vocal fold vibrations. Using a physical vocal fold model, Chan et al. (1997) showed that the lowest threshold pressure was achieved when the medial surface shape is rectangular. As the medial surface becomes convergent or divergent, higher phonation threshold pressure was

observed. Alipour and Scherer (2000) employed a three-dimensional computational model to study the effects of medial vocal fold bulging on phonation. They showed that in addition to increasing the glottal flow resistance and decreasing the mean glottal volume velocity, increasing medial surface bulging significantly decreased the maximum glottal amplitude and area. Pickup and Thomson (2010) compared the dynamics of synthetic vocal fold models based on MRI data with the responses of M5-based models. The study showed that the MRI-based models, which have a convergent prephonatory intraglottal angle, can improve the model vibration by generating more typical convergent-divergent behavior and mucosal wave-like motion. A consistent finding was later reported by Pickup and Thomson (2011) which used a computation model to further investigate the effects of geometric parameters on vocal fold vibration. Pickup and Thomson (2011) also found that increasing the glottal thickness, as well as increasing the entrance and exit radii of the cover layer, can increase the inferior-superior phase difference. Lately, Wu and Zhang (2019) used a larynx-specific MRI-based vocal fold model to study the influence of medial surface shape on voice production. Their results showed that increasing the inferior or superior medial surface bulging greatly lowered the phonation threshold pressure and increased the closed quotient of vocal fold vibration. However, the degree of the effects was found to be subject specific and dependent on the cross-sectional geometry and its longitudinal variation.

1.4.3 Glottal aerodynamics

1.4.3.1 Glottal jet: Is it quasi-steady?

The separation of airflow from the glottal wall causes the formation of glottal jet. In the conventional Bernoulli-based description of glottal flow, this jet is assumed to be quasi-steady and have a constant diameter (Story and Titze, 1995), which means the unsteady acceleration and the convection term in the following one-dimensional momentum equation can be neglected, resulting in the uniform pressure distribution downstream of the flow separation point.

$$\rho \frac{\partial u}{\partial t} + \frac{1}{2}\rho \frac{\partial u^2}{\partial x} = -\frac{\partial p}{\partial x} \tag{1.5}$$

This assumption greatly simplifies the computational modeling of phonatory flow. However, are the temporal and spatial variations in the jet flow really small enough to be ignored? In their investigation of the dynamics of flow through glottis-like channels, Hofmans et al. (2003) showed that the pressure in the jet is significantly nonuniform because a quasi-steady jet cannot be established during a typical vocal fold vibration cycle. This argument is supported by the experimental study conducted by Krane et al. (2007). In fact, the glottal airflow accelerates in the opening phase and decelerates in the closing phase, which significantly changes the jet velocity during one glottal cycle. This is apparent from the waveform of the supraglottal jet velocity shown in Alipour et al. (1995) and Krane et al. (2007, 2010). The dramatic velocity changes in the waveform, especially in the closing phase, suggests the unsteady acceleration of the jet may be non-negligible. Also, visualization by particle image velocimetry (PIV) of glottal jet showed very strong three-dimensional features (Triep and Brücker, 2010; Krebs et al., 2012; Nielson et al., 2013). The jet was generally observed to be narrow in the lateral direction near the separation point and gradually expand as it travels downstream. In addition, Triep and Brücker (2010) showed that the aspect ratio of the jet cross section undergoes a temporal evolution, which eventually causes the axis switching of the jet. The results of these studies indicate that convective acceleration is present in the jet and the assumption of constant-diameter jet is debatable. The unsteady acceleration and convective acceleration in the jet were quantified by Krane et al. (2010) based on the experimentally measured velocity field. Their results did show that neither of the two accelerations are negligible. However, they found that the two accelerations are of commensurate magnitude but with opposite signs in the middle 40% of the cycle, resulting in the nearly zero jet inertia. The experimental results of Krane et al. (2010) are consistent with the previous conclusions of quasi-steady flow during the middle stages of the phonation cycle (Mongeau et al., 1997; Zhang et al., 2002; Vilain et al.; 2004). Krane et al. (2010) also emphasized that the jet inertia

should be considered for the rest of the cycle, especially in the late closing stages, during which the voice sound power and spectral features are determined.

The asymmetric characteristic of glottal jet has been observed and thoroughly investigated by many previous studies (Scherer et al., 2001; Erath and Plesniak, 2006a, b, 2010; Krane et al., 2007; Neubauer et al., 2007; Dreschsel and Thomson, 2008; Pickup and Thomson, 2009; Zheng et al., 2011a, b). The asymmetric flow causes nonuniform velocity and pressure distributions within the glottis. More specifically, flow velocity is higher on the side that the flow is deflected to than the other side, resulting in the relatively small pressure on the former side. The pressure difference between the flow attached and detached sides can vary from 5% to 25%, depending on the divergent angle of the glottis (Scherer et al., 2001; Shinwari et al., 2003; Erath and Plesniak, 2005). Moreover, the direction of jet deflection can change from toward one side to toward the other side within one phonatory cycle, causing the so-called jet "flip-flopping" phenomenon (Neubauer et al., 2007; Zheng et al., 2011, a). The asymmetric glottal jet and its time-varying dynamics may cause complex and unsteady pressure distributions downstream of the flow separation point, which could undermine the legitimacy of the quasi-steady glottal jet assumption.

1.4.3.2 Effects of glottal wall motion

In the general fluid-structure interactions, the fluid dynamics and the solid dynamics are coupled through the kinematic and dynamic conditions at the fluid-structure interface, and the two dynamics advance simultaneously. However, in the specific case where the time scale of the solid dynamics is much longer than that of the fluid dynamics, the quasi-static aeroelasticity approximation can be applied. In the approximation, the solid dynamics only provides the position of the interface, and the kinematic condition at the interface does not affect the fluid dynamics. This is the exact assumption behind the Bernoulli-based glottal flow models, and it seems justifiable under the condition of normal human phonation, where the glottal airflow velocity is approximate 40 m/s, and the vibration velocity of glottal wall is around 0.1 m/s (Zhang et al., 2002). The effect of the flow unsteadiness due to glottal wall movements were

investigated by Deverge et al. (2003) and Krane and Wei (2006) with experimental or theoretical methods. Their results suggested that the importance of glottal wall motion depends on the ratio of the volume flux displaced by the wall to the volume flux driven by the transglottal pressure. The wall motion effect can be important only during the short periods of flow initiation and shutoff, when the two fluxes become comparable. Following the studies of Krane and Wei (2006) and Krane et al. (2010), Ringenberg et al. (2021) found that the acceleration due to wall motion is smaller than the convective acceleration by a factor of f^{*2} (f^* is the reduced vibration frequency defined by Krane et al. (2010)), and the ratio of the above two fluxes is proportional to f^*. Since f^* is around 0.035 for a voice frequency around 100 Hz, the study suggested that the effect of wall motion on glottal flow is trivial in normal phonation. Zhao et al. (2002) investigated the effect of glottal wall motion on the sound production. They found that at a frequency of 125 Hz, the monopole sound, which is due to the volume flux displaced by the wall motion, is only about 20% of the dominant dipole source and is negligible in speech production. However, they also found that the strength of the monopole sound is proportional to frequency and can become significant at high frequencies above 400 Hz. Using one-dimensional mass and momentum conservation equations, Deguchi and Hyakutake (2009) theoretically examined the effect of wall motion on glottal flow in a frequency range of 16-500 Hz. They showed that the unsteady effect induced by vocal fold wall motion can become comparable to the Bernoulli effect at 500 Hz. These studies suggest that the effect of glottal wall motion is frequency dependent, and applying the quasi-static aeroelasticity approximation in human voice research may be questionable when the frequency exceeds a certain value.

CHAPTER 2 NUMERICAL METHOD

Numerical simulation of human phonation requires to model both the vocal fold dynamics and the glottal flow dynamics, as well as the interaction between them. In this study, a 3D linear viscoelastic solid solver based on the finite-element method is employed to solve the vocal fold dynamics. An immersed boundary method based incompressible Navier-Stokes solver and a Bernoulli flow model are used to simulate the airflow dynamics. The explicit coupling scheme is used for the fluid-structure interaction simulation. For more detailed descriptions of the solvers, readers are referred to Mittal et al. (2008), Zheng (2009), Zheng et al. (2010, 2011a, b) and Xue (2011).

2.1 Viscoelastic solver for solid mechanics

2.1.1 Constitutive model

The previous chapter has shown that vocal fold tissues exhibit elastic and viscous characteristics, as well as a nearly linear stress-strain relationship at low strains. Considering the deformation of vocal folds is small during normal phonation, the linear viscoelasticity constitutive model can be used to describe the biomechanical response of vocal folds (Alipour et al., 2000). In the current study, the Kevin-Voigt model is employed, and the linear constitutive equation is as follows:

$$\boldsymbol{\sigma} = \boldsymbol{C\varepsilon} + \boldsymbol{A\dot{\varepsilon}} \tag{2.1}$$

where $\boldsymbol{\sigma}$, $\boldsymbol{\varepsilon}$, and $\boldsymbol{\dot{\varepsilon}}$ are the second-order stress, strain, and strain rate tensor, respectively. $\boldsymbol{C}$ and $\boldsymbol{A}$ are the fourth-order elasticity and viscosity tensor, respectively. The transversely isotropic elasticity is used to represent the anisotropic mechanical properties of vocal fold tissues. The stress-strain relationship for the transversely isotropic elasticity is given by

$$\begin{Bmatrix}\sigma_{xx}\\ \sigma_{yy}\\ \sigma_{zz}\\ \sigma_{yz}\\ \sigma_{zx}\\ \sigma_{xy}\end{Bmatrix} = \begin{bmatrix} \frac{1-\nu_{pz}\nu_{zp}}{E_pE_z\Delta} & \frac{\nu_p+\nu_{zp}\nu_{pz}}{E_pE_z\Delta} & \frac{\nu_{zp}+\nu_p\nu_{zp}}{E_pE_z\Delta} & 0 & 0 & 0 \\ \frac{\nu_p+\nu_{pz}\nu_{zp}}{E_zE_p\Delta} & \frac{1-\nu_{zp}\nu_{pz}}{E_zE_p\Delta} & \frac{\nu_{zp}+\nu_{zp}\nu_p}{E_zE_p\Delta} & 0 & 0 & 0 \\ \frac{\nu_{pz}+\nu_p\nu_{pz}}{{E_p}^2\Delta} & \frac{\nu_{pz}(1+\nu_p)}{{E_p}^2\Delta} & \frac{1-{\nu_p}^2}{{E_p}^2\Delta} & 0 & 0 & 0 \\ 0 & 0 & 0 & 2G_{zp} & 0 & 0 \\ 0 & 0 & 0 & 0 & 2G_{zp} & 0 \\ 0 & 0 & 0 & 0 & 0 & \frac{E_p}{1+\nu_p} \end{bmatrix} \begin{Bmatrix}\varepsilon_{xx}\\ \varepsilon_{yy}\\ \varepsilon_{zz}\\ \varepsilon_{yz}\\ \varepsilon_{zx}\\ \varepsilon_{xy}\end{Bmatrix} \quad (2.2)$$

where x-y plane is the plane of isotropy, z is the direction perpendicular to the plane of isotropy. E_p and ν_p are the Young's modulus and Poisson's ratio in the plane of isotropy. E_z, G_{zp} and ν_{zp} are the Young's modulus, shear modulus and Poisson's ratio in the z direction, respectively. ν_{pz} and ν_{zp} are related by

$$\frac{\nu_{pz}}{E_p} = \frac{\nu_{zp}}{E_z} \quad (2.3)$$

The expression for Δ is $\Delta = \frac{(1+\nu_p)(1-\nu_p-2\nu_{pz}\nu_{zp})}{{E_p}^2E_z}$. The factor of 2 before the shear moduli in the stiffness matrix comes from the difference between shear strain and engineering shear strain, where $\gamma_{xy} = \varepsilon_{xy} + \varepsilon_{yx} = 2\varepsilon_{xy}$, etc.

The stress-strain rate relationship for the viscous component in the constitutive equation is given as,

$$\begin{Bmatrix}\sigma_{xx}\\ \sigma_{yy}\\ \sigma_{zz}\\ \sigma_{yz}\\ \sigma_{zx}\\ \sigma_{xy}\end{Bmatrix} = \eta \begin{Bmatrix}\dot{\varepsilon}_{xx}\\ \dot{\varepsilon}_{yy}\\ \dot{\varepsilon}_{zz}\\ \dot{\varepsilon}_{yz}\\ \dot{\varepsilon}_{zx}\\ \dot{\varepsilon}_{xy}\end{Bmatrix} \quad (2.4)$$

where η is the viscosity of the tissue.

2.1.2 Governing differential equation and FEM formulation

The governing equation of motion for the solid dynamics is given by

$$\sigma_{ij,j} + f_i = \rho\ddot{d}_i \quad (2.5)$$

where the indices i and j range from 1 to 3, f is the body force, ρ is the tissue density, and d is the displacement.

In the FEM formulation, the weak form of the governing equation is first derived. Eq. 2.5 is written in the weighted residual integral form. The residual, R_i, is defined as:

$$R_i = \sigma_{ij,j} + f_i - \rho\ddot{d}_i \tag{2.6}$$

When the solution to Eq. 2.5 is accurate, $R_i = 0$. Through multiplying R_i by a weight function w and integrating over the domain v, the weighted residual form is obtained as follows:

$$\int_v w\sigma_{ij,j}dv + \int_v wf_i dv - \int_v w\rho\ddot{d}_i dv = 0 \tag{2.7}$$

By considering chain rule, $w\sigma_{ij,j}$ becomes

$$w\sigma_{ij,j} = (w\sigma_{ij})_{,j} - w_{,j}\sigma_{ij} \tag{2.8}$$

Substituting Eq. 2.8 into Eq. 2.7 and rearranging the terms in the new equation gives

$$\int_v w\rho\ddot{d}_i dv = \int_v (w\sigma_{ij})_{,j}dv - \int_v w_{,j}\sigma_{ij}dv + \int_v wf_i dv \tag{2.9}$$

By applying divergence theorem to the first term on the right side, Eq. 2.9 becomes

$$\int_v w\rho\ddot{d}_i dv + \int_v w_{,j}\sigma_{ij}dv = \int_s w\sigma_{ij}n_j ds + \int_v wf_i dv \tag{2.10}$$

Since the surface traction $t_i = \sigma_{ij}n_j$, the surface integral term $\int_s w\sigma_{ij}n_j ds = \int_s wt_i ds$. Therefore, the weak form of the governing equation can be written as follows

$$\int_v w\rho\ddot{d}_i dv + \int_v w_{,j}\sigma_{ij}dv = \int_s wt_i ds + \int_v wf_i dv \tag{2.11}$$

Next, the shape function $N_{i\alpha}$ (α denotes the element node number) is chosen as the weight function for applying the Galerkin method. Eq. 2.11 becomes

$$\int_v \rho N_{i\alpha}\ddot{d}_i dv + \int_v N_{i\alpha,j}\sigma_{ij} dv = \int_s N_{i\alpha} t_i ds + \int_v N_{i\alpha} f_i dv \tag{2.12}$$

The constitutive equation is $\sigma_{ij} = C_{ijkl}\varepsilon_{kl} + A_{ijkl}\dot{\varepsilon}_{kl}$. For the case of small strain, the strain-displacement relationship is $\varepsilon_{ij} = \frac{1}{2}(d_{i,j} + d_{j,i})$. Substituting the strain-displacement relationship into the constitutive equation gives

$$\sigma_{ij} = \frac{1}{2}C_{ijkl}d_{k,l} + \frac{1}{2}C_{ijkl}d_{l,k} + \frac{1}{2}A_{ijkl}\dot{d}_{k,l} + \frac{1}{2}A_{ijkl}\dot{d}_{l,k} \tag{2.13}$$

Since C_{ijkl} and A_{ijkl} are symmetric in ij and kl, and k and l are dummy indices and can be swapped, Eq. 2.13 becomes

$$\sigma_{ij} = C_{ijkl}d_{k,l} + A_{ijkl}\dot{d}_{k,l} \tag{2.14}$$

The displacement at an arbitrary point can be written as $d_i = D^\alpha N_{i\alpha}$, where D^α is the nodal displacement. $d_{i,j}$ can be written as $d_{i,j} = D^\alpha N_{i\alpha,j} = D^\alpha B_{ij\alpha}$. Therefore, $\sigma_{ij} = C_{ijkl}B_{kl\alpha}D^\alpha + A_{ijkl}B_{kl\alpha}\dot{D}^\alpha$. Substituting the expression of d_i and σ_{ij} into Eq. 2.12 and rearranging the terms gives

$$\begin{aligned}(\int_v \rho N_{i\alpha}N_{j\beta}dv)\ddot{D}^\beta + (\int_v A_{ijkl}B_{ij\alpha}B_{kl\beta}dv)\dot{D}^\beta + (\int_v C_{ijkl}B_{ij\alpha}B_{kl\beta}dv)D^\beta \\ = \int_s N_{i\alpha}t_i ds + \int_v N_{i\alpha}f_i dv\end{aligned} \tag{2.15}$$

where $\int_v \rho N_{i\alpha}N_{j\beta}dv$ is the mass matrix and is denoted by $M_{\alpha\beta}$, $\int_v A_{ijkl}B_{ij\alpha}B_{kl\beta}dv$ is the damping matrix and is denoted by $C_{\alpha\beta}$, $\int_v C_{ijkl}B_{ij\alpha}B_{kl\beta}dv$ is the stiffness matrix and is denoted by $K_{\alpha\beta}$, $\int_s N_{i\alpha}t_i ds$ is the traction force and here is denoted by F_α^t, $\int_v N_{i\alpha}f_i dv$ is the body force and here is denoted by F_α^b. Therefore, the partial differential equation is converted to the second-order ordinary differential equation which is shown as follows:

$$M_{\alpha\beta}\ddot{D}^\beta + C_{\alpha\beta}\dot{D}^\beta + K_{\alpha\beta}D^\beta = F_\alpha^t + F_\alpha^b \tag{2.16}$$

The M, C and K matrices are obtained through performing numerical integration over the elements, and like many other FEM programs, the isoparametric coordinates which employ area-fraction as coordinates are used.

When calculating the traction force, only the normal stress is considered since the shear stress is small under normal phonation ($Re \approx 3000$). The equivalent nodal force is calculated via the integration based on shape functions. For the tetrahedral element, the expression is as follows:

$$F_i^t = A(\frac{p_i}{6} + \frac{p_j}{12} + \frac{p_k}{12}) \tag{2.17}$$

where A represents the area of surface triangular element, i, j and k denote the nodal indices, and p is the surface pressure at the corresponding node.

A second-order Newmark method is used to discretize Eq. 2.16 in time, resulting in the following discretized equation:

$$\begin{aligned}\left(K + \frac{1}{\beta\Delta t^2}M + \frac{\gamma}{\beta\Delta t}C\right)D^{n+1} \\ = F^{n+1} + M\left[\frac{1}{\beta\Delta t^2}D^n + \frac{1}{\beta\Delta t}\dot{D}^n + \left(\frac{1}{2\beta} - 1\right)\ddot{D}^n\right] \\ + C\left[\frac{\gamma}{\beta\Delta t}D^n + \left(\frac{\gamma}{\beta} - 1\right)\dot{D} + \left(\frac{\gamma}{2\beta} - 1\right)\Delta t\ddot{D}_n\right]\end{aligned} \tag{2.18}$$

where β and γ are two parameters in the Newmark method, and for $\beta = 0.25$ and $\gamma = 0.5$, the method is unconditionally stable and is second-order accurate. Eq. 2.18 is solved using the banded LU decomposition with the Cuthill-Mckee and Gibbs-Poole-Stockmeyer algorithms being used to permute the matrix and reduce the bandwidth to produce the banded matrix.

2.1.3 Contact modeling

Vocal fold contact occurs when the two folds collide with each other to close the glottis. In normal phonation, vocal fold contact usually takes place around the central plane of the glottis due to the left-right symmetric vibrations of the two folds. In chapter 3, vocal fold contact is modeled as follows: two

symmetric contact planes which create a small gap near the central plane are selected, the two vocal folds are forced to stop when they arrive at the location of their corresponding contact planes, and the velocity and acceleration of the nodes on the contact planes are simply set to zero. This is a very simple contact model which ignores the contact force applied on the vocal folds and is employed for the only purpose of simplifying and stabilizing the numerical solution. The small gap saved between the two contact planes is necessary for the success of the fluid solver. In chapter 4, the vocal fold contact is modeled by a penalty coefficient method, in which the contact pressure is calculated as follows:

$$p_{cont} = \gamma k_{max} \Delta d \tag{2.19}$$

where p_{cont} is the contact pressure, γ is the penalty coefficient, k_{max} is the maximum value in the global stiffness matrix K in Eq. 2.18, and Δd is the penetration depth. The contact pressure is proportional to the penetration depth and is always in the direction of preventing further penetration. In the FEA formulation, the contact pressure is converted to equivalent nodal forces which are added to the surface traction term on the right-hand side of Eq. 2.16.

2.2 Fluid dynamics solver

2.2.1 Navier-Stokes equations and finite difference scheme

In chapter 3 and 5, the governing equations of motion for glottal airflow are the 3D unsteady viscous incompressible Navier-Stokes equations:

$$\frac{\partial u_i}{\partial x_i} = 0 \tag{2.20}$$

$$\frac{\partial u_i}{\partial t} + \frac{\partial (u_i u_j)}{\partial x_j} = -\frac{1}{\rho}\frac{\partial p}{\partial x_i} + \nu \frac{\partial}{\partial x_j}\left(\frac{\partial u_i}{\partial x_j}\right) \tag{2.21}$$

where $i, j = 1, 2, 3$, u_i are the velocity components, ρ and ν are the air density and kinematic viscosity, and p is the pressure.

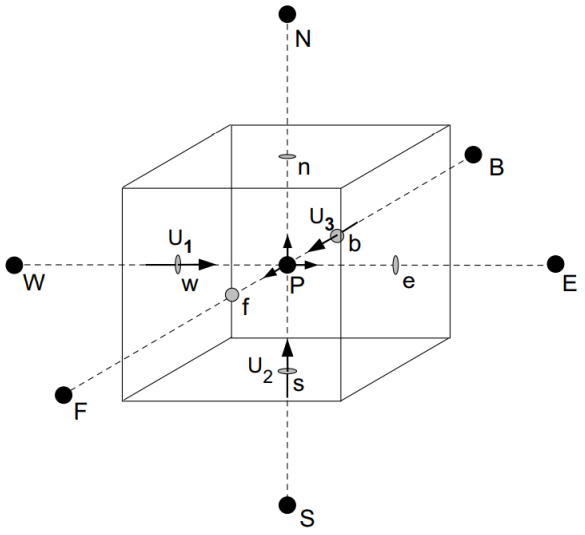

Figure 2.1 Schematic showing the naming convention and location of velocity components in the spatial discretization (from Mittal et al., 2008).

Eq. 2.21 are spatially discretized using a cell-centered, collocated arrangement of the primitive variables (u_i, p), and both the cell-center velocities, u_i, and the face-center velocities, U_i, are computed. The naming convention and location of velocity components is shown in Figure 2.1. The fractional step method of Van-Kan (1986) is used to integrate the equations in time, and it consists of three sub-steps.

In the first sub-step, an intermediate velocity u_i^* is obtained through solving a modified momentum equation. A second-order, Adams-Bashforth scheme and an implicit Crank-Nicolson scheme are respectively used to discretize the convective terms and the diffusion terms. The modified momentum equation is given by

$$\frac{u_i^* - u_i^n}{\Delta t} + \frac{1}{2}\left[3N_i^n - N_i^{n-1}\right] = -\frac{1}{\rho}\frac{\delta p^n}{\delta x_i} + \frac{1}{2}(D_i^* + D_i^n) \tag{2.22}$$

where $N_i = \frac{\delta(U_j u_i)}{\delta x_j}$ and $D_i = \nu \frac{\delta}{\delta x_j}(\frac{\delta u_i}{\delta x_j})$ are the convective and diffusive terms, respectively. $\frac{\delta}{\delta x}$ represents a second-order central difference. A line-SOR scheme is used to solve Eq. 2.22. Next, the intermediate face-center velocities U^* are obtained through taking the average of the corresponding values at the grid nodes. The averaging procedure is as follows

$$\tilde{u}_i = u_i^* + \Delta t \frac{1}{\rho}(\frac{\delta p^n}{\delta x_i})_{cc}$$

$$\widetilde{U}_1 = \gamma_w \tilde{u}_{1P} + (1-\gamma_w)\tilde{u}_{1W}$$

$$\widetilde{U}_2 = \gamma_s \tilde{u}_{2P} + (1-\gamma_s)\tilde{u}_{2S} \tag{2.23}$$

$$\widetilde{U}_3 = \gamma_b \tilde{u}_{3P} + (1-\gamma_b)\tilde{u}_{3B}$$

$$U_i^* = \widetilde{U}_i - \Delta t \frac{1}{\rho}(\frac{\delta p^n}{\delta x_i})_{fc}$$

where γ_w, γ_s and γ_b are the linear interpolation weights for the west, south and back face velocity components, respectively. cc and fc represent cell-center and face-center, respectively.

In the second sub-step, the following pressure correction equation is solved with the constraint that the final velocity u_i^{n+1} is divergence-free.

$$\frac{u_i^{n+1} - u_i^*}{\Delta t} = -\frac{1}{\rho}\frac{\delta p'}{\delta x_i} \tag{2.24}$$

This brings the following Poisson equation for the pressure correction

$$\frac{1}{\rho}\frac{\delta}{\delta x_i}\left(\frac{\delta p'}{\delta x_i}\right) = \frac{1}{\Delta t}\frac{\delta U_i^*}{\delta x_i} \tag{2.25}$$

The last sub-step is to update the pressure and velocity based on the obtained pressure correction

$$p^{n+1} = p^n + p'$$

$$u_i^{n+1} = u_i^* - \Delta t \frac{1}{\rho}(\frac{\delta p'}{\delta x_i})_{cc} \tag{2.26}$$

$$U_i^{n+1} = U_i^* - \Delta t \frac{1}{\rho}(\frac{\delta p'}{\delta x_i})_{fc}$$

2.2.2 Immersed boundary method

The immersed boundary method solves the fluid on an Eulerian coordinate and the structure on a Lagrangian coordinate. Different from the conventional body-conformal grid methods, the immersed boundary method eliminates the time-consuming re-meshing process at each time-step, which greatly simplifies and accelerates the solution procedure.

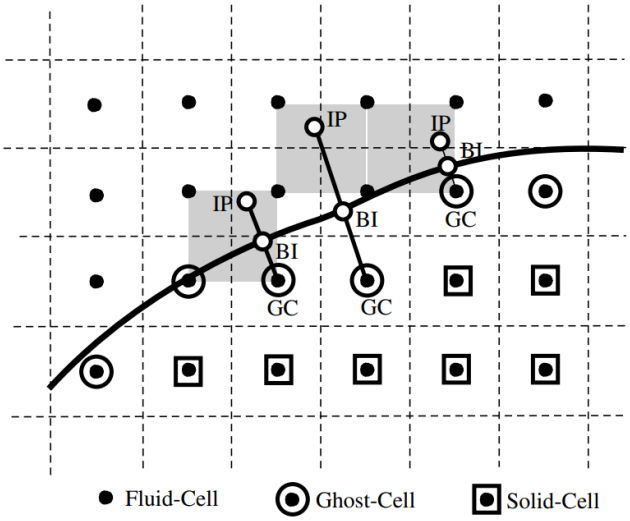

Figure 2.2 Schematic showing the ghost-cell methodology in 2D (from Mittal et al., 2008). The solid curved line and dashed straight lines represent the immersed boundary and the Cartesian grid, respectively. BI: boundary intercept; IP: image-point.

In the current study, a sharp interface immersed boundary method, which employs a multi-dimensional ghost-cell methodology to impose the boundary conditions on the immersed boundary, is used (Mittal et al., 2008). This method has the advantage of dealing with complex 3D stationary or moving boundaries. In this method, the surface of the immersed boundary is represented by an unstructured mesh with triangular elements, which is "immersed" into a Cartesian grid. The ghost-cell methodology is implemented as follows. First, the "solid cells" (cells inside the solid boundary) and "fluid cells" (cells outside the body) are identified in the Cartesian grid. For a specific cell, this is achieved through taking the dot-product between the surface normal vector of the triangular element closest to the cell node and the vector extending from this element to the cell node. The next step is to determine the "ghost cells" which are defined as the cells whose nodes are inside the solid and have at least one fluid neighbor (see Figure 2.2). From the node of each ghost cell, a line segment intersecting normal to the immersed boundary is extended to an "image-point" inside the fluid. The "image-point" is in a location such that the boundary intercept is midway between the ghost-node and the image-point. Next, a bilinear (for 2D) or a trilinear

(for 3D) interpolation is used to calculate the generic variable (ϕ) value at the image-point in terms of the variable values of the surrounding nodes.

$$\phi_{IP} = \sum \beta_i \phi_i \tag{2.27}$$

where i denotes the i_{th} surrounding node and ranges from 1 to 4 (for 2D) or 8 (for 3D). β is the interpolation weight. Along the line segment, the variable value at the ghost-cell is obtained using a linear interpolation which includes the prescribed boundary condition at the boundary intercept. For the velocity variables of Dirichlet boundary conditions, the expression is

$$\phi_{GC} + \sum \beta_i \phi_i = 2\phi_{BI} \tag{2.28}$$

For the Neumann boundary conditions needed in the Poisson equation, the expression is

$$\phi_{GC} - \sum \beta_i \phi_i = -\Delta l \left(\frac{\delta \phi}{\delta n}\right)_{BI} \tag{2.29}$$

where Δl is length of the line segment.

In the moving boundary problems, the position of the immersed boundary is first updated using the following equation

$$\frac{X_i^{n+1} - X_i^n}{\Delta t} = V_i^{n+\frac{1}{2}} \tag{2.30}$$

where X_i are the coordinates of the triangular element vertices, V_i is the vertex velocity, $n, n+\frac{1}{2}$ and $n+1$ are the time levels. Based on the updated immersed boundary location, the process of determining the ghost cells, image-points, body-intercepts and associated weights βs is repeated. Subsequently, the flow Eqs. 2.22-2.26 are marched by one time-step. When the immersed boundary moves to a new location, the cells inside the solid in the previous time-step (time level n) could become fluid cells in the current time-step (time level $n+1$), and these cells are called "fresh-cells". The intermediate velocity in the current time-step for the fresh-cells can be solved using the same interpolation approach as computing the ghost-cell variable values, but with the fresh-cell node being midway between the boundary intercept

and image point. Based on the intermediate velocity, the fresh-cell pressure can be obtained through solving the Poisson equation.

2.2.3 Bernoulli flow model

In chapter 4, the glottal airflow is modeled by a 1D quasi-steady Bernoulli flow which is assumed to separate from the glottal wall at the minimum glottal area. Downstream of the flow separation point, the pressure is assumed to be the atmospheric pressure (zero gauge pressure). Upstream of the flow separation point, the flow pressure (p_i) at a cross section i is computed by applying the following Bernoulli equation

$$p_i = p_{sub} + \frac{1}{2}\rho Q^2\left(\frac{1}{{A_{sub}}^2} - \frac{1}{{A_i}^2}\right) \tag{2.31}$$

where p_{sub} is the pressure at the inlet of subglottal tract (subglottal pressure), ρ is the air density, Q is the flow rate, A_{sub} and A_i are the cross-sectional area at the inlet of subglottal tract and i_{th} cross section respectively. By applying the Bernoulli equation between the inlet of subglottal tract and the minimum glottal area, the flow rate Q is given by

$$Q = \sqrt{\frac{2p_{sub}}{\rho}} A_{min} \tag{2.32}$$

where A_{min} is the minimum glottal area.

2.3 Fluid-structure coupling

In chapter 3 and 4, the solid solver is respectively coupled with the Navier-Stokes solver and the Bernoulli solver to perform the fluid-structure interaction simulations. For the coupling with the Navier-Stokes solver, the immersed boundary method is used to handle the boundary conditions at the fluid-structure interface. When solving the fluid dynamics, no-penetration, no-slip velocity boundary conditions and zero Neumann pressure boundary condition are applied, and the coordinates and velocities of the boundary intercept points can be interpolated from the values of the nodes on the solid surface. When solving the solid dynamics, the traction boundary condition is applied, and the nodal pressure at the interface is

interpolated from the values of the surrounding Cartesian nodes. For the coupling with the Bernoulli solver, an interface mesh which coincides with the solid surface mesh is used to transfer variable values between solid and fluid. The interface mesh tracks the position of the solid surface and only provides boundary position information when solving the fluid dynamics. A similar traction boundary condition as before is applied when solving the solid dynamics, but the nodal pressure is obtained through a linear interpolation between the pressure values of two adjacent cross sections.

Considering the density ratio between the air and vocal fold tissue is around 0.001, the solid solver and fluid solver are explicitly coupled (loose coupling). During each time step, the fluid dynamics is solved with boundary conditions determined from the current vocal fold positions and velocities. Next, the pressure on the vocal folds is obtained/updated through interpolating from flow field to vocal fold surface. Lastly, the solid solver evolves one time-step, and the vocal fold positions and velocities are updated, which provides the new boundary conditions for the fluid in next time step.

CHAPTER 3 EFFECT OF VOCAL FOLD STIFFNESS PARAMETERS ON VOICE PRODUCTION

The study described in this chapter was published in Journal of Voice, Vol 35, Wang, X., Jiang, W., Zheng, X., and Xue, Q., A computational study of the effects of vocal fold stiffness parameters on voice production, 327-e1, Copyright Elsevier (2021).

3.1 Introduction

Continuum-mechanics computer models can have a more realistic representation of the structure and material properties. Especially, a three-dimensional continuum vocal fold model allows to separate the transverse stiffness and longitudinal stiffness parameters; the latter is closely related to muscle activities. With this type of model, Titze and Talkin (1979) found that the fundamental frequency is primarily controlled by the muscular longitudinal stiffness, and the maximum acoustic power and vocal efficiency can be achieved through optimizing the ligament-body longitudinal stiffness ratio. Berry et al. (1994) found that lowering the transverse stiffness of the cover layer could result in subharmonic and chaotic vibrations. Very recently, Zhang (2015, 2016, 2017) performed a series of parametric studies on the effects of vocal fold stiffness and geometric parameters on vocal fold vibrations and voice acoustics by using one-layer and two-layer vocal fold structures. The studies showed consistent results across the models that, except for the obvious effects of stiffness on the fundamental frequency, glottal opening and glottal flow rate, the stiffness parameters had small effects on the vocal fold vibration and acoustics parameters including the closed quotient, sound pressure level, vertical phase difference and spectral shape of the acoustics. Interestingly, the studies further found that while the transverse stiffness still had a noticeable effect on the parameters, the longitudinal stiffness, which is primarily adjusted by the laryngeal muscles, had very minor impacts. Instead, the vertical thickness of vocal fold had a more significant impact than the stiffness parameters. Therefore, their results implied that the laryngeal muscles control the onset, phonation frequency and flow rate by regulating the stiffness conditions and control the vibration pattern

and acoustics primarily by regulating the vertical thickness of the vocal fold. It is a different argument from Hirano's body-cover theory of phonation and other previous studies which showed strong impact of the stiffness parameters on vocal fold vibrations and acoustics. The reason for the differences is unclear. One limitation of the model of Zhang (2015, 2016, 2017) was that the glottal flow dynamics was simplified to be one-dimensional and quasi-steady. Other studies have shown significantly unsteady and three-dimensional features in the glottal flow, such as viscous dissipation, dynamic flow separation and vena contracta, which could have strong effects on the flow-structure interactions and so on vocal fold vibrations and acoustics. In fact, we recently used a fully coupled three-dimensional flow-structure interaction model, which includes a Navier-Stokes equation based unsteady incompressible flow model and a finite-element method based continuum vocal fold model, to study the effect of cover layer depth on vocal fold vibrations. Because the elastic moduli of the cover and body layers were different in the model, changing the cover layer depth also changed the stiffness distribution inside the vocal fold. We found that such stiffness changes have a significant impact on vocal fold vibration patterns because it affects the excitation strength of eigenmodes. Another limitation in Zhang (2015, 2016, 2017) was that the transverse stiffness in the cover and body layers in the two-layer model was assumed to be identical, presumably based on the numerical finding that the difference of the transverse stiffness between the body and cover layers was small across a large range of CT/TA activation conditions. Considering that vocal fold vibration primarily occurs in the transverse plane, the transverse stiffness would have a large impact on vibrations. In fact, Berry et al. (1994) showed that a slight change in the cover transverse stiffness can induce significant changes in vocal fold vibration patterns. Besides, if the spring stiffness in the lumped mass model corresponds to the transverse stiffness, past studies have shown that different body-cover stiffness ratios resulted in very different vibration patterns (Story and Titze, 1995). Therefore, there is a need to further study the impact of stiffness of vocal fold tissues by considering the full complexities of glottal flow dynamics and body-cover transverse stiffness differences.

Clinically, understanding the effect of vocal fold tissue stiffness on voice production is important for both diagnosis and treatment. Many pathological conditions, such as vocal fold paralysis, atrophy and scar, alter the stiffness of vocal fold tissues. Such an understanding can help to clarify how these pathological conditions affect voice production and inform what treatment is needed to restore the normal voice. Surgical procedures, such as injection of biomaterials, are often used to treat voice disorders associated with vocal fold structure and property changes, i.e., to restore the elasticity of vocal fold. To estimate/predict the function and optimal placement of biomaterials, a good understanding of the effects of the stiffness changes on voice production is needed.

This work uses a three-dimensional continuum vocal fold model coupled with a Navier-Stokes equation based unsteady incompressible glottal flow model to understand the effect of stiffness parameters of vocal fold tissues on voice production. The vocal fold is modeled as a three-layer structure consisting of the cover, ligament and body layers. All the three layers are modeled as transversely isotropic materials for which the stiffness parameters include the transverse elastic modulus and longitudinal elastic modulus. Parametric simulations are performed by systematically varying the transverse and longitudinal stiffness parameters of each layer. Important aerodynamic and vocal fold vibration parameters, including the fundamental frequency f_o, peak and average glottal flow rate, open quotient, maximum medial-lateral displacement, vertical displacement, divergent angle and closing velocity are calculated and the effect of the stiffness parameters on these output parameters are analyzed.

3.2 Computational method

Figure 3.1(a) shows the three-dimensional geometry of the vocal fold, larynx, supraglottal and subglottal tract. The shape of the larynx is created based on a laryngeal CT scan of a 30-year-old male subject (Zheng et al., 2009). The supraglottal vocal tract is constructed by superimposing the neutral vocal-tract area function proposed by Story (2005) onto a realistic airway center line measured from in-vivo MRI scans (Story et al., 1996). The lengths of the supraglottal vocal tract and subglottal tract are 17.4 cm and 3.05

cm, respectively. The two vocal folds are symmetric about the glottal center plane. An artificial gap which is about 0.2 mm is enforced during the simulation to avoid the failure of the flow simulations.

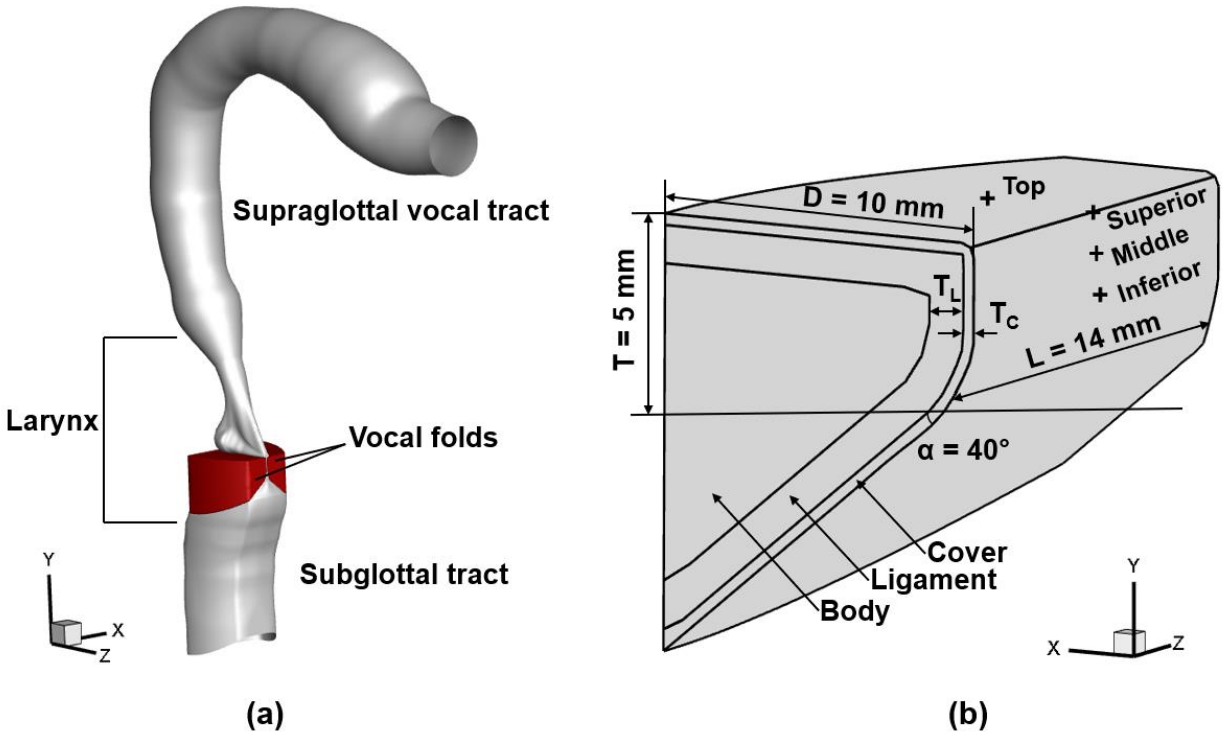

Figure 3.1 (a) The three-dimensional geometry of the vocal fold, larynx, supraglottal and subglottal tract. (b) The dimensions and inner layers of the vocal fold model.

Figure 3.1 (b) shows the dimensions and inner layers of the vocal fold model. The shape is generated based on the mathematical description by Titze and Talkin (1979). The values of the geometric parameters are L = 1.4 cm, T = 0.5 cm, D = 1.0 cm, α = 40°. The thickness of the cover (T_C) and ligament (T_L) layers are assumed to be constant along the longitudinal direction, and their values are taken from the averaged value of the histological measurements (Hirano, 1981), which are 0.33 mm and 1.11 mm, respectively. All the three layers of the vocal fold are modeled as linear elastic, transversely isotropic and nearly incompressible materials. A transversely isotropic material is described by five independent parameters which are the transverse Young's modulus (E), longitudinal Young's modulus (E'), longitudinal shear modulus (G'), transverse Poisson ratio (v) and longitudinal Poisson ratio (v'). Table 3.1 lists the values of these material properties of the three layers of the baseline model. These values are carefully determined

based on the experimental and numerical simulation data reported in previous studies (Min et al., 1995; Alipour et al., 2000; Chan et al., 2007; Cook et al., 2008; Chhetri et al., 2011; Xue et al., 2012). The Poisson ratios are determined by applying the incompressibility assumption (Cook et al., 2008).

Constant pressure boundary conditions are applied at the inlet of the subglottal tract and the outlet of the vocal tract, which are 1 kPa and 0 kPa, respectively. No-penetration, no-slip boundary conditions are applied on the vocal tract walls. For the vocal fold model, zero displacement boundary conditions are imposed on the lateral, anterior and posterior surfaces, while the traction boundary conditions are applied on the medial, top and inferior surfaces. The entire simulation is carried out on a non-uniform 64 × 256 × 192 Cartesian volume grid, and the highest grid density is provided around the glottis. Each vocal fold is discretized by 28997 tetrahedral elements. A small time-step corresponding to 1.014 μs is used for the simulation, which results in around 6000 time-steps per vibration cycle in the baseline case. The simulation is conducted on the XSEDE COMET cluster using 128 processors, and for the baseline case, each vibration cycle takes about 1.5 days. The spatial resolution and temporal resolution are determined based on our previous numerical simulations with similar configurations (Zheng et al., 2010; Zheng et al., 2011a; Xue et al., 2012; Xue et al., 2014).

Table 3.1 Material properties of the vocal fold tissues in the baseline model.

	ρ (g/cm^3)	E (kPa)	ν	E’ (kPa)	G’ (kPa)	ν’	η (poise)
Cover	1.043	2.014	0.9	40	10	0.0	5.0
Ligament	1.043	3.306	0.9	66	40	0.0	7.5
Body	1.043	3.990	0.9	80	20	0.0	12.5

Notes: ρ, tissue density; E, transverse Young’s modulus; ν, transverse Poisson ratio; E’, longitudinal Young’s modulus; G’, longitudinal shear modulus; ν’, longitudinal Poisson ratio; η, damping coefficient.

3.3 Results and discussion

3.3.1 Baseline model

Figure 3.2(a) shows the time history of the glottal flow rate in stable cycles in the baseline model to demonstrate the sustained vibrations. Figure 3.2(b) shows the waveform of the glottal flow rate over one

vibration cycle. The waveform is characterized by a slow rise, a quick drop and a flat portion corresponding to the opening, closing and closure of the glottis, respectively. Note that the non-zero value at the flat portion is leakage due to the enforced artificial gap between the two vocal folds even when the glottis was supposed to be closed. The leakage is about 28% of the maximum flow rate and the enforced artificial gap is about 31% of the maximum glottal opening. Although the incomplete glottal closure was not observed to affect the general convergent-divergent vibratory profile in our simulations, it could influence the vibrations in other aspects or under different conditions. For example, an incomplete glottal closure may result in an early opening of the glottis because the intraglottal pressure does not reduce to zero after the two vocal folds reach the minimum glottal gap. This pressure acts on the medial surfaces of the two vocal folds during the closed phase and could result in an earlier opening of the glottis. In addition, a severe incomplete glottal closure could also result in breathy voice (Sodersten and Lindestad, 1990), lower maximum SPL values and limited capacity to increase the voice intensity (Schneider and Bigenzahn, 2003).

Table 3.2 lists several important parameters calculated from the waveform of the glottal flow rate, including the fundamental frequency f_o, peak flow rate, average flow rate, open quotient and speed quotient. f_o is obtained by taking the Fourier transform of the glottal flow. The average flow rate is calculated by taking the arithmetic mean of the flow rate. The open quotient and speed quotient are defined as the time of the open phase over that of one vibrational cycle and the time of glottal opening over that of glottal closing, respectively. The peak flow rate, open quotient and speed quotient are the averaged values over the stable cycles. These parameters are found well within the reported physiological ranges (Holmberg et al., 1988; 1989; Titze, 2000).

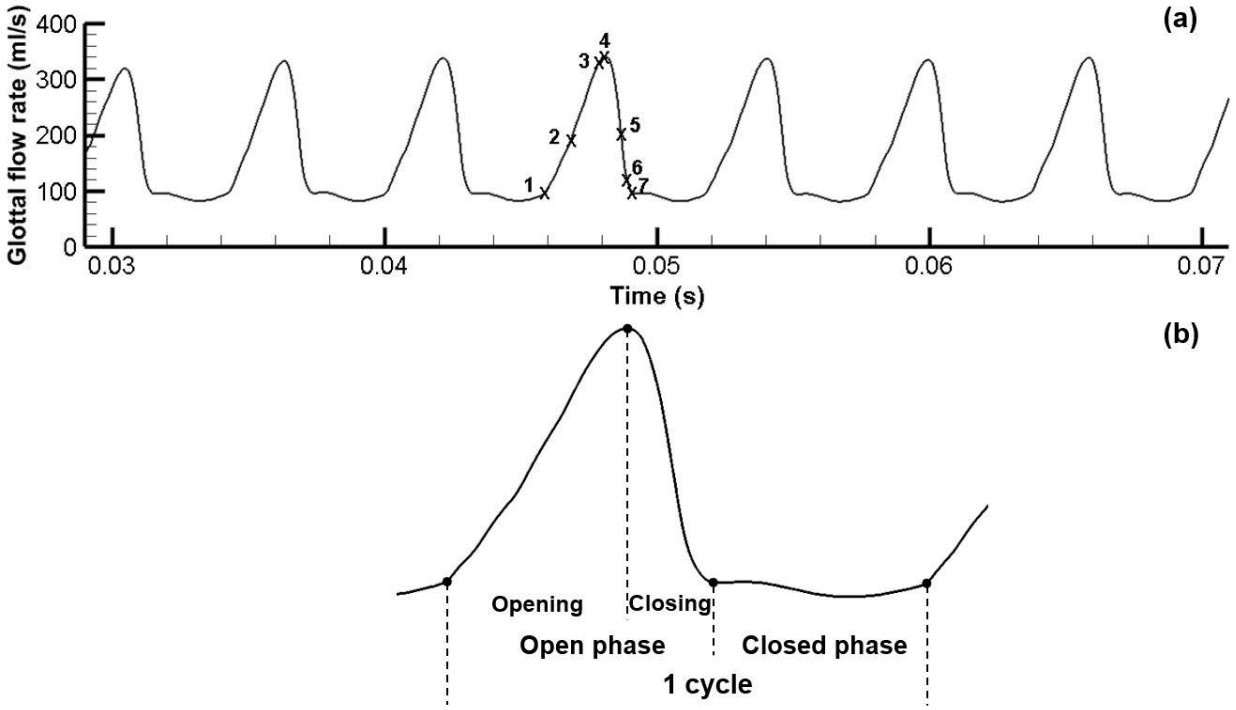

Figure 3.2 (a) The time history of the glottal flow rate of the baseline case. (b) The waveform of the glottal flow rate over one cycle.

Figure 3.3 shows the top view and mid-coronal profile of the vocal folds at seven time instants over a vibration cycle. The seven time instants are marked on Figure 3.2(a). A phase lag between the airflow and glottal opening can be observed by comparing time instant 3 and 4, which separately corresponds to the maximum glottal opening and maximum flow rate, as well as time instant 6 and 7, which represent the start of the minimum glottal opening and start of the minimum flow rate, respectively. It needs to be pointed out that the data for plotting vocal fold profiles were output every 200 steps in our numerical simulation, the time instant when the bottom parts of the vocal folds just contact is missed. Time instant 6 is closest to that time instant. The phase lag is due to the air inertia effect which delays the build-up and decline of the glottal flow rate. In Figure 3.3, the vocal folds show a typical alternating convergent-divergent vibration pattern. During the opening phase, the vocal folds are gradually pushed apart by the airflow from the inferior and form a convergent glottal shape, while during the closing phase, the vocal folds adduct with the inferior leading the superior due to the restoring force, and a divergent glottal shape

is built. This type of vibration is believed to be essential for the self-sustained oscillations of the vocal folds because it generates a temporally asymmetric pressure inside the glottis which leads to positive net energy transfer from the airflow to vocal fold tissues (Titze, 2000).

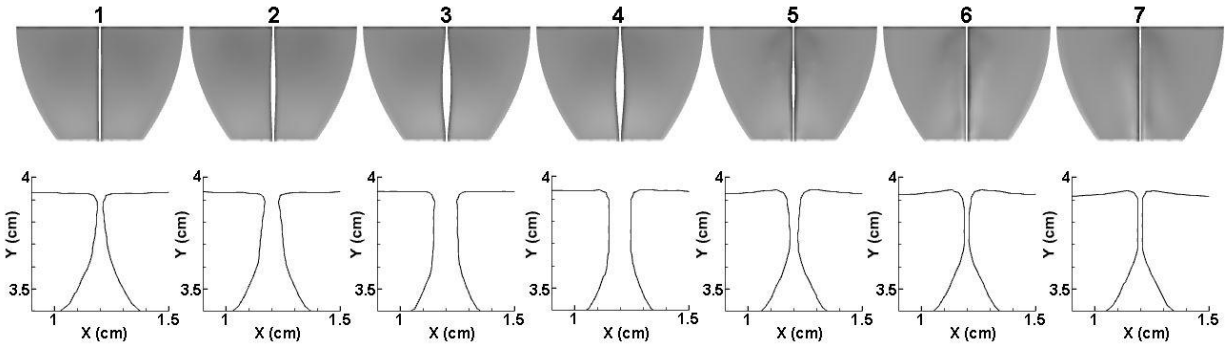

Figure 3.3 The top view (top row) and mid-coronal profile (bottom row) of the vocal folds at seven time instants over one vibration cycle. The 7 time instants are marked on Figure 2(a), in which time instant 3 corresponds to the maximum glottal opening and time instant 6 is the start of the minimum glottal opening.

In summary, the simulation of the baseline model captures a typical glottal flow waveform, flow rate, and vocal fold vibratory dynamics, indicating a reasonable representation of human phonation. It serves as a good foundation for the following parametric simulations.

Table 3.2 Several parameters calculated from the waveform of the glottal flow rate of the baseline case.

f_o (Hz)	Peak flow rate (ml/s)	Average flow rate (ml/s)	Open quotient	Speed quotient
169.6	336.6	159.6	0.56	1.97

3.3.2 Parametric simulation conditions

In the parametric simulations, the transverse elastic modulus (E) and longitudinal elastic modulus (E') of each tissue layer are systematically varied by changing one parameter and fixing all other parameters at a time. For each parameter, five values (0.25, 0.5, 1.0, 2.0 and 4.0 times of the baseline value) are used, which is the result of the compromise between computational time and covering the physiological range of each parameter as much as possible. The longitudinal shear modulus (G') is assumed to vary proportionally to the longitudinal Young's modulus (E'). For the convenience of description, the subscripts

C, L and B are used to distinguish the cover, ligament and body layers. For example, E_C and E_L represent the transverse stiffness of the cover and ligament layer, respectively. Each simulation is kept running until the vibration becomes sustained and stable. Data processing and analysis are performed based on the stable cycles. Because vocal fold vibration becomes chaotic in some cases, the statistical data is difficult to calculate. These cases are not included in the analysis.

3.3.3 Effects on natural frequencies

Figure 3.4 plots the relative change of the four lowest natural frequencies of the vocal fold versus the relative change of each stiffness parameter. The relative change is calculated by using the baseline parameters as the reference values. The modal number in the figure is arranged in a way that the natural frequency increases with the modal number. It shows that, among all the stiffness parameters, the ligament longitudinal stiffness has the largest effect on the frequencies, and the cover transverse stiffness almost has no effect. The effect of the longitudinal stiffness is much more significant than the transverse stiffness in each layer.

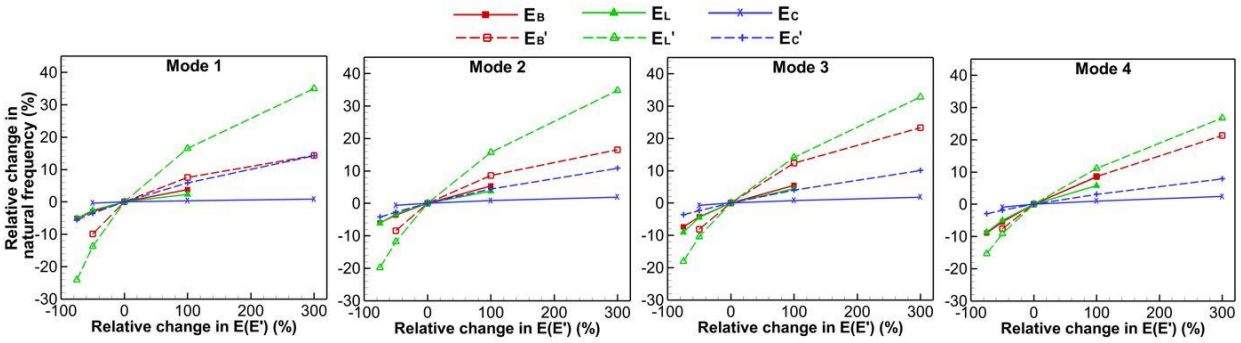

Figure 3.4 Relative change of the natural frequencies versus relative change of the stiffness parameters of different layers.

Interestingly, it is noticed that the effect of the longitudinal stiffness is more significant on the lower frequencies than on the higher frequencies and this phenomenon is most visible in the ligament layer. For example, a 375% (-75% to 300%) increase in the ligament longitudinal stiffness increases the natural frequency of mode 1, mode 2, mode 3 and mode 4 by 59.2%, 54.7%, 50.8% and 42.1%, respectively. It is

similar for the body and cover layers but with smaller differences. Because the fundamental frequency is close to the lower natural frequencies during vibrations, the results imply that varying the longitudinal stiffness is more effective in controlling the fundamental frequency. In contrast, the transverse stiffness has a larger impact on higher natural frequencies, although the impact is generally minor. For example, a 175% (-75% to 100%) increase in the body transverse stiffness increases the natural frequency of mode 1, mode 2, mode 3 and mode 4 by 8.9%, 11.4%, 12.9% and 17.4%, respectively.

3.3.4 Effects on fundamental frequency f_o

Figure 3.5 plots the relative change of f_o versus the stiffness parameters. As expected, f_o exhibits positive relationships with all the stiffness parameters. For each layer, the longitudinal stiffness has a larger impact than the transverse stiffness, although the degree of difference is small in the cover and body layers. In the ligament layer, a 150% (-50% to 100%) change in the longitudinal and transverse stiffness results in about 20.5% and 3.6% change in f_o, while it is 3.8% and 1.5% in the cover layer and 16.0% and 10.9% in the body layer. The longitudinal stiffness playing a more important role in controlling f_o agrees with the physiological observation that the higher-pitched phonation is achieved primarily through the activation of the cricothyroid (CT) muscle which greatly tenses the vocal fold in the longitudinal direction (Flanagan et al., 1976; Atkinson, 1978; Roubeau et al., 1997). Among all the stiffness parameters, the ligament longitudinal stiffness has the most dominant effect on f_o. By comparing different layers, it is found that, in the transverse plane, the body layer stiffness has the most significant effect, while in the longitudinal direction the ligament layer stiffness has the most significant effect. In both the transverse plane and longitudinal direction, the cover layer stiffness has the smallest impact. These different effects on f_o indicates that there exists different f_o adjusting patterns between the transverse plane and longitudinal direction. It is worth mentioning that the f_o control pattern in the transverse plane to some extent agrees with Zhang (2009), who employed a body-cover isotropic vocal fold model and found that the body stiffness is more effective in controlling phonation frequency when the body-cover stiffness ratio is

smaller than 10. Considering the maximum body-ligament transverse stiffness ratio is not larger than 10 and the transverse Young's modulus in this study is similar to the stiffness parameter in an isotropic model, the consistency between our result and that of Zhang (2009) is not surprising.

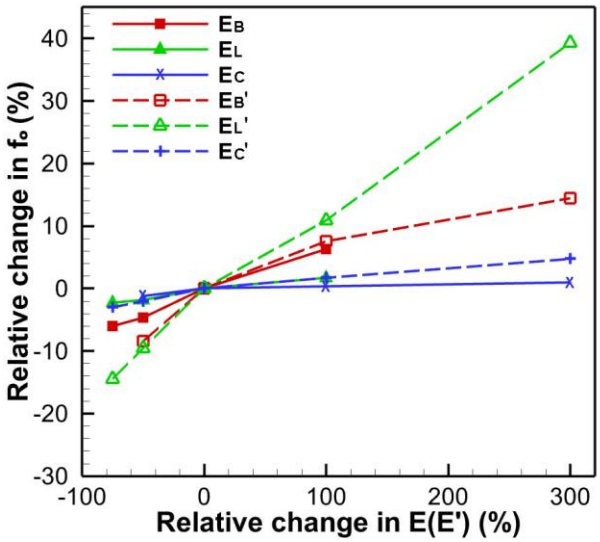

Figure 3.5 Relative change of f_o versus relative change of the stiffness parameters of different layers.

3.3.5 Effects on flow rate

Figure 3.6 and 3.7 plot the relative change of the average and peak flow rate versus the stiffness parameters. In general, the flow rates decrease with increasing stiffness. Similar to the effect on the f_o, the ligament longitudinal stiffness is most effective in controlling the flow rates. It is also interesting to note that, in contrast to the minor effect of the cover stiffness on f_o, the cover longitudinal stiffness has a much larger impact on the flow rates. An inconsistent change of the peak flow rate and average flow rate is also observed as the cover transverse stiffness decreases from the baseline value to 0.5 times of that, the peak flow rate decreases by 6.7% while the average flow rate slightly increases by 1.0%. A close look at the flow rate waveform reveals that this inconsistency is because the maximum glottal opening decreases but the open phase is longer.

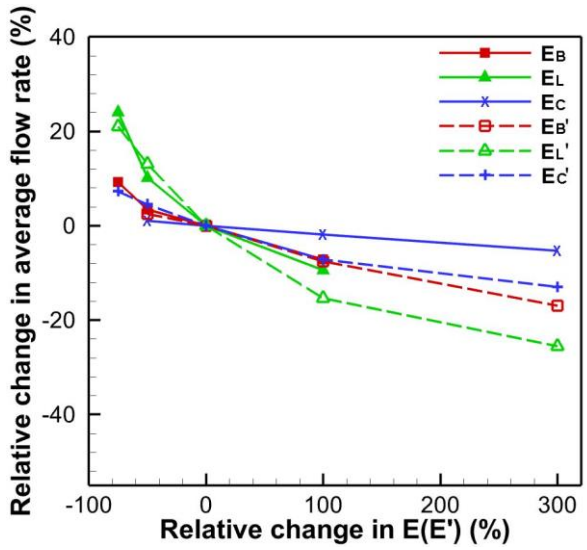

Figure 3.6 Relative change of the average flow rate versus relative change of the stiffness parameters of different layers.

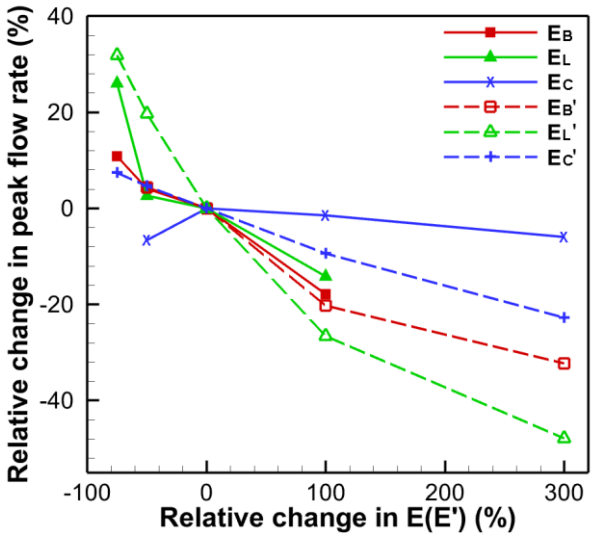

Figure 3.7 Relative change of the peak flow rate versus relative change of the stiffness parameters of different layers.

3.3.6 Effects on open quotient

Figure 3.8 plots the absolute change of the open quotient versus the relative change of each stiffness parameter. The absolute change is calculated by subtracting the open quotient value of the baseline case from that of each case. In contrast to the consistent effects of the stiffness parameters on f_o and flow rates, the effect of these stiffness parameters on open quotient is more complex and shows inconsistency between the layers. For example, the open quotient generally increases with the increasing body layer stiffness but decreases with the increasing cover and ligament layer stiffness, although it slightly decreases when the body longitudinal stiffness increases from 2.0 to 4.0 times of the baseline value. It is also noted that some of the stiffness parameters, including the ligament transverse stiffness, cover transverse stiffness and body longitudinal stiffness, have a much larger impact on the open quotient when their values are small, i.e., below the baseline value.

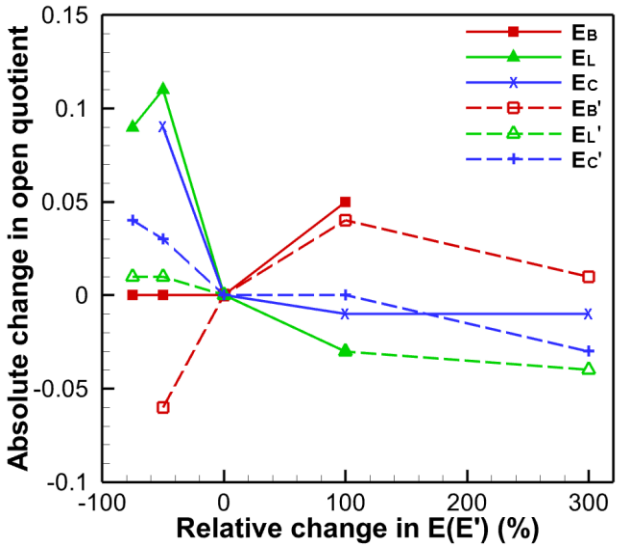

Figure 3.8 Absolute change of the open quotient versus relative change of the stiffness parameters of different layers.

3.3.7 Effects on vocal fold vibration

3.3.7.1 Medial-lateral displacement

Figure 3.9 and 3.10 show the relative change of the maximum medial-lateral displacements of the superior and inferior edge [Figure 3.1(b)] of the mid-coronal plane of the vocal fold versus the stiffness parameters. Very like the flow rates, increasing the stiffness parameters decreases the medial-lateral displacement, and the ligament layer stiffness parameters have the most significant influence. Further analysis shows that the influence on the superior and inferior displacements are at different degrees. For example, a 375% (-75% to 300%) change in E_L' and E_C' decreases the superior displacement by 129.0% and 44.2%, respectively, while it decreases the inferior displacement by 73.3% and 31.6%, respectively, which separately represents 55.7% and 12.6% difference of the influence on the superior and inferior displacement. It needs to be pointed out that the above finding does not apply to the lower range of cover longitudinal stiffness, i.e., relative change from -75% to 0, in which both the superior and inferior displacements vary a little, and the influence difference is not evident. A significant difference in the influence (15.8%) is also observed as E_L decreases from the baseline value to 0.5 times of that. Varying

other stiffness parameters also have different influences between the superior and inferior edges, but the differences are relatively small. The different influences on the superior and inferior edges could contribute to the change of convergent and divergent angles, which will be discussed later.

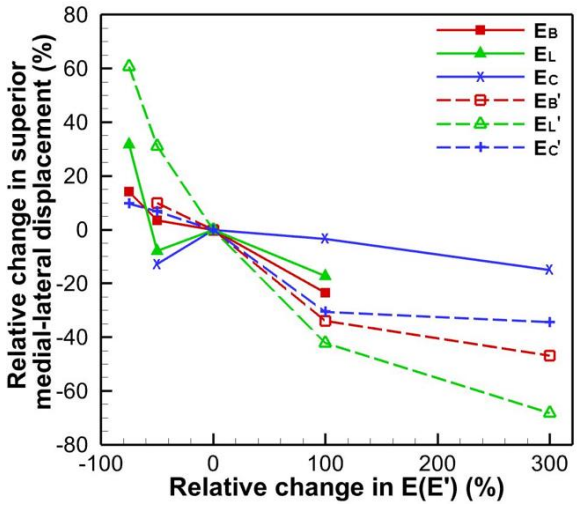

Figure 3.9 Relative change of the superior medial-lateral displacement versus relative change of the stiffness parameters of different layers.

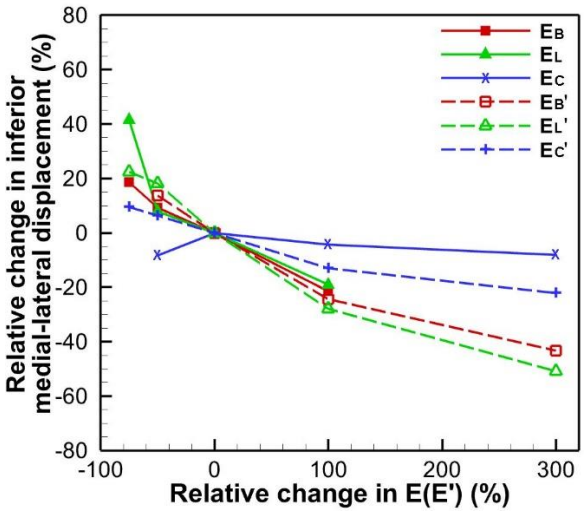

Figure 3.10 Relative change of the inferior medial-lateral displacement versus relative change of the stiffness parameters of different layers.

Figure 3.11 shows the relative change of the maximum vertical displacement of the point on the top surface [Figure 3.1(b)] of the vocal fold versus the stiffness parameters. One major observation is that the transverse stiffness parameters have a much smaller effect than the longitudinal stiffness parameters. As an illustration, a 175% (-75% to 100%) increase in E_L decreases the maximum vertical displacement by 28.4%, whereas increasing E_L' with the same percentage results in a decrease of 101.2%. This significant effect of the longitudinal stiffness on the vertical displacement is consistent with Zhang (2016), which reported that the large vertical deformation of the vocal fold can be effectively suppressed by increasing the anterior-posterior stiffness.

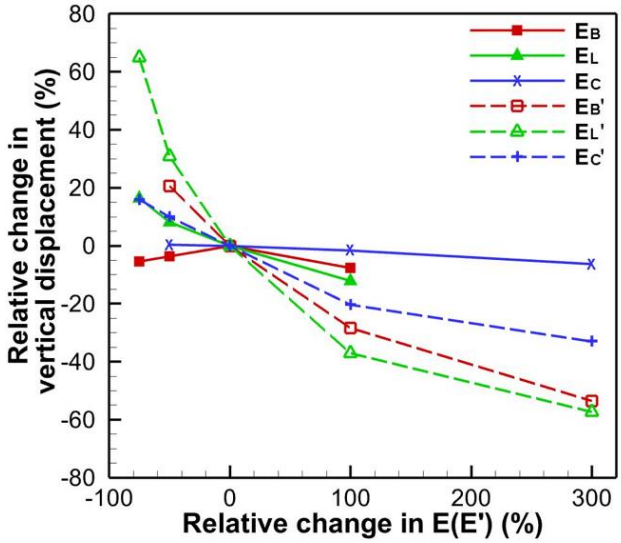

Figure 3.11 Relative change of the vertical displacement versus relative change of the stiffness parameters of different layers.

3.3.7.2 Divergent angle

Figure 3.12 shows the relative change of the maximum divergent angle during glottal closing versus stiffness parameters. The maximum divergent angle is calculated as follows, on each vocal fold, three points (superior, middle and inferior) located on the medial edge of the mid-coronal plane are selected. At each time instant, the angle between the lines passing through superior and middle point on left and right vocal fold are calculated, as well as the angle between the lines passing through superior and inferior point. The relatively larger angle is chosen as the output value, and the maximum divergent angle is determined by choosing the maximum value from all time instants during the closing phase. As can be seen in Figure 3.12, increasing E_L' from the smallest to the largest value causes the most significant decrease of the divergent angle – from 91.3% to -82.7%. It is also observed that increasing E_C' from the minimum to the maximum value causes the divergent angle to vary in a relatively wider range than the other cases (except E_L'). In fact, increasing E_L' and E_C' were found to induce a more significant decrease of the superior medial-lateral displacement than the inferior medial-lateral displacement in the previous

analysis. Therefore, the important effect of E_L' and E_C' on the maximum divergent angle is very likely due to their different influences on the superior and inferior displacement. This correlation can be further verified by checking the cases of increasing E_C and E_B. In Figure 3.12, as E_C increases from the baseline value to 4.0 times of that, the maximum divergent angle has a decrease of 61.4%, and the corresponding decreases of the superior and inferior medial-lateral displacement are 14.9% and 8.1% respectively. The relatively larger decrease of the superior medial-lateral displacement leads to the decrease of the maximum divergent angle. As E_B increases from 0.25 to 1.0 times the baseline value, the maximum divergent angle increases by 56.7%, and the corresponding decreases of the superior and inferior medial-lateral displacement are 14.2% and 18.7% respectively. A larger decrease of the inferior medial-lateral displacement results in the increase of the maximum divergent angle. It needs to be pointed out that the divergent angle is formed by the inferior part of the vocal fold leading the superior part during the closing phase, the closing phase difference between the inferior and superior part also plays an important role in forming the divergent angle. However, in the current cases, the maximum divergent angle seems to be closely related to the different influences on the superior and inferior medial-lateral displacement caused by varying the vocal fold stiffness.

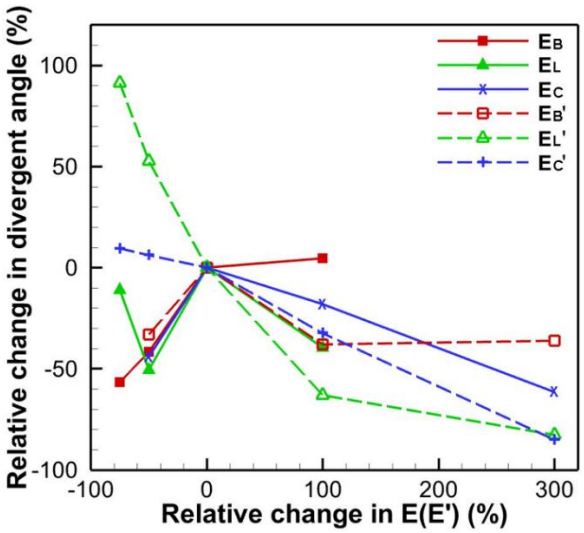

Figure 3.12 Relative change of the divergent angle versus relative change of the stiffness parameters of different layers.

3.3.7.3 Maximum closing velocity

The closing velocity of the vocal fold is believed to determine the deceleration of the glottal flow, as well as the aerodynamic parameters that are related to voice quality, such as MFDR. The point located at the inferior medial edge of the mid-coronal plane [Figure 3.1(b)] is selected to compare the closing velocity of the vocal fold in different cases. Figure 3.13 shows the relative change of the maximum closing velocity versus the stiffness parameters. For the increasing longitudinal stiffness, a general decrease of the maximum closing velocity is observed, and the decrease is more significant with increasing E_L'. The decrease of the maximum closing velocity with the increase of the longitudinal stiffness may seem counterintuitive, because the increase of the stiffness is supposed to increase the restoring force and thus increase the closing velocity. But it should be noted that the lateral displacement is decreased, and so is the strain in the medial-lateral direction. Therefore, the final restoring force should be the balanced result

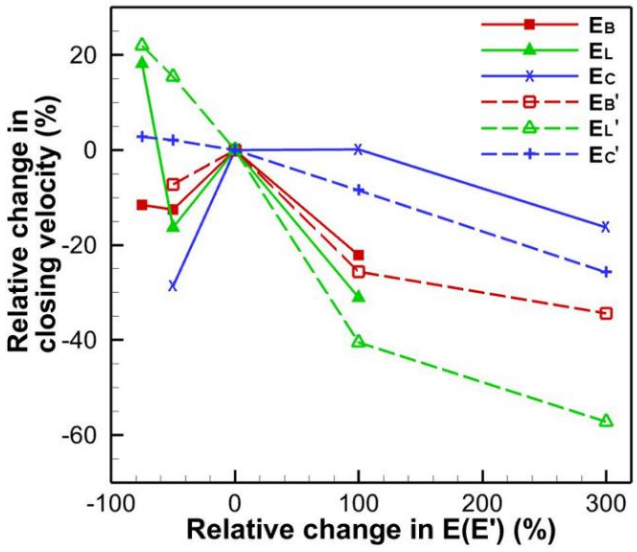

Figure 3.13 Relative change of the closing velocity versus relative change of the stiffness parameters of different layers.

of the increased stiffness and the decreased strain. The decreased maximum closing velocity observed here is probably because the effect of the latter exceeds the former. On the other hand, if the increased stiffness plays a more important role, an increased restoring force or closing velocity is also possible with

increasing the stiffness. This might explain the rise of the maximum closing velocity observed in Figure 3.13 when the value of E_B' is small. For increasing the transverse stiffness, the effect is more complex and non-monotonic.

3.3.8 Sensitivity analysis on all the output measures

To evaluate the sensitivity of the output parameters to each stiffness parameter, sensitivity coefficients are calculated and listed in Table 3.3. The sensitivity coefficient is defined as the relative (absolute) change of the output measure divided by the relative change of each stiffness parameter, which is essentially the slope of the curves in Figure 3.5 to 3.13. For each stiffness parameter, the sensitivity coefficient is the root-mean-square values of the piece-wise slopes of the corresponding curves in Figure 3.5 to 3.13. The three highest values in each column of Table 3.3 are highlighted, and they indicate that the output measure is most sensitive to these stiffness parameters, or in other words, these stiffness parameters can cause significant changes in the output measure. Here, it needs to be pointed out that if the variation of the output measure is monotonic, a larger sensitivity coefficient would suggest a wider range of change. On the other hand, if the variation of the output measure is non-monotonic, a very sharp local change could also result in a large sensitivity coefficient, but it does not necessarily mean a wide range of change in the output measure. Take the divergent angle for example, the largest sensitivity coefficient corresponds to varying E_L, and this is due to the jagged variation of divergent angle, however, Figure 3.12 clearly shows varying E_L' causes the widest range of change in divergent angle. This consideration is required for interpreting the data in Table 3.3. Overall, the ligament stiffness parameters including both the transverse and longitudinal stiffness have the most dominant effect on most output measures. In comparison, the body and cover stiffness are only dominant in a few output measures. It is also noticed that while the cover layer stiffness overall has a very small effect on the output measures, varying the cover transverse stiffness can cause significant changes in the open quotient and closing speed, indicating that it still has important effects on vocal fold vibrations.

Table 3.3 Sensitivity coefficients between stiffness parameters of different layers and acoustic and vibration output measures (boldface values in each column indicate that the output measure is most sensitive to these stiffness parameters).

	f_o	Flow rate average/peak	Open quotient	Medial-lateral displacement superior/inferior	Vertical displacement	Divergent angle	Closing velocity
E_B	**0.073**	**0.146/0.192**	0.0003	**0.286/0.271**	0.073	**0.594**	0.194
E_L	0.026	**0.345/0.550**	**0.0014**	**0.927/0.792**	**0.224**	**1.103**	**0.841**
E_C	0.014	0.019/0.079	**0.0010**	0.154/0.099	0.017	0.535	**0.334**
E_B'	**0.109**	0.059/0.132	**0.0007**	0.231/0.218	**0.298**	0.440	0.172
E_L'	**0.164**	**0.222/0.344**	0.0002	**0.706/0.251**	**0.772**	**0.988**	**0.290**
E_C'	0.030	0.081/0.092	0.0004	0.177/0.113	0.187	0.227	0.065

3.4 Summary

The purpose of this study is to investigate the effect of the stiffness parameters of vocal fold tissues on glottal flow and vocal fold vibrations during voice production. A three-dimensional cover-ligament-body continuum model of human vocal fold coupled with the Navier-Stokes equation based glottal flow model is used to simulate the flow-induced vocal fold vibrations. Parametric simulations are performed by systematically varying the transverse and longitudinal stiffness parameter of each layer of the vocal fold. The effects of these variations on the important voice quality related parameters, which include f_o, peak flow rate, average flow rate, open quotient, glottal opening, vertical displacement, divergent angle and closing velocity, are carefully analyzed. The sensitivity of the output measures to each stiffness parameter is quantified. Overall, these output measures show variant sensitivities to the stiffness parameters in different layers. The main findings are summarized below.

1) f_o is primarily controlled by the longitudinal stiffness, and to a much lesser degree by the transverse stiffness. Among all the stiffness parameters, the longitudinal stiffness of the ligament has the most significant effect on f_o. This f_o control pattern is consistent with the physiological observation that the great longitudinal tension provided by the contraction of CT muscle is necessary for the higher-pitched phonation.

2) The glottal flow rates generally decrease as the stiffness parameters increase, and the ligament stiffness parameters (both transverse and longitudinal) have the most significant effect among the three layers. It is also interesting to notice that, in contrast to the minor effect of the cover stiffness on f_o, the cover longitudinal stiffness has a much larger impact on the flow rates.
3) The stiffness of different tissue layers shows inconsistent effect on the open quotient. In general, the open quotient increases with the increasing body layer stiffness and decreases with the increasing cover and ligament layer stiffness.
4) Very like the flow rates, increasing the stiffness parameters decreases the medial-lateral displacement of the vocal fold, and the ligament layer stiffness parameters have the most significant influence. Furthermore, the stiffness parameters show different degrees of influence on the superior and inferior parts of the vocal fold. Increasing the longitudinal stiffness of the outer layer (ligament and cover) is found to decrease the superior maximum medial-lateral displacement more significantly than the inferior.
5) The vertical displacement of the vocal fold decreases as the stiffness increases. Furthermore, it is much more significantly influenced by the longitudinal stiffness than the transverse stiffness.
6) The stiffness parameters show a complex effect on the divergent angle and the complexity is closely related to their different degrees of influence on the medial-lateral displacement of the superior and inferior edge of the vocal fold. Overall, varying the ligament and cover longitudinal stiffness can cause a greater range of changes in the divergent angle than other stiffness parameters.
7) On the whole, the maximum closing velocity decreases as the stiffness parameters increase. However, non-monotonic relationships are also observed. These are mainly due to the combined effect of the stiffness on the medial-lateral displacement which determines the strain and further determines the elastic restoring force.

8) In terms of the overall sensitivity, the ligament stiffness parameters including both the transverse and longitudinal stiffness have the most dominant effect on most output measures. In particular, varying the ligament longitudinal stiffness can cause the widest range of change in most output measures. Moreover, while the cover layer stiffness overall has small effects on the output measures, the cover transverse stiffness still has important effects on the open quotient and glottal closing speed.

One limitation of the current study is the linear material assumption, which is unrealistic because vocal fold tissues exhibit great nonlinearity in experiments (Min et al., 1995; Kelleher et al., 2013a). Considering the maximum strain in some cases in the current study is around 30%, future studies might need to include the nonlinear material models to represent the deformation of the vocal fold more precisely. Another limitation is that the study is based on one set of the geometric parameters of the vocal fold and inner structures and one subglottal pressure. Different choices of these parameters may result in different system behaviors. For example, changing the subglottal pressure or the thickness of the inner layers might result in different vibration patterns which may lead to different observations in the output parameters. Additionally, the same vocal tract shape throughout the study may limit the findings to one specific speech sound. In order to draw a more general conclusion, a larger parametric space will be needed in future studies.

CHAPTER 4 INFLUENCE OF FIBER ORIENTAITON OF CONUS ELASTICUS IN VOCAL FOLD MODELING

4.1 Introduction

The conus elasticus, also known as the cricothyroid ligament, cricothyroid or cricovocal membrane, denotes the fibroelastic layer that lies under the mucosal of the subglottic region of the larynx (Williams et al., 1989). While the median part of the conus elasticus connects the anterior arch of cricoid cartilage to the inferior rim of thyroid cartilage and does not directly affect vocal fold vibration, the lateral parts of the conus elasticus extend from the upper rim of the cricoid cartilage to the inferior edge of the vocal ligament, of which the structure could have influence on vocal fold vibration. Due to their firm anchorage to the cricoid cartilage, the lateral parts of the conus elasticus are believed to be important for restricting the vertical movements of the vocal folds (Reidenbach, 1996).

The tissue biomechanical anisotropy is known to be closely related to the orientation of tissue fibers (Namani et al., 2012; Kelleher et al., 2013a). The conus elasticus is generally considered to be composed of densely arranged elastic and collagen fibers (Prithishkumar and Felicia, 2014; Sato, 2018). As the fiber orientations in the conus elasticus are primarily along the caudal-cranial direction (Reidenbach, 1996), the anisotropy of the conus elasticus is believed to be different from that of the vocal ligament, in which the fibers are aligned predominantly in the anterior-posterior direction (Gray et al., 2000; Miri et al., 2012). However, the role of this material anisotropy difference in vocal fold vibration and voice production has seldom been investigated. While the true material anisotropy of conus elasticus is naturally included in the in vivo/excised larynx, most of current computer or physical (synthetic) vocal fold models have largely overlooked this material anisotropy difference. The conus elasticus layer was either simply not considered or integrated with the ligament layer with the same material anisotropy (Alipour et al., 2000; Tao and Jiang, 2006; Luo et al., 2008; Murray and Thomson, 2012; Murray et al., 2014; Migimatsu and Tokuda, 2019). A number of mechanics studies have shown that fiber orientation can play a significant role in determining the mechanical properties of anisotropic materials (Namani et al., 2012; Martin and

Boardman, 1993; Lee et al., 2010). Therefore, it is of interest to study how the fiber-orientation-related material anisotropy would affect the stiffness of conus elasticus and further affect the biomechanical response of vocal folds during the interaction with glottal aerodynamic forces. Furthermore, if the material anisotropy of conus elasticus has a significant influence on vocal fold vibration, it is worth investigating how it will affect the aerodynamic and acoustic parameters that reflect voice type and quality.

In this work, we study the influence of material anisotropy of conus elasticus on the flow-induced vibration of a three-dimensional computational vocal fold model. One focus of this work is to understand the relationship between the mechanical anisotropy and the structural stiffness of the conus elasticus. It is expected that such knowledge would help explain the vibratory characteristics of vocal fold observed in the fluid-structure interaction simulation. Output measures including vocal fold motion, flow rate waveform, aerodynamic and acoustic parameters are analyzed to determine the effect of conus elasticus on vocal fold vibration and voice production. The findings of this study could provide deeper insights into the mechanism of vocal fold vibration and contribute to the development of more accurate voice simulation models.

4.2 Methods

4.2.1 Vocal fold and conus elasticus model

In this study, the geometric shape of the vocal fold is based on the mathematical description by Titze and Talkin (1979). The dimensions and layered structure of the vocal fold model are shown in Figure 4.1(a). In the model, the cover layer is the outermost layer covering the entire vocal fold. The ligament and conus elasticus are located between the cover and body layer. The ligament layer starts at the superior margin of the medial surface and connects to the conus elasticus at the inferior margin of the medial surface. The conus elasticus goes through the subglottic region and ends at the inferior aspect of the lateral surface of vocal fold. The thickness of the cover and ligament layers are determined from the averaged value of the

histological measurements (Hirano, 1981), which are 0.33 mm and 1.11 mm, respectively. The thickness of the conus elasticus in the coronal plane is assumed to be the same as that of the ligament. Along the anterior-posterior direction, the thickness of the cover, ligament and conus elasticus are assumed to be invariant.

All the layers in the model are modelled as nearly incompressible and transversely isotropic linear elastic materials. For the cover, ligament and body layer, the transverse isotropy is in the coronal plane, and the material properties of the three layers are taken from previous studies (Alipour et al., 2000; Zheng et al., 2011a). To investigate the effect of anisotropic direction of the conus elasticus, two vocal fold models with different anisotropic directions in the conus elasticus are generated: In model 1, the fiber orientation in conus elasticus is consistent with the histologic observation which is aligned with the inferior angle of the vocal fold (40 degrees with respect to the horizontal plane in current model); in model 2, it is assumed that the fiber orientation in conus elasticus is the same as that in the ligament layer, which is along the anterior-posterior direction. Due to the lack of measurements of the material properties of conus elasticus and the fact that the conus elasticus has similar composition of elastic and collagen fibers as the ligament (Gray et al., 2000; Prithishkumar and Felicia, 2014), the material parameters of conus elasticus are assumed to be the same as those in the ligament layer. The adopted material parameters of each vocal fold layer are listed in Table 4.1.

Table 4.1 Material properties of each vocal fold layer

	ρ (g/cm^3)	E (kPa)	ν	E’ (kPa)	G’ (kPa)	ν’	η (poise)
Cover	1.043	2.014	0.9	40	10	0.0	5.0
Ligament	1.043	3.306	0.9	66	40	0.0	7.5
Conus elasticus	1.043	3.306	0.9	66	40	0.0	7.5
Body	1.043	3.990	0.9	80	20	0.0	12.5

Notes: ρ, tissue density; E, transverse Young’s modulus; ν, transverse Poisson ratio; E’, longitudinal Young’s modulus; G’, longitudinal shear modulus; ν’, longitudinal Poisson ratio; η, damping coefficient.

4.2.2 Fluid-structure interaction model and simulation setup

The vocal fold vibrations are assumed to be symmetric about the glottal midplane so that only the left vocal fold is simulated in the current study. The glottal midplane is set at x = 0.0 mm, which creates a nearly zero initial glottal gap. The anterior, posterior and lateral surfaces of the vocal folds are fixed to mimic the attachment to the thyroid and arytenoid cartilage, while the remaining surfaces are free to move. The vocal fold is discretized into 9767 ten-node quadratic tetrahedral elements, and grid independence is achieved with this mesh. Along the flow direction, the glottis is discretized into 100 equidistant sections where the Bernoulli's equation is solved and glottal pressures are evaluated. Air density of 1.225 kg/m^3 is used. For each vocal fold model, five FSI simulations with five different subglottal pressures (0.2, 0.4, 0.6, 0.8 and 1.0 kPa) are conducted. A small time-step of 2.27e^{-5} s is used for all the simulations. Each simulation runs for a period of 0.5 s, which is sufficient to pass the transient state and reach the limit-cycle steady-state vibration.

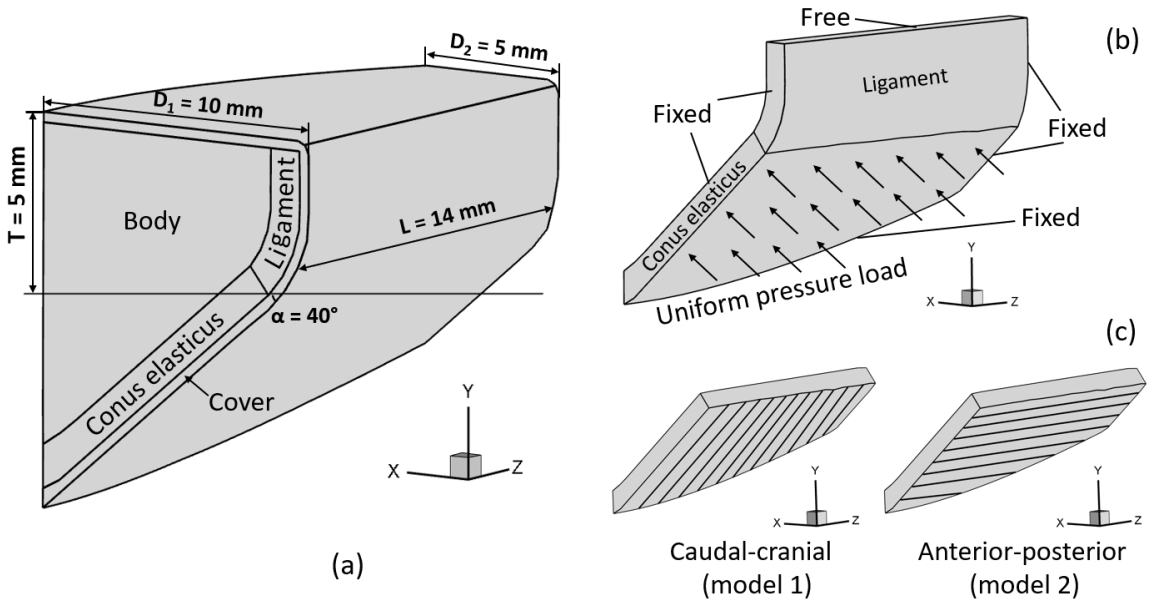

Figure 4.1 (a) The dimensions and layered structure of the vocal fold in the simulation. (b) Loading and boundary conditions in the finite element model of conus elasticus. (c) Caudal-cranial (left) and anterior-posterior (right) fiber orientation in the conus elasticus.

4.3 Results and discussion

4.3.1 Effect of material anisotropy on stiffness of conus elasticus

The conus elasticus is cranially connected to the bottom of the vocal ligament. Its stiffness will affect the pliability of the ligament by providing constraints at the junction of the two tissues, which can further change the overall stiffness of the vocal fold and influence the vocal fold vibrations. As tissue is much stiffer in the fiber direction than in the transverse plane, different arrangement of the fiber orientation in the conus elasticus will alter its stiffness. Therefore, it is of interest to first evaluate the effect of material anisotropy on the stiffness of conus elasticus itself.

The conus elasticus and ligament layer are separated from the vocal fold model. The ligament layer is included for providing a more realistic boundary condition for the cranial surface of the conus elasticus. Static finite element analysis is performed with a uniform pressure load applied on the medial surface of the conus elasticus (Figure 4.1[b]) to evaluate its stiffness through the force-displacement response. Such pressure loading would deform the conus elasticus in a similar way that the subglottal pressure deforms the conus elasticus during phonation. The anterior and posterior surface of the conus elasticus and ligament are fixed, as well as the bottom surface of the conus elasticus. These boundary conditions are consistent with those in the full vocal fold model. The fiber orientations in the two conus elasticus models are illustrated in Figure 4.1(c), where the caudal-cranial and anterior-posterior orientation correspond to model 1 and model 2, respectively. Figure 4.2(a) shows the deformation of the two models under a pressure load of 0.4 kPa. The maximum deflection of both models occurs below the cranial surface of the conus elasticus, where the conus elasticus bulges outward into a cone shape. The cranial edge of the conus elasticus in model 1 is deflected more in the sideways direction than that in model 2. Force-displacement relationships are obtained for both models by varying the pressure load from 0.2 to 1.0 kPa with an increment of 0.2 kPa, in which the force is the total resultant force of the applied pressure load, and the displacement is the average displacement of the entire conus elasticus. The force-displacement

relationships for the two models are shown in Figure 4.2(b). Both models exhibit linear relations between force and displacement due to the linear elastic materials used in this study. The slopes of the force-displacement curves indicate the stiffness of conus elasticus under this uniform pressure load. The slopes corresponding to model 1 and model 2 are 0.523 N/cm and 0.613 N/cm, respectively, suggesting the anterior-posterior fiber orientation makes the stiffness of conus elasticus 17.2% larger than that of the caudal-cranial fiber orientation. To compare the deflection at the junction of the conus elasticus and ligament between the two models under the applied pressure load, force versus the average displacement of the cranial surface of the conus elasticus are shown in Figure 4.2(c). The relations between the force and displacement are also linear. The average displacement of the cranial surface of the conus elasticus in model 1 is 54.2% larger than that in model 2 throughout all the loads applied.

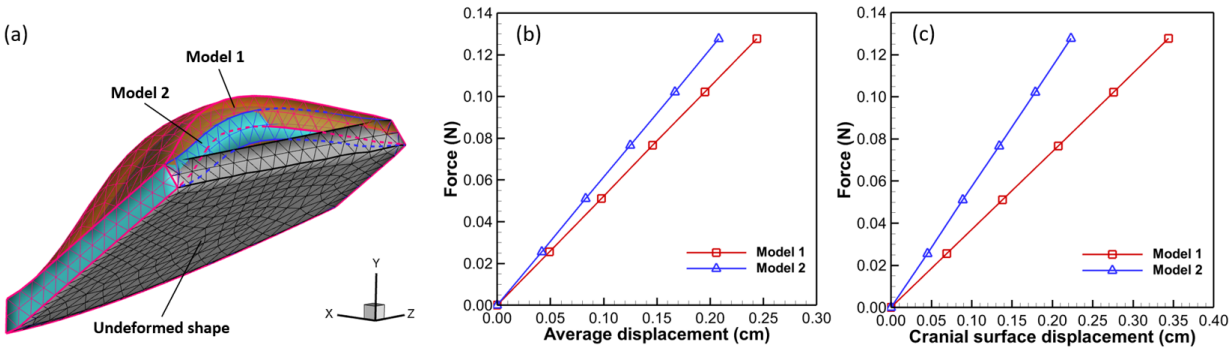

Figure 4.2 (a) Deformation of the two conus elasticus models under the pressure load of 0.4 kPa. Dashed lines indicate the location of the cranial edges that are not visible in the figure. (b) Force versus average displacement of the entire conus elasticus under uniform pressure loads. (c) Force versus average displacement of the cranial surface of the conus elasticus under uniform pressure loads.

As conus elasticus is integrated into the vocal fold model, its stiffness will affect the pliability of the tissues connected to it and the overall vocal fold stiffness. With the caudal-cranial fiber orientation, the cranial edge of conus elasticus is more flexible in the lateral direction, thus providing less resistance at the junction of the conus elasticus and ligament and allowing larger vibrations of vocal fold in the lateral direction. It suggests that the caudal-cranial arrangement of the fibers in the conus elasticus may be

important not only for preventing the excessive vertical motion but also for promoting the lateral vibration of vocal folds.

4.3.2 Vocal fold vibration

To further evaluate the effects of anisotropic direction of conus elasticus on vocal fold vibration, FSI simulations were conducted on the two vocal fold models with five different subglottal pressures (0.2, 0.4, 0.6, 0.8 and 1.0 kPa). Figure 4.3(a) shows the mid-coronal profiles of the two models under 1.0 kPa subglottal pressure at five different phases within one oscillatory cycle. The medial surface of both vocal fold models exhibits a convergent shape during glottal opening (t/T=0.25) and a divergent shape during glottal closing (t/T=0.68), which are the typical vibratory characteristics of vocal fold during normal phonation. The third phase (t/T=0.46) is around the time at which the maximum flow rate of model 1 is reached. To compare the kinematics between the two models, the medial-lateral displacements are extracted at three points (denoted by the markers in Figure 4.3[a]) located at the superior and inferior aspect of the medial surface and the junction of conus elasticus and ligament, respectively. The corresponding phase averaged displacement waveforms of the two models at 1.0 kPa subglottal pressure are plotted in Figure 4.3(b). The superior, inferior and junction displacement amplitudes of model 1 are larger than those of model 2, suggesting a larger overall vibration amplitude with model 1. At the junction of conus elasticus and ligament, the displacement amplitude of model 1 (0.363 mm) is 24.3% larger than that of model 2 (0.292 mm). The inferior displacement amplitude of model 1 is 0.503 mm, while the corresponding value of model 2 is 0.429 mm, decreasing by 14.7%. Compared to the junction and inferior displacement amplitude, the difference in superior displacement amplitude between the two models is relatively small. The superior displacement amplitude of model 2 is 6.6% smaller than that of model 1. The displacement amplitude vs subglottal pressure for the two models is plotted in Figure 4.4 for the inferior and superior point, separately. Overall, model 1 has a larger vibration amplitude than model 2 at both inferior and superior aspects with a larger difference occurring at the inferior aspect, suggesting that

the overall vibration amplitude of the vocal fold increases with the decrease of the stiffness of conus elasticus and the effect is more prominent at the inferior aspect where conus elasticus is close. The difference of displacement amplitude between model 1 and model 2 increases with the increasing of subglottal pressure.

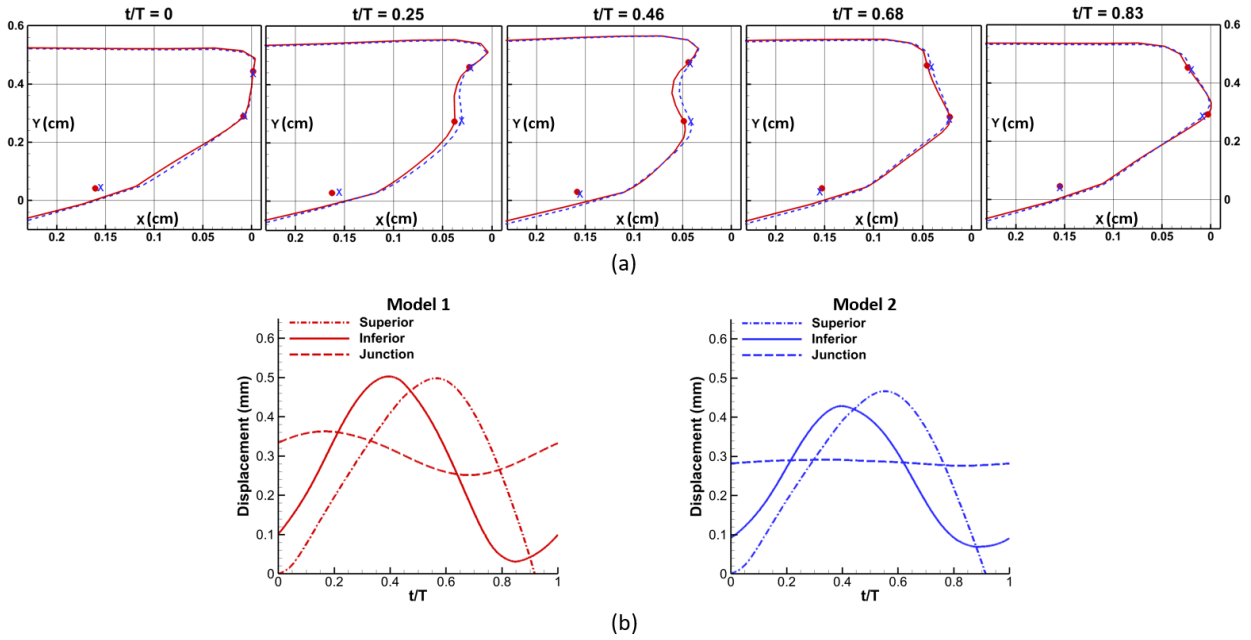

Figure 4.3 (a) Mid-coronal profiles of model 1 (red solid line) and model 2 (blue dashed line) at five phases within one oscillatory cycle (Due to the much smaller displacement difference between the two models as compared to the vertical dimension of vocal fold, the x-axis scale is enlarged to better show the difference). From top to bottom, the red dot and blue cross markers indicate the superior and inferior aspect of the vocal fold and the junction of the conus elasticus and ligament. (b) Superior (dash-dot line), inferior (solid line) and junction (dash line) medial-lateral (x-direction) displacement waveforms over steady-state cycles of the two models. Left, model 1; right, model 2. Data correspond to the cases of 1.0 kPa subglottal pressure.

Proper orthogonal decomposition (POD) method is exploited to further disclose the vocal fold dynamics of the two models (Berry et al., 1994; Vampola et al., 2016). In all current cases, the first two POD modes account for more than 99% of the total kinetic energy, therefore the dominant coherent vibration patterns can be sufficiently identified by just using these two modes. The mid-coronal profile of the first two modes of model 1 and model 2 at 1.0 kPa subglottal pressure are shown in Figure 4.5(a). The first mode mostly captures the medial-lateral motion while the second mode captures the propagation of mucosal wave

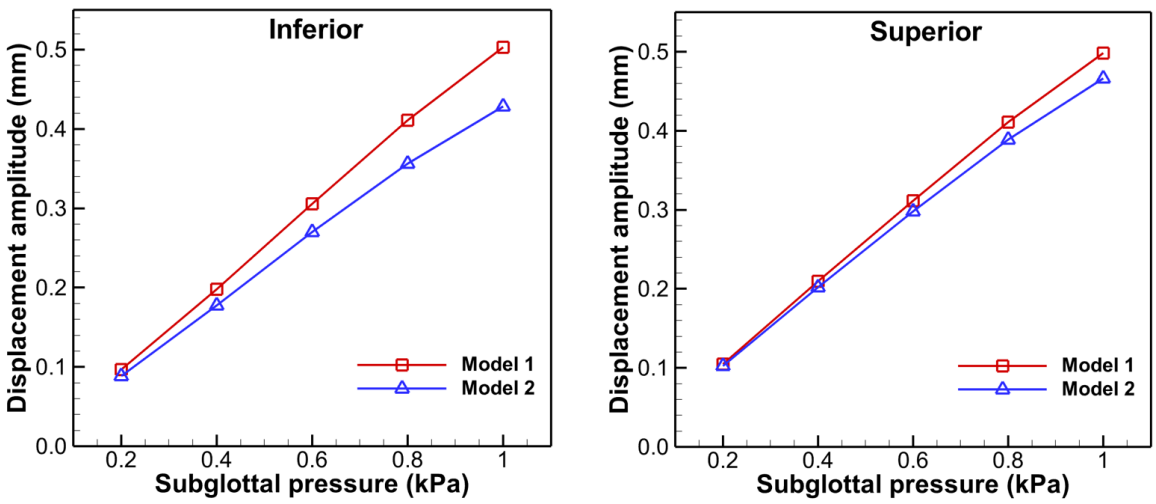

Figure 4.4 Medial-lateral displacement amplitude at inferior and superior vocal fold versus subglottal pressure.

along the medial surface. To quantify the mode similarity between model 1 and 2, dot-products are conducted between the corresponding normalized POD modes from the two models. A value of 1 represents two identical modes while a value of 0 represents two orthogonal modes. Figure 4.5(b) shows the mode dot-products at different subglottal pressures for the first and second modes, respectively. The overall mode similarities for all the cases are above 0.86, indicating that POD modes between the two models are highly similar and the material anisotropy of conus elasticus has little effect on the POD modes. The energy distribution between the two modes at different subglottal pressures are further plotted in Figure 4.6(a). In general, with the increase of subglottal pressure, the energy percentage of the first mode increases while the energy percentage of the second mode decreases. This observation indicates that for both models, the lateral motion becomes relatively stronger with the increase of subglottal pressure. It is also found that compared with model 2, the energy percentage of the second mode is about 14% higher in model 1, suggesting that model 1 has a relatively stronger mucosal type of motion. Figure 4.6(b) shows the maximum divergent glottal angle of the two models as a function of subglottal pressure. The maximum divergent glottal angle is observed to be larger in model 1 than model 2, and the difference increases from 0.18 degrees at 0.2 kPa subglottal pressure to 1.42 degrees at 1.0 kPa subglottal pressure, confirming that

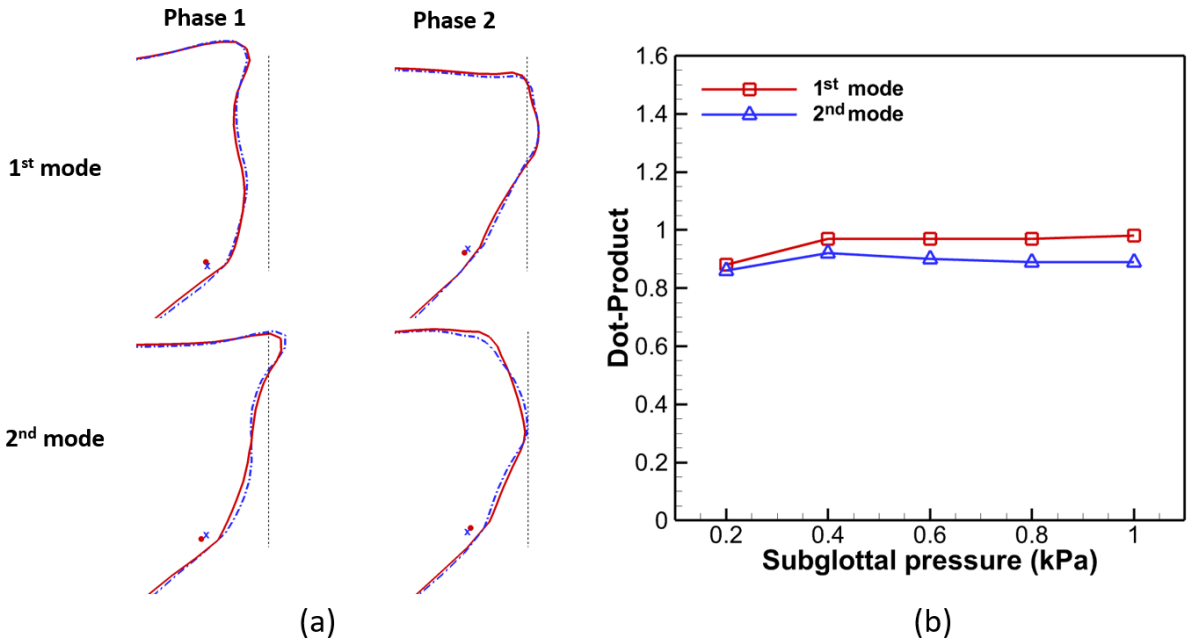

Figure 4.5 (a) The mid-coronal profile of the first two POD modes at two-extreme phases for model 1 (red solid line) and model 2 (blue dashed line) at 1.0 kPa subglottal pressure. The red dot and blue cross marker denote the junction of the conus elasticus and ligament in model 1 and model 2, respectively. The vertical dotted line (x = 0.0 mm) is the reference line for comparing mode shapes at different phases. (b) Dot-products of the first and second POD modes between model 1 and 2 at different subglottal pressures.

a stronger mucosal type of motion occurs with model 1. These results suggest that a decrease in the stiffness of conus elasticus would promote a stronger mucosal wave motion of the vocal fold. It needs to be pointed out that the increase of the maximum divergent angle with subglottal pressure in Figure 4.6(b) does not contradict the decrease of the energy percentage of the second mode observed in Figure 4.6(a), as the absolute value of the mode energy can still increase with subglottal pressure, resulting in the increase of the maximum divergent angle.

4.3.3 Flow rate waveform and aerodynamic measures

The phase averaged flow rate waveforms of the two models at 1.0 kPa subglottal pressure are shown in Figure 4.7(a). The flow rate of the two models are almost identical during the opening stage. While model 2 reaches its maximum flow rate of 268.1 ml/s at t/T = 0.402, the flow rate in model 1 continues increasing until t/T = 0.444, leading to a higher peak flow rate of 298.2 ml/s, which is 11.2% higher than model 2. For

all other different subglottal pressures, the flow rate is consistently higher in model 1 with increasing difference with the increase of subglottal pressure. The variation of peak flow rate with subglottal pressure for the two models are shown in Figure 4.7(b). Figure 4.7(c) shows the variation of average flow rate with subglottal pressure for the two models. Compared with the peak flow rate, the difference in average flow rate between the two models is small. Similar to the peak flow rate, the average flow rate difference increases with the increase of subglottal pressure.

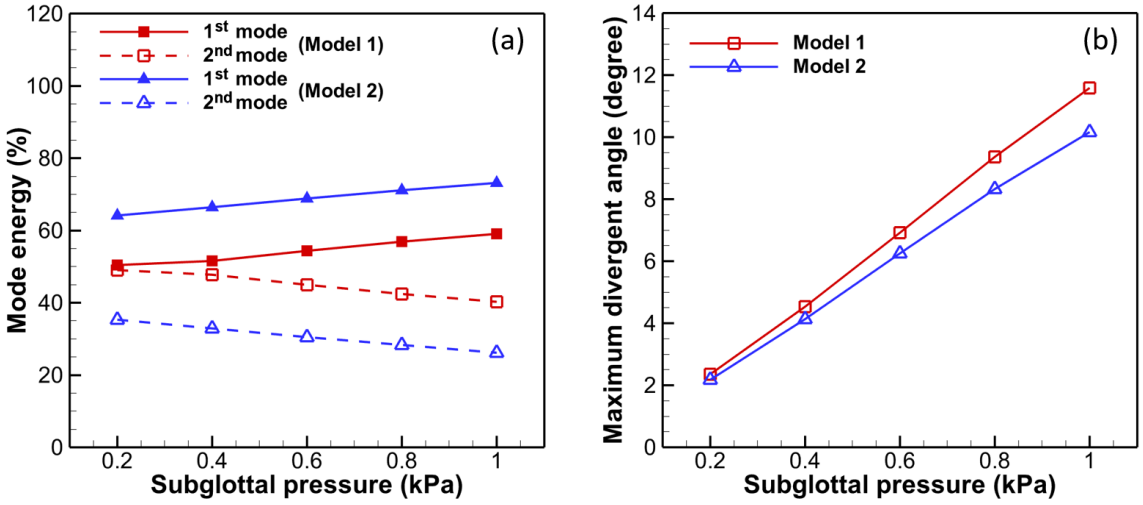

Figure 4.6 (a) Mode energy of the first two POD modes versus subglottal pressure. (b) Maximum divergent glottal angle versus subglottal pressure.

In the flow rate waveforms, the rising phase corresponds to a convergent glottal shape, under which the superior aspect of the glottis dictates the minimum glottal area and therefore glottal flow rate, while the falling phase is accompanied by a divergent glottal shape, under which the inferior aspect of the glottis dictates the minimum glottal area and flow rate. The peak flow rate occurs at the moment when the superior glottis has the same opening area as the inferior, and this moment is usually during the closing phase of inferior aspect of vocal fold and the opening phase of superior aspect of vocal fold. A detailed analysis of vocal fold vibration reveals that the maximum medial-lateral displacement of inferior aspect of vocal fold happens around 0.396T for both of the two models at 1.0 kPa subglottal pressure. However, the peak flow rate occurs at different phases: 0.444T for model 1 and 0.402T for model 2, suggesting it

takes a longer time for the inferior aspect to recoil back to the same glottal area of the superior aspect in model 1. This is likely due to the softer inferior aspect (smaller stiffness) and larger medial-lateral displacement at the inferior aspect in model 1, which prolongs the opening phase and leads to a higher peak glottal flow rate.

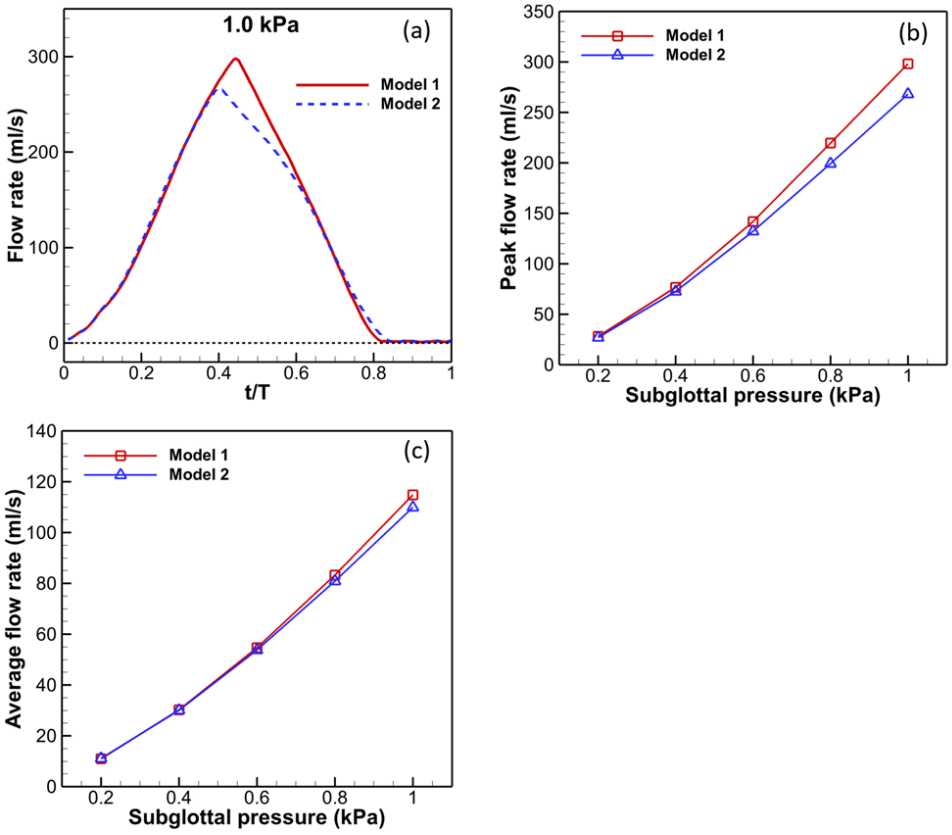

Figure 4.7 (a) Flow rate waveforms over steady-state cycles at 1.0 kPa subglottal pressure. (b) Variation of peak flow rate with subglottal pressure. (c) Variation of average flow rate with subglottal pressure.

Figure 4.8(a) and (b) show the open quotient and speed quotient of each model versus the subglottal pressure, respectively. The relationships of open quotient and speed quotient with subglottal pressure are non-monotonic. Overall, the open quotient is nearly constant with a small fluctuation for each model. The open quotient of model 2 is consistently larger than that of model 1, with the largest difference of 3.7% observed at subglottal pressures of 0.8 and 1.0 kPa. In contrast to the small difference in open quotient between the two models, the speed quotient of model 1 is significantly higher than that of model

2 for all the simulated subglottal pressures, resulting from the relatively longer opening phase and shorter closing phase of the glottis in model 1. Thus, it suggests that the caudal-cranial fiber orientation in the conus elasticus can slightly decrease the open quotient and significantly increase the speed quotient.

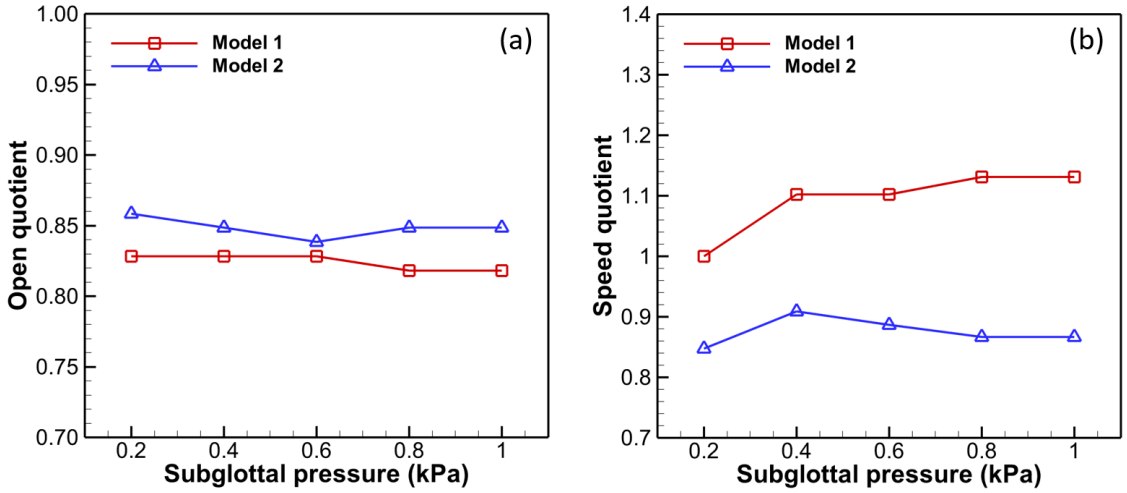

Figure 4.8 (a) Open quotient versus subglottal pressure. (b) Speed quotient versus subglottal pressure.

4.3.4 Spectrum analysis and acoustic measures

Figure 4.9 presents the plot of fundamental frequency (f_o) versus the subglottal pressure of the two models. f_o is determined from the largest harmonic (also the first harmonic in this study) in the spectrum of flow rate waveform (Figure 4.10[a]), as no glottal tract is included in this study. As the linear elastic material is adopted, the frequencies almost do not change with the subglottal pressure. f_o of model 2 is higher than that of model 1, with a minimum difference of 3.7% at 0.2 kPa and a maximum difference of 6.6% at 1.0 kPa. This finding is not surprising considering the conus elasticus in model 2 is stiffer than that in model 1.

Spectra of the glottal source contain important acoustic cues to the variation of voice quality (Klatt and Klatt, 1990). Figure 10(a) presents the spectrum of flow rate waveform of model 1 at 1.0 kPa subglottal pressure. This spectrum is obtained by performing fast Fourier transform on the last 0.5 s period of the

flow rate waveform. Important acoustic measures including H1-H2, H1-H4 (spectral amplitude difference between the first two harmonics, first and fourth harmonics) and spectral slope are extracted from the

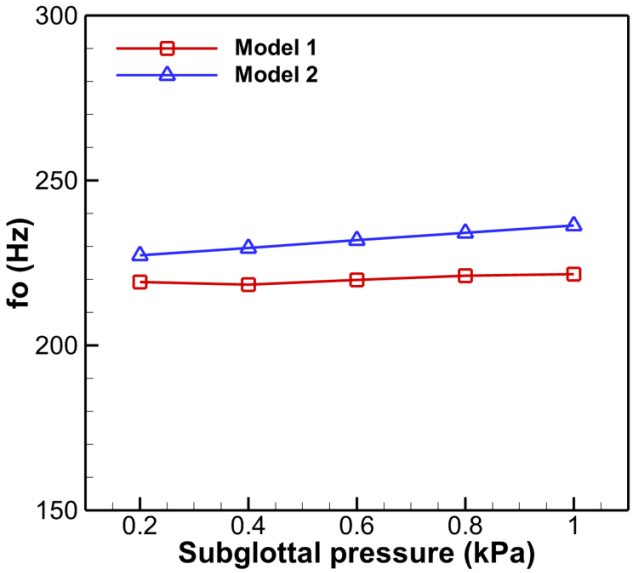

Figure 4.9 f_o versus subglottal pressure.

spectrum. The spectral slope is obtained by using the linear regression in MATLAB to fit the harmonics between 0 and 2k Hz. The slope of the fitted line is multiplied by $\log_{10}$ of 2 to get the slope in dB per octave. Figure 4.10(b) illustrates the spectral peaks and the fitted line of model 1 at 1.0 kPa subglottal pressure. H1-H2, H1-H4 and spectral slope of the two models versus subglottal pressure are shown in Figure 4.11. As can be seen, both H1-H2 and H1-H4 of model 2 are apparently higher than model 1. Take the values at 1.0 kPa subglottal pressure for example, H1-H2 and H1-H4 of model 2 are respectively, 21.6% and 7.6% greater than that of model 1. Considering model 2 has comparatively smaller flow deceleration rate (Figure 4.7[a]), the phenomenon observed here is consistent with the finding of Holmberg et al. (1995), where they pointed out that relatively gradual vocal fold closures induced a higher amplitude of the first harmonic. In addition, Holmberg et al. (1995) also found from their measurements that a higher H1-H2 is mostly accompanied by a small adduction (closed) quotient. Our results show the same relationship as open quotient of model 2 is larger than that of model 1. In Figure 4.11, it is also observed that at most subglottal pressures, model 2 has steeper spectral slopes. A steeper spectral slope indicates

a weak excitation of higher-order harmonics, which is compatible with the lower flow deceleration rate observed in model 2.

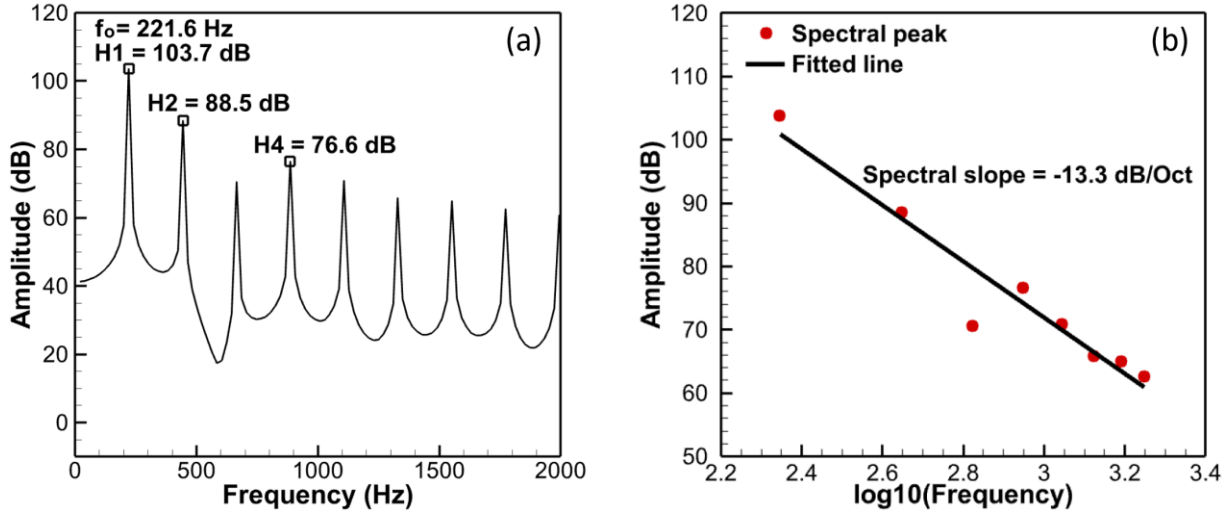

Figure 4.10 (a) Spectrum of flow rate waveform of model 1 at 1.0 kPa subglottal pressure. f_o, fundamental frequency; H1, the first harmonic; H2, the second harmonic; H4, the fourth harmonic. (b) Spectral slope is calculated from the line fitting the spectral peaks. Spectral slope in dB/Oct is obtained by multiplying the slope of fitted line by log10 of 2.

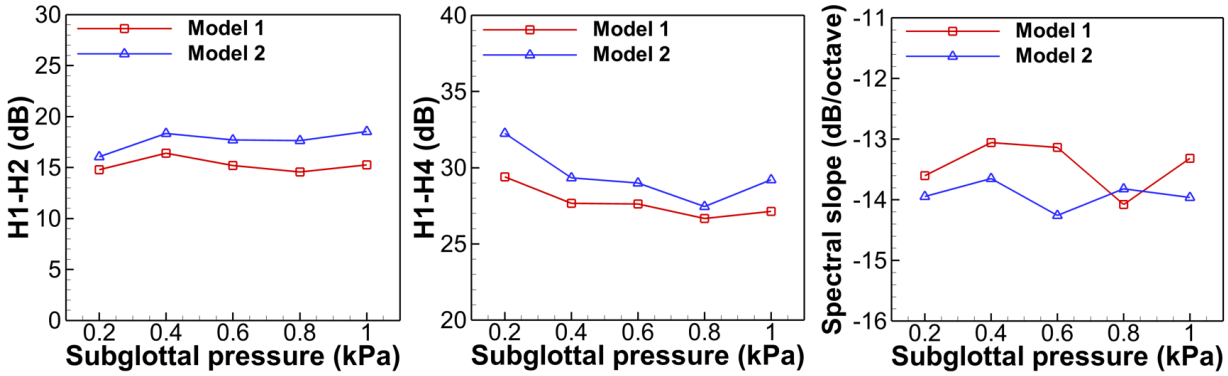

Figure 4.11 Spectral amplitude difference between the first two harmonics (H1-H2), first and fourth harmonics (H1-H4) and spectral slope versus subglottal pressure.

4.4 Summary

In this study, two vocal fold models with two different fiber orientations in the conus elasticus are built to examine the influence of the direction of material anisotropy on vocal fold vibrations, as well as

aerodynamic and acoustic measures of voice production. It is demonstrated that the separate conus elasticus model with a realistic fiber direction, caudal-cranial fiber orientation yields smaller structural stiffness and larger deflection at the junction of the conus elasticus and ligament than the model with the anterior-posterior fiber orientation, which yields a greater vibration amplitude and stronger mucosal wave of the vocal fold. Different directions of material anisotropy in the conus elasticus between the two models also induce differences in aerodynamic and acoustic measures. The flow rate waveform with a caudal-cranial fiber orientation in the conus elasticus is observed to have a larger peak flow rate and a higher speed quotient, owing to the increased inferior vibration amplitude which prolongs the opening phase. Spectrum analysis of the flow rate waveform has shown a decreased fo of vocal fold vibration when a realistic conus elasticus is considered. Finally, the vocal fold model with a realistic conus elasticus generates a voice with relatively smaller first harmonic amplitude and spectral slope. It is worth pointing out that compared with the spectral slope variations in human voice, the difference of spectral slope between the two models in current study is small and may not be perceived.

One limitation of this study is that the glottal flow is assumed to be inviscid and quasi-steady. Neglecting the viscous and inertia effects may be acceptable in the middle stage of the phonation cycle, but it can cause non-negligible errors when the glottis is nearly closed. Another limitation of this study is that the material parameters of conus elasticus are simply assumed to be the same as those of vocal ligament, due to the lack of measurement of the mechanical properties of conus elasticus. In addition, the anterior, posterior and lateral surface of the vocal fold are fixed to mimic the attachment to the cartilages, which can only be thought of as an approximate treatment because in the real situation, the attachment of vocal fold tissues to the cartilages is complex. For example, the anterior part of conus elasticus was reported to be only attached to the thyroid cartilage midway between the notch and the caudal border of the thyroid cartilage (Williams et al., 1989; Sato, 2018). From the caudal border of the thyroid cartilage down, the conus elasticus extends to the upper rim of the cricoid cartilage and is not attached to any cartilage

anteriorly. To simulate the boundary conditions applied on the vocal fold more truthfully, a more realistic, complete larynx model which includes the cartilages is needed. Furthermore, including material and geometric nonlinearities in vocal fold models may predict more realistic vibrations of vocal folds.

CHAPTER 5 EXAMINING THE VALIDITY OF THE QUASI-STEADY ASSUMPTION FOR GLOTTAL FLOW

5.1 Introduction

The quasi-steady flow assumption (QSFA) is one of the major approximations widely adopted in the study of biomechanics of phonation. The assumption is that the unsteady flow behaviors have negligible effects on glottal flow dynamics and vocal fold vibration. More specifically, the QSFA approximates time-varying glottal flow with a series of steady flow solutions on their 'time frozen' glottal channel configurations. Based on that, the classic Myoelastic-aerodynamic theory of phonation employs the Bernoulli law to elucidate the energy transfer mechanism between the glottal flow and vocal fold vibration (van den Berg, 1958; Titze, 2006). Flow-pressure relationships during vocal fold vibration have been studied by conducting experiments on static glottal configurations (Scherer and Guo, 1990; Scherer et al., 2001). Computer vocal fold models, e.g., the two-mass model (Ishizaka and Flanagan, 1972), the three-mass model (Story and Titze, 1995), as well as some recent continuum models (Alipour et al., 2000; Zhang, 2016; 2017), are often coupled with steady flow solutions to approximately model the flow-structure interaction during phonation.

The adoption of the QSFA was mainly justified based on two observations of typical human phonation (Ishizaka and Flanagan, 1972): (1) the dimensions of the glottis, length (L) and thickness (T), are small relative to the "wavelength" defined as glottal particle velocity (v_g) divided by fundamental frequency (f_o). This yields one nondimensional inequality condition of $S_t \ll 1$, where S_t is Strouhal number defined as $S_t = f_o L/v_g$. If f_o, L and v_g are adopted as the characteristic frequency, length and velocity, respectively, in the process of nondimensionalization, the unsteady flow term in the Navier-Stokes equation becomes $S_t \frac{\partial \vec{V}}{\partial t}$, which can be dropped out from the equation through order analysis; (2) the velocity of glottal airflow, v_g, is much larger than the oscillation speed of the vocal fold, v_v. This yields the second nondimensional inequality condition of $V_r \gg D$, where $V_r = v_g/v_v$ is the reduced velocity, and

$D = \xi_o/L$ is the displacement number. ξ_o is the vocal fold vibration amplitude defined as $\xi_o = v_v/f_o$. Through nondimensional order analysis on the boundary conditions, this inequality condition enables the static glottal wall approximation.

The validity of the QSFA in human phonation was examined by several experimental studies (Mongeau et al., 1997; Pelorson, 2001; Zhang et al., 2002; Deverge et al., 2003; Vilain et al., 2004). The QSFA model showed good agreements on pressure and flow rate predictions with experimental measurements during the majority of a vibration cycle; yet significant deviations were observed at the stages of flow initiation and shutoff. Krane and Wei (2006) performed a theoretical study in which the order of magnitude of the unsteady and convective acceleration terms in the momentum equation were assessed. Based on the relative importance of each term, they showed that the glottal flow dynamics can be split into an unsteady-effect dominated flow-initiation/shutoff interval and a convective-acceleration dominated quasi-steady interval. Subsequently, Krane et al. (2010) combined the theoretical work with experimentally measured flow velocity data and calculated the waveshapes of the unsteady and convective accelerations. They reported that while the glottal jet inertia was found to be nearly zero during the middle 40% of the open phase, the unsteady acceleration could dominate the flow for the rest of the cycle.

The nondimensional inequality condition of $S_t \ll 1$ indicates that the validity of QSFA is frequency dependent. The well-accepted value of S_t for glottal flow is on the order of 10^{-2}, estimated based on the frequency range in normal speech phonation (~100-200 Hz). Since S_t is proportional to frequency, a high vibration frequency would lead to a large S_t. In addition, the oscillation speed of vocal fold largely increases at high vibration frequencies, which probably undermines the justification of large V_r. The above two considerations suggest the need of further testing the validity of the QSFA at high phonation frequencies (in the 300-1000 Hz range).

Another factor that could affect the validity of the QSFA is the air inertia effect in the vocal tract. Kucinschi and Scherer (2006) revealed that the inertance of air in the trachea and vocal tract generates a dynamic pressure at glottal inlet and outlet which interacts with the unsteady glottal flow and can cause the flow rate and transglottal pressure to vary with frequency. Clearly, whether considering the air inertia effects or not in the validation of the QSFA will yield different results.

Different glottal shapes can generate quite different glottal flow dynamics (Schere et al., 2001; Triep and Brucker, 2010), which also likely affect the validity of the QSFA. The effects of glottal geometry on the adequacy of the quasi-steady predictions were found to be insignificant in previous studies (Zhang et al., 2002; Deverge et al., 2003). However, only simplified and symmetric glottal shapes were investigated. Under pathological vocal fold vibrations, the glottal geometry can be asymmetric and involve longitudinal wave propagations (Švec et al., 2000; Neubauer et al., 2001). It is not known if the above conclusion remains valid when these much more complex glottal shapes are considered. Furthermore, although the glottal areas were varying with time to produce the flow waveform, the glottal shapes in previous studies stayed unchanged throughout the glottal cycle. As noted by Krane et al. (2010), investigating the dynamics of phonatory flow with an unchanging glottal shape leaves out the unsteady effects associated with the motion of flow separation point. Therefore, glottal shapes that can produce the alternating convergent-divergent motion and the real-life traveling of a separation point are necessary for the validation of the QSFA.

This study aims to use the numerical method to further investigate the range of validity of the QSFA by considering voice frequency range, air inertia in the vocal tract and complexity of glottal shapes. To separate the effects of flow unsteadiness ($\frac{\partial \vec{V}}{\partial t}$) and vocal fold vibration, three different types of simulation were performed. The first is the dynamic simulation which obtains the unsteady solution of Navier-Stokes equations with prescribed glottal wall motion. The solution contains the effects of both unsteady flow and vocal fold motion. The second is the pseudo static simulation which obtains the steady solution of Navier-

Stokes equations of each glottal wall shape with the glottal wall velocity incorporated. The solution does not include unsteady flow effects but contains the effect of glottal wall motion. The third is the quasi-steady simulation which obtains the steady solution of Navier-Stokes equations of each glottal wall shape without incorporating the glottal wall velocity. The solution eliminates the effects of both unsteady flow and glottal wall motion. To test the dependence of the QSFA on frequency, two vibration frequencies, 100 and 500 Hz, are simulated. Additionally, the effect of air inertia in the supraglottic vocal tract is tested by implementing the simulations with and without a vocal tract. Two sets of asymmetric glottal shapes, which incorporate different vertical and longitudinal wave patterns on the right and left vocal folds, are used to address the effect of glottal shapes on the assumption. The validity of the QSFA is assessed by quantifying the differences of glottal flow and glottal wall pressure between the dynamic and quasi/pseudo simulations. A momentum budget analysis is also conducted to evaluate the magnitude of unsteady acceleration in glottal airflow and its effect on the quantified flow and wall pressures.

5.2 Methods

5.2.1 Glottal shapes derived from normal modes of vibration

The two distinct sets of time-dependent glottal shapes were produced following the surface-wave approach in Smith and Titze (2018). Based on the approach, the wave-type motion of vocal fold surface can be described by a combination of (m,n) modes, where m and n represent the number of half wavelengths along the anterior-posterior direction and inferior-superior direction, respectively. For normal phonation, the most dominant modes are the (1,0) and (1,1) modes, where (1,0) represents the medial-lateral motion and (1,1) represents the convergent-divergent motion. Higher mode numbers represent more complex wave motions, which can occur in voice disorders. For example, a (2,1) mode represents two half wavelengths in the anterior-posterior direction and half wavelength in the inferior-superior direction.

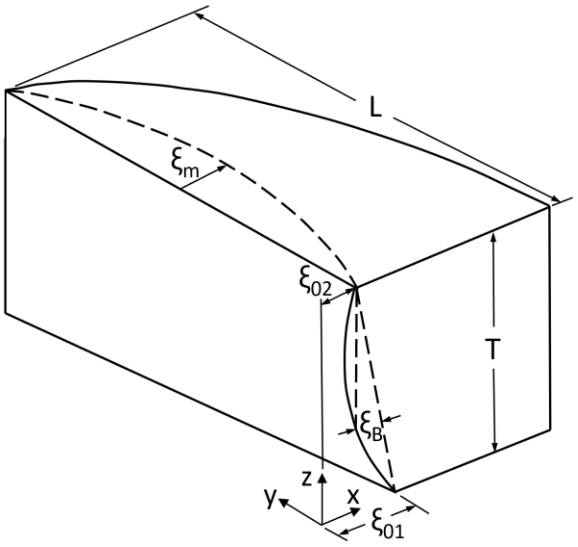

Figure 5.1 Diagram of the pre-phonatory configuration of the right medial surface.

In this study, the two sets of asymmetric glottal wall dynamics were obtained through a combination of the (1,0) mode (left) and (2,1) mode (right) in the first set and (1,1) mode (left) and (2,0) mode (right) in the second set. The pre-phonatory medial surface on the right side (Figure 5.1) was defined by a glottal half-width ξ_{0R} as a function of anterior-posterior (y) and inferior-superior (z) directions, which is:

$$\xi_{0R}(y,z) = (1 - y/L)[\xi_{0R2} + (\xi_{0R1} - \xi_{0R2} - 4\xi_{BR}z/T)(1 - z/T)] \quad (5.1)$$

where L and T are respectively the length and thickness of the vocal fold, ξ_{0R1} and ξ_{0R2} are respectively the inferior and superior glottal half-widths at the vocal process, ξ_{BR} is a surface bulging parameter which controls the vertical curvature of the medial surface. The subscript R denotes the right medial surface. A same equation with a subscript L exists for the pre-phonatory glottal width of the left medial surface (ξ_{0L}). For the right medial surface, the modal displacement ξ_R at any instant of time (t) is defined as:

$$\xi_R(y,z,t) = \xi_{mR}\sin(m\pi y/L)\,[sin\omega t - n(\omega/c)(z - z_{mR})cos\omega t] \quad (5.2)$$

where ξ_{mR} is the modal displacement amplitude, ω is the angular frequency, c is the mucosal wave speed, z_{mR} is the inflection point for the vertical half wavelength. An equivalent equation exists for the modal displacement of the left medial surface (ξ_L). The overall three-dimensional glottal shape at any moment

of time was obtained by superimposing the modal displacements on the pre-phonatory shape of the medial surfaces:

$$g(y,z,t) = \xi_{0R}(y,z) + \xi_{0L}(y,z) + \xi_R(y,z,t) + \xi_L(y,z,t) \quad (5.3)$$

For this study, the same vocal fold thickness, T, and length, L, were adopted from Smith and Titze (2018), which are 0.8 and 1.5 cm, respectively. A uniform (parallel surface) initial posturing configuration was used, and the parameter values (in cm) are:

$$\xi_{0R1} = 0.10, \xi_{0R2} = 0.10, \xi_{BR} = 0.005, \xi_{0L1} = 0.10, \xi_{0L2} = 0.10, \xi_{BL} = 0.005.$$

Vibration frequency of the vocal folds, *f,* was chosen to be 100 Hz and 500 Hz. The angular frequency, *ω,* and the mucosal wave speed, *c,* were defined as $2\pi f$ and $T\pi f$, respectively. For both the right and left medial surfaces, the modal displacement amplitude ξ_m was 0.1 cm, and the inflection point z_m was defined as $T * (0.6 - 0.02\xi_B)$.

For both the (1,0)-(2,1) and (1,1)-(2,0) set of glottal wall motions, sixteen sequential static glottal shapes were extracted from one period of vibration. For this period, T, the sixteen shapes were evenly spaced between t/T = 0 and 0.9375 with an increment of 0.0625. A view of the sixteen shapes of each set can be obtained from the contact patterns shown in Figure 5.2. The position of the glottal wall in dynamic simulations was specified at each instant of time through cubic spline interpolation between the initial sixteen shapes. The vibration velocity of each point on the glottal wall was calculated by dividing the moving distance between two adjacent moments by the time increment. As a result, the progression from shape to shape in dynamic simulations was kept continuous, as in normal vocal fold vibration. For pseudo static simulations, each of the sixteen shapes corresponded to an independent numerical simulation, and the wall velocity, which is the same as the velocity in dynamic simulations under the same shape, was applied on the glottal wall. Steady flow solutions were obtained for each glottal shape. For quasi-steady simulations, each glottal shape also corresponded to an independent simulation, but there was no velocity on the wall. Steady flow solutions were also obtained for each glottal shape.

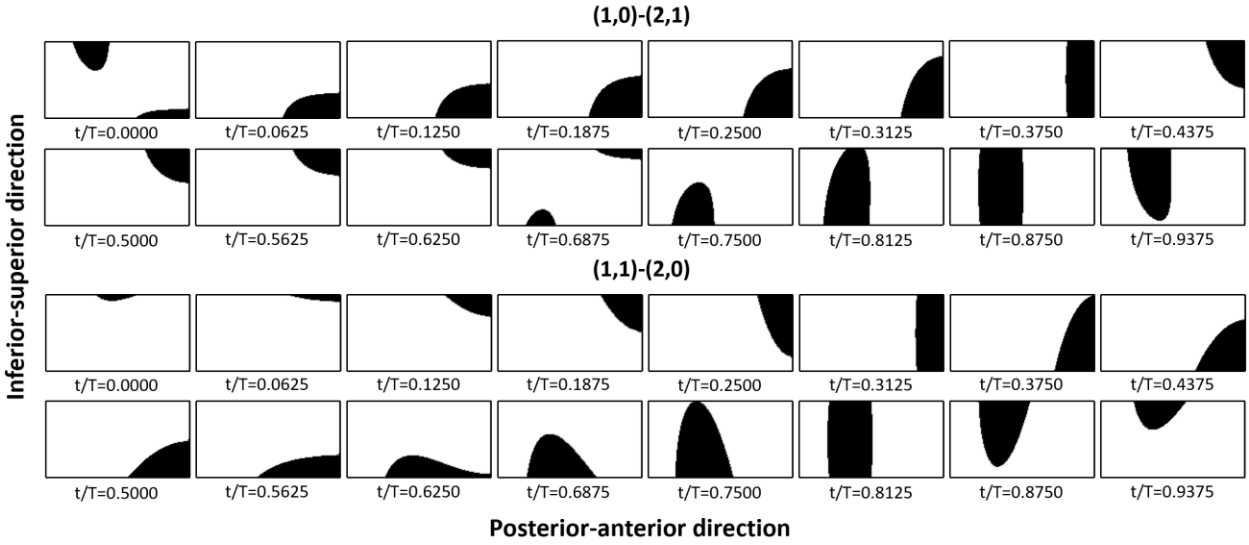

Figure 5.2 Contact patterns of the sixteen shapes of (1,0)-(2,1) and (1,1)-(2,0) set of glottal wall motions. For each phase, contact area (black) and open glottis (white) are marked on the rectangular medial surface.

5.2.2 Simulation setup and case summary

The computational model setup is shown in Figure 5.3. Figure 5.3(a) is the setup with a supraglottic tract. The vocal fold surfaces were immersed into a 1.5 cm × 21.8 cm × 1.5 cm rectangular computational domain which consists of a 17-cm-long supraglottic vocal tract, a 3.6-cm-long trachea and a pair of 0.8-cm-thick vocal folds. The computational domain was discretized by a 128 × 128 × 96 (x × y × z) Cartesian grid, and grid independent solutions were achieved with such resolution. Dirichlet boundary conditions for pressure of 1.0 kPa and 0 kPa were applied at the inlet and exit of the computational domain, respectively. No-slip and no-penetration boundary conditions were imposed on the trachea and vocal tract walls. Figure 5.3(b) is the setup without the supraglottic tract. The vocal folds were located at the bottom center of a cubic computational domain which is 30 cm × 30 cm × 30 cm. A 256 × 128 × 128 (x × y × z) Cartesian grid was used to discretize the computational domain. The grid around the vocal folds is 164 × 72 × 92 (x × y × z) which provides similar grid resolution as the setup with the supraglottic tract. Similar to the former setup, constant pressure boundary conditions were applied at the entrance of the trachea and top surface of the

computational domain, which are 1.0 kPa and 0 kPa, respectively. The left, right, front, back surface, as well as the bottom surface outside the trachea area, were treated as no-slip and no-penetration boundaries. For both setups, the surface of each vocal fold was discretized into 71400 triangular elements with 0.01 cm resolution. In dynamic simulations, a 0.006 cm gap was saved between the left and right medial surface in contact area, which is necessary for the success of airflow simulation.

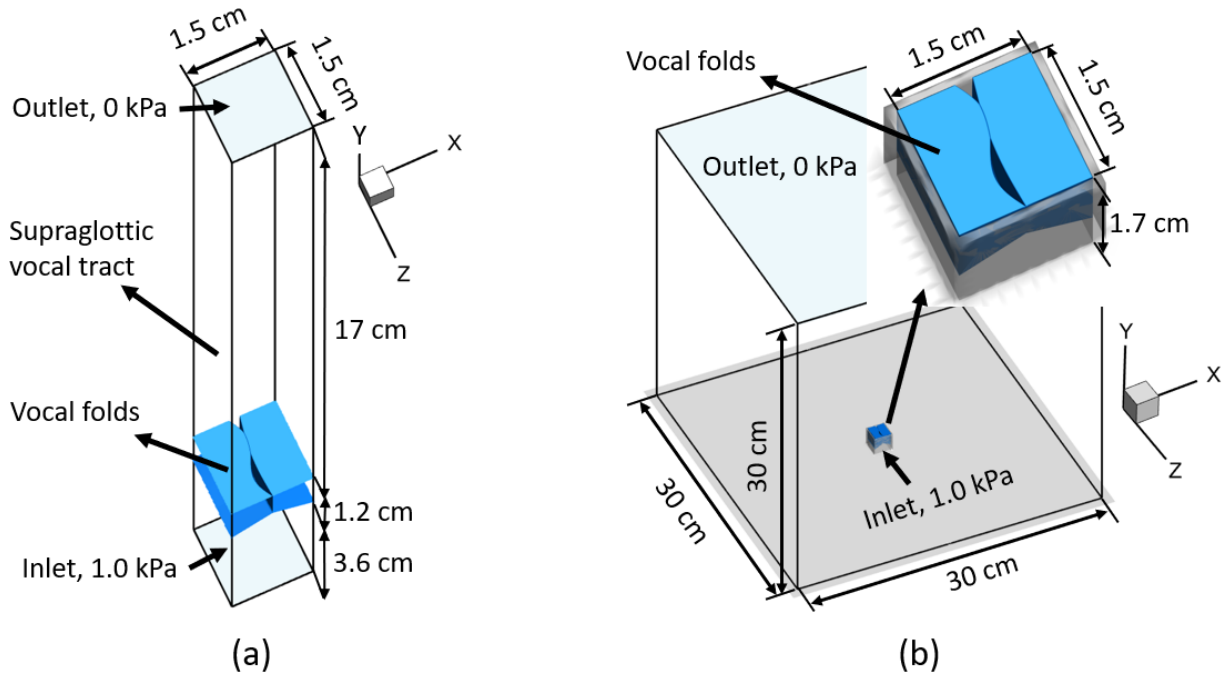

Figure 5.3 Computational domain and vocal fold model: (a) Setup with supraglottic vocal tract (b) Setup without supraglottic vocal tract. The glottal shape corresponds to (1,0)-(2,1) set) at t/T=0.0000.

In the dynamic simulations, the time steps of 1.184×10^{-6} s and 1.016×10^{-6} s were used for the case of 100 Hz and 500 Hz vibration frequency, respectively. Such time steps were chosen for (1) outputting the results at the specific 16 phases and (2) complying with the Courant-Friedrichs-Lewy condition. The same time steps were used for pseudo static simulations. A time step of 8.45×10^{-7} s was used for quasi-steady simulations. For a few cases that experienced numerical convergence problems, the time step was appropriately reduced. Time independence was achieved for all three types of simulations. Glottal airflow was simulated for two cycles in each dynamic simulation. For each pseudo static or quasi-steady simulation, glottal airflow was calculated until a sustained steady state is reached. The simulations were

run on the XSEDE COMET cluster which uses Intel's Xeon Processor E5-2600 v3 family. 128 processors were used for each dynamic or pseudo static simulation, and 64 processors were used for each quasi-steady simulation. The computational time of the dynamic simulation range from about 2 to 10 days, depending on the vibration frequency and number of grids. The average computational time for a pseudo static or quasi-steady simulation is around 1 day.

The total number of simulations included in this study is 200 with 8 dynamic, 128 pseudo static and 64 quasi-steady cases. For each of the two sets of glottal wall motions, setups were implemented with two different supraglottic configurations (with and without vocal tract) under two different vibration frequencies (100 and 500 Hz), which adds up to the 8 dynamic cases. Based on each dynamic case, 16 pseudo static simulations, each with a frozen glottal wall velocity extracted from the dynamic simulation, were built, resulting in the 128 pseudo static cases. Since there are no velocities on the glottal wall and the simulations are frequency-independent, the number of quasi-steady cases reduces to 64.

5.3 Results

5.3.1 Glottal flow rate comparison

For each of the two sets of glottal shapes, glottal flow rates were compared between the dynamic and quasi-steady/pseudo static simulation over one vibration cycle under different simulation setups. In Figure 5.4, flow rate waveforms of quasi-steady and pseudo static simulations were obtained using the "spline" function in MATLAB to interpolate over the 16 discrete flow rate points. Percent errors of the flow rate of the pseudo static and quasi-steady assumptions were calculated by comparing to and normalizing using the values in the dynamic simulations in the corresponding configurations. For the (1,0)-(2,1) set at 100 Hz frequency, flow rate waveforms of the three different simulations exhibit great similarities when the vocal tract was not considered. The cycle-averaged absolute percent errors of the quasi-steady and pseudo static simulations are 6.5% and 7.2%, respectively. When the vocal tract was included, flow rate waveforms of the quasi-steady and pseudo static case deviate from that of the dynamic

case with larger phase difference. The cycle-averaged absolute percent errors increase to 14.7% and 15.0%, respectively. Significant departures of flow rate from those of the dynamic case are observed at 500 Hz frequency. For the setup without a vocal tract, the cycle-averaged absolute percent errors between the quasi-steady/pseudo static and dynamic simulation are 24.1% and 22.2%, respectively, while for the setup with a vocal tract, the errors are 46.1% and 44.5%, respectively. These data show that the flow rate errors nearly tripled as vibration frequency increased to 500 Hz. In addition, including the vocal tract in the simulation nearly doubled the errors, as compared with the simulation without the vocal tract. The same observation applies to the flow rate waveforms of (1,1)-(2,0) set. In the bottom row of Figure 5.4, the cycle-averaged absolute percent errors between the quasi-steady/pseudo static and dynamic simulation are, from left to right, 8.0% and 6.5%, 22.7% and 20.4%, 15.7% and 13.6%, 41.2% and 37.5%, respectively. The nearly identical error levels of the two different sets of glottal shapes indicates the accuracy of flow rate calculation by quasi-steady or pseudo static approximation is insensitive to glottal shapes.

In Figure 5.4, the magnitudes of flow rate errors are observed to vary during a vibration cycle. During the opening phase, the flow rate errors increase after the flow initiation until a peak value is reached and then gradually decrease to nearly zero around the maximum flow rate. The same trend of first increasing and then decreasing is observed for the flow rate errors during the closing phase. This variation of errors is more prominent at 500 Hz frequency. The nonuniform error levels indicates the importance of unsteady effects varies within one vibration cycle.

Figure 5.4 also shows that the flow rate waveforms of quasi-steady and pseudo static case almost overlap with each other for all cases at 100 Hz frequency, which suggests that glottal wall motion has little effect on flow rates at low vibration frequency. However, at 500 Hz frequency, discrepancies between the two flow rate waveforms are observed, especially during the flow deceleration stage.

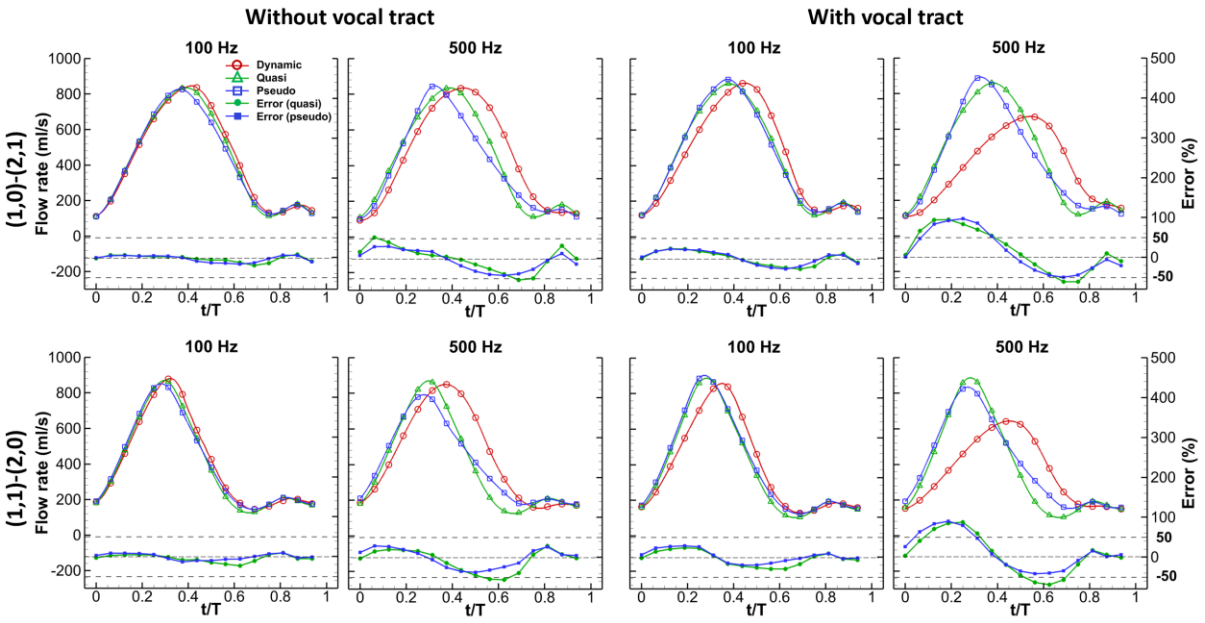

Figure 5.4 Flow rate waveforms of the three simulations and percent errors of flow rate between quasi-steady/pseudo static and dynamic simulation over one vibration cycle.

5.3.2 Strouhal number and velocity ratio analysis

To investigate how the variations of flow rate errors are associated with the quasi-steady assumption, the Strouhal number of each dynamic case was calculated. In the equation $S_t = f_o L/v_g$, f_o and L are respectively the vibration frequency and thickness (0.8 cm) of the vocal fold. v_g is the cycle-averaged upstream flow velocity (taken at glottis inlet) in each case. The value of S_t for each dynamic case is listed in Table 5.1. At 100 Hz vibration frequency, S_t is of the order of 0.03 for all cases, while S_t increases nearly fivefold (approximately 0.15) at 500 Hz frequency. Note that the value of S_t at 100 Hz frequency falls well within the range (0.01 to 0.1) accepted for the QSFA (Vilain et al., 2004; Krane et al., 2010), whereas the value of S_t at 500 Hz frequency is beyond the range. Therefore, the very large errors of flow rate at 500 Hz frequency are due to the breakdown of the QSFA. In other words, the unsteady effects of glottal flow play an important role at high phonation frequency and cannot be neglected.

Table 5.1 Value of Strouhal number for each dynamic case

	100 Hz		500 Hz	
	Without vocal tract	With vocal tract	Without vocal tract	With vocal tract
(1,0)-(2,1)	0.032	0.031	0.154	0.165
(1,1)-(2,0)	0.031	0.03	0.145	0.145

The ratio of flow velocity to vocal fold vibration velocity, v_g/v_v, was also calculated at each of the 16 phases for the dynamic cases. In the calculation v_g/v_v, v_g is the mean flow velocity at glottis inlet, and v_v is the averaged vibration velocity of the medial surface on the side of mode with one half wavelength in the longitudinal direction. Figure 5.5 plots v_g/v_v versus t/T during the open phase for each case, and the absolute percent errors of flow rate between the quasi-steady/pseudo static and dynamic simulations are displayed in the form of error bars at each data point (length of error bars denote the magnitude of errors). Figure 5.5 shows that v_g/v_v is generally in an inverse relationship with flow rate error. For t/T < 0.2 and t/T > 0.5, v_g/v_v is relatively small but accompanied by a large flow rate error. While in the interval 0.2 < t/T < 0.5, v_g/v_v has a high value, but the error is significantly reduced. It is observed that, in the cases with the frequency of 500 Hz and the supraglottic tract included, this inverse relationship is not as evident as other cases, probably because the air inertia effect in the vocal tract dominates the flow rate error. Figure 5.5 also shows that for the same phase, v_g/v_v has a much smaller value at 500 Hz than at 100 Hz, which is consistent with the above inverse relationship since flow rate errors are much larger at 500 Hz. These observations verify that a large velocity ratio between glottal flow and vocal fold vibration is necessary for the legitimacy of quasi-steady assumption. Lastly, it is also noted that the variations of flow rate errors with v_g/v_v show a similar trend for quasi-steady and pseudo static case.

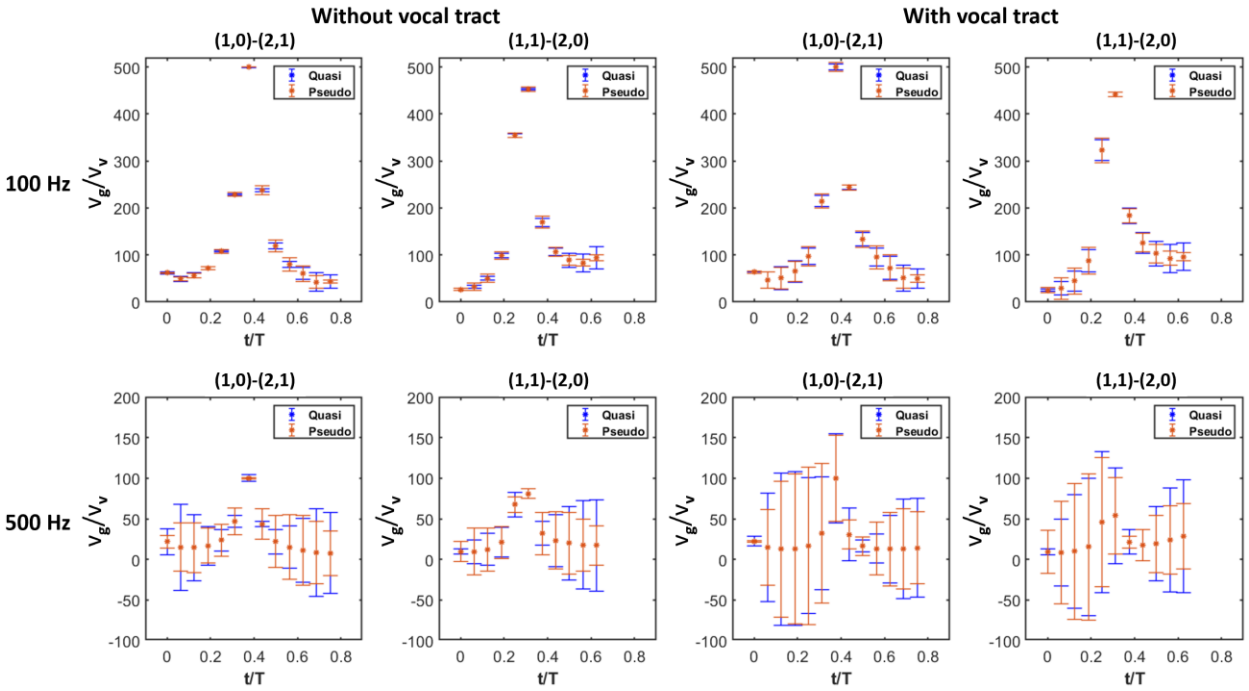

Figure 5.5 v_g/v_v versus t/T during the open phase with the flow rate errors being presented in the form of error bars. Maximum values of v_g/v_v for (1,0)-(2,1) set at 100 Hz are much larger than 500. They are limited at 500 for a better view of other v_g/v_v values.

5.3.3 Errors of important parameters

To understand how the QSFA influences voice outcomes, errors of several important aerodynamic parameters between quasi-steady/pseudo static and dynamic cases were calculated and listed in Table 5.2. Errors of peak flow are minor ($\leqslant$5.35%) for all cases at 100 Hz, and they are also observed to be small at 500 Hz without a vocal tract. A significant decrease in peak flow of dynamic cases is noticed at 500 Hz with a vocal tract, which is consistent with the experimental and numerical results of Kucinschi and Scherer (2006). Kucinschi and Scherer (2006) pointed out that the decrease of peak flow with the increase of vibration frequency is due to the inertive effects of the air in the trachea and vocal tract, which resist the change in flow rate and grow with frequency. This also explains why there are no significant errors of peak flow at 500 Hz without a vocal tract. Similar to the peak flow, small errors ($\leqslant$3.94%) of mean flow are observed at 100 Hz. Errors of mean flow appear more significant at 500 Hz, with a maximum error of 14.07%. In Table 5.2, phase shift was calculated as the phase difference of the peak flow between the

quasi-steady/pseudo static case and dynamic case. Remarkable phase shifts are observed for all cases, and the negative sign before each number indicates that the occurrence of peak flow in the dynamic case is delayed compared with that in the quasi-steady/pseudo static case. By comparing the phase shift values between the case with and without the vocal tract, it is noted that the values significantly increase when a vocal tract is considered. The increased value of the phase shift is due to the air inertia in the vocal tract, which skews the waveform of flow rate to the right. Additionally, as the frequency increases from 100 Hz to 500 Hz, a significant increase in phase shift is also observed for all cases. For both the (1,0)-(2,1) and (1,1)-(2,0) set, errors of MFDR of quasi-steady case are relatively small at both 100 Hz and 500 Hz without a vocal tract, while large errors of MFDR are observed for pseudo static case, which could be attributed to the relatively large phase shift of pseudo static case. Errors of MFDR seem to increase only for quasi-steady case when a vocal tract is considered. No strong correlation between errors of MFDR and frequency is observed in Table 5.2. In general, applying the QSFA yields small errors in flow rate prediction but large errors in phase shift prediction. Moreover, the errors grow substantially when a vocal tract is included or at high vibration frequency. The pseudo static assumption, which considers the effect of glottal wall motion, does not predict a better flow rate waveform, it actually produces larger errors in phase shift and MFDR predictions.

Table 5.2 Errors of important aerodynamic parameters between quasi-steady/pseudo static and dynamic cases

Frequency	Parameters	Without vocal tract		With vocal tract	
		(1,0)-(2,1) quasi/pseudo	(1,1)-(2,0) quasi/pseudo	(1,0)-(2,1) quasi/pseudo	(1,1)-(2,0) quasi/pseudo
100 Hz	Peak flow (%)	-1.18 / -2.18	-1.18 / -3.28	-0.33 / 2.51	3.49 / 5.35
	Mean flow (%)	-2.24 / -3.94	-3.36 / -1.74	-0.61 / -0.39	-2.14 / 2.61
	Phase shift (°)	-11.34 / -20.37	-8.82 / -13.93	-25.48 / -27.83	-24.89 / -28.04
	MFDR (%)	0.49 / -16.55	-9.51 / -20.48	-13.83 / -22.94	-13.19 / -15.81
500 Hz	Peak flow (%)	0.09 / 1.46	2.44 / -6.79	28.12 / 32.95	37.71 / 29.57
	Mean flow (%)	-5.58 / -10.26	-11.01 / -8.56	14.07 / 10.68	4.46 / 10.75
	Phase shift (°)	-21.27 / -44.63	-27.62 / -34.76	-59.11 / -77.93	-61.89 / -66.04
	MFDR (%)	0.79 / -34.05	-2.18 / -25.48	12.84 / -16.15	17.69 / -3.63

5.3.4 Errors of glottal pressure

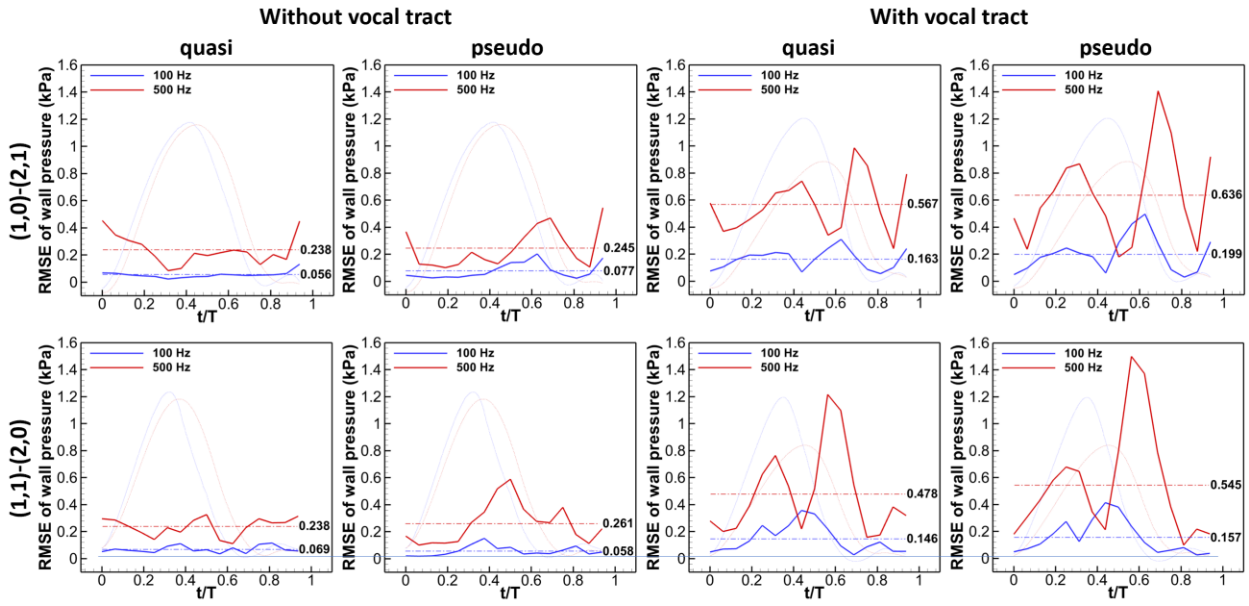

Figure 5.6 RMSE of glottal wall pressure between quasi-steady/pseudo static and dynamic case over one vibration cycle.

Figure 5.6 shows the root-mean-square error (RMSE) of glottal wall pressure between the quasi-steady/pseudo static and the dynamic case over one vibration cycle. For the same element node on the medial surface, pressure differences between quasi-steady/pseudo static and dynamic case were first calculated, RMSE was then obtained by taking the square root of the arithmetic mean of the squares of the pressure differences. In Figure 5.6, RMSE of glottal wall pressure is small at 100 Hz, but significantly increases as frequency increases to 500 Hz. For example, the cycle-averaged RMSE is between 0.056 and 0.077 kPa for all cases at 100 Hz without a vocal tract, while this value is between 0.238 and 0.261 kPa for all cases at 500 Hz without a vocal tract, which represents a nearly fourfold increase. By comparing the cases with and without the vocal tract, a significant increase in RMSE of glottal wall pressure is also observed. Figure 5.6 also shows that RMSEs of glottal wall pressure are almost on the same level for the two different sets, suggesting glottal shapes do not have an important effect on the errors of glottal pressure. The difference of cycle-averaged RMSE of glottal wall pressure between the quasi-steady and pseudo static case is small, although a more significant error is commonly observed during the flow

deceleration stage for pseudo static cases. The trend of errors of glottal pressure is consistent with the observations made above, that is, errors are small at low frequency without a vocal tract, and increasing vibration frequency or including a vocal tract remarkably increases the errors.

5.3.5 Momentum budget analysis

To evaluate the unsteady acceleration in glottal airflow and assess the relative importance of unsteady effects, a momentum budget analysis was conducted for each case. The Navier-Stokes momentum equation was nondimensionalized as follows:

$$\frac{\partial \vec{V}}{\partial t} + (\vec{V} \cdot \vec{\nabla})\vec{V} = -\vec{\nabla}P + \frac{1}{Re}\vec{\nabla}^2\vec{V} \tag{5.4}$$

where $\frac{\partial \vec{V}}{\partial t}$ is the unsteady acceleration term, $(\vec{V} \cdot \vec{\nabla})\vec{V}$ is the convection term, $\vec{\nabla}P$ and $\frac{1}{Re}\vec{\nabla}^2\vec{V}$ are the pressure and shear stress terms, respectively. To quantify and compare the magnitude of different terms, the glottis was taken as the control volume (CV), and a volume integral was performed for the convection, pressure and shear stress term within the control volume. By utilizing the divergence theorem, the volume integrals were converted to the surface integrals over the boundary of the glottis, which were evaluated in the flow direction (y-direction) as follows:

$$\iiint_{CV} (\vec{V} \cdot \vec{\nabla})\vec{V} dV = \oiint_{\partial V} \vec{V}(\vec{V} \cdot \vec{n} dS) = -\sum V_{1i} \cdot Q_{1i} + \sum V_{2i} \cdot Q_{2i} \quad (Convection) \tag{5.4}$$

$$\iiint_{CV} \vec{\nabla}P dV = \oiint_{\partial V} P\vec{n}\, dS = -\sum P_{1i} \cdot A_{1i} + \sum P_{2i} \cdot A_{2i} + \sum P_{wi} \cdot A_{wi} \quad (Pressure) \tag{5.5}$$

$$\begin{aligned} \iiint_{CV} \frac{1}{Re}\vec{\nabla}^2\vec{V} dV &= \frac{1}{Re} \oiint_{\partial V} (\nabla V_x,\ \nabla V_y, \nabla V_z)\vec{n} dS \\ &= \frac{1}{Re}\left[\sum \left(\frac{\partial V}{\partial x} dydz\right)_i + \sum \left(\frac{\partial V}{\partial z} dxdy\right)_i\right] \quad (Shear) \end{aligned} \tag{5.6}$$

where the subscript 1, 2 and w denote the glottal inlet, outlet and wall, respectively, the subscript *i* (*1, 2, 3, …*) denotes the grid number through the corresponding area. Because the simulation results were not

output at every time step, the volume integral of the unsteady acceleration term was obtained through balancing the momentum equation, that is,

$$\iiint_{CV} \frac{\partial \vec{V}}{\partial t} dV = -\iiint_{CV} (\vec{V} \cdot \vec{\nabla})\vec{V} dV - \iiint_{CV} \vec{\nabla} P dV + \iiint_{CV} \frac{1}{Re} \vec{\nabla}^2 \vec{V} dV \quad (Unsteady) \qquad (5.7)$$

Figure 5.7 shows the variation of the convection, pressure, shear and unsteady term over one vibration cycle in each dynamic case. At 100 Hz, the magnitude of the unsteady acceleration term is markedly smaller than that of the convection and pressure term, but still of the same order of magnitude as the later and is non-negligible, which explains the approximate 10% errors of flow rate between the quasi-steady/pseudo static and dynamic cases. The unsteady acceleration term increases significantly and at about the same level as the convection and pressure term at 500 Hz. For the cases without a vocal tract and with a vocal tract, the greatest unsteady acceleration respectively increases about fivefold and fourfold, as compared with that at 100 Hz. This suggests that the unsteady term grows with frequency.

Within one vibration cycle, the variation of the unsteady acceleration shows an approximately sinusoidal shape in Figure 5.7. The term is positive during the opening phase and negative during the closing phase, indicating the flow first accelerates and then decelerates in a cycle. For both the flow acceleration and deceleration phase, the magnitude of the unsteady term first rises and subsequently declines, and a peak appears at roughly the middle stage of the phase. The term is nearly zero only in a brief moment around the transition between flow acceleration and deceleration, as well as a short time period at early opening (100 Hz) and near the closed phase (100 and 500 Hz). In most of a glottal cycle, the unsteady term is not trivial at both low and high frequencies. This observation differs from the previous finding that the flow unsteadiness is important merely at the short instants of flow initiation and shutoff, but agrees with the results of Ringenberg et al. (2021) given the transglottal and dynamic pressures computed by the unsteady Bernoulli equation in their study are about equal just during a short time interval around the maximum glottal opening. A similar sinusoidal variation is also noted in the flow rate errors in Figure 5.4. By

overlapping the horizontal axis of the unsteady acceleration plot with that of the flow rate error plot and appropriately adjusting the magnitude of the variables, it is observed that the variation of flow rate errors closely follows that of the unsteady acceleration, suggesting the former is highly associated with the later. By comparing the cases without a vocal tract and with a vocal tract in Figure 5.7, it is observed that including a vocal tract has little effect on the magnitude of unsteady term at 100 Hz. However, a substantial increase in the errors was formerly observed in the cases with a vocal tract. It indicates that the increased errors are not induced by the unsteady effect of flow acceleration in glottis, but due to the inertance effect of the air column in the vocal tract or the interaction between the two effects. At 500 Hz, the magnitude of the unsteady acceleration is significantly smaller in the cases with a vocal tract, owning to the air inertia in the tract which resists the change of flow velocity. The decreased flow unsteadiness confirms that the larger errors in the cases with a vocal tract cannot be simply caused by the flow acceleration in the glottis.

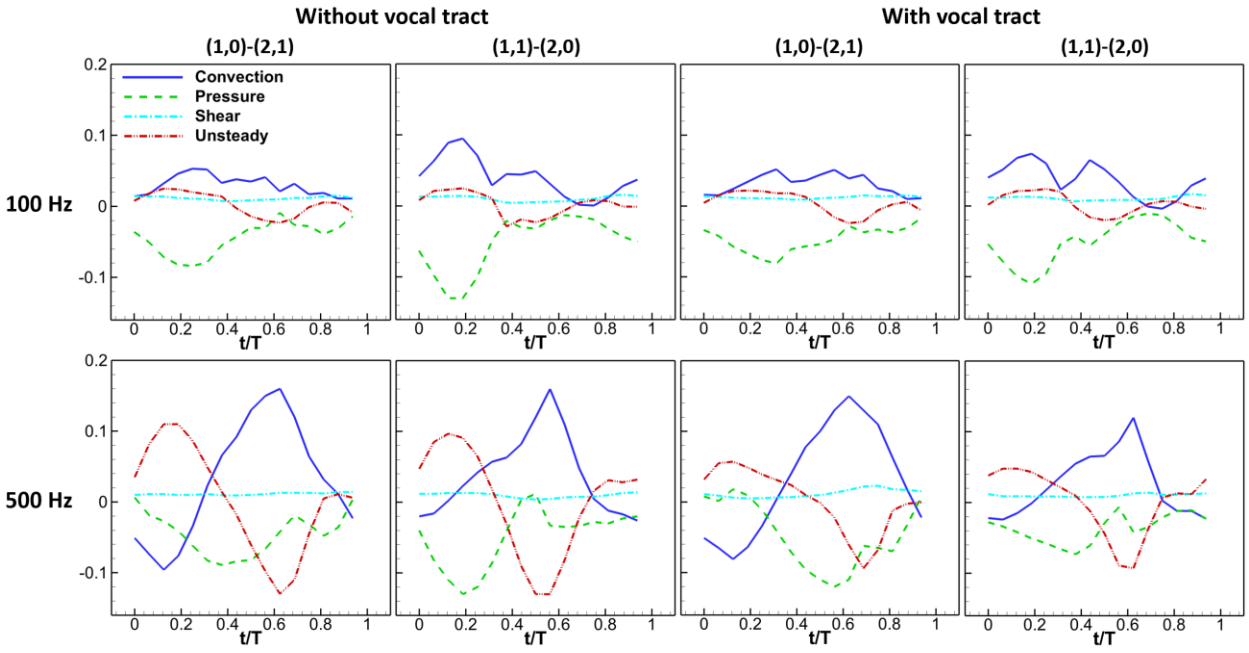

Figure 5.7 Variation of convection, pressure, shear and unsteady term over one vibration cycle in each dynamic case.

Figure 5.8 shows a comparison of convection, pressure and shear term between the dynamic and quasi-steady/pseudo static cases. For the three different simulations, the magnitude and variation of the three terms exhibit great similarities at 100 Hz, suggesting the dynamics of the glottal flow are very similar among the three simulations, and the contributions of unsteady acceleration are small. As the frequency increases to 500 Hz, significant deviations of the convection, pressure and shear term among the three simulations are observed. For the convection term, the quasi-steady predictions are radically different from the dynamic simulations, in both the magnitude and variation trend. However, the variation of pseudo static predictions closely follows that of the dynamic simulations, and their magnitudes are within the same level, which indicates that the displaced flow by wall motion might play an important role in determining the convective acceleration. Compared with the dynamic simulations, the pressure term predicted by adopting the two assumptions are quite different in general, indicating large prediction errors in the transglottal pressure force and vocal fold drag at high frequency. For the shear term, the quasi-steady predictions are nearly the same as the dynamic simulations, while the pseudo static predictions are completely incorrect during the early opening and sometimes the closed phase.

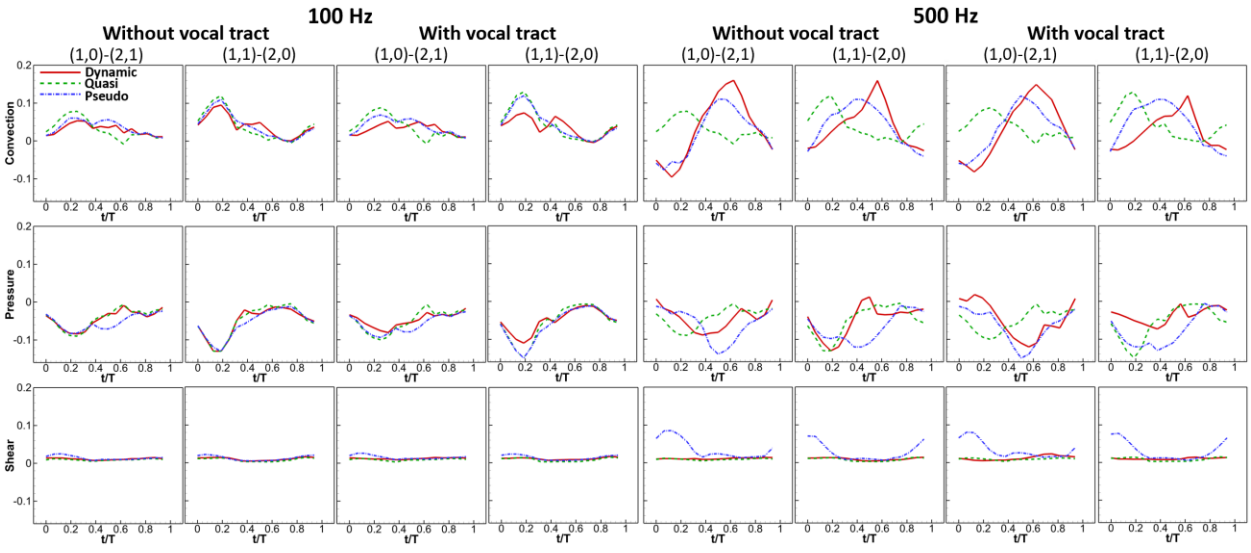

Figure 5.8 Comparison of convection, pressure and shear term between the dynamic and quasi-steady/pseudo static cases.

5.4 Discussion and conclusions

In this study, the validity of the quasi-steady assumption is numerically tested by examining the errors of glottal flow and glottal pressures under two sets of irregular glottal shapes. Dynamic, quasi-steady and pseudo static simulations are designed, and simulation setups with and without the supraglottic tract are considered. Vibration frequency of the vocal folds is chosen to be 100 and 500 Hz. The results show that applying the assumption at 100 Hz frequency yields small errors in flow rate and glottal pressure. The momentum budget analysis further reveals that the dynamics of the quasi-steady glottal flow highly resembles that of an unsteady glottal flow. In contrast, the assumption breaks down at 500 Hz, which is reflected in the severe errors in the predicted flow rate and glottal pressure, as well as the significant difference in the terms of the momentum equation between quasi-steady/pseudo static and dynamic simulations. In the current study, the unsteady acceleration is found to be non-negligible (as compared to the convective acceleration) even at low vibration frequency, which agrees with the experimental results made by Krane et al. (2010). This reveals that the glottal flow is essentially unsteady. However, in the frequency range of normal speech phonation (~ 100 Hz-200 Hz), the contribution of the unsteady acceleration is small, and the quasi-steady approximation gives fair predictions in the glottal flow and pressure. Therefore, the assumption can be deemed valid. As the frequency increases, there should be a point where the errors in the predicted flow are not acceptable, and the assumption no longer stands. Based on current results, the assumption is apparently not suitable for the study of voice production above 500 Hz.

Similar to the studies of Mongeau et al. (1997), Zhang et al. (2002), and Vilain et al. (2004), the errors in glottal flow are found to be nonuniform throughout the vibration cycle. Previous results showed more significant errors at the early opening and late closing phase while very small errors at the middle interval of the cycle. This variation of errors in the cycle is believed to be related to the importance of glottal wall motion (Krane et al., 2006). At the short stages of flow initiation and shutoff, the velocity of the glottal

wall can be considerable compared to that of glottal flow, and the unsteady effects of wall motion become appreciable. In contrast, at the stages around the maximum glottal opening, the wall velocity is much smaller than the flow velocity, and the unsteady effects of wall motion are no longer important. In current study, the two peaks of flow rate errors during the open phase do not take place strictly in the stages of flow initiation and shutoff, but some time after the initiating point and before the closure point. The reason for this difference is that the two sets of glottal shapes never generate a fully closed glottis (see Figure 5.2), and the flow does not actually develop from zero velocity and end with a complete stop. Nevertheless, the velocity ratio (v_g/v_v) analysis still supports the previous argument. Generally, the errors of flow rate are larger at the stages where the relative magnitude of airflow velocity and vocal fold velocity is small, while the errors become small as the relative magnitude increases.

This study also shows that the inertance effect of the air column in the vocal tract can cause a significant increase in the errors of quasi-steady flow. As shown in the momentum budget analysis, including vocal tract inertance does not increase the unsteadiness of glottal flow, which suggests the increased errors are caused by a mechanism different from that of increasing frequency. Using lumped-element models of the vocal tract inertance, Titze (2006, Chapter 5) has shown that there exist nonlinear interactions between the vocal tract inertance and glottal airflow, and the effects of the interactions are skewing the flow rate waveform to the right and reducing the peak flow. In the current study, these effects are also observed in the results of dynamic simulations. Nevertheless, such effects will not appear in the quasi-steady simulations because the flow is steady, and there are no interactions between the glottal flow and air inertia in the vocal tract.

The two different sets of glottal shapes used in this study give consistent results, suggesting the glottal shapes have little effect on the validity of the quasi-steady assumption. Compared to those in previous studies, the glottal geometries in the current study are much more complex, which not only show the alternating convergent-divergent characteristic of vocal fold vibration, but also include the left-right

asymmetry and longitudinal wave propagation that could occur under irregular vocal fold vibrations. Even though the glottal shapes are quite different, many findings in the current study are similar to those reported in previous studies, which also supports the conclusion that the effects of glottal geometry on the quasi-steady assumption are insignificant.

The very similar results of the quasi-steady and pseudo static simulation at 100 Hz indicates that the effect of glottal wall motion is not important, and deviations from the dynamic simulations are mainly induced by the local acceleration ($\frac{\partial \vec{V}}{\partial t}$). As discussed above, the importance of glottal wall motion depends on the relative magnitude of the glottal wall velocity to the flow velocity. At a relatively low vibration frequency, the wall velocity cannot be comparable with the flow velocity except at the instants when flow is just initiated or is close to shut-off, thus the unsteady effects of glottal wall motion cannot be significant for most of the cycle. This is consistent with Deverge et al. (2003) and Krane and Wei (2006) who pointed out that the flow unsteadiness due to wall movement is not important. The study of Xi et al. (2018) also shows that the primary effect of the glottal wall motion is on the secondary flow in the transverse direction which involves flow separation, swirling flows and vortex shedding, while the instability of the main flow in the streamwise direction is largely affected by the unsteady acceleration of flow itself. In contrast, the glottal wall motion can become important at high frequencies, as shown by the difference between the flows predicted in the quasi-steady and pseudo static simulations at 500 Hz, especially during the flow deceleration stage. Additionally, the momentum budget analysis suggests that the displaced flow by glottal wall motion may contribute significantly to the convective acceleration.

One limitation of this study is that the shape of the supraglottic tract is not realistic. The real epilarynx airway is narrower, especially around the vestibular fold area (Saldias et al., 2021). Since the vocal tract inertance is reversely proportional to the cross-sectional area (Titze, 2001), it can be expected that the inertance effect will be more prominent if a more lifelike supraglottic tract is considered. The other limitation is that only two sets of glottal shapes were studied, which is due to the high computational cost

of the numerical simulation. As a result, the conclusion that the quasi-steady assumption is insensitive to glottal shapes is limited to the two sets of glottal shapes used in this study. To draw a more general conclusion, future studies need to include a wider range of glottal shapes (e.g., involving higher vibration modes; changing the prephonatory glottal configuration from uniform to divergent or convergent).

CHAPTER 6 SUMMARY AND CONCLUSIONS

Human voice production arises from the multiphysics interaction between vocal fold biomechanics and airflow dynamics. The voice outcomes are highly dependent on the biomechanical properties of vocal fold and the characteristics of glottal flow. The purposes of this dissertation are to use computational methods: (1) to understand the cause-and-effect relationship between vocal fold stiffness and voice production; (2) to further examine the range of validity of the quasi-steady assumption for glottal flow under complex glottal shapes. The main contributions are summarized as follows:

A three-dimensional flow-structure interaction model of voice production is used to investigate the effect of the stiffness parameters of vocal fold layers on voice production. The vocal fold is modeled as a three-layer structure consisting of the cover, ligament and body layers. All the three layers are modeled as transversely isotropic materials for which the stiffness parameters include the transverse elastic modulus and longitudinal elastic modulus. The results show that, in addition to the obvious monotonic effects on the fundamental frequency, flow rate and glottis opening, the stiffness parameters also have significant and non-monotonic effects on the divergent angle, open quotient and closing velocity. It is further found that the longitudinal stiffness parameters generally have more significant impacts on glottal flows and vocal fold vibrations than the transverse stiffness parameters. The sensitivity analysis shows that, among all the stiffness parameters, the transverse and longitudinal stiffness of the ligament layer have the most dominant effect on most output measures.

While conus elasticus is generally considered a part of continuation of the vocal ligament, histological studies have revealed different fiber orientations that fibers are primarily aligned in the caudal-cranial direction in the conus elasticus and in the anterior-posterior direction in the vocal ligament. In chapter 4, two continuum vocal fold models are constructed with two different fiber orientations in the conus elasticus: the caudal-cranial direction and the anterior-posterior direction. Flow-structure interaction simulations are conducted at different subglottal pressures to investigate the effects of fiber orientation

in the conus elasticus on vocal fold vibrations, aerodynamic and acoustic measures of voice production. The results reveal that including the realistic fiber orientation (caudal-cranial) in the conus elasticus yields smaller structural stiffness and larger deflection at the junction of the conus elasticus and ligament and subsequently leads to a greater vibration amplitude and larger mucosal wave amplitude of the vocal fold. The smaller stiffness of conus elasticus also causes a larger peak flow rate and higher speed quotient. Furthermore, the voice generated by the vocal fold model with a realistic conus elasticus has a lower fundamental frequency, smaller first harmonic amplitude and smaller spectral slope.

Another important contribution of this dissertation is a further examination of the quasi-steady assumption for glottal flow by considering different vocal fold vibration frequencies, irregular glottal shapes and vocal tract inertance effect. To separate the effects of flow unsteadiness and glottal wall motion, three different types of numerical simulation, namely the dynamic, pseudo static and quasi-steady simulation, are performed. The dependence of the assumption on frequency is tested under a normal speech frequency (100 Hz) and a high fundamental frequency (500 Hz). The two sets of asymmetric glottal shapes used in chapter 5 incorporate different vertical and longitudinal wave patterns on the right and left sides, as well as complex contact patterns which generate various flow channels such as split, merged, multichannel, etc. The air inertia effect in the supraglottic vocal tract is tested by implementing the simulations with and without a vocal tract. The results show that at the normal speech frequency, the flow rate and glottal pressure predicted by the quasi-steady assumption have small errors, and the dynamics of the quasi-steady flow is very similar to that of the unsteady flow. However, at the high fundamental frequency, the assumption induces very larges errors in the predicted flow rate and glottal pressure and is deemed invalid. The air inertia effect in the supraglottic tract also has a significant effect on the validity of the assumption, due to the nonlinear interaction between the vocal tract inertance and unsteady glottal flow. Finally, the effects of glottal shapes and glottal wall motion on the assumption are found to be insignificant.

In the computational modeling of voice production, assumptions are frequently made to reduce the computational effort, which include linear material assumption for vocal fold tissues, incompressible flow assumption, inviscid flow assumption, etc. Some of these assumptions are adopted in the current work and should be regarded as one of the limitations in this research. However, properly applying these assumptions can produce solid results. For example, if the research focus is on vocal fold dynamics, adopting the incompressible flow assumption may be reasonable because it might not have an important effect on the flow pressure exerted on the vocal folds.

Finally, some future works are needed to improve the current research. In chapter 5, the glottal wall motion is prescribed, and the simulation is one way coupling, which is different from the situation in real voice production. If the deformations of the vocal folds are large enough, the volume flux produced by the prescribed methodology can be very different from that by the fully coupled approach, which could affect the validity of the quasi-steady assumption. In the future, fluid-structure interaction simulations should be performed to investigate how important this effect could be. Also in chapter 5, the quasi-steady assumption is deemed valid at 100 Hz but invalid at 500 Hz fundamental frequency. Future work should consider the intermediate frequencies between 100 Hz and 500 Hz and find the critical frequency point that makes the errors produced by the assumption not acceptable. This could provide guidance for appropriately applying the quasi-steady assumption in phonation research.

BIBLIOGRAPHY

Alipour, F., & Scherer, R. C. (2000). Vocal fold bulging effects on phonation using a biophysical computer model. *Journal of Voice*, *14*(4), 470-483.

Alipour, F., & Titze, I. (1995). Combined simulation of two dimensional airflow and vocal fold vibration. *Status and Progress Report, National Center for Voice and Speech*, *8*, 9-14.

Alipour, F., Berry, D. A., & Titze, I. R. (2000). A finite-element model of vocal-fold vibration. *The Journal of the Acoustical Society of America*, *108*(6), 3003-3012.

Alipour, F., Scherer, R. C., & Patel, V. C. (1995). An experimental study of pulsatile flow in canine larynges.

Alipour-Haghighi, F., & Titze, I. R. (1991). Elastic models of vocal fold tissues. *The Journal of the Acoustical Society of America*, *90*(3), 1326-1331.

Atkinson, J. E. (1978). Correlation analysis of the physiological factors controlling fundamental voice frequency. *The journal of the Acoustical Society of America*, *63*(1), 211-222.

Berry, D. A., & Titze, I. R. (1996). Normal modes in a continuum model of vocal fold tissues. *The Journal of the Acoustical Society of America*, *100*(5), 3345-3354.

Berry, D. A., Herzel, H., Titze, I. R., & Krischer, K. (1994). Interpretation of biomechanical simulations of normal and chaotic vocal fold oscillations with empirical eigenfunctions. *The Journal of the Acoustical Society of America*, *95*(6), 3595-3604.

Chan, R. W., Fu, M., Young, L., & Tirunagari, N. (2007). Relative contributions of collagen and elastin to elasticity of the vocal fold under tension. *Annals of biomedical engineering*, *35*(8), 1471-1483.

Chan, R. W., Titze, I. R., & Titze, M. R. (1997). Further studies of phonation threshold pressure in a physical model of the vocal fold mucosa. *The Journal of the Acoustical Society of America*, *101*(6), 3722-3727.

Chhetri, D. K., Zhang, Z., & Neubauer, J. (2011). Measurement of Young's modulus of vocal folds by indentation. *Journal of Voice*, *25*(1), 1-7.

Cook, D. D., Nauman, E., & Mongeau, L. (2008). Reducing the number of vocal fold mechanical tissue properties: Evaluation of the incompressibility and planar displacement assumptions. *The Journal of the Acoustical Society of America*, *124*(6), 3888-3896.

Deguchi, S., & Hyakutake, T. (2009). Theoretical consideration of the flow behavior in oscillating vocal fold. *Journal of biomechanics*, *42*(7), 824-829.

Deverge, M., Pelorson, X., Vilain, C., Lagrée, P. Y., Chentouf, F., Willems, J., & Hirschberg, A. (2003). Influence of collision on the flow through in-vitro rigid models of the vocal folds. *The Journal of the Acoustical Society of America*, *114*(6), 3354-3362.

Drechsel, J. S., & Thomson, S. L. (2008). Influence of supraglottal structures on the glottal jet exiting a two-layer synthetic, self-oscillating vocal fold model. *The Journal of the Acoustical Society of America*, *123*(6), 4434-4445.

Erath, B. D., & Plesniak, M. W. (2005). An experimental investigation of velocity fields in divergent glottal models of the human vocal tract. *The Journal of the Acoustical Society of America*, *118*(3), 2025-2025.

Erath, B. D., & Plesniak, M. W. (2006a). The occurrence of the Coanda effect in pulsatile flow through static models of the human vocal folds. *The Journal of the Acoustical Society of America*, *120*(2), 1000-1011.

Erath, B. D., & Plesniak, M. W. (2006b). An investigation of jet trajectory in flow through scaled vocal fold models with asymmetric glottal passages. *Experiments in fluids*, *41*(5), 735-748.

Erath, B. D., & Plesniak, M. W. (2010). An investigation of asymmetric flow features in a scaled-up driven model of the human vocal folds. *Experiments in Fluids*, *49*(1), 131-146.

Erath, B. D., Zañartu, M., & Peterson, S. D. (2017). Modeling viscous dissipation during vocal fold contact: The influence of tissue viscosity and thickness with implications for hydration. *Biomechanics and modeling in mechanobiology*, *16*(3), 947-960.

Erath, B. D., Zañartu, M., Peterson, S. D., & Plesniak, M. W. (2011). Nonlinear vocal fold dynamics resulting from asymmetric fluid loading on a two-mass model of speech. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, *21*(3), 033113.

Finkelhor, B. K., Titze, I. R., & Durham, P. L. (1988). The effect of viscosity changes in the vocal folds on the range of oscillation. *Journal of Voice*, *1*(4), 320-325.

Flanagan, J. L., Rabiner, L. R., Christopher, D., Bock, D. E., & Shipp, T. (1976). Digital analysis of laryngeal control in speech production. *The Journal of the Acoustical Society of America*, *60*(2), 446-455.

Flanagan, J., & Landgraf, L. (1968). Self-oscillating source for vocal-tract synthesizers. *IEEE Transactions on Audio and Electroacoustics*, *16*(1), 57-64.

Geng, B., Pham, N., Xue, Q., & Zheng, X. (2020). A three-dimensional vocal fold posturing model based on muscle mechanics and magnetic resonance imaging of a canine larynx. *The Journal of the Acoustical Society of America*, *147*(4), 2597-2608.

Gray, S. D., Alipour, F., Titze, I. R., & Hammond, T. H. (2000). Biomechanical and histologic observations of vocal fold fibrous proteins. *Annals of Otology, Rhinology & Laryngology*, *109*(1), 77-85.

Herzel, H., & Knudsen, C. (1995). Bifurcations in a vocal fold model. *Nonlinear Dynamics*, *7*(1), 53-64.

Herzel, H., Berry, D., Titze, I., & Steinecke, I. (1995). Nonlinear dynamics of the voice: signal analysis and biomechanical modeling. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, *5*(1), 30-34.

Hirano, M. (1974). Morphological structure of the vocal cord as a vibrator and its variations. *Folia phoniatrica et logopaedica*, *26*(2), 89-94.

Hirano, M. (1981). The structure of the vocal folds. *Vocal fold physiology*, 33-41.

Hirano, M. (1985). Cover-body theory of vocal cord vibration. *Speech science*, 1-46.

Hofmans, G. C. J., Groot, G., Ranucci, M., Graziani, G., & Hirschberg, A. (2003). Unsteady flow through in-vitro models of the glottis. *The Journal of the Acoustical Society of America*, *113*(3), 1658-1675.

Holmberg, E. B., Hillman, R. E., & Perkell, J. S. (1988). Glottal airflow and transglottal air pressure measurements for male and female speakers in soft, normal, and loud voice. *The Journal of the Acoustical Society of America*, *84*(2), 511-529.

Holmberg, E. B., Hillman, R. E., & Perkell, J. S. (1989). Glottal airflow and transglottal air pressure measurements for male and female speakers in low, normal, and high pitch. *Journal of Voice*, *3*(4), 294-305.

Holmberg, E. B., Hillman, R. E., Perkell, J. S., Guiod, P. C., & Goldman, S. L. (1995). Comparisons among aerodynamic, electroglottographic, and acoustic spectral measures of female voice. *Journal of Speech, Language, and Hearing Research*, *38*(6), 1212-1223.

Hunter, E. J., Titze, I. R., & Alipour, F. (2004). A three-dimensional model of vocal fold abduction/adduction. *The Journal of the Acoustical Society of America*, *115*(4), 1747-1759.

Ishizaka, K., & Flanagan, J. L. (1972). Synthesis of voiced sounds from a two-mass model of the vocal cords. *Bell system technical journal*, *51*(6), 1233-1268.

Ishizaka, K., & Kaneko, T. (1968). On equivalent mechanical constants of the vocal cords. *THE JOURNAL OF THE ACOUSTICAL SOCIETY OF JAPAN*, *24*(5), 312-313.

Jiang, W., Farbos de Luzan, C., Wang, X., Oren, L., Khosla, S. M., Xue, Q., & Zheng, X. (2022). Computational modeling of voice production using excised canine larynx. *Journal of Biomechanical Engineering*, *144*(2).

Kelleher, J. E., Siegmund, T., Du, M., Naseri, E., & Chan, R. W. (2013a). Empirical measurements of biomechanical anisotropy of the human vocal fold lamina propria. *Biomechanics and modeling in mechanobiology*, *12*(3), 555-567.

Kelleher, J. E., Siegmund, T., Du, M., Naseri, E., & Chan, R. W. (2013b). The anisotropic hyperelastic biomechanical response of the vocal ligament and implications for frequency regulation: A case study. *The Journal of the Acoustical Society of America*, *133*(3), 1625-1636.

Klatt, D. H., & Klatt, L. C. (1990). Analysis, synthesis, and perception of voice quality variations among female and male talkers. *the Journal of the Acoustical Society of America*, *87*(2), 820-857.

Krane, M. H., & Wei, T. (2006). Theoretical assessment of unsteady aerodynamic effects in phonation. *The Journal of the Acoustical Society of America*, *120*(3), 1578-1588.

Krane, M. H., Barry, M., & Wei, T. (2010). Dynamics of temporal variations in phonatory flow. *The Journal of the Acoustical Society of America*, *128*(1), 372-383.

Krane, M., Barry, M., & Wei, T. (2007). Unsteady behavior of flow in a scaled-up vocal folds model. *The Journal of the Acoustical Society of America*, *122*(6), 3659-3670.

Krebs, F., Silva, F., Sciamarella, D., & Artana, G. (2012). A three-dimensional study of the glottal jet. *Experiments in fluids*, *52*(5), 1133-1147.

Kucinschi, B. R., Scherer, R. C., DeWitt, K. J., & Ng, T. T. (2006). An experimental analysis of the pressures and flows within a driven mechanical model of phonation. The Journal of the Acoustical Society of America, 119(5), 3011-3021.

Lee, D. J., Kim, M. W., Kim, S. Y., Lee, S. H., & Youn, J. R. (2010). Three dimensional flow simulation and structural analysis on stiffness of fiber reinforced anisotropic parts. *Korea-Australia Rheology Journal*, *22*(2), 95-103.

Luo, H., Mittal, R., & Bielamowicz, S. A. (2009). Analysis of flow-structure interaction in the larynx during phonation using an immersed-boundary method. *The Journal of the Acoustical Society of America*, *126*(2), 816-824.

Luo, H., Mittal, R., Zheng, X., Bielamowicz, S. A., Walsh, R. J., & Hahn, J. K. (2008). An immersed-boundary method for flow–structure interaction in biological systems with application to phonation. *Journal of computational physics*, *227*(22), 9303-9332.

Martin, R. B., & Boardman, D. L. (1993). The effects of collagen fiber orientation, porosity, density, and mineralization on bovine cortical bone bending properties. *Journal of biomechanics*, *26*(9), 1047-1054.

Mendelsohn, A. H., & Zhang, Z. (2011). Phonation threshold pressure and onset frequency in a two-layer physical model of the vocal folds. *The Journal of the Acoustical Society of America*, *130*(5), 2961-2968.

Migimatsu, K., & Tokuda, I. T. (2019). Experimental study on nonlinear source–filter interaction using synthetic vocal fold models. *The Journal of the Acoustical Society of America*, *146*(2), 983-997.

Min, Y. B., Titze, I. R., & Alipour-Haghighi, F. (1995). Stress-strain response of the human vocal ligament. *Annals of Otology, Rhinology & Laryngology*, *104*(7), 563-569.

Miri, A. K., Tripathy, U., Mongeau, L., & Wiseman, P. W. (2012). Nonlinear laser scanning microscopy of human vocal folds. *The Laryngoscope*, *122*(2), 356-363.

Mittal, R., Dong, H., Bozkurttas, M., Najjar, F. M., Vargas, A., & Von Loebbecke, A. (2008). A versatile sharp interface immersed boundary method for incompressible flows with complex boundaries. Journal of computational physics, 227(10), 4825-4852.

Mongeau, L., Franchek, N., Coker, C. H., & Kubli, R. A. (1997). Characteristics of a pulsating jet through a small modulated orifice, with application to voice production. *The Journal of the Acoustical Society of America*, *102*(2), 1121-1133.

Murray, P. R., & Thomson, S. L. (2012). Vibratory responses of synthetic, self-oscillating vocal fold models. *The Journal of the Acoustical Society of America*, *132*(5), 3428-3438.

Murray, P. R., Thomson, S. L., & Smith, M. E. (2014). A synthetic, self-oscillating vocal fold model platform for studying augmentation injection. *Journal of Voice*, *28*(2), 133-143.

Namani, R., Feng, Y., Okamoto, R. J., Jesuraj, N., Sakiyama-Elbert, S. E., Genin, G. M., & Bayly, P. V. (2012). Elastic characterization of transversely isotropic soft materials by dynamic shear and asymmetric indentation. *Journal of biomechanical engineering*, *134*(6).

Neubauer, J., Mergell, P., Eysholdt, U., & Herzel, H. (2001). Spatio-temporal analysis of irregular vocal fold oscillations: Biphonation due to desynchronization of spatial modes. *The Journal of the Acoustical Society of America*, *110*(6), 3179-3192.

Neubauer, J., Zhang, Z., Miraghaie, R., & Berry, D. A. (2007). Coherent structures of the near field flow in a self-oscillating physical model of the vocal folds. *The Journal of the Acoustical Society of America*, *121*(2), 1102-1118.

Nielson, J. R., Daily, D. J., Truscott, T. T., Luegmair, G., Döllinger, M., & Thomson, S. L. (2013, November). Simultaneous tracking of vocal fold superior surface motion and glottal jet dynamics. In *ASME International Mechanical Engineering Congress and Exposition* (Vol. 56222, p. V03BT03A039). American Society of Mechanical Engineers.

Oren, L., Dembinski, D., Gutmark, E., & Khosla, S. (2014). Characterization of the vocal fold vertical stiffness in a canine model. *Journal of Voice*, *28*(3), 297-304.

Pelorson, X. (2001). On the meaning and accuracy of the pressure–flow technique to determine constriction areas within the vocal tract. Speech communication, 35(3-4), 179-190.

Pham, N., Xue, Q., & Zheng, X. (2018). Coupling between a fiber-reinforced model and a Hill-based contractile model for passive and active tissue properties of laryngeal muscles: A finite element study. *The Journal of the Acoustical Society of America*, *144*(3), EL248-EL253.

Pickup, B. A., & Thomson, S. L. (2009). Influence of asymmetric stiffness on the structural and aerodynamic response of synthetic vocal fold models. *Journal of biomechanics*, *42*(14), 2219-2225.

Pickup, B. A., & Thomson, S. L. (2010). Flow-induced vibratory response of idealized versus magnetic resonance imaging-based synthetic vocal fold models. *The Journal of the Acoustical Society of America*, *128*(3), EL124-EL129.

Pickup, B. A., & Thomson, S. L. (2011). Identification of geometric parameters influencing the flow-induced vibration of a two-layer self-oscillating computational vocal fold model. *The Journal of the Acoustical Society of America*, *129*(4), 2121-2132.

Prades, J. M., Dumollard, J. M., Duband, S., Timoshenko, A., Richard, C., Dubois, M. D., ... & Peoc'h, M. (2010). Lamina propria of the human vocal fold: histomorphometric study of collagen fibers. *Surgical and radiologic anatomy*, *32*(4), 377-382.

Prithishkumar, I. J., & Felicia, C. (2014). Histology of the cricothyroid membrane: a clinical perspective. *Indian Journal of Otolaryngology and Head & Neck Surgery*, *66*(1), 316-319.

Reidenbach, M. M. (1996). The attachments of the conus elasticus to the laryngeal skeleton: Physiologic and clinical implications. *Clinical Anatomy: The Official Journal of the American Association of Clinical Anatomists and the British Association of Clinical Anatomists*, *9*(6), 363-370.

Ringenberg, H., Rogers, D., Wei, N., Krane, M., & Wei, T. (2021). Phase-averaged and cycle-to-cycle analysis of jet dynamics in a scaled up vocal-fold model. *Journal of fluid mechanics*, *918*.

Roubeau, B., Chevrie-Muller, C., & Saint Guily, J. L. (1997). Electromyographic activity of strap and cricothyroid muscles in pitch change. *Acta oto-laryngologica*, *117*(3), 459-464.

Ruty, N., Pelorson, X., Van Hirtum, A., Lopez-Arteaga, I., & Hirschberg, A. (2007). An in vitro setup to test the relevance and the accuracy of low-order vocal folds models. *The Journal of the Acoustical Society of America*, *121*(1), 479-490.

Saldías, M., Laukkanen, A. M., Guzmán, M., Miranda, G., Stoney, J., Alku, P., & Sundberg, J. (2020). The vocal tract in loud twang-like singing while producing high and low pitches. Journal of voice, 35(5), 807.e1-807.e23.

Sataloff, R. T. (2017). *Voice science*. Plural Publishing.

Sato, K. (2018). *Functional histoanatomy of the human larynx*. Springer.

Scherer, R. C., Shinwari, D., De Witt, K. J., Zhang, C., Kucinschi, B. R., & Afjeh, A. A. (2001). Intraglottal pressure profiles for a symmetric and oblique glottis with a divergence angle of 10 degrees. *The Journal of the Acoustical Society of America*, *109*(4), 1616-1630.

Scherer, R.C. and Guo, C.G., 1990. Laryngeal modeling: Translaryngeal pressure for a model with many glottal shapes. In First International Conference on Spoken Language Processing.

Schneider, B., & Bigenzahn, W. (2003). Influence of glottal closure configuration on vocal efficacy in young normal-speaking women. *Journal of Voice*, *17*(4), 468-480.

Schwarz, R., Döllinger, M., Wurzbacher, T., Eysholdt, U., & Lohscheller, J. (2008). Spatio-temporal quantification of vocal fold vibrations using high-speed videoendoscopy and a biomechanical model. *The Journal of the Acoustical Society of America*, *123*(5), 2717-2732.

Shinwari, D., Scherer, R. C., DeWitt, K. J., & Afjeh, A. A. (2003). Flow visualization and pressure distributions in a model of the glottis with a symmetric and oblique divergent angle of 10 degrees. *The Journal of the Acoustical Society of America*, *113*(1), 487-497.

Smith, S. L., & Hunter, E. J. (2014). A viscoelastic laryngeal muscle model with active components. *The Journal of the Acoustical Society of America*, *135*(4), 2041-2051.

Smith, S. L., & Titze, I. R. (2018). Vocal fold contact patterns based on normal modes of vibration. Journal of biomechanics, 73, 177-184.

Södersten, M., & Lindestad, P. Å. (1990). Glottal closure and perceived breathiness during phonation in normally speaking subjects. *Journal of Speech, Language, and Hearing Research*, *33*(3), 601-611.

Steinecke, I., & Herzel, H. (1995). Bifurcations in an asymmetric vocal-fold model. *The Journal of the Acoustical Society of America*, *97*(3), 1874-1884.

Story, B. H. (2005). A parametric model of the vocal tract area function for vowel and consonant simulation. *The Journal of the Acoustical Society of America*, *117*(5), 3231-3254.

Story, B. H., & Titze, I. R. (1995). Voice simulation with a body-cover model of the vocal folds. *The Journal of the Acoustical Society of America*, *97*(2), 1249-1260.

Story, B. H., Titze, I. R., & Hoffman, E. A. (1996). Vocal tract area functions from magnetic resonance imaging. *The Journal of the Acoustical Society of America*, *100*(1), 537-554.

Švec, J. G., Horáček, J., Šram, F. and Veselý, J. (2000). Resonance properties of the vocal folds: In vivo laryngoscopic investigation of the externally excited laryngeal vibrations. The Journal of the Acoustical Society of America, 108(4), 1397-1407.

Tao, C., & Jiang, J. J. (2006). Anterior-posterior biphonation in a finite element model of vocal fold vibration. *The Journal of the Acoustical Society of America*, *120*(3), 1570-1577.

Titze, I. R. (1973). The human vocal cords: a mathematical model. *Phonetica*, *28*(3-4), 129-170.

Titze, I. R. (1974). The human vocal cords: A mathematical model. *Phonetica*, *29*(1-2), 1-21.

Titze, I. R. (1976). On the mechanics of vocal-fold vibration. *The Journal of the Acoustical Society of America*, *60*(6), 1366-1380.

Titze, I. R. (2000). Principles of voice production (second printing). *Iowa City, IA: National Center for Voice and Speech*, 229-233.

Titze, I. R. (2001). Acoustic interpretation of resonant voice. Journal of voice, 15(4), 519-528.

Titze, I. R., & Story, B. H. (2002). Rules for controlling low-dimensional vocal fold models with muscle activation. *The Journal of the Acoustical Society of America*, *112*(3), 1064-1076.

Titze, I. R., & Strong, W. J. (1975). Normal modes in vocal cord tissues. *The Journal of the Acoustical Society of America*, *57*(3), 736-744.

Titze, I. R., & Talkin, D. T. (1979). A theoretical study of the effects of various laryngeal configurations on the acoustics of phonation. *The Journal of the Acoustical Society of America*, *66*(1), 60-74.

Titze, I. R., Luschei, E. S., & Hirano, M. (1989). Role of the thyroarytenoid muscle in regulation of fundamental frequency. *Journal of Voice*, *3*(3), 213-224.

Titze, I.R. and Alipour, F., 2006. The myoelastic aerodynamic theory of phonation. National Center for Voice and Speech.

Tokuda, I. T., Horáček, J., Švec, J. G., & Herzel, H. (2007). Comparison of biomechanical modeling of register transitions and voice instabilities with excised larynx experiments. *The Journal of the Acoustical Society of America*, *122*(1), 519-531.

Tokuda, I. T., Zemke, M., Kob, M., & Herzel, H. (2010). Biomechanical modeling of register transitions and the role of vocal tract resonators. *The Journal of the Acoustical Society of America*, *127*(3), 1528-1536.

Tran, Q. T., Gerratt, B. R., Berke, G. S., & Kreiman, J. (1993). Measurement of Young's modulus in the in vivo human vocal folds. *Annals of Otology, Rhinology & Laryngology*, *102*(8), 584-591.

Triep, M., & Brücker, C. (2010). Three-dimensional nature of the glottal jet. *The Journal of the Acoustical Society of America*, *127*(3), 1537-1547.

Vampola, T., Horáček, J., & Klepáček, I. (2016). Computer simulation of mucosal waves on vibrating human vocal folds. *Biocybernetics and biomedical engineering*, *36*(3), 451-465.

Van den Berg, J. (1958). Myoelastic-aerodynamic theory of voice production. Journal of speech and hearing research, 1(3), 227-244.

Van Kan, J. J. I. M. (1986). A second-order accurate pressure-correction scheme for viscous incompressible flow. *SIAM journal on scientific and statistical computing*, *7*(3), 870-891.

Vilain, C. E., Pelorson, X., Fraysse, C., Deverge, M., Hirschberg, A., & Willems, J. (2004). Experimental validation of a quasi-steady theory for the flow through the glottis. *Journal of sound and vibration*, *276*(3-5), 475-490.

Wang, X., Jiang, W., Zheng, X., & Xue, Q. (2021). A computational study of the effects of vocal fold stiffness parameters on voice production. Journal of Voice, 35(2), 327-e1.

Williams, P., Warwick, R., Dyson, M., and Bannister, L. (1989). Gray's Anatomy, 17th ed, Churchill Livingstone, Edinburgh, London, Melbourne, New York.

Wong, D., Ito, M. R., Cox, N. B., & Titze, I. R. (1991). Observation of perturbations in a lumped-element model of the vocal folds with application to some pathological cases. *The Journal of the Acoustical Society of America*, *89*(1), 383-394.

Wu, L., & Zhang, Z. (2019). Voice production in a MRI-based subject-specific vocal fold model with parametrically controlled medial surface shape. *The Journal of the Acoustical Society of America*, *146*(6), 4190-4198.

Xi, J., Si, X. A., Dong, H., & Zhong, H. (2018). Effects of glottis motion on airflow and energy expenditure in a human upper airway model. European Journal of Mechanics-B/Fluids, 72, 23-37.

Xue, Q. (2011). *Computational modeling and analysis of fluid-structure interaction during phonation in healthy and diseased larynges*. The Johns Hopkins University.

Xue, Q., Mittal, R., Zheng, X., & Bielamowicz, S. (2010). A computational study of the effect of vocal-fold asymmetry on phonation. *The Journal of the Acoustical Society of America*, *128*(2), 818-827.

Xue, Q., Mittal, R., Zheng, X., & Bielamowicz, S. (2012). Computational modeling of phonatory dynamics in a tubular three-dimensional model of the human larynx. *The Journal of the Acoustical Society of America*, *132*(3), 1602-1613.

Xue, Q., Zheng, X., Mittal, R., & Bielamowicz, S. (2014). Subject-specific computational modeling of human phonation. *The Journal of the Acoustical Society of America*, *135*(3), 1445-1456.

Yang, A., Berry, D. A., Kaltenbacher, M., & Döllinger, M. (2012). Three-dimensional biomechanical properties of human vocal folds: Parameter optimization of a numerical model to match in vitro dynamics. *The Journal of the Acoustical Society of America*, *131*(2), 1378-1390.

Yang, A., Lohscheller, J., Berry, D. A., Becker, S., Eysholdt, U., Voigt, D., & Döllinger, M. (2010). Biomechanical modeling of the three-dimensional aspects of human vocal fold dynamics. *The Journal of the Acoustical Society of America*, *127*(2), 1014-1031.

Yang, A., Stingl, M., Berry, D. A., Lohscheller, J., Voigt, D., Eysholdt, U., & Döllinger, M. (2011). Computation of physiological human vocal fold parameters by mathematical optimization of a biomechanical model. *The Journal of the Acoustical Society of America*, *130*(2), 948-964.

Yin, J., & Zhang, Z. (2013). The influence of thyroarytenoid and cricothyroid muscle activation on vocal fold stiffness and eigenfrequencies. *The Journal of the Acoustical Society of America*, *133*(5), 2972-2983.

Yin, J., & Zhang, Z. (2014). Interaction between the thyroarytenoid and lateral cricoarytenoid muscles in the control of vocal fold adduction and eigenfrequencies. *Journal of biomechanical engineering*, *136*(11).

Yin, J., & Zhang, Z. (2016). Laryngeal muscular control of vocal fold posturing: Numerical modeling and experimental validation. *The Journal of the Acoustical Society of America*, *140*(3), EL280-EL284.

Zhang, K., Siegmund, T., & Chan, R. W. (2009). Modeling of the transient responses of the vocal fold lamina propria. *Journal of the mechanical behavior of biomedical materials*, *2*(1), 93-104.

Zhang, Z. (2009). Characteristics of phonation onset in a two-layer vocal fold model. *The Journal of the Acoustical Society of America*, *125*(2), 1091-1102.

Zhang, Z. (2010). Vibration in a self-oscillating vocal fold model with left-right asymmetry in body-layer stiffness. *The Journal of the Acoustical Society of America*, *128*(5), EL279-EL285.

Zhang, Z. (2015). Regulation of glottal closure and airflow in a three-dimensional phonation model: Implications for vocal intensity control. *The Journal of the Acoustical Society of America*, *137*(2), 898-910.

Zhang, Z. (2016). Cause-effect relationship between vocal fold physiology and voice production in a three-dimensional phonation model. *The Journal of the Acoustical Society of America*, *139*(4), 1493-1507.

Zhang, Z. (2017). Effect of vocal fold stiffness on voice production in a three-dimensional body-cover phonation model. *The Journal of the Acoustical Society of America*, *142*(4), 2311-2321.

Zhang, Z., & Hieu Luu, T. (2012). Asymmetric vibration in a two-layer vocal fold model with left-right stiffness asymmetry: Experiment and simulation. *The Journal of the Acoustical Society of America*, *132*(3), 1626-1635.

Zhang, Z., Mongeau, L., & Frankel, S. H. (2002). Experimental verification of the quasi-steady approximation for aerodynamic sound generation by pulsating jets in tubes. *The Journal of the Acoustical Society of America*, *112*(4), 1652-1663.

Zhao, W., Zhang, C., Frankel, S. H., & Mongeau, L. (2002). Computational aeroacoustics of phonation, Part I: Computational methods and sound generation mechanisms. *The Journal of the Acoustical Society of America*, *112*(5), 2134-2146.

Zheng, X. (2009). *Biomechanical modeling of glottal aerodynamics and vocal fold vibration during phonation* (Doctoral dissertation, The George Washington University).

Zheng, X., Bielamowicz, S., Luo, H., & Mittal, R. (2009). A computational study of the effect of false vocal folds on glottal flow and vocal fold vibration during phonation. *Annals of biomedical engineering*, *37*(3), 625-642.

Zheng, X., Mittal, R., & Bielamowicz, S. (2011b). A computational study of asymmetric glottal jet deflection during phonation. *The Journal of the Acoustical Society of America*, *129*(4), 2133-2143.

Zheng, X., Mittal, R., Xue, Q., & Bielamowicz, S. (2011a). Direct-numerical simulation of the glottal jet and vocal-fold dynamics in a three-dimensional laryngeal model. *The Journal of the Acoustical Society of America*, *130*(1), 404-415.

Zheng, X., Xue, Q., Mittal, R., & Beilamowicz, S. (2010). A coupled sharp-interface immersed boundary-finite-element method for flow-structure interaction with application to human phonation. *Journal of biomechanical engineering*, *132*(11).

BIOGRAPHY OF THE AUTHOR

Xiaojian Wang was born in Jiangsu, China in 1991. He attended the Shandong University in 2009 and graduated in 2013 with a Bachelor's degree in Civil Engineering. He then entered the Hydraulic Structure Engineering graduate program at the Hohai University in 2014. After receiving his Master's degree, he entered the Mechanical Engineering PhD program at the University of Maine in 2017. Xiaojian Wang is a candidate for the Doctor of Philosophy degree in Mechanical Engineering from the University of Maine in December 2022.