Theoretical and Numerical Analyses of Laryngeal Biomechanics: Towards Understanding Vocal Hyperfunction

by

Mohamed Ahmed Serry

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Doctor of Philosophy
in
Mechanical and Mechatronics Engineering

Waterloo, Ontario, Canada 2023

Examining Committee Membership

The following served on the Examining Committee for this thesis. The decision of the Examining Committee is by majority vote.

External Examiner:	Luc Mongeau Professor Dept. of Mechanical Engineering McGill University
Supervisor(s):	Sean D. Peterson Professor Dept. of Mechanical and Mechatronics Engineering University of Waterloo
Internal Members:	Stewart McLachlin Assistant Professor Dept. of Mechanical and Mechatronics Engineering University of Waterloo
	Naveen Chandrashekar Associate Professor Dept. of Mechanical and Mechatronics Engineering University of Waterloo
Internal-External Member:	Adil Al-Mayah Associate Professor Dept. of Civil and Environmental Engineering University of Waterloo

Author's Declaration

This thesis consists of material all of which I authored or co-authored: see Statement of Contributions included in the thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Statement of Contributions

This thesis consists, in part, of chapters adapted from research papers that were published, where I, Mohamed Ahmed Serry, and other coauthors were involved. In particular:

- The work presented in Chapter 3 is based on a research paper published in the Journal of Acoustical Society of America under the title "Physics of Phonation Offset: Towards Understanding Relative Fundamental Frequency Observations" [127], where Mohamed Ahmed Serry is the first author, and Cara E. Stepp and Sean D. Peterson are coauthors. M.A.S was the primary investigator in this study. M.A.S. and S.D.P. conceived of the study. M.A.S. developed the model. M.A.S. generated preliminary results and figures. All authors analyzed and interpreted results. M.A.S. wrote the first draft of the article. All authors reviewed and approved the manuscript.

- The work presented in Chapter 4 is based on a paper that has been published in the Biomechanics and Modelling in Mechanobiology journal under the title " Exploring the Mechanics of Fundamental Frequency Variation During Phonation Onset" [128], where Mohamed Ahmed Serry is the first author, and Cara E. Stepp and Sean D. Peterson are coauthors. M.A.S was the primary investigator in this study. M.A.S. and S.D.P. conceived of the study. M.A.S. developed the model. M.A.S. generated preliminary results and figures. M.A.S. wrote the first draft of the article. All authors reviewed and approved the manuscript.

- The work presented in Chapter 5 is adapted from a published research paper in the Biomechanics and Modelling in Mechanobiology journal under the title "Modeling the influence of the extrinsic musculature on phonation" [126], where Mohamed Ahmed Serry is the first author, and Gabriel A. Alzamendi, Matías Zañartu, and Sean D. Peterson are coauthors. M.A.S was the primary investigator in this study. M.A.S. and S.D.P. conceived of the study. M.A.S. and G.A. developed the model. M.A.S. generated preliminary results and figures. All authors analyzed and interpreted results. M.A.S. wrote the first draft of the article. All authors reviewed and approved the manuscript.

Abstract

Speech is a cornerstone in human communication and an irreplaceable tool for several academic, legal, and artistic careers. Producing intelligible speech is an extremely complex process, involving coupling between the air flow driven by the pressure built up in the lungs, the vibrating viscoelastic tissues in the larynx, namely the vocal folds, and the subglottal and supraglottal (vocal) tracts, including the nasal and oral passages. Therefore, the occurrence of a pathology in one of the organs responsible for voice production may result in deteriorated speech production and, consequently, a negative impact on the daily life and/or professional career of the individual. A specific class of voice disorders that is common among adults is vocal hyperfunction, which is associated with the misuse of vocal organs, resulting in inefficient voice production and, in some cases, vocal trauma. Researchers over the years have conducted clinical and numerical analyses of vocal hyperfunction and have developed assessment tools and therapeutic procedures for vocal hyperfunction; however, a comprehensive understanding of the underlying mechanisms of vocal hyperfunction remains unachieved. Fortunately, easily collected clinical measurements shed some light on potential mechanisms underlying vocal hyperfunction, and numerical and theoretical modelling campaigns of laryngeal biomechanics have shown some success in partially elucidating the biomechanics of voice production. Therefore, a potential route to pursue for a better understanding of the mechanics of vocal hyperfunction lays behind numerical and theoretical analyses guided by available clinical and experimental data.

The aim of this thesis is to explore and elucidate, through four research projects, some of the underlying mechanisms associated with voice production in general and vocal hyperfunction in particular, where we resort to 1) data collected using some promising assessment tools and standard clinical measurements, and 2) models of voice production and larynx biomechanics, where theoretical and numerical analyses are conducted guided by the aforementioned clinical measurements.

The first project analyses theoretically and numerically the underlying laryngeal factors altering fundamental frequency, for both healthy speakers and speakers with phonotraumatic vocal hyperfunction, during phonation offset, where clinical data of relative fundamental frequency are resorted to in modeling and analysis. We show that the clinically observed drop in fundamental frequency during phonation offset is potentially due to the decline in vocal fold collision forces, which is induced by increasing the glottal gap. We also show how the fundamental frequency drop rate can be modulated by the activation of certain laryngeal muscles, which we speculate to underlie the differences between healthy

and hyperfunctional speakers. Besides, we illustrate how certain manifestations of vocal hyperfunction can also affect the drop rate during phonation offset.

The second project extends the first one, where phonation onset is explored with similar numerical and theoretical approaches. We illustrate that, when all laryngeal and aerodynamic parameters are fixed in time, fundamental frequency tends to rise due to the increased vocal fold collision levels, and that matches with the clinical observations of the onset of initial or isolated vowels and, in some cases, vowels preceded by voiced consonants. On the other hand, we show through numerical simulations that the decline in fundamental frequency in the case of onset of vowels preceded by voiceless consonants requires involvement of laryngeal muscles, which we speculate to manifest the differences between healthy speakers and patients with vocal hyperfunction.

In the third project, we attempt to elucidate the influence of extrinsic laryngeal muscles on posturing mechanics and phonation, and link findings with clinical observations collected from patients with vocal hyperfunction. We show how the vocal fold tension and phonation fundamental frequency vary with varying the magnitude, direction, and location of the net pulling force exerted by the extrinsic laryngeal muscles. Using the previous analysis in combination with clinical data, we pinpoint potential roles of specific extrinsic muscles in modulating fundamental frequency and we suggest some potential roles for extrinsic laryngeal muscles in hyperfunctional phonation.

Finally, in the fourth project, we study the mechanics underlying curved and incomplete glottal closure configurations that are observed in some patients with vocal hyperfunction, where we develop and analyse a composite beam model for the vocal folds and we integrate it with a posturing model to enable exploring the effects of certain laryngeal maneuvers. The model predictive capability is adequate, matching clinical observations and simulations produced by high-fidelity models, yet providing useful insights into the underlying mechanism of curved glottal configurations due to its relative simplicity. Our analyses, based on the proposed model, show that the vocal fold layered structure and its interaction with the mechanical loading, resulting during laryngeal maneuvers, induce bending moments that result in different curved (convex and concave) vocal fold shapes that are associated with incomplete glottal closure patterns. We suggest, based on the conducted analyses, some potential laryngeal mechanisms that may be at play in patients with vocal hyperfunction.

Acknowledgements

I would like to thank all the patient people who supported me intellectually, socially, and emotionally throughout my PhD journey, including my mother and best friend, Fayzah Talaat, my caring father, Ahmed Moursi, my loving brothers Youssef and Yasser, my advisor and mentor, Dr. Sean Peterson, and my friends J. D., W. H., A. A., and F. H.

Table of Contents

List of Figures

List of Tables

List of Abbreviations

1-D one-dimensional 36, 51

1-DOF one degree-of-freedom 31, 32, 51

2-D one-dimensional 51

2-DOF two degrees-of-freedom 32

3-D three-dimensional 51

BCM body-cover model 32, 33, 36, 39, 40, 42, 45, 50, 51, 53, 83, 84, 90, 94, 95

CT cricothyroid 14, 16, 21, 22, 28, 39, 42, 46, 48, 49, 52, 61, 63, 73, 74, 78, 79, 83, 90, 92, 94–96, 98, 101, 102, 135, 136, 144, 146

DG digastric 18, 20

EMG electromyographic 20, 21, 24, 120

GH geniohyoid 18, 20, 118, 120

IA interarytenoid 14, 16, 46, 49, 94, 135–138, 140, 141, 143

LCA lateral cricoarytenoid 14, 16, 39, 42, 46, 49, 61, 63, 65, 66, 74, 77, 78, 83, 92, 94, 95, 135–138, 140, 141, 143

MH mylohyoid 18, 20

MTD muscle tension dysphonia 22–24, 29, 120, 121, 141–143

NPVH nonphonotraumatic vocal hyperfunction 22, 23, 29

OH omohyoid 18

PCA posterior cricoarytenoid 14, 16, 31, 39, 46, 49, 61, 63, 74, 78, 83, 94, 135, 136, 138, 140, 141, 143

PVH phonotraumatic vocal hyperfunction 22, 23, 52, 74, 77, 78, 144

RFF relative fundamental frequency 24, 25, 28, 52, 56, 58–60, 63, 66, 70, 72–74, 77–79, 144, 146

SH sternohyoid 18, 21

ST sternothyroid 18, 21, 109, 118

StyH stylohyoid 18, 20, 118, 120

TA thyroarytenoid 14, 16, 31, 39–42, 46, 49, 61, 79, 83, 90, 92, 94, 95, 98, 101, 102, 123, 125, 126, 133–138, 140, 141, 143, 144, 146

TBCM triangular body-cover model 45

TH thyrohyoid 18, 21, 23, 109, 110, 118, 120

VF vocal fold 9, 11–14, 21–24, 28, 29, 31, 32, 34, 35, 39–42, 45, 46, 48–53, 56, 60, 61, 63, 65, 66, 70, 73, 74, 77–81, 83–88, 90, 92, 94–96, 98, 103–110, 112, 114–116, 118, 120–123, 125–128, 130–136, 138, 140–146, 169, 171, 176

VH vocal hyperfunction 22–24, 52, 79, 104, 122, 146

WRA wave reflection analog 36, 39, 84

Chapter 1

Introduction

When it comes to the human population, speech remains the primary communication method, even with the rapid rise and expansion of technologies that allow communication through other means such as messaging. Speech is vital in the daily life of the general population and is central for individuals pursuing educational (e.g., lecturing and teaching), artistic (e.g., singing, and acting), and legal (e.g., lawyering) careers. Therefore, it is important to keep the organs responsible for voice production (also known as phonation) healthy and, in the unfortunate case of pathology, to diagnose and treat voice disorders successfully. Successful diagnosis and effective treatment of voice disorders are crucial for individuals and society as a whole, as it has been found that nearly 30% of adults experience a voice disorder in their lifetime, with 7.2% of the cases resulting in the individual's career being negatively impacted by the occurrence of the pathology [124]. Motivated by that, researchers have been striving for years to develop assessment/classification tools (see, e.g., [136, 92]) and treatments (see, e.g., [123, 169, 149]) for a wide variety of voice disorders.

From a physics point of view, voice production is a complex fluid-structure-acoustics interaction phenomenon that involves the lungs, the vocal folds, and the subglottal and supraglottal (vocal) tracts, where the pressure built up in the lungs drives air flow that passes between the vocal folds, inducing them to vibrate and collide with each other, which, in turn, generates acoustic waves that are transmitted and modified through the tracts, thus producing the unique voice characteristics of each individual. The importance of speech for the general population and the complexity of phonation have motivated researchers over the years to study the mechanisms associated with normal and pathological voice production. Such efforts include experimental investigations of vocal mechanisms of living humans (see, e.g., [79, 130, 173]), excised human/animal larynges (see, e.g., [90, 66, 106]), and physical models or replicas (see, e.g., [163, 89]). Besides the mentioned experimental works, there

have been several theoretical investigations (see, e.g., [155, 83, 55]) and numerical analyses (see, e.g., [139, 194]) of the mechanisms of voice production.

A prevalent class of voice disorders that is of interest is vocal hyperfunction, which is characterized by the misuse/overuse of vocal mechanisms during phonation [47, 48]. Untreated vocal hyperfunction may, amongst other presentations, lead to the formation of vocal lesions (e.g., polyps and nodules) [48]. It has been found that hyperfunctional disorders are among the most common voice disorders occurring in adults [11]. Therefore, there have been several efforts focusing on diagnosing/assessing [136, 171] and treating [169, 123, 149] vocal hyperfunctional disorders. However, there is generally a lack of comprehensive understanding of the mechanisms associated with vocal hyperfunction. This is due to the broad range of manifestations of vocal hyperfunction, where vocal mechanisms are suboptimally used, e.g., to compensate for some underlying issue in the vocal organs, which results in the development of a vocal pathology over a considerably long period of time. A better perception of vocal hyperfunction is therefore important as it will consequently enable developing better diagnostic/assessment tools and treatments.

Despite the lack of comprehensive understanding, there has been notable progress in elucidating the characteristics of vocal hyperfunction. For example, it has been observed in speakers with vocal hyperfunctional disorders that the frequency characteristics of voice during transient periods of phonation are different from those observed in healthy speakers [136]. Moreover, it has been found that some statistical features of hyperfunctional speech, estimated from week/day long measurements of neck acceleration, deviate from those produced by healthy speakers [171, 170]. Besides that, standard clinical inspections have revealed distinctive glottal characteristics in patients with vocal hyperfunction [99, 105]. In some cases, such observations resulted in developing robust assessment tools capable of distinguishing between healthy and hyperfunctional speakers [136, 138] and also quantifying the level of vocal misuse/overuse in the case of patients with vocal hyperfunction [93]. Voice researchers attribute the observed unique characteristics associated with hyperfunction to the abuse and misuse of vocal mechanisms (likely to compensate for an underlying issue in the vocal system); however, the driving mechanisms of vocal hyperfunction are still not fully understood. Further investigation of these characteristics, which is a major part of this thesis, can then potentially reveal some of the underlying mechanisms of vocal hyperfunction.

Typical and hyperfunctional voice production involve activation of several respiratory, laryngeal, and articulatory muscles, where neural motor controllers utilize feedback auditory and somatosensory information and send neural signals to speech-related muscles to correct the acoustic output in order to reach a desired target value [177]. Hence, the vocal system integrated with neural control from the brain can be thought of as a control system,

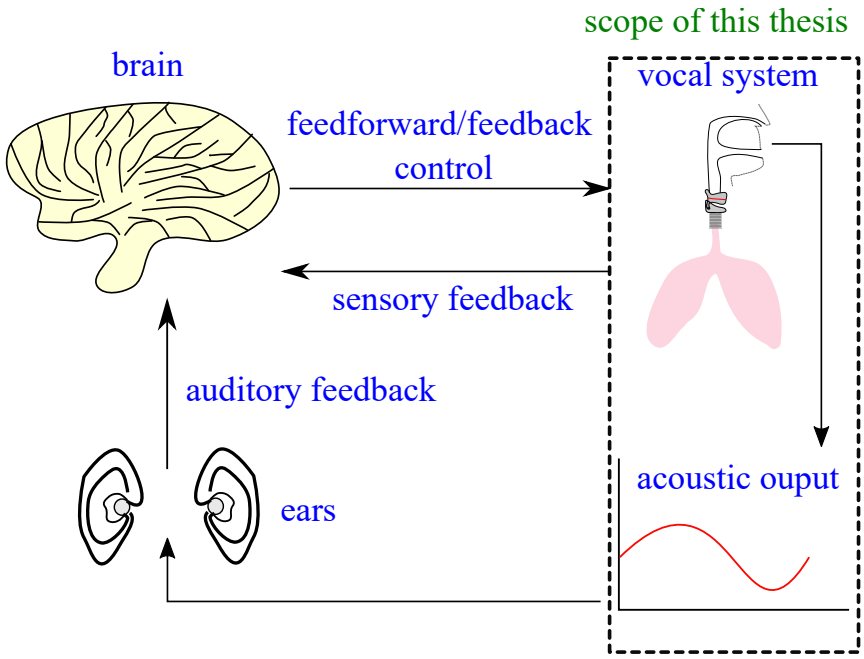


Figure 1.1: Neural control of the vocal system and the scope of the current thesis.

where the brain sends a feedforward neural signal (to speech-related muscles) that would 'ideally' drive the vocal system to produce the target acoustic output, and incorporates feedback control that corrects, based on sensory and auditory measurements, the feedforward signal when the actual acoustic output deviates from its target value (see Figure 1.1). It has been suggested that several voice disorders, including vocal hyperfunction, involve abnormal neural control of the vocal system [21, 177, 137].

In theory, conducting an extensive experimental/clinical investigation of patients with vocal hyperfunction may be preferred as that will provide hands-on exposure to and clear comprehension of the underlying mechanisms. However, and realistically speaking, experimental analysis in voice research is generally very challenging due to several reasons. For example, clinical measurements typically suffer from significant person-to-person variability in geometry and material properties of the vocal organs, which depends upon several factors, including sex, age, and pathological nature, if existent (see, e.g., [157, 28]). For example, the unstretched length of the vocal folds in males and females change over time with a length of about 3 millimeters (mm) at birth, in both males and females, which then increases to about 16 mm in the case of adult males and to about 10 mm in the case of

adult females [157]. Even when one speaker is considered in experimental analysis, variations in hydration levels of the vocal folds, fatigue, and/or other uncontrolled exogenous factors can lead to significant variability in measurements. Hence, controlling all factors associated with voice production in experimental setups is practically unachievable. Besides that, experimental measurements of some important speech-related variables, such as the lungs pressure, the glottal flow rate, and the vocal fold contact forces are typically difficult to measure directly. This often compels researchers to collect indirect measurements and then utilize them (e.g., by means of inverse filtering [120]) to estimate the desired variables, which may worsen the variability issue due to added uncertainty from the adopted conversion methods. Therefore, results from experimental measurements of phonation tend to be statistical in nature with significant standard deviations, due to accumulating sources of variability, which limit the applicability of such results. Furthermore, the ethical constraints on human and animal experimentation and the limited financial resources minimize the possibility of conducting extensive experimental research that provides a deep insight into the mechanisms of voice production.

These difficulties and limitations have driven voice researchers over the years to resort to the relatively less restricted and cheaper numerical and theoretical studies, where parameters of voice production can be controlled and analyzed thoroughly, with the trade off of reduced accuracy due to embedding simplifying approximations and assumptions in order to enable feasible analysis. Despite their limitations, numerical and theoretical analyses have shown the ability to successfully reveal some of the underlying mechanisms of voice production (see, e.g., [155, 83, 84, 10, 62]). In some cases, numerical studies have even shown success in elucidating some of the mechanisms associated with pathological voice production [186, 23]. It should be noted here that theoretical and numerical routes are not separate from the experimental route and, in fact, these routes complement each other as available clinical data are used in different steps of modeling (e.g., motivating model development, and validating associated assumptions) and, in other cases, experimental studies are conducted to verify results from numerical and theoretical analyses.

The goal of this thesis is to adopt a modelling paradigm, where theoretical and numerical tools are utilized in an attempt to understand some of the underlying mechanisms of laryngeal biomechanics in general, and the several manifestations of vocal hyperfunction in particular. We focus herein on analysing the mechanics (e.g., contact forces, aerodynamics, stresses, strain, etc.) and laryngeal-muscular executions associated with voice-production and laryngeal posturing, without incorporating neural control, and explore the variations in the kinematic and acoustic outputs due to alterations in such factors (see Figure 1.1). We aim at addressing, at least in part, the following vocal-hyperfunction-related issues, which are not fully understood:

(1) the context-specific temporal variations of fundamental frequency during transient phonation and the discrepancy between healthy speakers and speakers with vocal hyperfunction,

(2) the relevance of neck muscles in normal and hyperfunctional phonation, and

(3) incomplete and curved glottal closure patterns and their underlying mechanisms.

Our modelling and analyses in this thesis will be often guided by clinical and experimental observations and previous numerical and theoretical studies.

1.1 Objectives

The objectives of this thesis are:

(1) exploring by means of theoretical and numerical analyses the effects of laryngeal and biomechanical factors on transient phonation, especially phonation offset (Chapter 3) and phonation onset (Chapter 4), and how such factors are at play in the case of normal and hyperfunctional voice production, where this exploration is guided by clinical observations from the promising assessment tool of relative fundamental frequency [136];

(2) developing a phonation model that incorporates the net effects of extrinsic laryngeal muscles, where numerical and theoretical investigations are conducted on this model and implications on typical and hyperfunctional phonation are concluded based upon available clinical data; and

(3) constructing a simple yet insightful mechanical vocal fold model that is capable of predicting curved and incomplete glottal closure patterns that are observed clinically in some patients with vocal hyperfunction [99, 105] and elucidating the physical mechanisms underlying these curved shapes.

The remainder of the document is organized as follows: Chapter 2 contains necessary background which covers: (a) the anatomy of the vocal organs, and the basic mechanisms of voice production; (b) vocal hyperfunctional disorders, their categories and biomechanics, and relevant clinical observations; and (c) an elaborate discussion of laryngeal modelling, with a focus on the body-cover model [142]. Chapters 3 and 4 discuss the work conducted

to elucidate the laryngeal factors associated with healthy and hyperfunctional speakers during phonation offset and onset, respectively, with the help of clinical data using the assessment tool of relative fundamental frequency and that covers the first thesis objective. Chapter 5 introduces a phonation model that incorporates the net effect of extrinsic laryngeal musculature with accompanied analysis and implications regarding hyperfunctional phonation and that fulfills the second objective. In chapter 6, we introduce a vocal fold beam model that is capable of predicting some of the clinically observed curved and incomplete glottal closure patterns, and conduct extensive analysis on the mechanisms underlying this model, where we link these mechanisms to hyperfunctional patients, and that addresses the third objective. Finally, Chapter 7 concludes the thesis with a summary and suggested future work.

Chapter 2

Background

In this chapter, we introduce the preliminary background required to understand the framework of this thesis. In particular, we discuss the anatomy of the vocal organs, and the fundamentals of voice production. Moreover, we present the concept of vocal hyperfunctional disorders, their categories and biomechanics, a related promising diagnostic tool, namely relative fundamental frequency, where we present and discuss some relevant clinical observations associated with transient phonation, and demonstrate some laryngeal patterns (in particular, incomplete glottal closure) that are linked to vocal hyperfunction. Finally, we discuss various aspects of laryngeal modelling, where we provide a thorough description of a particular phonation model, namely the body-cover model [142], which we utilize extensively in this thesis, and shed some light on different aspects of laryngeal modelling, including modelling of acoustics and glottal posture.

2.1 Physiological Planes and Orientations

We start this chapter with a brief introduction of the anatomical conventions used in describing the relative location of tissues, cartilages, organs, and so on (see Figure 2.1). In general, three primary planes are used in anatomical description: (a) a sagittal plane, also referred to as median or medial plane, that divides the human/animal body into right and left halves, (b) a coronal plane that divides the body into front and back halves, (c) and a transverse plane that divides the body into upper and lower halves. Based on the aforementioned planes, the following primary directions are used in anatomical description of relative locations: (a) superior and inferior directions pointing upward and downward, respectively (with respect to the transverse plane), (b) anterior and posterior

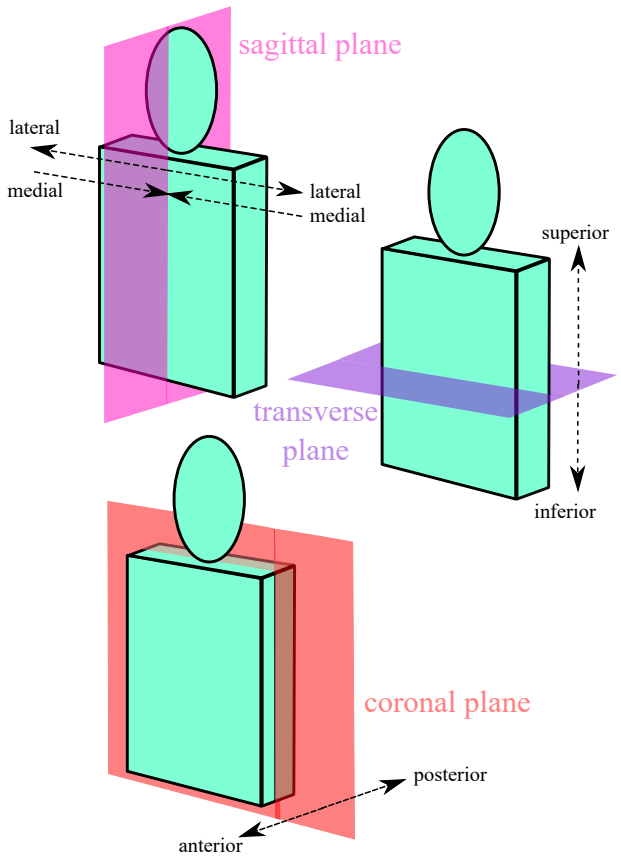


Figure 2.1: Anatomical planes and directions.

directions pointing forward and backward, respectively (with respect to the coronal plane), (c) and lateral and medial directions pointing away from and towards the sagittal plane, respectively. In the context of voice research, we say, for example, that the trachea is superior to the lungs as it is located above the lungs. We also say that the vocal folds are attached anteriorly at the thyroid cartilage meaning that the frontal ends of the vocal folds are attached to the mentioned cartilage. Moreover, we say that the vibrations of the vocal folds are primarily medial/lateral, meaning that the associated displacements are in a direction perpendicular to the medial plane.

2.2 Biomechanics of Voice Production

Herein, the fundamentals of voice production and the primary tissues/organs involved in this process are introduced. Detailed discussion of the laryngeal musculature will be presented in the following sections. The general anatomical structure of the vocal organs can be seen in Figure 2.2, which consists of: (a) the lungs, which are the generators of the air pressure required for voice production, (b) the larynx (i.e., voice box) and vocal folds which are the primary sound producers, (c) and subglottal and supraglottal (vocal) tracts, which work as filtering and transmission components of the sound generated in the larynx; hence, the tracts provide the unique voice characteristics of each individual.

2.2.1 Larynx and vocal folds

The laryngeal structure consists of several cartilages including: (a) the cricoid cartilage, a ring-like structure that sits on top of the trachea and forms the base of the larynx, (b) the thyroid cartilage (an anterior cover of the larynx), which is located slightly anteriorly and superiorly relative to the cricoid cartilage and is attached to it posteriorly at the cricothyroid joint, (c) and two arytenoid cartilages resting posteriorly on the cricoid cartilage. Besides the aforementioned cartilages, the larynx also consists of auxiliary cartilages with other physiological roles (e.g., the epiglottis cartilage prevents food particles from entering the respiratory passage during swallowing).

The most important tissues involved in the process of voice production are the vocal fold (VF) tissues, which are two band-like structures that are attached anteriorly to a common point on the thyroid cartilage and posteriorly at the arytenoid cartilages. The opening between the VFs is called the glottis. Figure 2.3 provides a superior schematic view of the mentioned three primary cartilages and the VFs. We refer the readers to, e.g., [147] for an overview of the larynx anatomy.

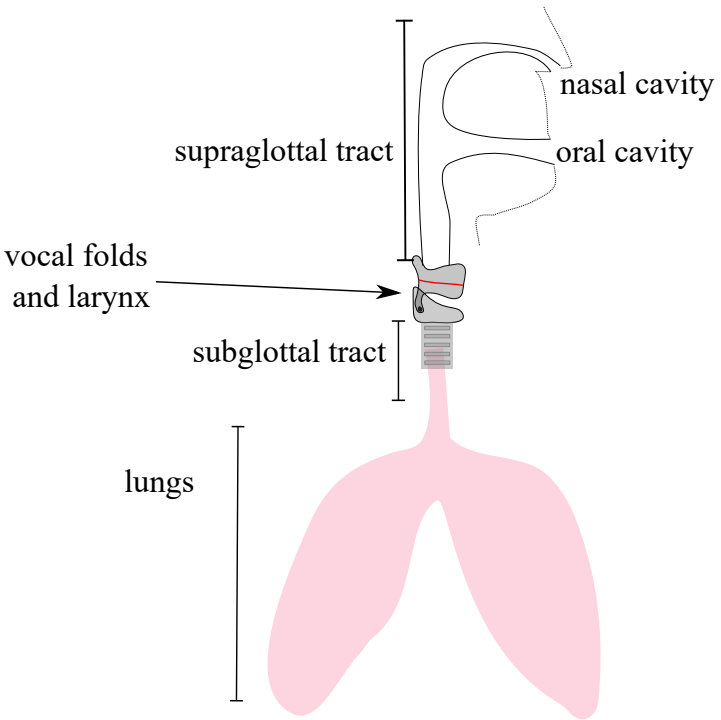


Figure 2.2: General anatomical structure of the vocal organs.

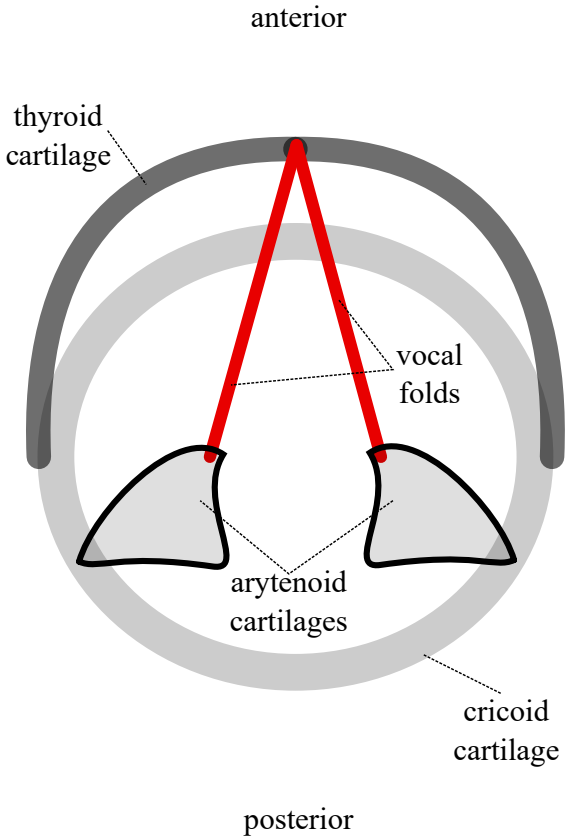


Figure 2.3: Superior schematic view of the thyroid, cricoid, and arytenoid cartilages, and the vocal folds.

Each VF consists of several layers [191], including (a) an inner layer consisting of the muscle fibers of the thyroarytenoid muscle, with the muscle fibers aligning primarily along the anterior-posterior direction, (b) a soft tissue layer of the lamina propria, (c) and an outer-most epithelium layer. The lamina propria can be further divided into three layers consisting primarily of fibrous proteins, namely elastin and collagen. These layers are: (a) a superficial layer, (b) an intermediate layer, (c) and a deep layer. The layers of VFs and their compositions give the VFs their viscoelastic characteristics which are very essential in the process of voice production [13].

When modelling the VFs, the complex layered structure is typically abstracted using fewer representing layers. For example, in some models, the VF is often assumed to consist of two layers, a cover layer and a body layer [49, 142], where the cover layer corresponds to

the epithelium and the superfacial and intermediate layers of the lamina propria, and the body corresponds to the deep lamina propria layer and the thyroarytenoid muscle . A more refined VF abstraction divides each VF into three layers [50, 164]: the mucosa (epithelium and superfacial lamina propria layer), the ligament (intermediate and deep lamina propria layers), and the thyroarytenoid muscle.

2.2.2 Essential physics of phonation

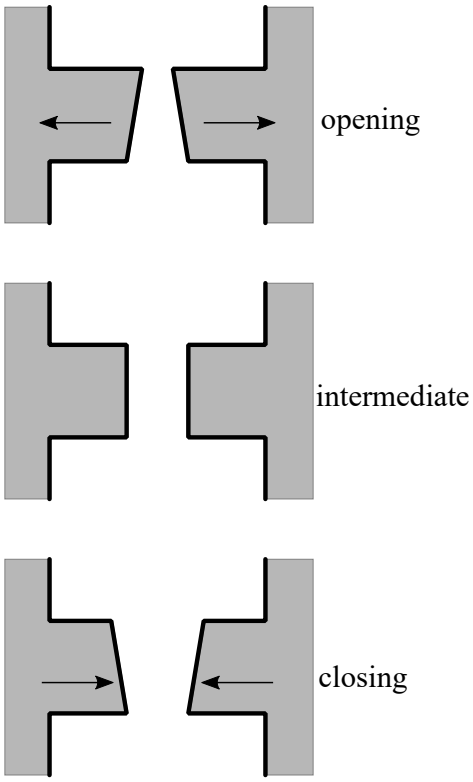


Figure 2.4: Coronal schematic views of the the different phases of phonation cycle during self-sustained oscillations.

In this section, we discuss briefly the fundamentals of voice production. This process can be thought of as a fluid-structure interaction problem, where the air flow induced by

the lungs provides the energy to the VFs sufficient to initiate and sustain their vibrations. On the other hand, the vibrations of the VFs alter the aerodynamics of the glottal flow. We summarize the process below.

At rest, when laryngeal muscles are inactive, the vocal folds are apart, or abducted, making the glottis shape approximately triangular. Prior to phonation, the folds are brought together (i.e., adducted) to almost full closure, which obstructs the glottis. Pressure is then built up in the lungs, by activating the diaphragm, forcing the air from the lungs through the obstructed glottis. The air driven by the lungs pushes the adducted VFs apart while the restoring elastic forces in the VFs work to bring them back to their resting adducted positions. Once the lungs pressure reaches a critical value, known as phonation threshold pressure [155], aerodynamic energy is transferred to the VFs, overcoming their tissue damping, which consequently leads the VFs to start vibrating and colliding (when vibration amplitude is sufficiently large) causing a mucosal wave that propagates superiorly along the medial surfaces of VFs and creating a time-varying glottal geometry switching between closed and open configurations. During sustained oscillations, the VFs transition from a convergent configuration (with respect to the superior direction) to a parallel configuration during the opening phases, then shift gradually to a divergent shape during the intermediate phases, and finally return back to the convergent configuration during the closing phases [96] (see Figure 2.4).

The obstructed glottal flow associated with the VF sustained vibrations comprises a source that induced pressure waves propagating through the supraglottal and subglottal tracts, which amplify and filter the acoustic waves, causing a modification to the frequency content of the acoustic wave. The amplification and filtering processes provide the unique sound characteristics of each individual. This acoustic interaction with the tracts also alters the dynamics of VF oscillations [159]. The sound heard during phonation comprises the radiated pressure waves from the oronasal (oral and nasal) cavities. The vowel, e.g., /a/ or consonant, e.g., /v/, sounds produced during phonation are determined by the geometry of the supraglottal tract, including the oral and nasal passages, which is determined by the positioning of the vocal articulators such as the lips, tongue, and teeth [143, 108]. Depending on the presence or absence of VF oscillations, consonant sounds can be categorized into (a) voiced consonant sounds, which incorporate VF vibrations (e.g., /b/ and /v/), and unvoiced consonant sounds, wherein VFs are abducted and not vibrating and sound is produced by the interaction between the lungs-induced air flow and supraglottal tract (e.g., /f/, /p/, and /s/), see [108].

In the case of healthy voice production, VF vibrations are quite periodic during sustained phonation, whereas in pathological conditions, vibrations of VFs can be aperiodic or even chaotic. These distinctive characteristics of voice signals have motivated several

numerical and experimental works in the literature to model and classify these pathological oscillatory behaviors [62, 189, 61].

An important quantity that is central in voice research is phonation fundamental frequency. Phonation fundamental frequency, which is often denoted f_o, can be deduced from acoustic, glottal area, or VF vibration signals. This quantity will be critical in our investigations and discussions thereafter. In the case of a periodic signal, fundamental frequency is simply the inverse of the signal period. On the other hand, this definition is not well-posed when we consider initial or final stages of phonation, wherein the acoustic (or glottal area) signal amplitude and oscillation rate is time-varying. For such transient cases, fundamental frequency is estimated for each oscillation cycle using signal peaks (detailed definitions are given in Chapters 3 and 4), and such an approach is typically adopted in clinical studies (see, e.g., [136]).

2.3 Intrinsic Laryngeal Muscles

The posture of the laryngeal cartilages, and subsequently the glottal geometry and VF tension levels, are primarily controlled by the intrinsic laryngeal muscles (see Figure 2.5). The naming of these muscles, and all muscles in general, reflects their origin and insertion points. For example, the cricothyroid muscle originates at the cricoid cartilage and its attachment point is at the thyroid cartilage. The intrinsic muscles consist of the:

(a) cricothyroid (CT),

(b) thyroarytenoid (TA),

(c) lateral cricoarytenoid (LCA),

(d) posterior cricoarytenoid (PCA), and

(e) interarytenoid (IA) muscles.

The CT muscle is responsible for stretching the VFs, the TA, LCA, and IA muscles are responsible for adducting the VFs, with the TA muscle also shortening the VFs, and the PCA muscle is responsible for abducting the VFs (see, e.g., [4]). Figure 2.6 illustrates a superior view of the lines of action of the adducorty and abductory muscles and the resulting adduction and abduction motions, and Figure 2.7 provides a side view of the mechanism of which the CT muscle stretches the VFs.

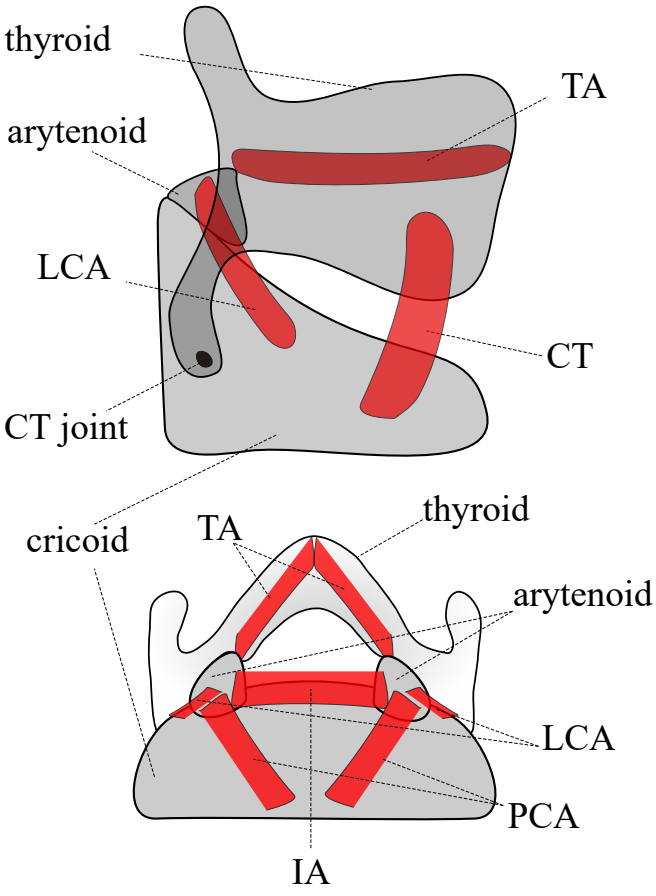


Figure 2.5: Schematic side (top) and posterior (bottom) views of the larynx with intrinsic laryngeal muscles: cricothyroid (CT), thyroarytenoid (TA), lateral cricoarytenoid (LCA), posterior cricoarytenoid (PCA), and interarytenoid (IA) muscles.

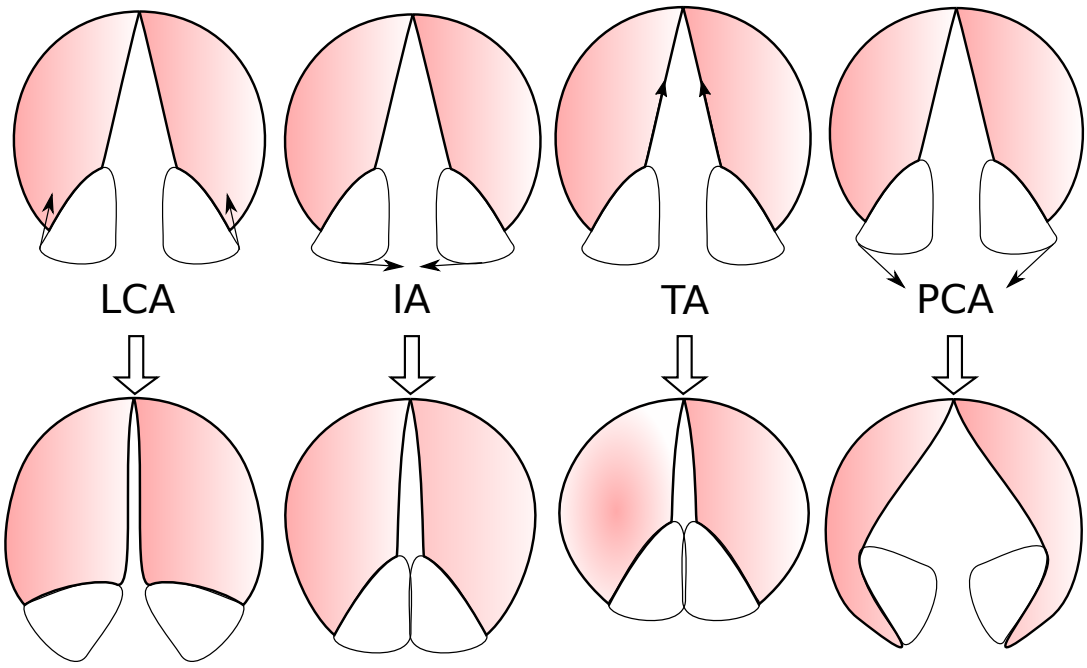


Figure 2.6: Schematic superior view illustrating the lines of action of some laryngeal muscles and associated adduction and abduction processes: activation of lateral cricoarytenoid (first column), interarytenoid (second column), thyroarytenoid (third column), and posterior cricoarytenoid (fourth column) muscles.

Intrinsic laryngeal muscles have been investigated extensively in many clinical and numerical studies in order to elucidate their roles in phonation. In general, it has been found that the CT muscle is the primary muscle modulating phonation fundamental frequency, where increasing its activation typically leads to increasing fundamental frequency [80, 17]. On the other hand, the role of the TA muscle in modulating fundamental frequency is more complex as it can either increase or decrease fundamental frequency, with controversy in the literature regarding its exact role (see the discussion in [101]). Furthermore, it has been found that activation of the LCA and IA muscles is positively correlated with fundamental frequency [19], whereas PCA activation exhibits negative correlation with fundamental frequency [18].

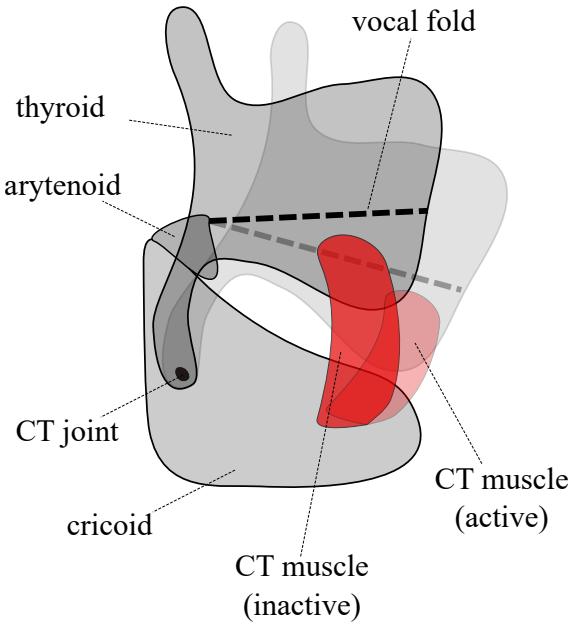


Figure 2.7: Schematic lateral view of the mechanics of vocal fold lengthening through the activation of the cricothyroid muscle.

2.4 Extrinsic Laryngeal Muscles

The extrinsic muscles, which surround the larynx, are involved in modulating the glottal posture, in addition to their roles in other physiological functions, such as head-neck movement [100] and swallowing [43]. Extrinsic muscles can be categorized into two groups, depending on their location relative to the hyoid bone (see [7] and Figures 2.8 and 2.9), a horseshoe-shaped bone that is located superior to the larynx.

Infrahyoid muscles (see [35]) refer to the group of neck muscles that are positioned below the hyoid bone (see Figure 2.8). These muscles include:

(a) the sternohyoid (SH), which connects the hyoid bone to the sternum,

(b) the sternothyroid (ST), which connects the thyroid cartilage to the sternum,

(c) the thyrohyoid (TH), which connects the thyroid cartilage to the hyoid bone, and

(d) the omohyoid (OH), which connects the scapula to the hyoid bone [148].

The thyrohyoid membrane, which consists of connective tissues, also connects the hyoid bone to the thyroid cartilage.

Suprahyoid muscles refer to the group of muscles that are located above the hyoid bone. This group includes:

(a) the anterior and posterior digastric (DG) muscles,

(b) the geniohyoid (GH),

(c) the mylohyoid (MH), and

(d) the stylohyoid (StyH) muscles (see Figure 2.9).

The GH and StyH muscles originate from the mandible and cranial base, respectively, and are inserted into the hyoid bone. The MH consists of anterior and posterior portions, which connect the mandible to the hyoid bone. The anterior and posterior branches of DG muscle arise from the mandible and cranial base, respectively, and share a common tendon with an insertion point in the hyoid bone [115]. Pearson et al. [115] investigated the influence of suprahyoid muscles on the movement of the hyoid bone in the superior-inferior and anterior-posterior directions using a human cadaver model. They concluded

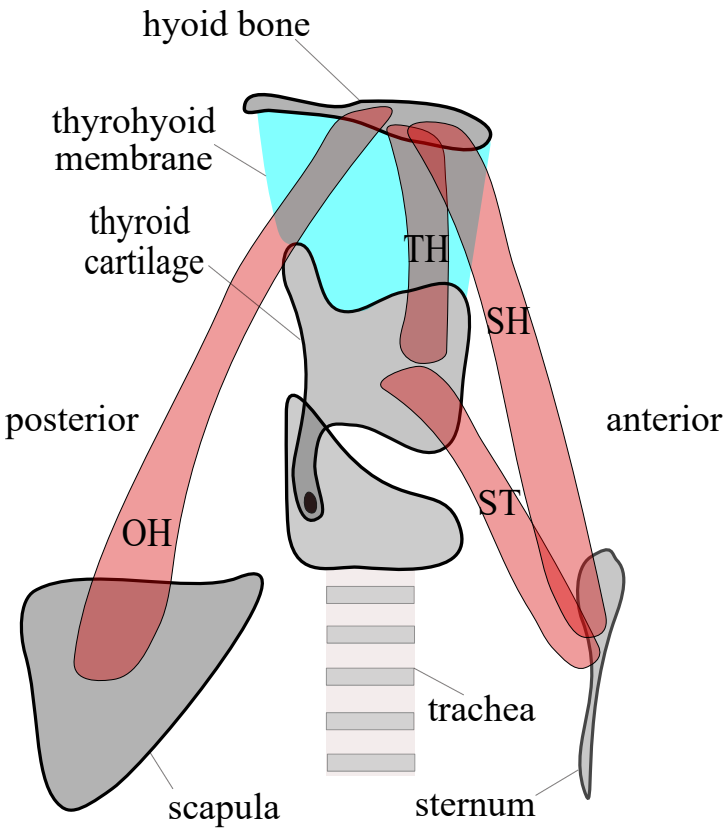


Figure 2.8: Lateral schematic view of the infrahyoid muscles: thyrohoid (TH), sternohoyid (SH), sternothyroid (ST), and omohyoid (OH) muscles: partially inspired by Figure 5 in [53].

that the MH, DG, StyH, and GH muscles have the likelihood to move the hyoid bone in the superior direction, with the MH muscle having the highest potential. Moreover, the MH, DG, and GH muscles are likely to move the hyoid bone in the anterior direction, with the GH muscle having the highest potential, whereas the StyH muscle has the tendency to move the hyoid bone in the posterior direction.

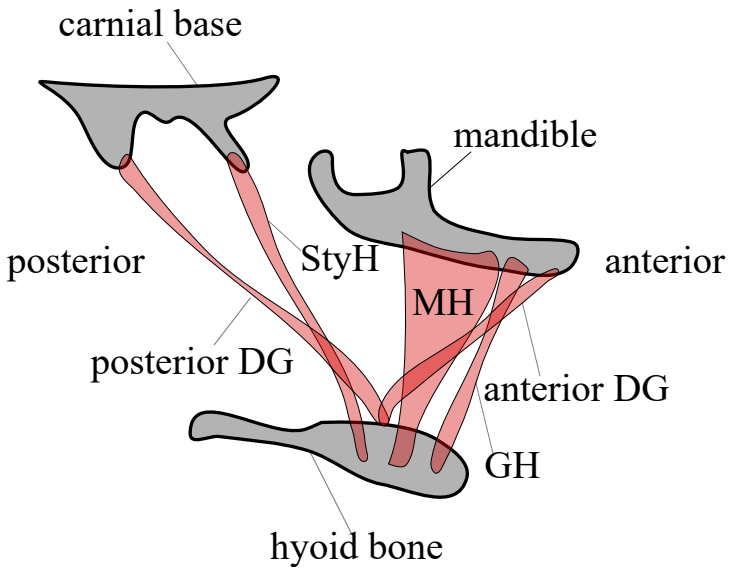


Figure 2.9: Lateral schematic view of the suprahyoid muscles: digastric (DG), geniohyoid (GH), mylohyoid (MH), and stylohyoid (StyH) muscles: inspired by Figure 2 in [115].

Extrinsic laryngeal muscles are also involved in the biomechanics of phonation. Kirzinger and Jurgens [68] recorded the electromyographic (EMG) activity of some extrinsic and intrinsic laryngeal muscles in squirrel monkeys during the production of calls of different voice characteristics, observing noticeable activation levels of various extrinsic laryngeal muscles that depended on the nature of the calls. Moisik and Gick [98] developed a three dimensional numerical posturing model of the larynx, which incorporates different intrinsic and extrinsic laryngeal muscles, with the aim of investigating laryngeal maneuvers that are likely to be associated with different glottal articulatory states (without considering how these laryngeal maneuvers influence phonation). Their findings suggested that certain

extrinsic muscles may play a role in some articulatory states (e.g., glottal configurations of aryepiglotto–epiglottal stops and fricatives).

The impact of extrinsic muscles on phonation remains poorly understood, with conflicting findings reported in the literature on the influence of specific muscles. For example, Atkinson [6] reported negative correlations between the activation of the SH and ST muscles (see Figure 2.8) and phonation fundamental frequency, f_o, based on statistical analysis of voice measures and EMG activity of laryngeal muscles. In contrast, Hong et al. [53] studied an *in vivo* canine laryngeal model and found that contraction of the SH and ST muscles corresponds to a rise in subglottal pressure, shortened CT distance (the distance between two fixed points on the circoid and thyroid cartilages, respectively), lengthened VFs, and increased f_o and vocal intensity. Moreover, they found that activation of the TH muscle corresponds to lower subglottal pressure, widener CT distance, shortened VFs, and lower f_o and vocal intensity. Roubeau et al. [121] reported, based on recorded EMG measurements in two singers during glissandos, that the relation between the activation of ST, SH, and TH and fundamental frequency is context-dependent, where at low and high f_o values, TH, ST, and SH muscles are active; however, their activity becomes less prominent at moderate f_o values. For a comprehensive review of the conflicting findings in the literature regarding the roles of extrinsic muscles, see [174].

Despite the controversy, there are some insightful investigations in the literature that shed some light into the biomechanics of extrinsic laryngeal muscle activation during phonation. In their review article, Vilkman et al. [174] suggested two models of fundamental frequency modulation that depend on relative anterior movement of the thyroid cartilage and rotation of the CT joint, requiring synergetic activation and deactivtion of intrinsic and extrinsic laryngeal muscles, in addition to other neck muscles. Honda et al. [52] conducted magnetic resonance imaging on three male subjects and recorded the positions of the articulators and the larynx during vowel production with different f_o values. They noted that in the high f_o range, the hyoid bone, which is attached to several extrinsic laryngeal muscles, moves horizontally while the larynx height remains relatively constant. In contrast, in the low f_o range, the entire larynx moves vertically and the cricoid cartilage rotates along the cervical lordosis (the spine curvature in the neck region). In addition, they found that as f_o decreases, the hyoid bone tends to move posteriorly and then inferiorly. They noticed that the rotation of the cricoid cartilage, in the direction corresponding to increasing the CT distance, is significant in the low f_o range. Finally, they suggested two mechanisms underlying the impact of extrinsic muscles on phonation at the ranges of high and low f_o ranges: In the high range, the extrinsic muscles attached to the superior margin of the hyoid bone pull the hyoid bone anteriorly leading to pulling the thyroid cartilage forward, causing increased VF tension. On the other hand, in the low f_o range,

the extrinsic muscles attached to the inferior margin of the hyoid bone may be contributing to lowering the larynx, leading to increased CT distance and lowered VF tension.

In a similar study, Hong et al. [54] conducted videofluoroscopic examinations of the movements of the hyoid bone, thyroid cartilage, and cricoid cartilage in addition to the CT distance during fundamental frequency elevation. They reported that vertical movements of the hyoid bone and the cricoid and thyroid cartilages positively correlate with fundamental frequency, whereas CT distance negatively correlates with it. Moreover, they found that the upward movement of the larynx has a more dominant effect on fundamental frequency elevation than the CT distance. They noticed that the extrinsic muscles contribute to pulling the hyoid bone and larynx upward, and that the larynx and trachea tilt forward when the fundamental frequency is higher. Moreover, the authors noticed that the hyoid bone and the thyroid cartilage move more antero-vertically than the cricoid cartilage during fundamental frequency elevation; they thus speculated that this motion may result in increased length and tension level of the VFs and, subsequently, increased f_o values.

2.5 Vocal Hyperfunction

In this section, we introduce a class of voice pathologies that is central in this thesis, namely vocal hyperfunction (VH), wherein we discuss its classifications, and associated characteristics. VH refers to the class of voice disorders associated with the abuse and/or misuse of the vocal mechanisms due to excessive and/or imbalanced muscular forces [47]. VH is associated with the most frequently occurring voice disorders, including benign vocal fold lesions (e.g., nodules, polyps) and laryngitis due to the misuse/overuse of VFs [11]. In general, VHs can be classified into two categories [48]:

(a) phonotraumatic vocal hyperfunction (PVH), previously known as adducted hyperfunction [47], which is characterized by vocal misuse associated with chronic tissue trauma (e.g., nodules and polyps) on the medial surfaces of the VFs,

(b) and nonphonotraumatic vocal hyperfunction (NPVH) (sometimes referred to as muscle misuse dysphonia, primary muscle tension dysphonia (MTD), or functional dysphonia [99, 123, 122]), which is associated with chronic voice-related symptoms in the absence of trauma or any structural, neurological, or psychological abnormalities.

Studies on subjects with PVH and NPVH have shown that their biomechanics are different. For example, glottal aerodynamic studies have shown that patients with PVH tend to

generate relatively higher subglottal pressure in order to produce normal sound pressure levels [27]. Increased subglottal pressure, in combination with the close proximity of the VFs observed in subjects with PVH [47], results in higher contact forces during phonation. This consequently causes trauma to the VF tissue. On the other hand, patients with NPVH display high subglottal pressure but the glottis, for such cases, is not completely closed; hence, contact forces are not high and trauma to the VFs is improbable [27]. The NPVH subclass is curious and opaque, due to the highly variable characteristics and diverse potential underlying mechanisms, wherein laryngeal damage is absent. For that subclass, there are various symptoms including, tight neck muscles, incomplete glottal closure, and abnormal fundamental frequency [99, 105].

The prevalence of VH among patients with voice disorders has motivated researchers to study the associated underlying mechanisms [186] and seek to develop robust assessment tools [136] and treatments [123]. Hillman et al. [48] have formalized, based on previous clinical studies, a general framework for VH and its subclasses. They hypothesized that VH is typically triggered by increased vocal demands, insufficient recovery, and/or emotional stress, where the vocal mechanisms are not used optimally due to such triggers, leading, in the case of PVH, to formation of lesions. In return, lesions may worsen the use of the vocal mechanisms which, if not treated, lead to a vicious cycle that deteriorates the vocal systems. In the case of NPVH, the misuse of the vocal mechanism does not cause trauma, and is speculated to be primarily driven by psychological distress, leading to aphonia (loss of voice) or dysphonia (difficulty of speaking) if the vocal mechanism is not corrected by means of voice therapy. Interestingly, clinical evidence suggested that patients with VH may be suffering from abnormal neural motor control of the vocal system [137] (see the discussion on neural motor control in Chapter 1). This has motivated researchers recently to develop neurocomputational models in order to gain refined insights into the neural motor control nature in normal and pathological speech [177].

2.5.1 Vocal hyperfunction and extrinsic laryngeal musculature

As highlighted in the above discussion, vocal hyperfunction, including PVH and NPVH (or MTD), is typically linked to excessive activation of neck muscles. Various investigations in the literature have highlighted the relevance of these muscles in hyperfunctional patients. Angsuwarangsee and Morrison [5] assessed the tension levels of extrinsic laryngeal musculature in patients with MTD using a palpation technique and tension grading system. The authors reported a strong relationship between the TH muscle tension and MTD, especially in subjects who suffer from gastroesophageal reflux. They suggested that this correlation is due to the TH muscle tensing to induce a supraglottal compression to

block the stomach acid from reaching the oral cavity. Lowell et al. [82] used a radiographic technique to measure the hyoid bone position, laryngeal position, and the hyolaryngeal space during phonation in patients with MTD, in comparison with control participants. They reported that normalized vertical hyoid and laryngeal positions during phonation are significantly higher for patients with MTD than for control participants. O'Keeffe et al. [111] utilized perilaryngeal-cranial surface EMG measurements, collected from healthy subjects and vocal hyperfunctional patients, to construct functional muscle networks [110] for each subject/patient and computed different network parameters under different vocal conditions. They reported notable differences in these parameters between healthy speakers and speakers with VH, where these differences are condition-specific, making functional muscle networks a promising assessment tool.

2.5.2 Relative fundamental frequency analysis

In recent years, voice researchers have strived to develop acoustic measures to diagnose and classify voice disorders without the need to conduct extensive clinical examinations of the larynx [61, 60, 46]. Such tools are important, especially when diagnosing non-phonotraumatic voice disorders, where no physiological abnormalities in the larynx are observed [122, 123]. One promising tool that has gained research interest recently due to its classification capabilities is relative fundamental frequency (RFF) of phonation surrounding voiceless consonants [36, 136, 45]. This classification capability of RFF is attributed to the acoustic and physiological differences between healthy and pathological speakers that manifest significantly during transient phonation periods surrounding voiceless consonants [44].

In clinical evaluations, RFF analysis is typically conducted on acoustic signals from microphone recordings [176, 136], or on acceleration signals from neck measurements [77], where phonation cycles surrounding voiceless consonants are detected either manually, which is the standard, or semi-automatically [75, 175]. Typically, twenty cycles are considered in the analysis, with 10 cycles for the offset part of phonation (i.e., the period corresponding to decaying VF oscillations) prior to the voiceless consonant and 10 cycles for the onset part (i,e, the period corresponding to rising VF oscillations) following the voiceless consonant (see Figure 2.10).

The fundamental frequency of these cycles is determined from the signal peaks and the RFF, a normalized measure of frequency in the form of semitones, is computed as

$$\mathbf{RFF}(f) = 12\log_2\left(\frac{f}{f_{\mathbf{ref}}}\right), \tag{2.1}$$

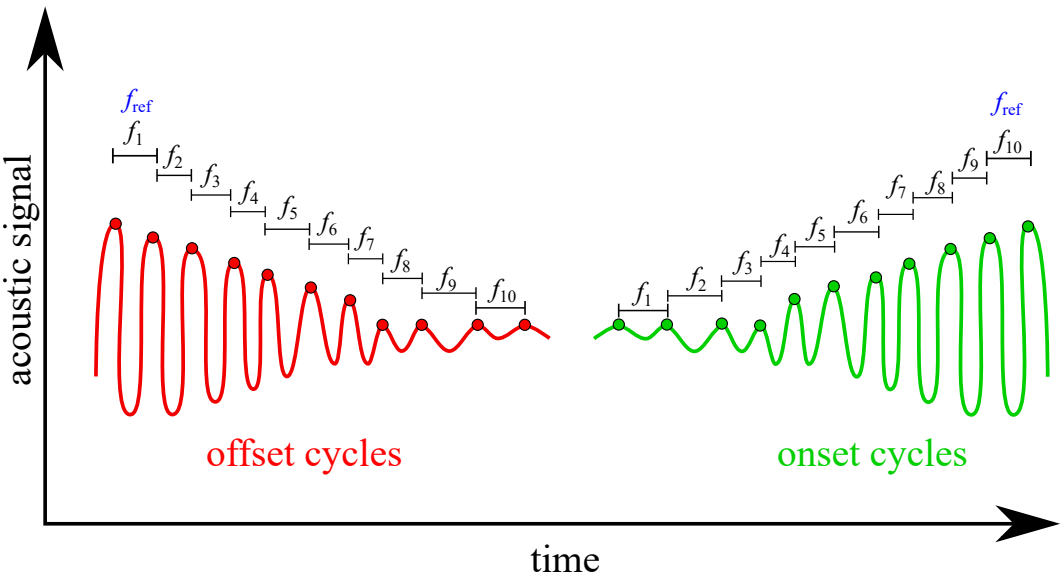


Figure 2.10: Schematic representation of an acoustic signal used in relative fundamental frequency analysis.

where f denotes the fundamental frequency, and $f_{\mathbf{ref}}$ denotes the fundamental frequency of the reference cycle. For phonation offset, $f_{\mathbf{ref}}$ is the first cycle, whereas the tenth cycle is used for phonation onset (see Figures 2.10 and 2.11).

2.5.3 Phonation onset and offset: Clinical observations

Herein, we present some clinical observations associated with phonation fundamental frequency and RFF during phonation onset and offset for various phonetic contexts (see Figure 2.12).

In general, speakers exhibit a drop in fundamental frequency and RFF during phonation offset, with unhealthy adult speakers and aged speakers exhibiting sharper drops in RFF during offset in comparison to healthy adult speakers and young speakers, respectively [136, 176].

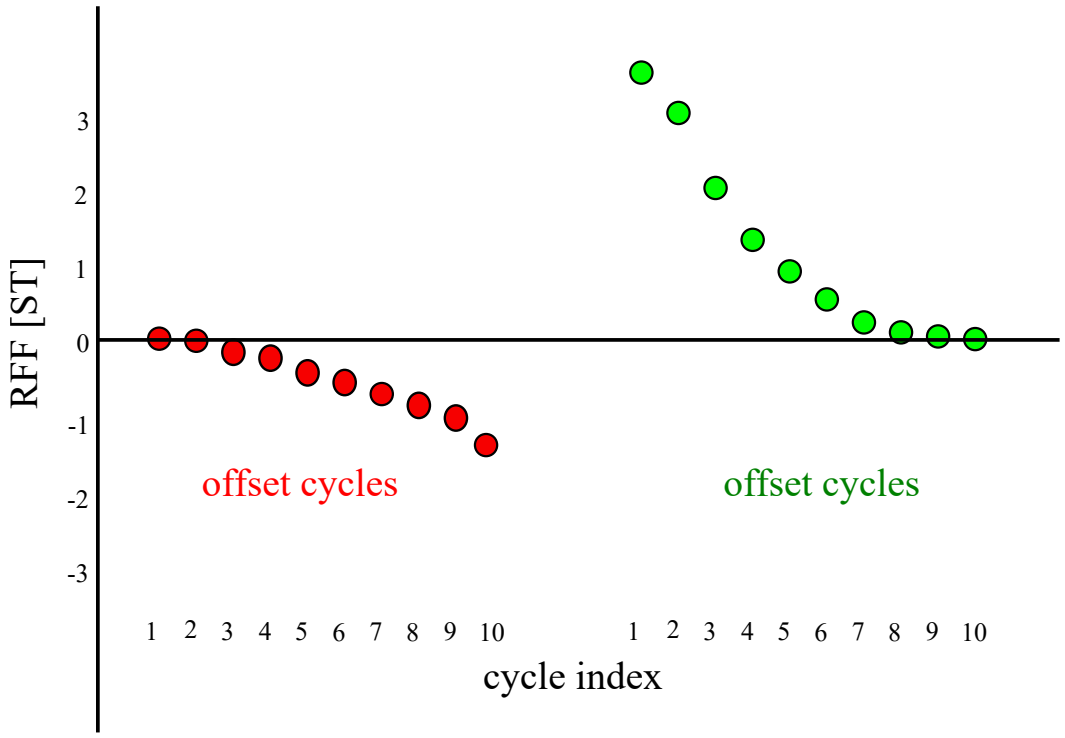


Figure 2.11: A schematic of a typical relative fundamental frequency plot

On the other hand, phonation onset exhibits a variety of fundamental frequency patterns depending on phonetic context and vocal health, as shown schematically in Figure 2.12. . In the case of isolated and initial vowels, fundamental frequency typically exhibits a gradual increase until a sustained phonation frequency is attained [97, 132]. When the vowel is preceded by a voiceless consonant, as in /pa/, an initial spike in fundamental frequency followed by a gradual decay is observed [109, 80], where the onset period has been found to be dependent on the language and the acoustic nature of the voiceless consonant (unaspirated vs. aspirated) [30]. With such gestures, speakers with healthy voices exhibit higher initial (relative) fundamental frequency values compared to speakers with vocal hyperfunction. When the vowel is preceded by a voiced consonant, such as /ba/, there is less agreement in the literature regarding the temporal evolution of fundamental frequency, with some studies observing a gradual increase [97, 51], but others finding inconsistent patterns between and within speakers [109]. Moreover, it has been found that fundamental frequency patterns in the case of vowels preceded by voiced consonants are

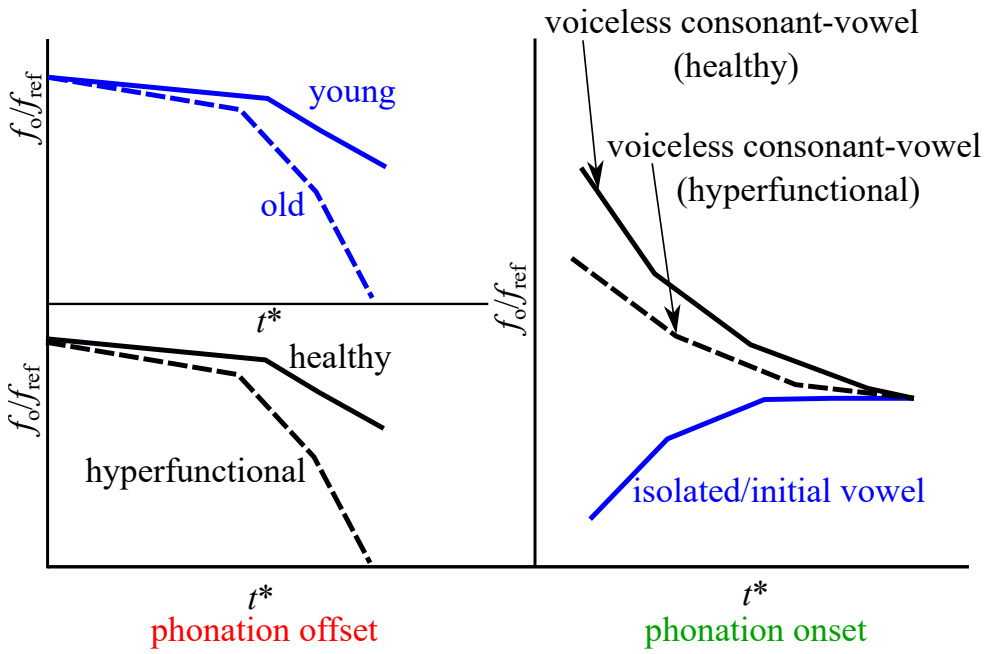


Figure 2.12: Schematic representation of fundamental frequency f_o, normalized by reference fundamental frequency f_{ref} (*e.g.*, fundamental frequency of the tenth onset/first offset cycle [136]) as a function of normalized time t^* (*e.g.*, number of onset/offset cycles [136]) for different phonetic contexts and vocal health.

context-dependent [42, 67]. Regardless, empirical evidence suggests that onset frequency in vowels preceded by voiceless consonants is higher than that in vowels preceded by voiced consonants [109].

Efforts to explain the patterns of phonation offset and onset

The physics of phonation offset is extremely complex due to its transient nature and the involvement of various laryngeal, aerodynamic, and acoustic factors. During offset, laryngeal muscles abduct the VFs and alter their mechanical properties (vibrating mass, stiffness, etc), which induces a gradual transition from a collision to non-collision regime [25]. Similar to offset, the mechanics of phonation onset is extremely complex with the VFs being adducted and the vibrations of VFs being initiated when VFs are sufficiently close depending on the subglottal pressure induced by the lungs. Moreover, adduction of the VFs involves activation of several laryngeal muscles, which subsequently changes the geometry and mechanical properties of the VFs.

A number of studies have attempted to elucidate the physics of phonation offset. Smith and Rob [132] conducted an experimental analysis on utterances of different aerodynamic characteristics and suggested that laryngeal and aerodynamic factors are equally important in varying frequency during phonation offset. They further suggested that the drop in RFF during offset may be attributed to a decrease in the VF stiffness. Jaiswal [59] investigated clinically the role of the CT muscle during offset and found that the CT activation increase during offset is not consistent among all speakers. Watson [176] analyzed the differences in RFF trends between young and adult speakers and hypothesized that the drop in RFF during offset is attributed to bio-mechanical factors and, in particular, the decrease in collision forces associated with abduction.

Several underlying factors have been hypothesized to drive the observed fundamental frequency patterns during phonation initiation, including laryngeal muscle tension, aerodynamics, and vocal fold contact[1]. Smith and Rob [132] empirically investigated onset fundamental frequency patterns of vowels preceded by fricatives and stop consonants, in addition to isolated vowels. They speculated that the rise of onset fundamental frequency in the case of isolated vowels is due to a rise in VF tension. They further suggested that laryngeal muscle tension is a predominant factor in the case of vowels preceded by voiceless consonants. Lofqvist et al. [80] investigated cricothyroid muscle activation during phonation onset using electromyography and found correlation between increased muscle

[1]For a thorough discussion of some of the mechanisms through which aerodynamics and muscle tension are hypothesized to alter onset fundamental frequency, see the seminal work of Hombert et al. [51].

activation and the higher fundamental frequency observed during the onset of vowels preceded by voiceless consonants. Lofqvist et al. [81] estimated the glottal flow characteristics from oral flow measurements and found that peak glottal flow is higher in vowels preceded by voiceless consonants in comparison with voiced consonants, indicating a correlation with the observed higher initial fundamental frequency. Moreover, Lofqvist et al. [81] found that the glottal flow characteristics differ between aspirated voiceless consonants and their unaspirated counterparts.

2.5.4 Incomplete glottal closure

Incomplete glottal closure is a characteristic of NPVH as mentioned previously. Typically, the VFs are brought together to almost full closure prior to the onset of phonation, and the interaction between the fully-adducted VFs and the air flow driven by the lungs results in VF vibrations and subsequently acoustic wave propagation, the basis of voiced speech. In some scenarios, complete glottal closure is not attained, which can result in inefficient voice production [186], and, in some cases, stress concentration in the VFs that may lead to VF trauma [23]. Hence, incomplete glottal closure is sometimes linked to disorders that are associated with inefficiencies in or damage to the vocal mechanism, including Parkinson's disease [41][2] and NPVH or MTD [99].

Incomplete glottal closure[3] comes in various patterns [99, 105, 133] as shown schematically in Figure 2.13, including bowed shape: a glottal pattern where the left and right VFs geometries are typically convex with a gap at the mid-membranous portion; posterior glottal opening: a pattern, where glottal closure is achieved anteriorly and mid-membranously with a remaining gap at the posterior margin; and hourglass glottal configuration: a pattern with anterior and posterior gaps and potential VF contact mid-membranously. In addition, there exist other incomplete glottal closure patterns, sharing similarities with those mentioned above, such as spindle-shaped glottis, and anterior opening (see [119, 133]).

There are some findings in the literature that shed light into the potential mechanisms underlying the patterns of incomplete glottal closure. For example, Morrison and Rammage [99] suggested that, based on inspections of cadaver larynges, posterior glottal

[2]Parkinson's disease is a relatively prevalent disorder, affecting the human central, peripheral, and enteric nervous systems [12].

[3]In our analysis of incomplete glottal closure in Chapter 6, a glottal configuration/pattern refers to resting glottal shape prior to phonation initiation. In clinical settings, incomplete glottal closure configurations are typically determined as the shapes corresponding to maximum glottal closure during phonation [133, 105]. Our definition herein isolates laryngeal factors by dismissing the dynamics of VF vibrations.

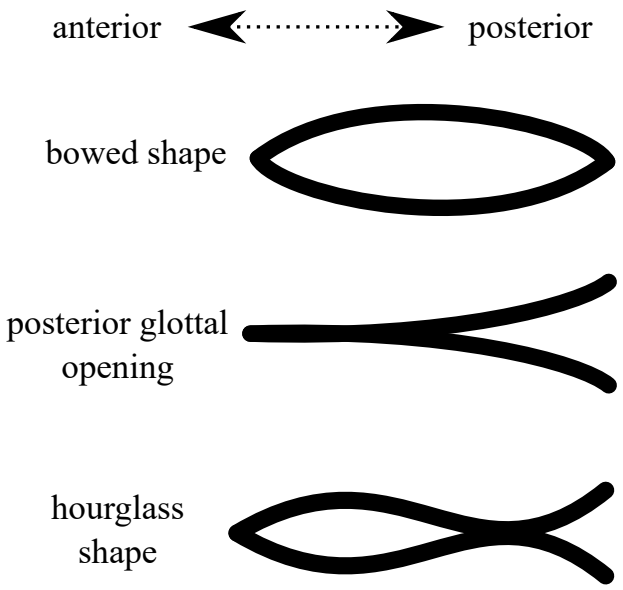


Figure 2.13: Schematic diagram (superior view) of some incomplete and curved glottal closure patterns that are observed clinically.

opening is associated with excessive activation of the PCA muscle. In experimental investigations of excised canine models, it was found that activating the TA muscle, while keeping other adductory laryngeal muscles inactive, leads to anterior and mid-membranous glottal closure, whereas the posterior glottis stays open [20, 15], thus resulting in a closure pattern similar to posterior glottal opening. On the other hand, when the TA muscle is relaxed, and other adductory muscles are activated, closure is achieved only at the posterior margins of the VFs [20, 15], thus leading to a bowed configuration as seen in Figure 2.13. Complete glottal closure was attained with the coactivation of all adductory muscles [20, 15]. These findings are in alignment with simulation results from high-fidelity numerical models [182, 183]. Recently, research interest has also been directed towards investigating how activating laryngeal muscles alters the VF medial surfaces [117].

2.6 Laryngeal Modelling

Important note: The defined parameters and variables introduced herein are valid only within this section (local variables). In each chapter, relevant laryngeal parameters and parameters will be redefined with potentially different notations.

As has been shown in the previous sections, the anatomy of the larynx and the mechanics healthy and pathological phonation are quite complex. Besides that, experimental investigation of phonation and laryngeal biomechanics is very challenging due to several sources of variability, including (1) anatomical aspects of vocal organs (e.g., VF length, thickness, and depth), which vary between individuals, and (2) phonation conditions (e.g., muscle activation levels, subglottal pressure, hydration level, etc), which are not consistent even within subjects. Therefore, the modelling paradigm provides a feasible alternative that enables analyzing the mechanics of phonation under controlled conditions.

2.6.1 One degree-of-freedom model

We start our discussion with one of the simplest phonation models, namely the one degree-of-freedom (1-DOF) VF model of Titze [155], which was developed to investigate the underlying physics of small amplitude oscillations of VFs during phonation. The model is a simple spring-mass-damper system, abstracting one of the VFs and assuming symmetry with respect to the medial plane (i.e., the movement of the other fold is a reflection of that of the first fold), where the aerodynamic forces from the glottal flow are modelled using a Bernoulli nonlinear forcing term.

Titze conducted a stability analysis on the 1-DOF model, showing that the phonation threshold pressure is proportional to the neutral glottal gap and tissue damping. These relations have been found to agree reasonably with experimental measurements from a physical model of the VFs [163]. Interestingly, it has been observed that the minimum pressure needed to initiate phonation (onset threshold pressure) is larger than the minimum pressure needed to sustain phonation (offset threshold pressure) [163]. Motivated by this observation, Lucero [83] adapted the 1-DOF model of Titze and verified theoretically the experimental observations of onset and offset threshold pressures. Further theoretical studies have adopted/modified the 1-DOF model of Titze to investigate further the conditions of phonation initiation [84, 86, 87]. Despite the simplicity of the 1-DOF model, it has enabled a refined understanding of the mechanisms of phonation, which illustrates the importance of theoretical analysis in voice research. In fact, we conduct similar theoretical investigations in this thesis, where we adopt similar 1-DOF models in Chapters 3 and 4 to gain refined insights into the physics of phonation onset and offset.

2.6.2 Lumped-element models: The Story-Titze model

1-DOF models have been used extensively in theoretical analyses of phonation. Such models are, however, not suitable for numerical simulations of phonation due, in part, to the restrictive assumptions embedded in their derivations. The first well-established model to be used in phonation simulations is the symmetric lumped element model of Ishizaka and Flanagan [58], which subdivides the mucosal layer into two masses (that is, two degrees-of-freedom (2-DOF) model) and thus is able to capture both the translational and rotational modes of VF vibration observed clinically. Since its publication, the two mass model has been adopted and modified by several researchers, resulting in a significant number of lumped-element phonation models [26]. Herein, we focus on a particular well-established lumped element model, namely the body-cover model (BCM) of Story and Titze [142], which we have adopted in our investigations in this thesis. Due to its relative simplicity and capability to capture the essential physics of VFs dynamics, the BCM has been adopted by several researchers in various voice studies [186, 3, 24, 40]. The BCM assumes symmetric vibration of vocal folds, with respect to the medial plane; hence, one VF is considered during its derivation. As its name indicates, the BCM adopts the body-cover approximation of the layered VF structure discussed in Section 2.2.1, where it consists of two small cover masses, m_1 and m_2 , and a large body mass m_b. These masses are connected via springs and dampers to model tissue viscoelasticity. The schematic diagram of the BCM is given in Figure 2.14. In the discussion below, we refer to m_1 as the lower mass and m_2 as the upper mass. The numerical values of the BCM parameters can be

found in [142, 185].

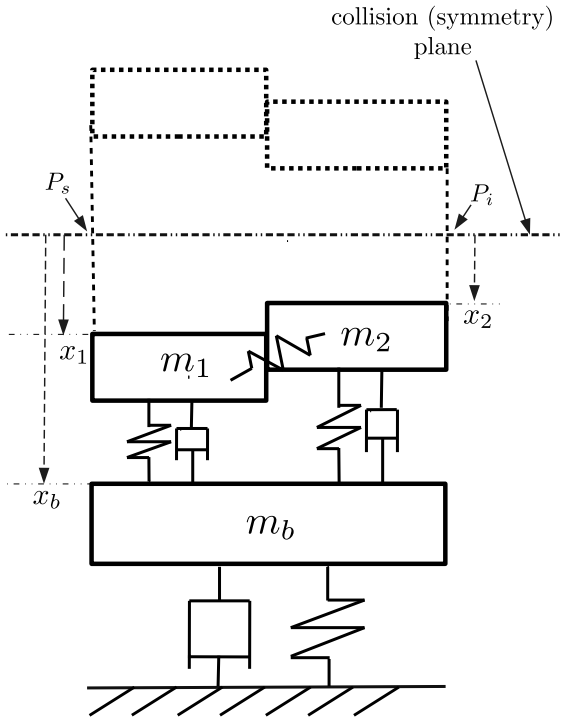


Figure 2.14: Schematic diagram of the body-cover model.

The momentum balances on the BCM masses are:

$$
\begin{aligned}
F_1 &= m_1\ddot{x}_1 = F_{k,1} + F_{d,1} + F_{k,c} + F_{e,1} + F_{\text{col},1}, \\
F_2 &= m_2\ddot{x}_2 = F_{k,2} + F_{d,2} - F_{k,c} + F_{e,2} + F_{\text{col},2}, \\
F_b &= m_b\ddot{x}_b = F_{k,b} + F_{d,b} - \left[F_{k,1} + F_{k,2} + F_{d,1} + F_{d,2}\right],
\end{aligned}
$$

where x_1, x_2 and x_b denote the lateral displacements, with respect to the medial plane, of the lower mass, upper mass, and the body mass, respectively. $F_{k,1}$, $F_{k,2}$, $F_{k,b}$ and $F_{k,c}$ denote the spring forces associated with the lower mass, upper mass, body mass, and the elastic coupling force between the two cover masses, respectively, $F_{d,1}$, $F_{d,2}$, $F_{d,b}$ denote the damping forces on the lower mass, upper mass, and body mass, respectively, $F_{\text{col},1}$ and

$F_{\text{col},2}$ denote the collision forces on the lower and upper masses respectively, and $F_{e,1}$ and $F_{e,2}$ denote the aerodynamic forces on the lower and upper cover masses, respectively.

The springs connecting the masses are assumed to be cubic to capture the nonlinearity of the VF tissues; hence, the spring forces $F_{k,1}$, $F_{k,2}$, $F_{k,b}$ and $F_{k,c}$ are given by the relations

$$\begin{aligned}
F_{k,1} &= -k_1\left(\{x_1 - x_{0,1}\} - \{x_b - x_{0,b}\} + \eta_1[\{x_1 - x_{0,1}\} - \{x_b - x_{0,b}\}]^3\right),\\
F_{k,2} &= -k_2\left(\{x_2 - x_{0,2}\} - \{x_b - x_{0,b}\} + \eta_2[\{x_2 - x_{0,2}\} - \{x_b - x_{0,b}\}]^3\right),\\
F_{k,b} &= -k_b\left(\{x_b - x_{0,b}\} + \eta_b\{x_b - x_{0,b}\}^3\right),\\
F_{k,c} &= -k_c\left(\{x_1 - x_{0,1}\} - \{x_2 - x_{0,2}\}\right),
\end{aligned}$$

where $x_{0,1}$, $x_{0,2}$ and $x_{0,b}$ denote the neutral positions of the lower, upper, and body masses, respectively, and k_1, k_2, k_b, η_1, η_2, and η_b are stiffness parameters.

The damping forces $F_{d,1}$, $F_{d,2}$, and $F_{d,b}$ are assumed to be linear, given by the relations

$$\begin{aligned}
F_{d,1} &= -d_1(\dot{x}_1 - \dot{x}_b),\\
F_{d,2} &= -d_2(\dot{x}_2 - \dot{x}_b),\\
F_{d,b} &= -d_b\dot{x}_b,
\end{aligned}$$

where d_1, d_2, and d_b are associated damping coefficients.

Collision forces are modelled using cubic springs, where they are applied on the lower and upper masses when they surpass the symmetry plane; that is,

$$F_{\text{col},1} = \begin{cases} -k_{\text{col},1}\left(x_1 + \eta_{\text{col},1}x_1^3\right), & x_1 < 0,\\ 0, & x_1 \geq 0, \end{cases}$$

$$F_{\text{col},2} = \begin{cases} -k_{\text{col},2}\left(x_2 + \eta_{\text{col},2}x_2^3\right), & x_1 < 0,\\ 0, & x_1 \geq 0, \end{cases}$$

where $k_{\text{col},1}$, and $k_{\text{col},2}$, $\eta_{\text{col},1}$ and $\eta_{\text{col},2}$ are associated stiffness parameters.

The aerodynamic forces on the lower and upper masses, $F_{e,1}$ and $F_{e,2}$, are associated with the pressure distribution induced by the glottal flow, where they are given by the relations

$$\begin{aligned}
F_{e,1} &= P_1 L_{\text{vf}} T_1,\\
F_{e,2} &= P_2 L_{\text{vf}} T_2,
\end{aligned}$$

where L_{vf} denote the VF length, T_1 and T_2 denote the thicknesses of the lower and upper masses,respectively, and P_1 and P_2 are the aerodynamic pressures on the lower and upper masses, respectively. Herein, it is assumed that the pressure distribution over each cover mass is uniform.

The aerodynamic forces can be deduced by assuming an idealized glottal flow that is divided into two regions based on the geometrical configuration of the cover masses, namely a Bernoulli flow region and a jet regime with an approximately uniform pressure. Let a_1 and a_2 denote the glottal areas at lower and upper masses, respectively, which are given by

$$a_1 = \max\{0, 2L_{\text{vf}}x_1\} \tag{2.2}$$

and

$$a_2 = \max\{0, 2L_{\text{vf}}x_2\}. \tag{2.3}$$

In the derivation of the glottal area formulas, the symmetry assumption is implemented; hence, the use of the factor 2 in Equations (2.2) and (2.3). The pressures over the lower and upper masses are computed as follows: let P_s and P_i denote the subglottal and supraglottal pressures, respectively, then, the pressures over the cover masses are computed as follows:

(a) If $0 < a_2 \leq a_1$ (divergent configuration), the flows over m_1 and m_2 are Bernoulli and jet, respectively, hence, $P_1 = P_s - (P_s - P_i)(a_2/a_1)^2$, $P_2 = P_i$.

(b) If $0 < a_1 \leq a_2$ (convergent configuration), the flows over m_1 and m_2 are both jet, hence, $P_1 = P_2 = P_i$.

If collision occurs for one of the cover masses, then the pressures over the masses are computed as follows:

(a) If $0 < a_2$ and $a_1 = 0$, then $P_1 = 0$, $P_2 = P_i$.

(b) If $0 < a_1$ and $0 = a_2$, then $P_1 = P_s$, $P_2 = 0$.

(c) Finally, if $a_1 = a_2 = 0$, then $P_1 = P_2 = 0$.

Acoustic modeling: Wave reflection analog

The computations of the aerodynamic forces depend on the subglottal and supraglottal pressure values, P_s and P_i. To evaluate these pressures and to simulate sound production, acoustic models of the subglottal and supraglottal vocal tracts should be integrated with

the BCM. Herein, we briefly present one of the most well-established acoustic modelings framework in speech research, namely the one-dimensional (1-D) wave reflection analog (WRA) [144, 78], which was first proposed by Kelly and Lochbaum [64]. In the WRA framework, the geometry of the tracts is assumed to be time invariant, where it is approximated using connected cylinders of equal length and the airflow in each cylinder is assumed to obey the 1D acoustic wave equations. Consequently, the pressure wave in each cylinder is decomposed linearly into forward and backward waves transmitting with the speed of sound. Several acoustic models have been proposed as extensions or variations of the WRA, including the Webster's Equation [167], 1-D models with conic segments [146], and time-varying segments [145], and two-dimensional models[102]. The WRA, when embedded in sustained phonation simulations, provides reasonable accuracy while requiring feasible computational cost, making it an attractive tool in speech research (see, e.g., [146]).

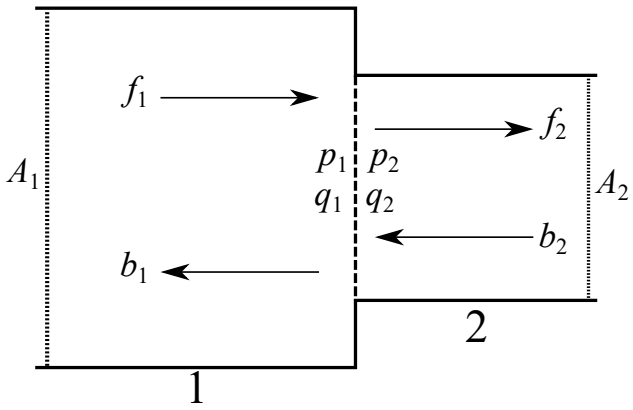


Figure 2.15: Two tube junction.

Let us further discuss the WRA modelling framework. Consider a junction point connecting two consecutive cylinders (cylinder 1 and cylinder 2) as seen in Figure 2.15. Then, the pressure in cylinder 1 at the left side of the junction can be written as $p_1 = f_1 + b_1$ where f_1 is forward traveling and b_1 is backward traveling pressure waves. Similarly, the

pressure in the second tube at the right side of the junction can be written as $p_2 = f_2 + b_2$. Let A_1 and A_2 denote the cross-sectional areas of cylinders 1 and 2, respectively, and ρ_a and c_a denote the density and speed of sound of air in the tracts, respectively. The flow rates at the left and right sides of the junction are given by[4]

$$q_1 = \frac{A_1}{\rho_a c_a}(f_1 - b_1),$$
$$q_2 = \frac{A_2}{\rho_a c_a}(f_2 - b_2),$$

respectively. By considering the continuity of pressure ($p_1 = p_2$), we have

$$f_1 + b_1 = f_2 + b_2. \tag{2.4}$$

Moreover, the continuity of the flow rate ($q_1 = q_2$) at the junction yields

$$\frac{A_1}{\rho_a c_a}(f_1 - b_1) = \frac{A_2}{\rho_a c_a}(f_2 - b_2). \tag{2.5}$$

By combining the continuity conditions given by Equations (2.4) and (2.5), we can write the outward pressures, leaving the junction, namely b_1 and f_2, in terms of the inward pressures entering the junction, namely f_1 and b_2, using the scattering equations

$$b_1 = \left(\frac{A_1 - A_2}{A_1 + A_2}\right) f_1 + \left(\frac{2A_2}{A_1 + A_2}\right) b_2,$$
$$f_2 = \left(\frac{2A_1}{A_1 + A_2}\right) f_1 + \left(\frac{A_2 - A_1}{A_1 + A_2}\right) b_2.$$

Similar equations will be also defined for all junction points of the discretized tracts.

If the junction is associated with the glottis, we can decompose the subglottal and supra-glottal pressures P_s and P_i as $P_s = P_{s,+} + P_{s,-}$ and $P_i = P_{i,+} + P_{i,-}$. The continuity of the glottal flow, with respect to pressure and flow rate, implies that

$$P_s - P_i = 2(P_{s,+} - P_{i,-}) - \frac{\rho_a c_a}{A^*} q_g, \tag{2.6}$$

[4]The flow rate formulas for one-dimensional linear acoustic wave propagation can be deduced from the facts that (1) the pressure p and the flow rate q in a cylinder, with a cross-sectional area A, satisfy the momentum equation $\rho_a \partial q/\partial t = -A\partial p/\partial x$, where x denotes the position through the cylinder length, and t denotes time, and (2) the pressure and flow rate obey the wave equation, propagating with the speed of sound c_a.

where q_g is the glottal flow rate and $1/A^* = 1/A_s + 1/A_i$. The glottal flow is assumed to satisfy the Bernoulli relation

$$P_s - P_i = \frac{k_t}{2}\rho\frac{|q_g|q_g}{a_g^2}, \tag{2.7}$$

where k_t is a trans-glottal pressure coefficient, and $a_g = \min\{a_1, a_2\}$ is the glottal area. By combining Equations (2.6) and (2.7), we obtain

$$\frac{k_t}{2}\rho\frac{|q_g|q_g}{a_g^2} + \frac{\rho c}{A^*}q_g = 2(P_{s,+} - P_{i,-}).$$

The above equation can be solved in terms of the incident pressures $P_{s,+}$ and $P_{i,-}$, yielding (see [88])

$$q_g = \pm\left(\frac{a_g c_a}{k_t}\right)\left(\frac{-a_g}{A^*} + \left[\left(\frac{a_g}{A^*}\right)^2 + \left(\frac{4k_t}{c_a^2\rho_a}|P_{s,+} - P_{i,-}|\right)\right]^{\frac{1}{2}}\right).$$

In the above flow rate equation, the plus sign is used when $P_{s,+} > P_{i,-}$ and the negative sign is used otherwise [154]. The output pressures $P_{s,-}$ and $P_{i,+}$ can then be given in terms of the inward pressures and the flow rate using the relations

$$P_{s,-} = P_{s,+} - \frac{\rho_a c_a}{A_s}q_g,$$
$$P_{i,+} = P_{i,-} + \frac{\rho_a c_a}{A_i}q_g.$$

At the terminal junction corresponding to the upper end of the supraglottal tract (e.g., mouth opening), a condition in terms of radiated pressure p_{rad} and flow rate q_{rad} are prescribed. For example, the pressure radiation can be modelled using a piston in a sphere, which has an approximate solution (with a resistance-inductance representation), given by (see, e.g., [184, 29])

$$\frac{\mathrm{d}q_{\text{rad}}}{\mathrm{d}t} = \frac{1}{R_{\text{rad}}}\frac{\mathrm{d}p_{\text{rad}}}{\mathrm{d}t} + \frac{p_{\text{rad}}}{L_{\text{rad}}},$$

where R_{rad} and L_{rad} denote the radiation resistance and inductance, respectively. Moreover, at the terminal subglottal junction attached to the lungs, a condition in terms of the lungs P_L, forward f_l, and backward b_l pressures are provided. For example, the following relation is typically used:

$$f_l = P_L - r_l b_l,$$

where r_l is a specified coefficient associated with the lungs junction. These conditions at the terminal ends of the subglottal and supraglottal tracts and the glottis provide the boundary conditions needed to solve for the acoustic pressures.

Now, let us assume that the length of each tube is δ and the time step is $\tau = \delta/c_a$, then after one time step, the output pressures cover the lengths of their associated cylinders and hence become input pressures to the adjacent junctions. Hence, the iterative use of the scattering equations, in combination with redefining the output pressures to be input pressures at neighboring junctions after each time step, enables finding the pressure profile at junction points at discrete time instances that are multiples of the given time step. We need to note that in the procedure above, redefined output pressures are typically multiplied by attenuation factors to account for viscous losses [184]. Simulations of the BCM model, coupled with WRA, enables generating a time profile of the pressure radiated from the vocal tract, corresponding to sound, at discrete time instances.

Muscle activation rules

Parameters of the BCM and similar lumped-element models (e.g., mass, stiffness coefficients, damping coeffieicnts, etc.) are abstract and not directly related to real VFs. Therefore, to enable physiologically relevant analysis, there is a need to connect the abstract BCM parameters to physiological parameters of the larynx, including the geometrical and mechanical properties of laryngeal tissues and the activation levels of intrinsic laryngeal muscles. Titze and Story [164] proposed a set of rules that establish a connection between the physiological laryngeal properties and the abstract BCM parameters. This was done by relating the parameters of the BCM to the passive geometrical and mechanical properties of the VFs, in addition to stresses and strains induced by activating the laryngeal muscles. The levels of activation in muscles are given in terms of a set of normalized parameters (with values between 0 and 1). In the following, we state some of these rules to clarify their machinery.

The rules in [164] account for the activation levels of four laryngeal muscles, the CT, TA, and LCA/PCA muscles, which are represented by the normalized parameters (with values between zero and one), ν_{ct}, ν_{ta}, and ν_{lca}, respectively. The term ν_{lca} is set to also account for PCA activation, by considering the extended range [-1,1], where the negative portion corresponds to active PCA muscle and inactive LCA muscle (PCA muscle is antagonistic to the LCA muscles). Hence, these rules are not capable of modelling coactivation of the LCA and PCA muscles.

The rules idealize the layered VF structure, wherein each VF is set to constitute of three discrete layers, the mucosa, ligament, and TA muscle (see Section 2.2.1). In addition, the rules assume that the VF strain ϵ (which is identical in each VF layer) is given by the relation

$$\epsilon = G(R\nu_{\text{ct}} - \nu_{\text{ta}}) - H\nu_{\text{lca}},$$

where G, R, and H are parameters that can be estimated from clinical measurements. Consequently, the VF length after muscle activation, L_{vf}, is given by

$$L_{\text{vf}} = (1+\epsilon)L_{\text{vf},0},$$

where $L_{\text{vf},0}$ is the resting VF length. Similarly, the vocal fold thickness T_{vf} is given by

$$T_{\text{vf}} = \frac{T_{\text{vf},0}}{1+0.8\epsilon},$$

where $T_{\text{vf},0}$ is the resting VF thickness. The division by the factor $1+0.8\epsilon$ accounts for material compressibility/incompressibility, where most of the length change is assumed to be absorbed by trickiness change (increase in length will result in decrease in thickness and vise versa).

As the BCM constitutes of only two layers (cover and body), the rules have to further relate the properties of the mucosa, ligament, and TA muscle tissue to the cover and body layers. For example, the depths of the cover and body layers, D_c and D_b, are computed to be

$$D_c = \frac{D_{\text{muc}} + 0.5D_{\text{lig}}}{1+0.2\epsilon},$$

and

$$D_b = \frac{\nu_{\text{ta}} D_{\text{mus}} + 0.5D_{\text{lig}}}{1+0.2\epsilon},$$

where D_{muc}, D_{lig}, and D_{mus} denote the resting depths of the mucosa, ligament, and TA muscle, respectively.

The cover layer of the BCM consists of two masses, and that necessitates introducing the nodal point parameter z_n, which divides the cover layer into lower and upper masses. This term is given by the relation

$$z_n = (1+\nu_{\text{ta}})T/3.$$

The cover and body masses are then given by

$$\begin{aligned} m_1 &= \rho_t L_{\text{vf}} T_{\text{vf}} D_c z_n / T, \\ m_2 &= \rho_t L_{\text{vf}} T_{\text{vf}} D_c (1 - z_n/T), \\ m_b &= \rho_t L_{\text{vf}} T_{\text{vf}} D_b, \end{aligned}$$

where ρ_t is the tissue density.

Another important rule in the Titze and Story framework [164] is the adduction rule, which determines the resting position of the upper cover mass, which is regulated through the relation

$$x_{0,2} = 0.25L_{\mathrm{vf},0}(1 - 2\nu_{\mathrm{lca}}). \tag{2.8}$$

The resting position of the lower cover mass, $x_{0,1}$, can then be estimated from $x_{0,2}$ and a specified relation for $x_{0,2} - x_{0,1}$ (e.g., for a rectangular glottis, $x_{0,2} - x_{0,1} = 0$).

The stiffnesses in the cover and body layers are also estimated in a similar fashion. For example, from the strain ϵ (and muscle activation ν_{ta} in the case of the TA muscle), the stresses in the three VF layers, namely, σ_{muc}, σ_{lig}, and σ_{mus} are estimated utilizing experimentally motivated constitutive relations. By means of weighted averages, the stresses in the body and cover layers, σ_b and σ_c are then estimated. After that, the stiffnesses of the cover and body masses are estimated from the stresses and the passive mechanical properties. For example, the body stiffness is given by the relation using the relations

$$k_b = 2\mu_b(L_{\mathrm{vf}}T_{\mathrm{vf}}/D_b)T_{\mathrm{vf}} + \pi^2\sigma_b(D_b/L_{\mathrm{vf}})T_{\mathrm{vf}},$$

where μ_b denotes the shear modulus in the body layer. Note that the stiffness above is a linear combination of a passive (stress -independent) term and a stress-dependent one. The passive term is deduced by adopting the following rationale. Consider a linearly elastic element (rectangular cuboid) with dimensions l, t, and d, and Young's modulus along the d direction 2μ, where μ is the shear modulus, undergoing uniform stress σ in the d direction. The corresponding displacement is assumed to be linear of the form $u_x = xu_0/d$, where x is the position along the d length. Then, in terms of the displacement, the stress is given by

$$\sigma = 2\mu\frac{\mathrm{d}u_x}{\mathrm{d}x} = 2\mu\frac{u_0}{d}.$$

The corresponding elastic force, considering the cross-sectional area $A = l \times t$, is then

$$F = \sigma A = 2\mu \times l \times t\frac{u_0}{d}.$$

Hence, the stiffness of the line element is computed as

$$K = F/u_0 = 2\mu \times l \times t/d.$$

The stress-dependent stiffness term is derived by relating the fundamental frequency of a harmonic oscillator with mass M and stiffness K to the fundamental frequency of a

corresponding vibrating string with length l, width d, height t, stress σ, and density ρ. That is,

$$\frac{1}{2\pi}\sqrt{\frac{K}{M}} = \frac{1}{2l}\sqrt{\frac{\sigma}{\rho}} \Rightarrow K = \left(\frac{\pi}{l}\right)^2 \frac{M}{\rho},$$

but M can be written as $M = \rho l t d$; hence,

$$K = \pi^2 \rho \times d \times t/l.$$

Let us now explain generally the influence of the muscle activation parameters on the lumped parameters of the BCM according to the muscle activation rules [164]. First, and as can be seen from Equation (2.8), the neutral gap of the BCM, modulated by the resting position $x_{0,2}$, depends on the activation of the LCA muscle, where increasing ν_{lca} leads to decreasing the neutral gap (i.e., VF adduction). Next, we plot the BCM masses and stiffnesses as functions of ν_{ta} and ν_{ct}, where we set $\nu_{\text{lca}} = 0.5$. Figure 2.16 shows that the BCM masses depend primarily on the activation of the TA muscle, where increasing ν_{ta} leads to increasing the lower cover mass m_1 and the body mass m_b, and decreasing the upper cover mass m_2. In addition, Figure 2.17 illustrates the complex interaction between the BCM stiffnesses and the activation of the TA and CT muscles. In the case of the stiffness of the lower cover mass k_1, the stiffness is primarily modulated by the CT muscle activation, where increasing ν_{ct} leads to a rise in k_1. In the case of the stiffness of the upper cover mass k_2, increasing ν_{ta} leads to a drop in the stiffness, whereas increasing ν_{ct} leads to a rise in the stiffness. Interestingly, the interaction relation between the body stiffness k_b with muscle activation is quite complex, where increasing the activation of one of the muscles can produce either a rise or fall in stiffness, depending on the location of $(\nu_{\text{ta}}, \nu_{\text{ct}})$ in the parameter space. Finally, in the case of the coupling stiffness k_c, Figure 2.17 shows that increasing the CT muscle activation leads to a rise in the stiffness, whereas increasing the TA muscle activation leads to a rise in the stiffness followed by a drop.

As seen from the discussion above, the rules of Titze and Story [164] contains a significant number of crude approximations and assumptions, in addition to several adhoc relations. Despite that, this model is capable to capture, at least qualitatively, the effects of laryngeal muscle activation on the biomechanics of phonation, and in particular the influence of CT and TA muscle activation on phonation fundamental frequency. This motivates adopting these rules in our investigations in Chapters 3 and 4.

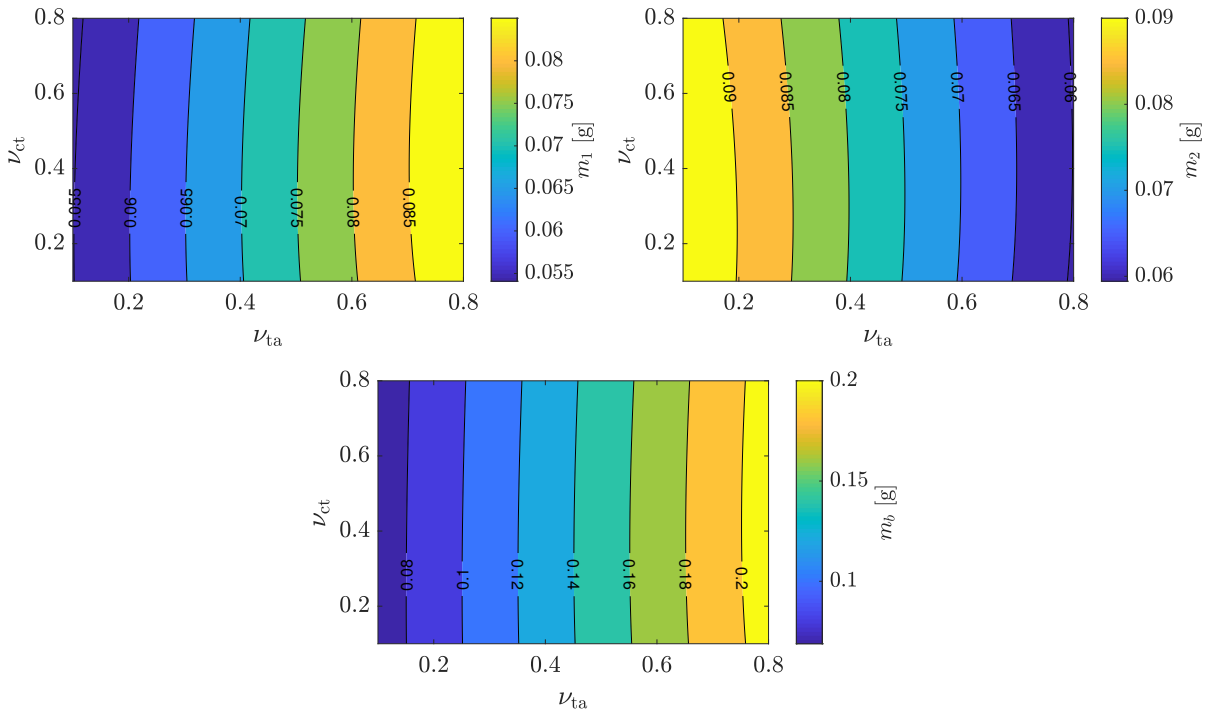


Figure 2.16: Contour plots of the BCM masses as functions of CT and TA muscle activation, where $\nu_{\text{lca}} = 0.5$.

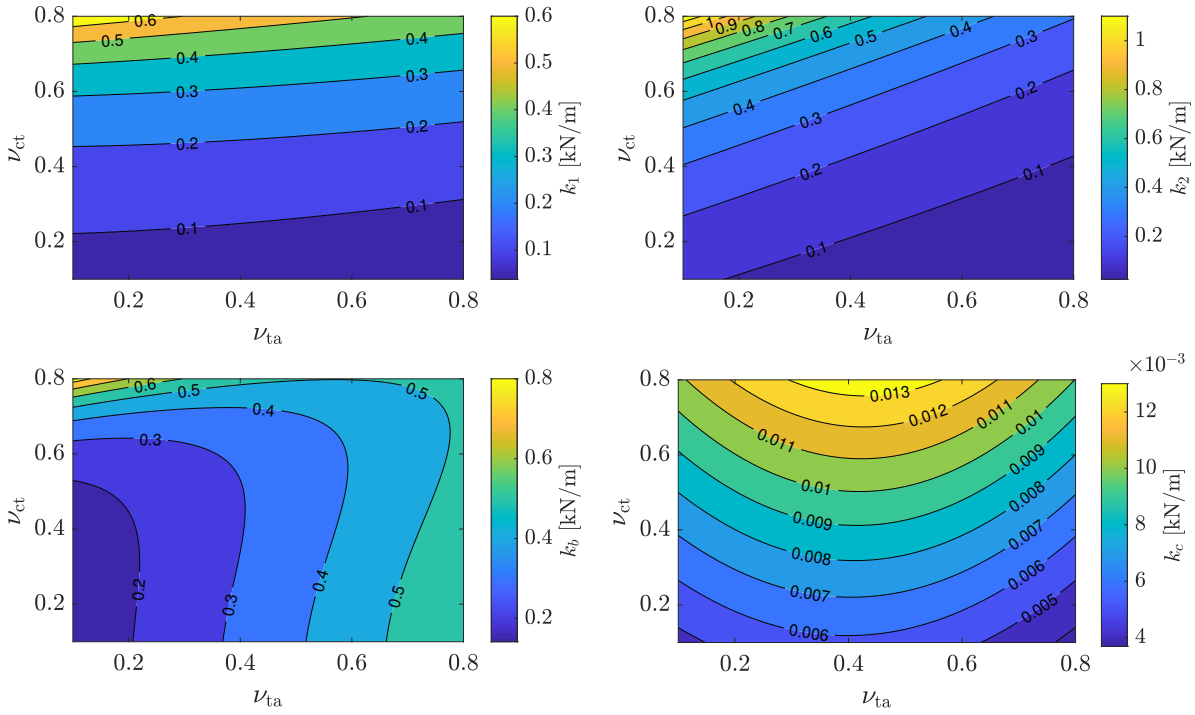


Figure 2.17: Contour plots of the BCM stiffnesses as functions of CT and TA muscle activation, where $\nu_{\text{lca}} = 0.5$.

2.6.3 Some modeling variations

In this section, we introduce some modeling variations that improve upon the BCM and the muscle activation rules.

Posterior glottal opening

Posterior glottal opening is a laryngeal phenomenon, wherein, the VFs are not fully closed posteriorly, resulting in air leakage during phonation. This phenomenon is often linked to voice disorders, including nonphonotraumatic vocal hyperfunction [48, 99]. In an effort to study the effect of posterior glottal opening on phonation, Zanartu et al. [186] introduced a modification to the BCM, adding a resting area to the glottal area, resulting in non-zero glottal flow during the closing portion of the phonation cycle. Incorporating this additional area requires modification to the flow equations, which are discussed thoroughly in [186]. Zanartu et al. [186] showed that incorporating posterior glottal area results in lower phonation efficiency, which necessitates increasing sublglottal pressure to compensate for the deficiency, leading to higher VF contact forces. The authors speculated that these higher contact forces may lead to VF trauma.

Triangular body-cover model

To capture more accurately the triangular glottal shape, the BCM (with posterior glottal opening) can be slightly modified, wherein the BCM masses are slightly tilted in a triangular configuration, resulting in the triangular body-cover model (TBCM), which was proposed by Galidno et al. [32] to have a more refined analysis of the influence of posterior glottal opening on the mechanics of phonation. The tilting of the BCM masses will result in changes in the glottal flow dynamics and yield a gradual change in VF contact forces, wherein contact pressure spans an increasing surface area of the cover masses as they continue penetrating the contact (medial) plane. However, the kinematics of the TBCM are essentially similar to those of the BCM; therefore, we omit detailed discussion of the TBCM equations. We adopt the TBCM in our investigation in Chapter 5.

Posturing model

As can be seen in Section 2.6.2, the muscle activation rules proposed in [164] depend heavily on crude simplifying assumptions and adhoc relations. In addition, these rules

lack the modeling capability of synergetically activating the LCA and PCA muscles or activating the IA muscle; hence, these rules need further refinement. Titze and Hunter [160] have proposed a two dimensional posturing model that relates VF elongation and configuration to laryngeal muscle activation in a physiologically more relevant manner, wherein the activation of the five intrinsic laryngeal muscles is related to the rotational and linear displacements of the cricothyroid joint and the arytenoid cartilages.

Define the index set

$$\mathcal{J} = \{\text{muc}, \text{lig}, \text{ta}, \text{lca}, \text{ia}, \text{pca}, \text{ct}\},$$

where muc, lig, ta, lca, ia, pca, ct refer to the mucosa, ligament, TA muscle, LCA muscle, IA muscle, PCA muscle, and the CT muscle, respectively. Consider the VF plane depicted in Figure 2.18, where an (x, y) coordinate system is established with an origin positioned on the medial plane such that the y axis is pointing anteriorly and the x axis is pointing medially. The posturing model in [160] assumes the symmetry with respect to the medial plane; hence, the cartilages and tissues on one side of the medial plane (right side) are considered. An important component in the posturing model is the vocal process, the anterior angle in each arytenoid, which serves as the posterior points of attachment for the VFs. Consider the right arytenoid, where, when at a cadaver state, the x axis passes by its vocal process. In other words, the cadaveric position of the right VF process is $(0, \bar{x})$. The cricoarytenoid joint which connects the arytenoid to the cricoid is another essential element of the posturing model, with cadaveric position $(x_{\text{caj}}, y_{\text{caj}})$, which is assumed to pass by the geometrical center of the arytenoid cartilage in cadaveric state. Let ξ, ψ, and θ denote the horizontal, vertical, and rotational displacements of the right arytenoid cartilage; hence, the position of the arytenoid cartilage after muscle coactivation is $(x_{\text{caj}} + \xi, y_{\text{caj}} + \psi)$. The governing equations (force and moment balances) driving the dynamics of the right arytenoid cartilage are

$$\begin{aligned}
M_{\text{ac}}\ddot{\xi} + d_x\dot{\xi} + k_x\xi &= \sum_{i\in\mathcal{J}} \alpha_i F_i, \\
M_{\text{ac}}\ddot{\psi} + d_y\dot{\psi} + k_y\psi &= \sum_{i\in\mathcal{J}} \beta_i F_i, \\
I_{\text{ac}}\ddot{\theta} + d_r\ddot{\theta} + k_r\theta &= \sum_{i\in\mathcal{J}} \gamma_i F_i,
\end{aligned}$$

where M_{ac} is the arytenoid mass, I_{ac} is the arytenoid mass moment of inertia, F_i denote the force generated by the i^{th} tissue from the index set J, α_i and β_i denote the associated directional cosines, γ_i denotes the associated directional moment arm, d_x, d_y and d_r are damping coefficients, and k_x, k_y, and k_r are stiffness coefficients.

The relative motion between the thyroid and cricoid cartilages with respect to the CT joint are modeled in a similar fashion, wherein the relative displacement decomposed

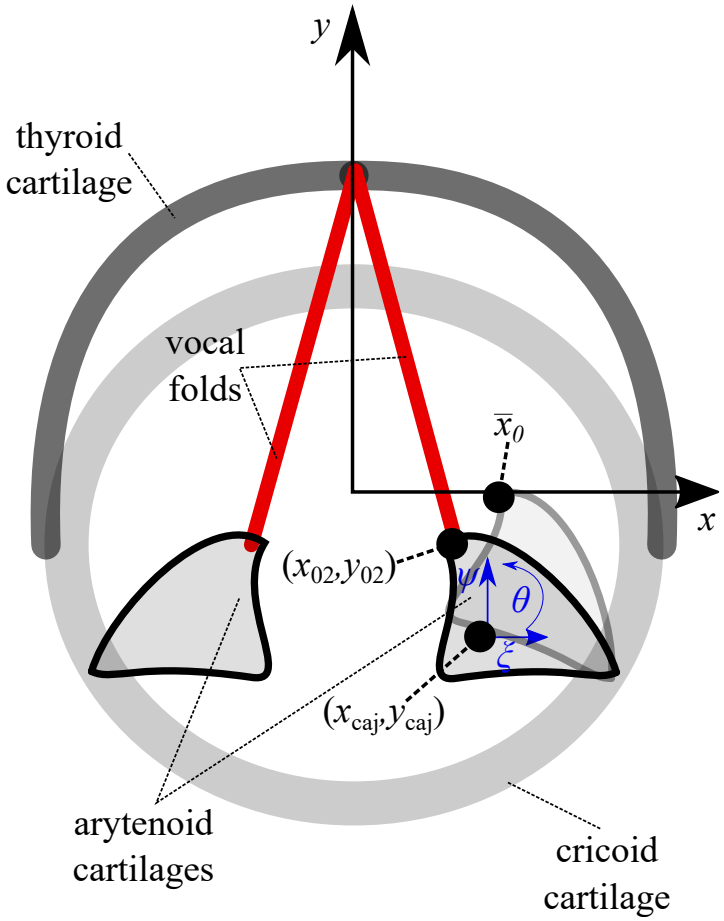


Figure 2.18: Superior schematic view of the glottis with the vocal folds and arytenoid cartilages

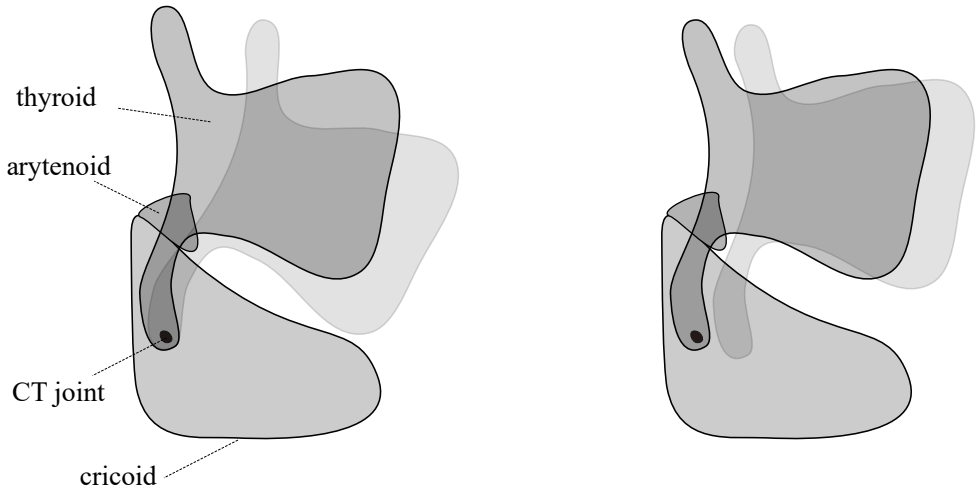


Figure 2.19: Rotation (left) and translation (right) of the cricothyroid joint.

into a transnational component along the VF plane with associated displacement φ, and a rotational component with associated angular displacement ϕ, see Figure 2.19. When modelling the dynamics of the CT joint movement, the movement of the arytenoid with respect to the cricoid is assumed to be negligible (i.e., the cricoid and the mounted arytenoid are considered as one unit). The associated governing equations are

$$M_{\mathrm{ctj}}\ddot{\varphi} + \kappa_t(t_t\dot{\varphi} + \varphi) = c_{\mathrm{ct}}F_{\mathrm{ct}} - (F_{\mathrm{ta}} + F_{\mathrm{lig}} + F_{\mathrm{muc}}),$$
$$M_{\mathrm{ctj}}\ddot{\phi} + \kappa_r(t_r\dot{\phi} + \phi) = r_{\mathrm{ct}}F_{\mathrm{ct}} - r_{\mathrm{ta}}(F_{\mathrm{ta}} + F_{\mathrm{lig}} + F_{\mathrm{muc}}),$$

where M_{ctj} and I_{ctj} are the cricothyroid joint mass and mass moment of inertia, respectively, cs_{ct} is the direction associated with the CT muscle force with respect to the translational movement of the CT joint, r_{ct} and r_{ta} are directional moment arms associated with the rotation about the CT joint, κ_t and κ_r are stiffness coefficients, and t_t and t_r are time constants associated with the viscous response of the CT joint. The displacement due to

rotation about the CT joint is approximately given by

$$\Delta L_r \approx r_{\text{ta}}\phi.$$

Moreover, the displacement due to translation along the VF plane is

$$\Delta L_t = \varphi.$$

Assuming the rigidity of the arytenoid cartilage and accounting for the displacements of the cricoid-arytenoid complex, the position of the vocal process $(x_{0,2}, y_{0,2})$ is then estimated as

$$\begin{aligned} x_{0,2} &= x_{\text{caj}} - (x_{\text{caj}} - \bar{x}_0)\cos(\theta) + y_{\text{caj}}\sin(\theta) + \xi, \\ y_{0,2} &= y_{\text{caj}}(1 - \cos(\theta)) + y_{\text{caj}}\sin(\theta) - (x_{\text{caj}} - \bar{x}_0)\sin(\theta) + \psi - (\Delta l_t + \Delta l_r). \end{aligned}$$

The VF elongation due to adduction is

$$\Delta l_a = -(y_{\text{caj}}(1 - \cos(\theta)) + y_{\text{caj}}\sin(\theta) - (x_{\text{caj}} - \bar{x}_0)\sin(\theta) + \psi).$$

The VF strain due to adduction is

$$\epsilon_a = \Delta L_a / L_{\text{vf},0}.$$

Similarly, the VF strains due to the translational and rotational displacements of the cricothyroid joint are

$$\epsilon_r = \Delta L_r / L_{\text{vf},0}$$

and

$$\epsilon_t = \Delta L_t / L_{\text{vf},0}.$$

The total VF strain ϵ is assumed to be a linear combination of the aforementioned strain. In other words,

$$\epsilon = \epsilon_a + \epsilon_t + \epsilon_r,$$

and that strain is identical to those in the mucosa, ligament, and TA muscle.

By considering the motion of the cricothyroid joint, the strain in the CT muscle can be estimated as

$$\epsilon_{\text{ct}} = -\frac{L_{\text{vf},0}}{L_c}\left(\epsilon_r r_{\text{ct}}/r_{\text{ta}} + \epsilon_t/c_{\text{ct}}\right),$$

where L_c is the resting length of the CT muscle. Unfortunately, measurements or estimates of strains in the LCA, IA , and PCA muscles are not available; therefore, and as noted by

the first author of [4] through verbal communication, the strains in the remaining muscles are given by adhoc relations of the form

$$\epsilon_i = g_i\theta, \;\; i \in \{\text{lca}, \text{ia}, \text{pca}\}.$$

The forces $F_i, i \in \mathcal{J}$, are computed according the the equation

$$F_i = A_i\sigma_i,$$

where A_i denotes the tissue cross-sectional area, and σ_i is the tissue stress. The stresses are computed according to the modified Kelvin constituive model, given by

$$t_s\dot{\sigma}_i + \sigma_i = \sigma_{a,i} + \sigma_{p,i} + Et_o\dot{\epsilon}_i,$$

where t_s and t_0 are time constants, $\sigma_{a,i}$ denotes the active stress in the i^{ctj} tissue/muscle, and $\sigma_{p,i}$ denotes the passive stress. The active stress is given by

$$t_a\dot{\sigma}_{a,i} + \sigma_i = a_i\sigma_m \max\{0, 1 - b(\epsilon_i - \epsilon_m)^2\},$$

where t_a is a time constant, b and ϵ_m are model parameters, σ_m is the maximum active stress, and a_i is a non-dimensional parameter between 0 and 1 that corresponds to the relative muscle activation levels. Note that for the ligament and mucosa, which do not consist of muscle tissues, the active stress is simply zero. In actual implementation, and due to limited measurements, the time constant t_0, was set to zero for all tissues, as communicated by the first author of [4]. The passive stress is given by a piecewise relation in the form

$$\sigma_{p,i} = \begin{cases} -\frac{\sigma_0}{\epsilon_1}(\epsilon_i - \epsilon_1), & \epsilon_i \le \epsilon_2, \\ -\frac{\sigma_0}{\epsilon_1}(\epsilon_i - \epsilon_1) + \sigma_2(\exp(B(\epsilon_i - \epsilon_2)) - 1 - B(\epsilon_i - \epsilon_2)), & \epsilon_i > \epsilon_2, \end{cases}$$

where $\sigma_0, \sigma_2, \; \epsilon_1, \; \epsilon_2, B$ are model parameters determined from stress-strain measurements.

2.6.4 Other phonation models

Despite the practicality of the BCM and its lumped-element variations in simulating various phonation scenarios, there is still a difficulty in establishing a connection between the parameters of such models and measured properties of real VFs. This flaw renders constructing subject-specific models (see, e.g., [180]), that can be used to further investigate

the phonation mechanisms in individuals, difficult. Therefore, there have been various attempts to develop one-dimensional (2-D) and three-dimensional (3-D) continuum models of phonation that can be easily related to physiological properties of the VFs (see, e.g., [1, 165, 194]). Moreover, there have been several complementing efforts to derive constitutive VF relations that can be utilized to develop more accurate continuum models [188, 193]. We note here that numerical simulations using continuum models are computationally costly and, therefore, the applicability of such models is limited.

In a parallel route, there have been several efforts to develop accurate models of the glottal flow. The dynamics of glottal flow are very complex due to the repetitive obstruction of the glottis during phonation, resulting in a discontinuity in the flow domain, and the large difference between the glottal and supra-glottal areas, which induce flow separation and turbulence. The simplest glottal flow model is the 1-D quasi-steady Bernoulli model which has been adopted in the 1-DOF model of Titze [155], the two mass model [58], and the BCM [142]. Despite its extreme simplicity, which dismisses the complex three dimensional nature of glottal flow, this model is capable, when integrated with a mechanical model of the VFs, of simulating the essential characteristics of phonation. Besides the Bernoulli model, several models have been proposed to simulate the glottal flow, including 1-D transient flow models [73, 70, 14], 2-D potential flow models [55, 65], and averaged and full Navier-Stokes models [22, 95]. Interestingly, when compared with the more complex 2-D Navier-Stokes model [22], the Bernoulli model exhibits reasonable accuracy in capturing the relevant aerodynamic features of glottal flow (e.g., pressure distribution over VFs) which, besides its simplicity, makes it an attractive model to be adopted in phonation simulations. For a comprehensive overview of glottal flow and developed flow models, we refer the readers to [96].

Chapter 3

Physics of Phonation Offset: Towards Understanding Relative Fundamental Frequency Observations

We saw in Sections 2.5.2 and 2.5.3 the utility of (relative) fundamental frequency measurements during transient phonation and their capability of distinguishing between healthy subjects and patients with VH. Despite the attempted efforts discussed in Section 2.5.3, our understanding of fundamental frequency or RFF patterns during phonation offset remains opaque, and that necessitates further extensive work, where the underlying mechanics are isolated and controlled.

This chapter discusses partially the first objective of this thesis, where theoretical and numerical models of phonation offset are developed and analysed to understand the biomechanical and laryngeal mechanisms altering fundamental frequency and RFF. We show how RFF temporal variation during phonation offset is modulated by the change in collision levels, associated with VF abduction, and the activation of laryngeal muscles (specifically CT muscle). The analysis is then carried over to explore different manifestations of PVH (increased subglottal pressure, VF cover mass, and adduction levels) and how they influence the drop rate of offset RFF.

The work presented in this chapter is adapted from a paper published in the Journal of Acoustical Society of America under the title "Physics of Phonation Offset: Towards Understanding Relative Fundamental Frequency Observations" [127], where Mohamed Ahmed Serry is the first author, and Cara E. Stepp and Sean D. Peterson are coauthors. M.A.S. and S.D.P. conceived of the study. M.A.S. developed the model. M.A.S. generated prelim-

inary results and figures. All authors analyzed and interpreted results. M.A.S. wrote the first draft of the article. All authors reviewed and approved the manuscript. This research was supported by the NIDCD of the NIH under awards P50DC015446 and R01DC015570. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

The organization of this chapter is as follows: Section 3.1 is an analytical examination of the driving factors affecting frequency during phonation offset using a single degree-of-freedom impact oscillator; Section 3.2 extends the analysis to a BCM of the vocal folds, enabling exploration of laryngeal factors, such as muscle activation and abduction rate; and the chapter concludes with a summary and discussion of future work in Section 3.3.

3.1 Insights From a Simple Impact Oscillator

We begin the examination of frequency variations during phonation offset via analysis of a simplified one degree-of-freedom impact oscillator. This model embeds the pertinent physical parameters (e.g., mass, stiffness, neutral gap, and collision elasticity) yet is analytically tractable. Impact oscillators arise in a variety of physical systems, including bouncing balls [103], vibratory plows [125], and rolling ships interacting with icebergs [37], and have been extensively studied numerically, experimentally, and analytically [129, 107, 57].

3.1.1 Impact oscillator model

Consider the impact oscillator model shown in Figure 3.1, where M is the mass, K is the body stiffness, $\delta \geq 0$ is the neutral gap (distance from the collision plane to the rest position of the mass), and $k_{\mathbf{col}}$ is the collision stiffness. We assume that the mass is initially at the rest position $\xi(0) = 0$ moving in the positive direction with kinetic energy E_0. Consequently, the initial velocity is given by $v_0 = \sqrt{2E_0/M}$. The kinetic energy E_0 corresponds to the aerodynamic energy transferred from the intra-glottal flow to the VFs. Herein, we implicitly implement a quasi-static assumption as it was observed clinically that VF abduction is slow relative to phonation frequency (the abduction period is about 80 ms, whereas the phonation period is less than 10 ms [25]). The governing equations of the impact oscillator are

$$M\ddot{\xi} + K\xi = 0, \qquad \xi(t) \geq -\delta, \tag{3.1}$$

$$M\ddot{\xi} + K\xi + k_{\mathbf{col}}(\xi + \delta) = 0, \qquad \xi(t) < -\delta. \tag{3.2}$$

To ensure the occurrence of collision after the oscillator crosses the rest position, we impose the condition $0 \leq \tilde{\delta} \leq 1$, where $\tilde{\delta} = \sqrt{K/(2E_0)}\delta$ is the normalized neutral gap.

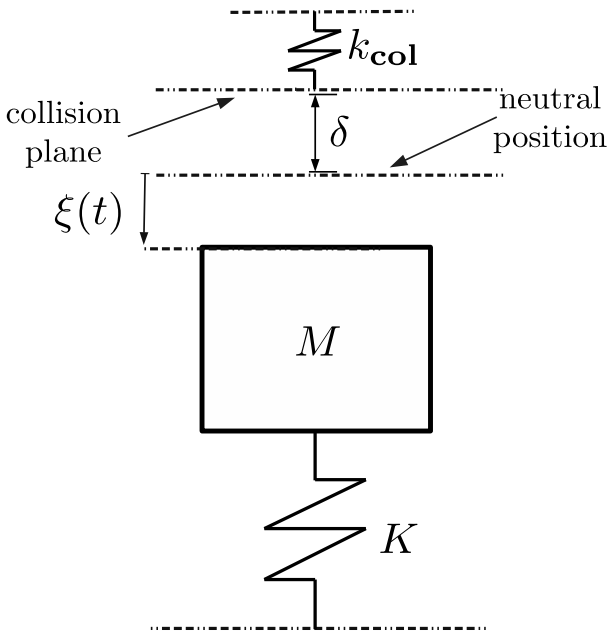


Figure 3.1: Schematic of the impact oscillator system.

3.1.2 Frequency analysis

The motion of the impact oscillator is periodic, as shown schematically in Figure 3.2. During each period, the impact oscillator spends t_c in the collision regime, t_f when $\xi(t) \geq 0$, and $2t_\delta$ between the neutral position and the collision plane, where t_δ is the time required to move from the neutral position to the collision plane (see Figure 3.2). The fundamental frequency of the oscillator is then given by $f = (t_c + t_f + 2t_\delta)^{-1}$. Substituting the closed-form relations of t_c, t_f, and t_δ (see Appendix A.1) yields

$$f = \frac{2f_n}{\frac{2}{\pi}\sqrt{\tilde{k}}\arctan\sqrt{\frac{\frac{1}{\tilde{\delta}^2}-1}{\tilde{k}}} + \frac{2}{\pi}\arcsin\tilde{\delta} + 1}, \tag{3.3}$$

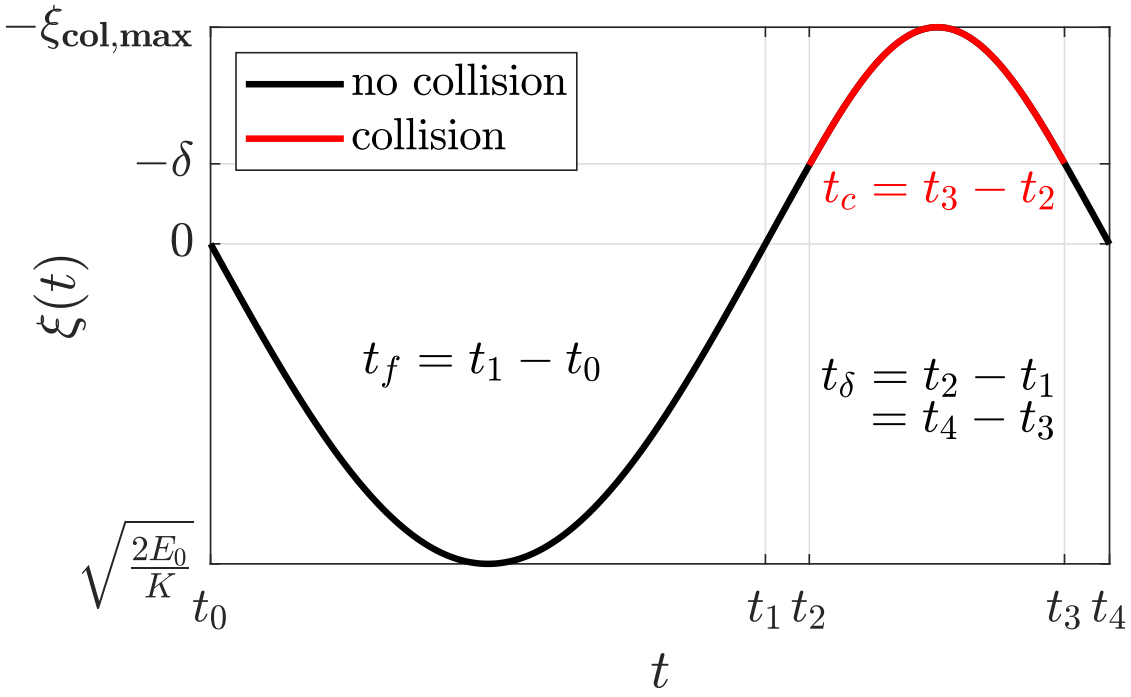


Figure 3.2: Periodic behaviour of the impact oscillator; $\xi_{\mathbf{col,max}} = \delta + \eta_{\mathbf{max}}$, where $\eta_{\mathbf{max}}$ is defined in Appendix A.3.

where $f_n = \sqrt{K/M}/(2\pi)$ is the natural frequency of the oscillator and $\tilde{k} = K/(K + k_{\mathbf{col}})$ is the normalized stiffness.

Note that the stiffness in the collision regime governed by Equation (3.2), $K + k_{\mathbf{col}}$, is larger than the stiffness in the non-collision regime (Equation (3.1)), K, while the mass remains unchanged. Hence, spending more time in the collision phase increases the system frequency. The fundamental frequency satisfies the bounds $f_n \leq f \leq 2f_n$ (see Appendix A.2), which indicates that fundamental frequency of the impact oscillator is larger than or equal to its natural frequency, with a theoretical upper bound being double the natural frequency. The limiting values of f can be inferred by considering the cases of (a) a rigid impact ($k_{\text{col}} \to \infty$) and no gap ($\delta = 0$), which results in a rectified version of the basic oscillator, and hence a fundamental frequency of $2f_n$; and (b) a large gap ($\tilde{\delta} > 1$) and/or infinitely compliant collision ($k_{\text{col}} = 0$), which would result in the basic oscillator without alteration, and hence a fundamental frequency of f_n.

3.1.3 Parameters influencing frequency

In this subsection, we illustrate the role of the body and collision stiffnesses and the neutral gap in altering fundamental frequency. Figure 3.3 shows that the normalized fundamental frequency f/f_n decreases with increasing $\tilde{k}$ and $\tilde{\delta}$. Physically, this indicates that fundamental frequency is proportional to collision stiffness $k_{\mathbf{col}}$ and inversely proportional to the neutral gap δ. The relation between f/f_n and $\tilde{\delta}$ also highlights the role of E_0, corresponding to aerodynamic energy transfer to the VFs, wherein an increase in the energy input leads to higher fundamental frequency. This agrees with the findings of previous studies which showed a positive correlation between fundamental frequency and trans-glottal pressure [156]. Equations (3.1) and (3.2) further state that increasing the body stiffness K increases f, whereas Equation (3.3) and the formula of f_n indicate that increasing the mass decreases the fundamental frequency as $f \propto M^{-0.5}$. Projecting these relations onto clinical data, we conclude the following: During offset, the VFs transition gradually from a collision to non-collision regime where the frequency in the collision regime is generally higher than the frequency in the non-collision regime (natural frequency). This shift explains the drop in fundamental frequency or RFF observed clinically. Moreover, activating different laryngeal muscles during offset changes the mechanical properties of the VFs (stiffness, mass, neutral gap, etc), which consequently alters profiles of fundamental frequency and RFF.

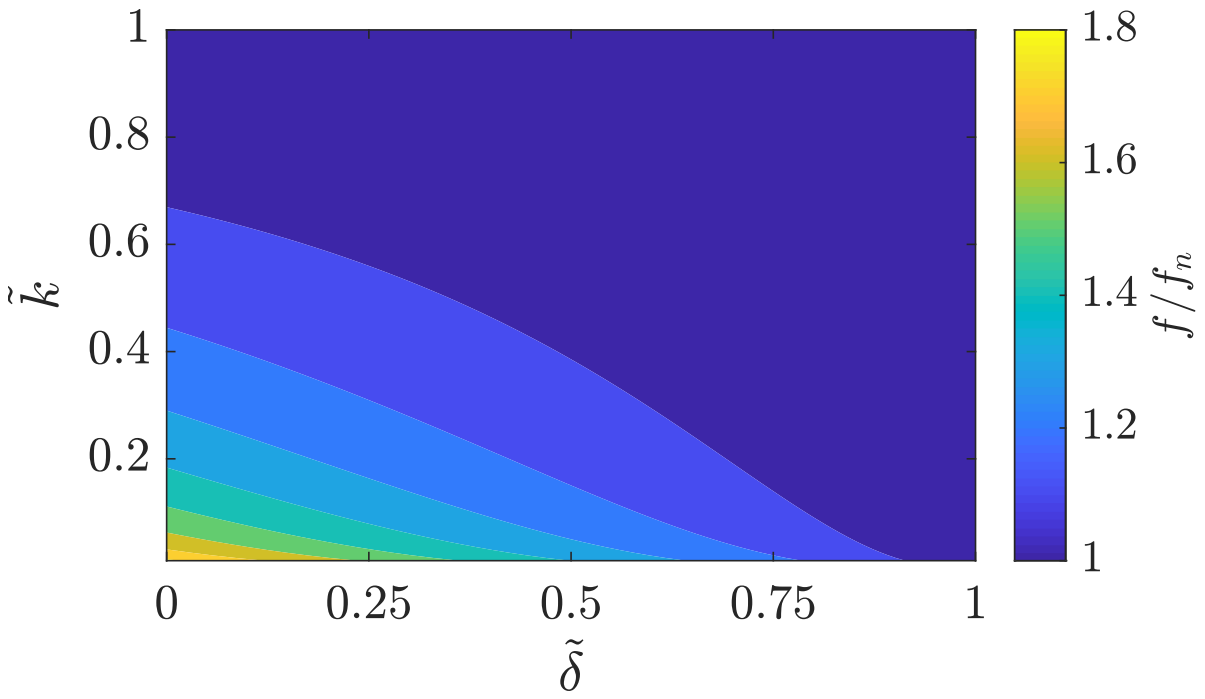


Figure 3.3: Contour plot of f/f_n as a function of the normalized neutral gap $\tilde{\delta}$ and normalized stiffness $\tilde{k}$.

3.1.4 Collision forces and frequency

In this subsection, we illustrate the correlation between fundamental frequency and collision forces. The correlation between collision forces and fundamental frequency was suggested previously by [176] to explain the clinical RFF observations of aged speakers. By direct computations (see Appendix A.3), we obtain the following for the maximum collision force, defined as

$$F_{\mathbf{col,max}} = F_{\mathbf{e,max}}(1-\tilde{k})\left(-\tilde{\delta}+\sqrt{\tilde{\delta}^2+\frac{1-\tilde{\delta}^2}{\tilde{k}}}\right), \tag{3.4}$$

where $F_{\mathbf{e,max}} = \sqrt{2E_0K}$ is the maximum elastic force in the non-collision regime. A contour

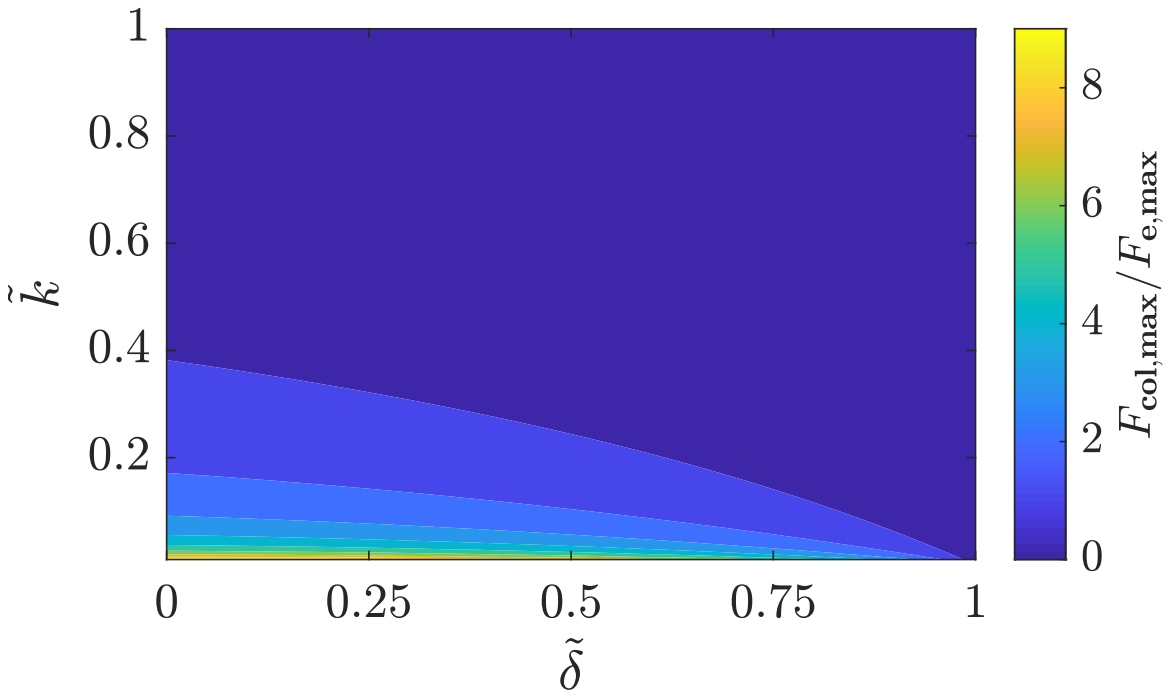


Figure 3.4: Contour plot of $F_{\mathbf{col,max}}/F_{\mathbf{e,max}}$ as a function of the normalized neutral gap $\tilde{\delta}$ and normalized stiffness $\tilde{k}$.

plot of $F_{\mathrm{col,max}}/F_{\mathrm{e,max}}$ versus $\tilde{\delta}$ and $\tilde{k}$ is presented in Figure 3.4. By comparing Figures 3.3 and 3.4, it can be seen that fundamental frequency and maximum collision force are correlated, as the effects of $\tilde{k}$ and $\tilde{\delta}$ on f/f_n are similar to their effects on $F_{\mathbf{col,max}}/F_{\mathbf{e,max}}$.

This suggests that experimental observations of RFF are correlated with the variation in collision forces. However, we note that the fundamental frequency does not depend solely on collision forces. This can be observed, for example, from the fact that the mass affects the fundamental frequency but it does not alter the maximum collision force, as seen in Equations (3.3) and (3.4). The correlation between fundamental frequency and collision forces will be further explored numerically in subsections 3.2.3 and 3.2.5.

3.1.5 Parameters influencing relative frequency

In the previous subsections, we analyzed the different factors altering fundamental frequency. In this subsection, we focus on fundamental frequency relative to the frequency at the fully adducted state, $f|_{\delta=0}$. Physically, this relates to RFF during phonation offset, as it compares the frequency in one state with that at the fully adducted state (e.g., during a sustained vowel). It can be seen from Figure 3.5 that the frequency ratio $f/f|_{\delta=0}$ is

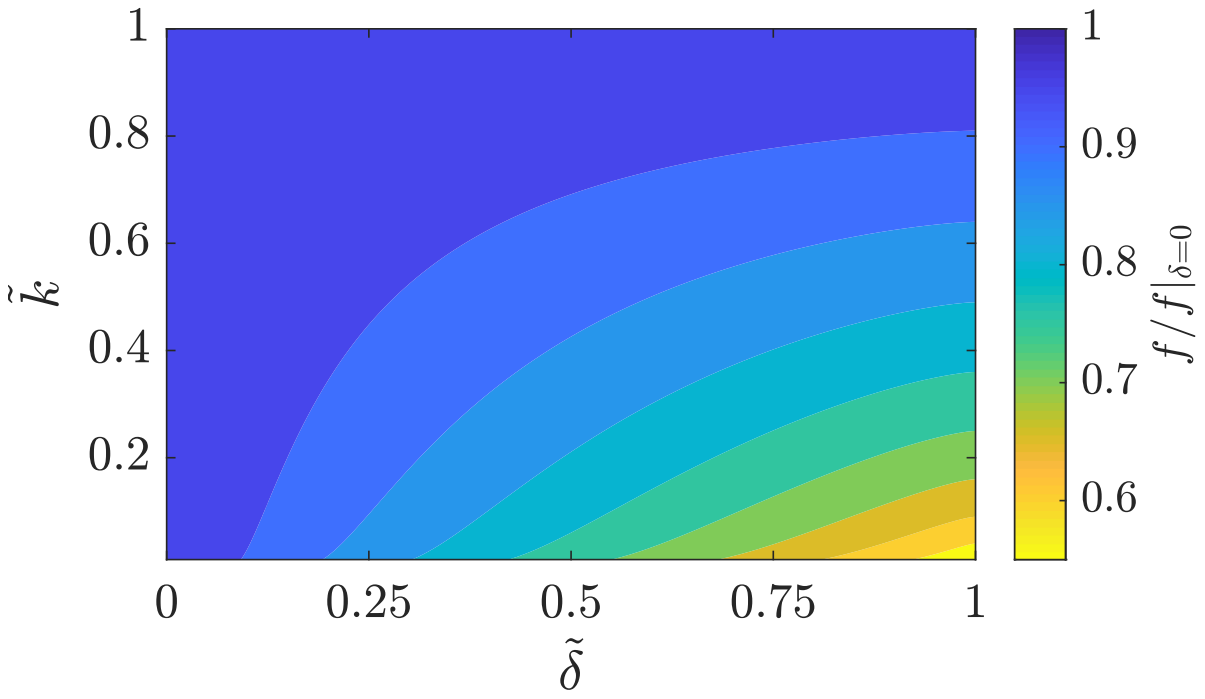


Figure 3.5: Contour plot of frequency ratio $f/f|_{\delta=0}$ as a function of the normalized neutral gap $\tilde{\delta}$ and normalized stiffness $\tilde{k}$.

always less than or equal to one, as the fundamental frequency is maximum at the fully adducted state $\delta = 0$. Moreover, the frequency ratio decreases with increasing $\tilde{\delta}$ due to reduced time in the collision phase as the neutral gap increases. This agrees with clinical data that shows RFF decreasing during offset as the VFs are abducted. Moreover, $f/f|_{\delta=0}$ increases as $\tilde{k}$ increases, which is attained by decreasing $k_{\mathbf{col}}$. For small collision stiffness, frequency is primarily determined by body stiffness K and thus, fundamental frequency is approximately equal to the natural frequency. In the case of real VFs, collision stiffness is significant [142, 135]. Hence, collision forces constitute a major factor altering the fundamental frequency. The role of body stiffness in altering relative frequency is complex as K affects $\tilde{\delta}$ and $\tilde{k}$ in opposing ways. However, by letting the collision stiffness $k_{\mathbf{col}}$ be dependent on the body stiffness, say $k_{\mathbf{col}} = \alpha K$ for some constant $\alpha > 0$ (see, e.g., [142, 135]), then $\tilde{k} = 1/(1+\alpha)$. Hence, the body stiffness does not alter the stiffness ratio $\tilde{k}$ when $k_{\mathbf{col}}$ depends on K linearly. However, the body stiffness K increases $\tilde{\delta}$ and consequently lowers the relative frequency values. This implies the correlation between increased stiffness of the VFs and sharper drops in RFF, supporting the findings from previous experimental studies [91].

3.2 Body-Cover Vocal Fold Simulations

Section 3.1 introduced a simplified theoretical analysis of phonation offset that ignored the physiological aspects of the VFs and their transient nature during such phonation periods. In this section, we adopt a more physiologically relevant model, namely the three-mass BCM [142] and carry out a numerical study exploring a variety of physical parameters, including muscle activation, abduction timing, and vibrating mass. This model is selected for its blend of relative simplicity coupled with demonstrated success in modeling several phonation scenarios [139, 186, 26].

3.2.1 Body-cover model

Figure 3.6 shows a schematic diagram of the BCM used in this study. This model, which embeds the essential physiological components of the VFs employed during modal voice, consists of two cover masses, m_1 and m_2, and a body mass, m_b, all connected via springs and dampers to model tissue viscoelasticity. The resting positions of the masses result in a a nearly rectangular glottal configuration with a neutral glottal gap given by x_0. The model assumes the motion of the VFs to be symmetric about the medial plane. Collision of the opposing folds is modelled by activating additional nonlinear spring forces applied

to the cover masses, where the spring force is proportional to the degree of overlap of the cover masses with the medial (collision) plane. We note here, unlike the implementation of [142], that the damping coefficients are not increased during collision. The model employs muscle activation rules to control the primitive model variables [164], wherein three muscle activation parameters, namely $a_{\mathbf{CT}}$, $a_{\mathbf{TA}}$, and $a_{\mathbf{LCA}}$, account for the activation of the CT, TA, and LCA/PCA muscles, respectively. Moreover, a sub-glottal tract [178] and a supra-glottal tract, configured to simulate the /i/ vowel [150], are included. The acoustics are modeled using wave reflection analog [64, 78, 140], where outward travelling pressure waves are multiplied by attenuation factors to account for losses, similar to implementations employed by [153] and [184]. The intra-glottal flow is simulated using the Bernoulli flow model. The BCM governing equations [142] are solved numerically using an explicit version of Newmark's method [104], which has previously been employed in studies involving discrete-time solutions of the BCM [33, 32, 39]. The numerical scheme can be described briefly as follows: let $\mathbf{x}_i$, $\mathbf{v}_i = \dot{\mathbf{x}}_i$, and $\mathbf{a}_i = \dot{\mathbf{v}}_i$ denote the vectors of estimated displacements, velocities, and accelerations, respectively, of the BCM masses at some time step i, where $\mathbf{a}_i$ is computed from the momentum balance given $\mathbf{x}_i$ and $\mathbf{v}_i$. At the next time step, $i+1$, the displacements and velocities are computed from the kinematic relations $\mathbf{x}_{i+1} = \mathbf{x}_i + \mathbf{v}_i\tau + \mathbf{a}_i(\tau^2/2)$ and $\mathbf{v}_{i+1} = \mathbf{v}_i + \mathbf{a}_i\tau$, where τ is a fixed time step size associated with the sampling frequency, herein set to 70 kHz. A convergence study indicated that steady state oscillation frequency and offset frequency patterns differed by less than 1% in comparison with simulations at a sampling frequency of 350 kHz, and was thus deemed sufficiently resolved for the present study.

The passive model parameters employed, including VF length, correspond to a male speaker [164]. The simulations span 1.1 seconds where all the model parameters are fixed during the first second to ensure that steady state sustained oscillations are achieved. After the initial one second, offset is initiated by varying muscle activation parameters in the remaining 0.1 second. Tables 3.1 and 3.2 list the variables and controlled parameters of interest in this study. Note that the maximum collision forces, $\max(F_{\mathbf{col},i})$, $i = 1, 2$, in Table 3.1 are estimated during the steady state oscillations.

3.2.2 Signal analysis

In this work, we consider the time-series of the glottal area, A_g, instead of acoustic signals in our frequency analysis. We are concerned with fundamental frequency, which is directly obtained from the glottal area waveform without the confounds of higher frequency content present in the acoustic signal. To facilitate instantaneous frequency estimation, we modify the glottal area signals by subtracting the nominal (resting) area $2Lx_0$, where

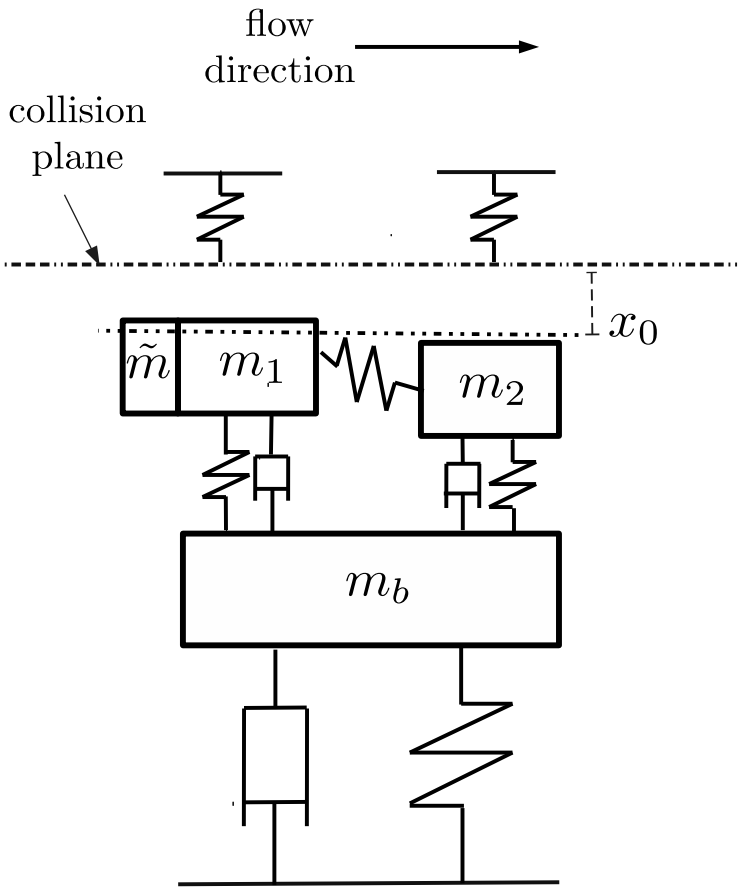


Figure 3.6: Schematic diagram of the body-cover model.

Table 3.1: Variables of interest in the current study

variable	definition
A_g	glottal area
$F_{\mathbf{col},1}$	collision force of lower cover mass
$F_{\mathbf{col},2}$	collision force of upper cover mass
$F_{\mathbf{col,max}}$	$\max\{\max(F_{\mathbf{col},1}), \max(F_{\mathbf{col},2})\}$
f	fundamental frequency
x_0	neutral position
t	time

Table 3.2: Controlled parameters and their default values

parameter	definition	default value (unit)
$a_{\mathbf{CT}}$	CT activation	0.2
$a_{\mathbf{TA}}$	TA activation	0.2
$a_{\mathbf{LCA}}$	LCA/PCA activation	$\in [0.3, 0.5]$
P_s	static sub-glottal pressure	1000 Pa
t_i	abduction initiation time	1020 ms
$T_{\mathbf{abd}}$	abduction period	80 ms

L is the length of the VF. We note here that the resting area is time-varying during offset simulations. Fundamental frequency is estimated from the peaks of the modified signal by identifying the times of two subsequent peaks, say t_{j-1} and t_j, and computing $f(t_j) = (t_j - t_{j-1})^{-1}$. We measure the differences between consecutive peaks and nadirs of the modified signal, which in general decay during phonation offset, and the last offset cycle is determined to be the cycle associated with a peak-to-nadir difference less than a threshold set to be 10% of the peak-to-nadir difference during sustained oscillations (see Figure 3.7). If such a criterion is not met by the end of the simulation time, the last offset cycle is selected to be the last cycle detected in the simulation. Finally, once frequencies are estimated, RFF is calculated using Equation (2.1), where the reference frequency $f_{\mathbf{ref}}$ is set to be the phonation fundamental frequency during the steady state oscillations.

3.2.3 Quasi-static analysis

We begin our numerical study with a quasi-static preliminary analysis to gain deeper insight into the effect of neutral gap, x_0, and collision forces on phonation fundamental fre-

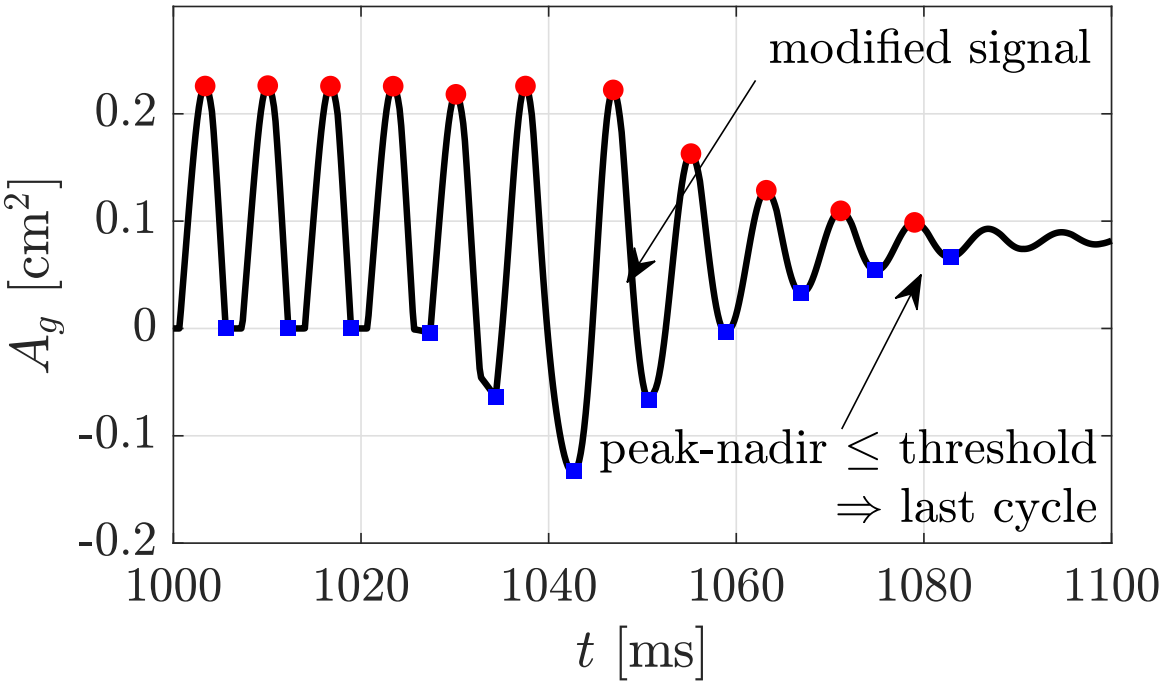


Figure 3.7: Modified glottal area profile with estimated location of the last offset cycle.

quency during sustained oscillations. Specifically, we record the neutral gap, fundamental frequency, and collision forces across several values of $a_{\mathbf{LCA}}$. Figure 3.8 presents the effect

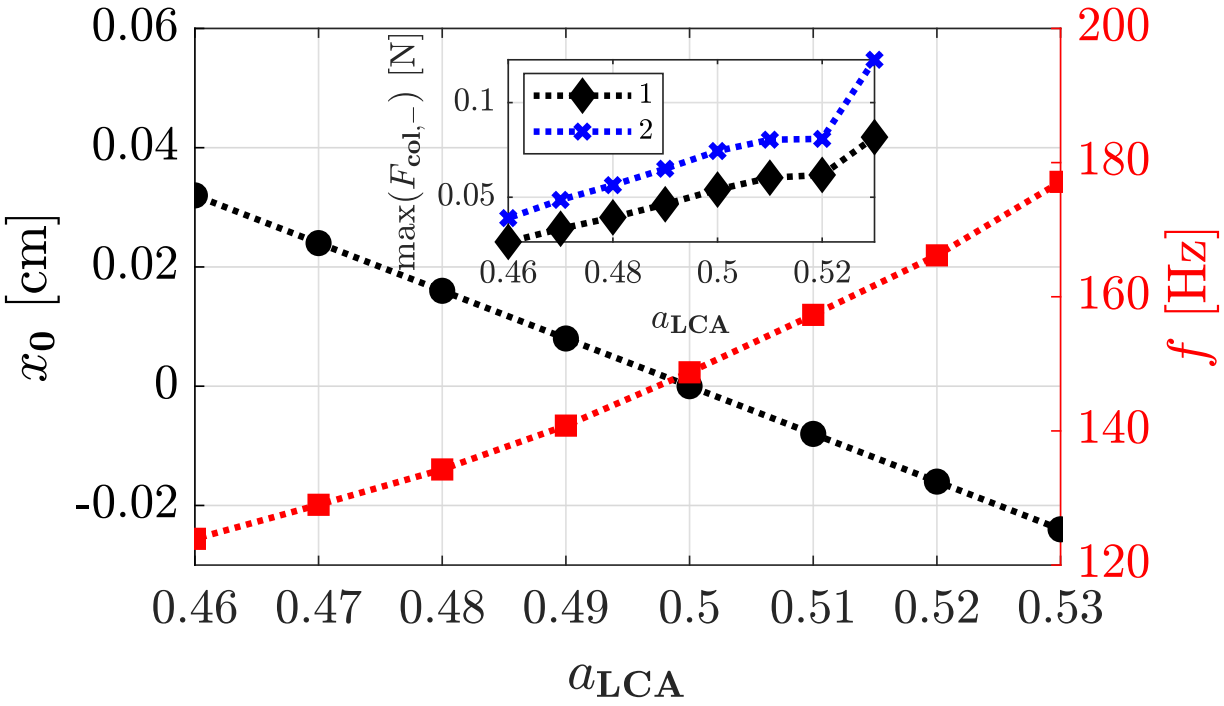


Figure 3.8: Neutral gap and phonation frequency vs Lateral cricoarytenoid activation and the associated maximum collision forces.

of LCA muscle activation on the neutral gap, showing that as LCA activates it adducts the folds, which leads to a higher fundamental frequency in accordance with the findings in Section 3.1. This follows the empirical relation introduced by [164], $x_0 = 0.25L_0(1-2a_{\mathbf{LCA}})$, where L_0 is the un-stretched length of the VFs. Included as an inset in Figure 3.8 are the maximum collision forces for m_1 and m_2 as a function of $a_{\mathbf{LCA}}$. This highlights that the increase in fundamental frequency is correlated with the increase in maximum collision forces. This observation supports our hypothesis and the findings from Section 3.1 that the variation in fundamental frequency, especially during offset, is associated with the change in collision forces. This will be revisited in Section 3.2.5, where transient abduction gestures are considered.

3.2.4 Offset simulations

In order to model offset with the BCM, we employ experimental measurements to calibrate and set an appropriate $a_{\mathbf{LCA}}$ profile. Specifically, we extract the glottal angle, $\theta_{g,\mathbf{exp}}$, as a function of time for a healthy male speaker during the offset portions of repeated /ifi/ utterances[1] [25]. We estimate the glottal area of the experiments, $A_{g,\mathbf{exp}}$, by assuming the glottis to be a circular sector, and thus $A_{g,\mathbf{exp}} = L_{\mathbf{male}}^2 \theta_{g,\mathbf{exp}}/2$, where $L_{\mathbf{male}} = 1.6$ cm is the typical length of a male VF. The LCA muscle activation profile is then tuned to match the experimental glottal area estimates. Reasonable abduction levels are obtained with an $a_{\mathbf{LCA}}$ profile that decreases from 0.5 to 0.3 over an offset period estimated to be 80 ms. This LCA profile will be used in all offset simulations unless otherwise stated. Figure 3.9 shows how the simulated glottal area, based on the tuned LCA profile, varies in time. Included as an inset are the empirical data , which exhibit similar glottal areas and abduction duration.

3.2.5 Exploring the role of collision on RFF

In this subsection, we examine our hypothesis that RFF curve shape is largely driven by changes in the degree of collision. During offset, the time instance corresponding to the cessation of collision is identified as the first period for which the nadir of the original glottal area signal is nonzero. Figure 3.10 presents RFF and collision force for the exemplar case $(a_{\mathbf{CT}}, a_{\mathbf{TA}}) = (0.3, 0.3)$. The RFF curve initially decays while collision forces are non-zero, followed by a minimum value, denoted by $\mathbf{RFF}(f_{\mathbf{min}})$, and subsequent increase in RFF, with last detected value being denoted $\mathbf{RFF}(f_{\mathbf{last}})$, when no collision forces are present. The figure shows correlation between the variation of the collision forces and the RFF profile, especially prior to collision cessation, which supports the main hypothesis of this work. After cessation of collision, the rise in RFF can be attributed to various transient effects including aerodynamics and abduction initiation time relative to phonation cycle. We note that this dip in the RFF profile has been observed experimentally within subjects, as seen in Figure 3.11, but it is unknown whether this rise is associated with the cessation of collision. Furthermore, such a dip has not been commonly found in the literature [176, 136, 45], perhaps because the presented data are typically averaged over and within several subjects, which might mask the transient effects.

[1] Empirical data, obtained under NIH award R01DC015570, were provided by the STEPP lab at Boston University. For details on the data collection methodology, including IRB approval information, see [25].

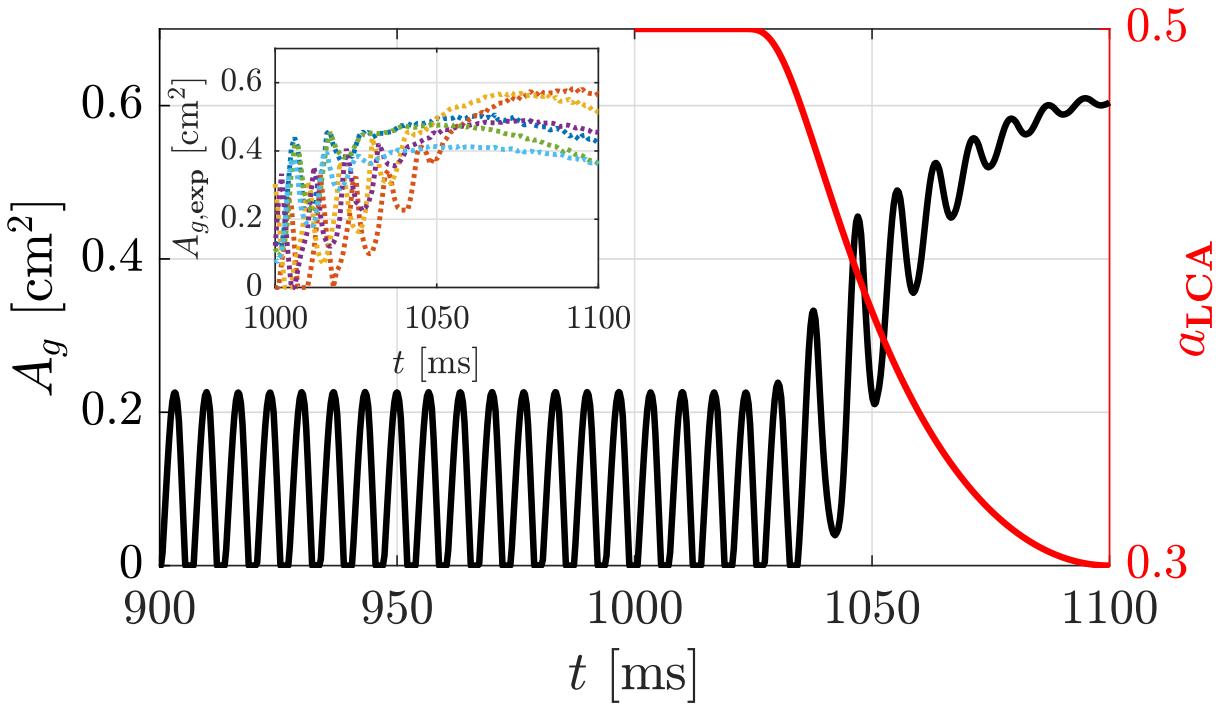


Figure 3.9: Estimated $a_{\mathbf{LCA}}$ profile and simulated glottal area based on empirical data of glottal area during the offset portion of the /ifi/ utterance. The inset shows empirical data extracted from a healthy male subject.

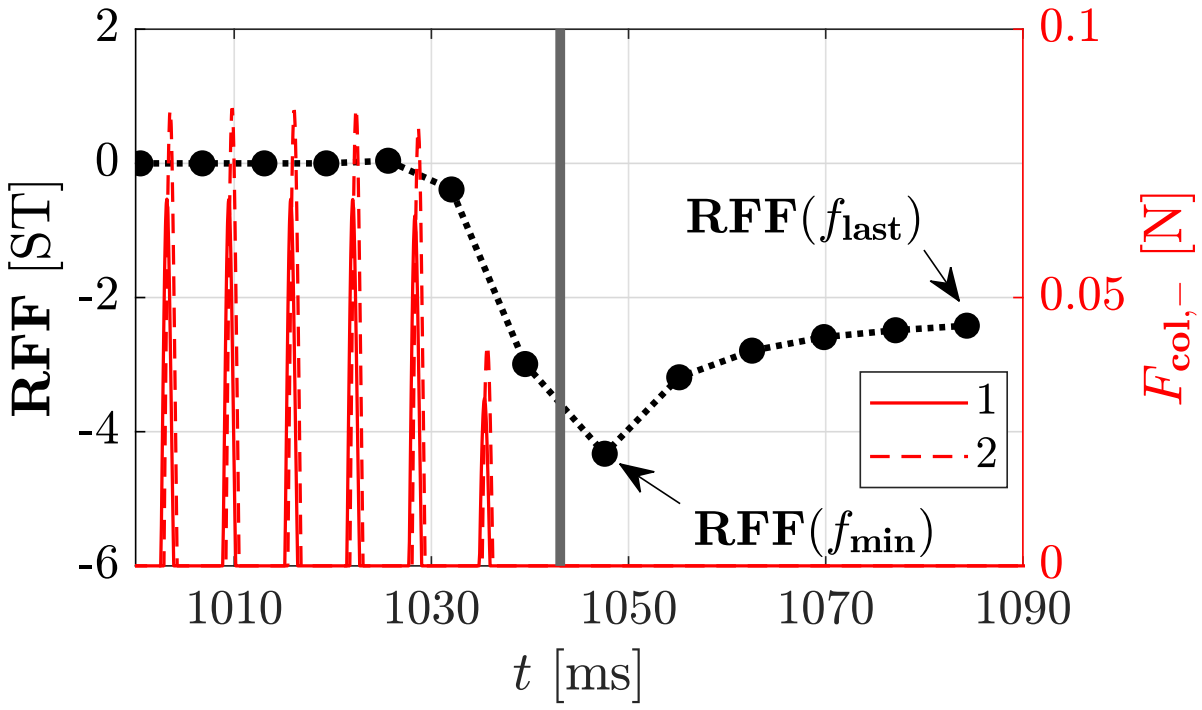


Figure 3.10: Relative fundamental frequency (RFF) and collision forces over time for $(a_{\mathbf{CT}}, a_{\mathbf{TA}}) = (0.3, 0.3)$. Vertical thick line indicates the detection of collision cessation.

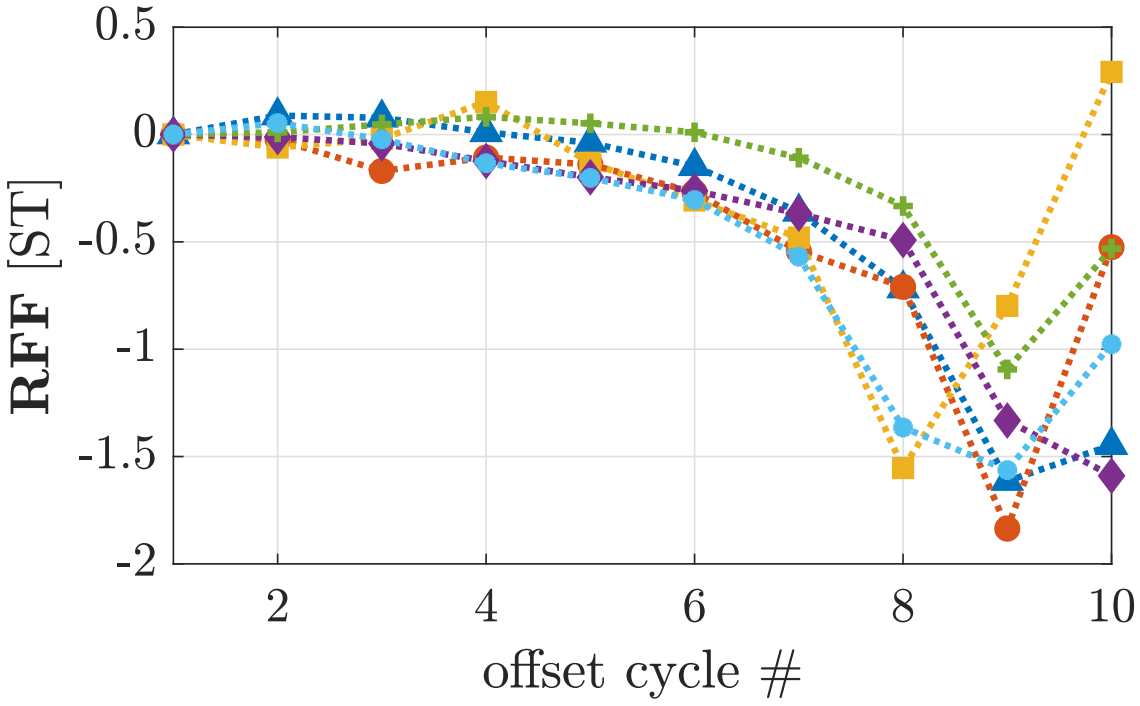


Figure 3.11: Relative fundamental frequency profiles, based on empirical data extracted from one male subject, during the offset portions of repeated /ifi/ utterances.

3.2.6 Abduction initiation time and period: sensitivity analysis

In this subsection, we study two possible sources of variability in RFF observations associated with transient effects, namely abduction initiation time and duration of the abduction period. In this study, the choice of abduction initiation time t_i was arbitrary (t_i = 1020 ms). Varying t_i will initiate abduction at different phases of phonation cycle, as shown in Figure 3.12, which impacts the vibration frequency. Therefore, it is important to analyze the robustness of RFF against these variations as it will contribute to gesture-to-gesture variability in RFF for a single subject, since relative abduction initiation time likely cannot be controlled with precision. Figure 3.13 shows that, in general, RFF profiles are

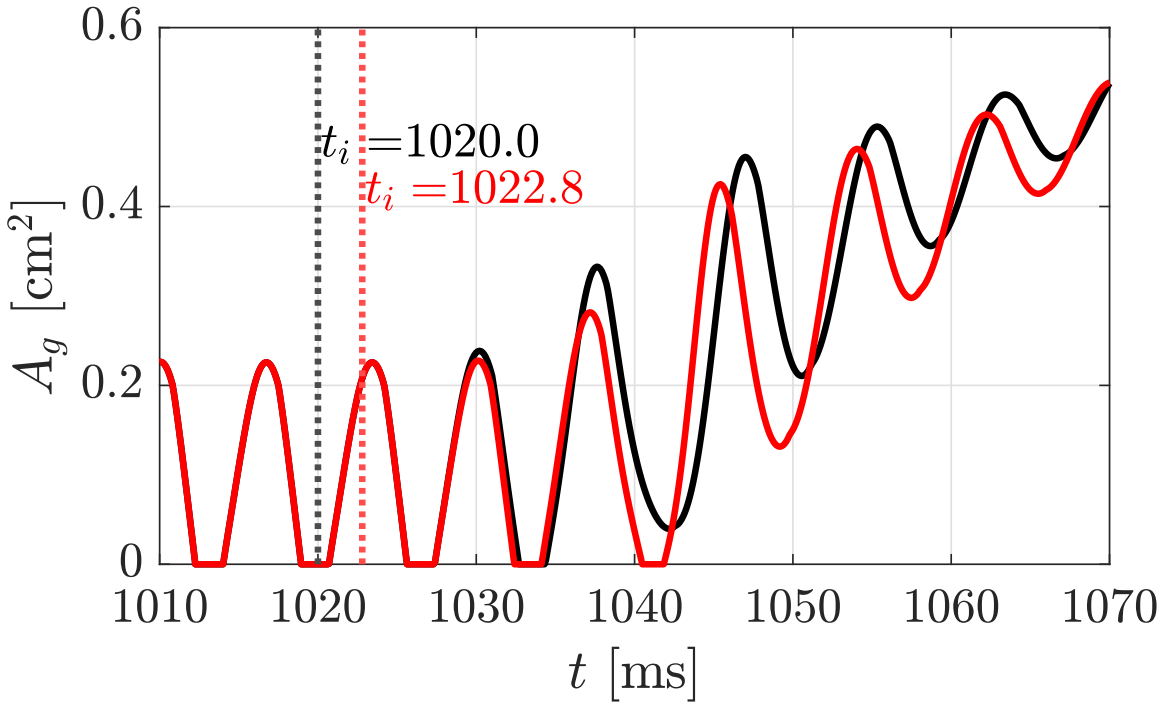


Figure 3.12: Abduction initiation at different stages of phonation cycle.

sensitive to the variation in abduction initiation time with the last offset cycles tending to be less sensitive. This sensitivity is due to the fact that fundamental frequency is affected by decaying transient effects, associated with the state of the VF system at the abduction initiation time, and once these effects vanish, fundamental frequency is equal to the natural frequency of the VFs system. This implies that abduction initiation can contribute to the

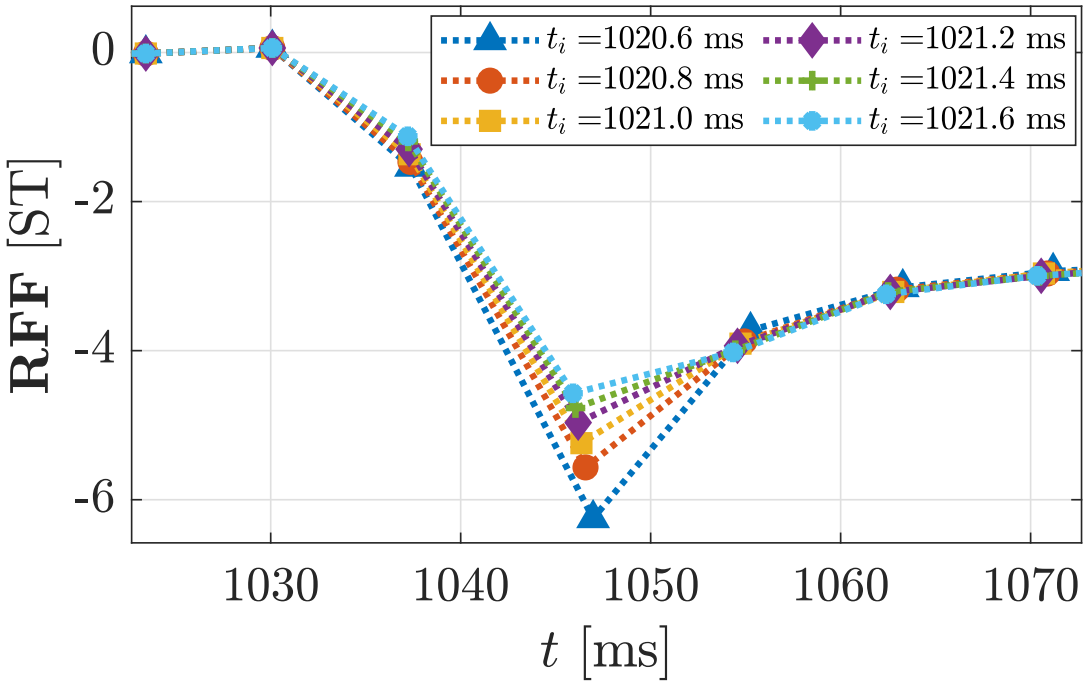


Figure 3.13: Effect of shifting abduction initiation on the relative fundamental frequency profile.

variability of intra-subject RFF experimental observations. Due to the robustness of the last offset cycles against varying abduction initiation time, we consider $\mathbf{RFF}(f_{\mathbf{last}})$ values in the remainder of this paper when studying different laryngeal factors.

In addition to variation in abduction initiation phase, the total duration of abduction may vary from gesture to gesture and subject to subject. In our analysis of empirical data in Section 3.2.4, we estimated the abduction period to be approximately $T_{\mathbf{abd}}$ = 80 ms. Here, we consider different values of the duration of the abduction period $T_{\mathbf{abd}}$, with fixed abduction initiation time. Figure 3.14 illustrates the sensitivity of the RFF profile to the

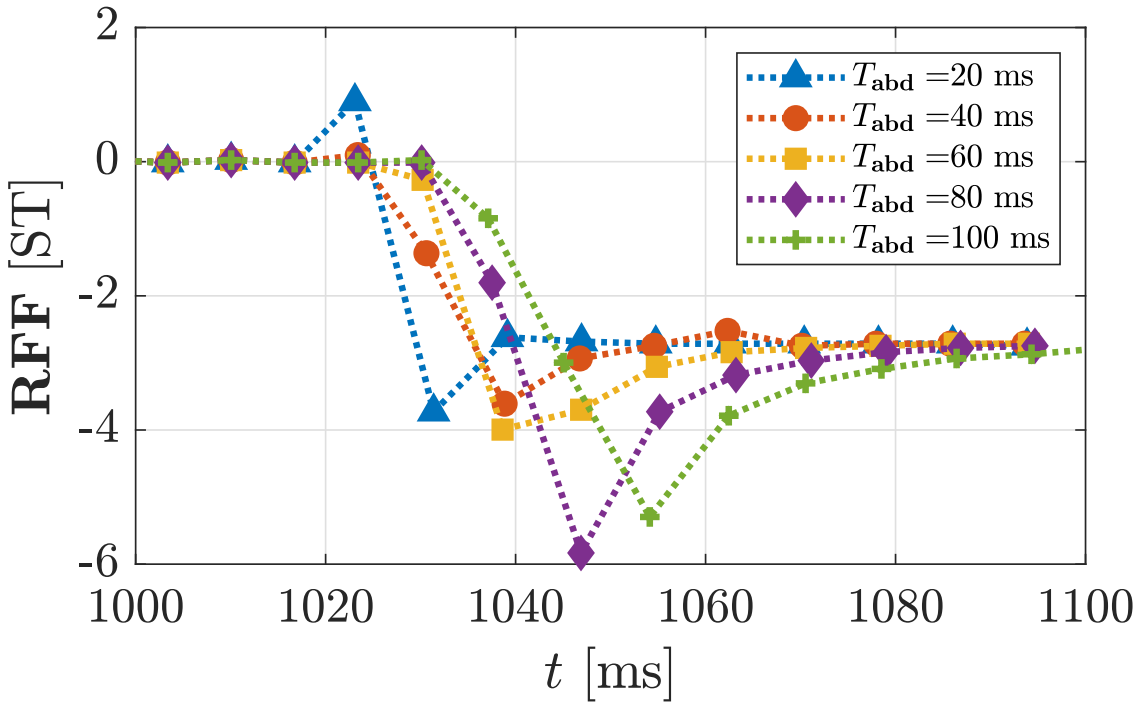


Figure 3.14: Sensitivity of the relative fundamental frequency profile to the duration of the abduction period.

duration of the abduction period. This sensitivity is due to the changed abduction rate and that, in addition to the effect of abduction initiation time, adds to the variability of the RFF observations clinically.

3.2.7 CT muscle as a compensatory mechanism

It is understood that the CT muscle is the primary regulator of frequency during different phonation scenarios [121]. In particular, it has been found in previous studies [164] that increasing the CT muscle activation increases fundamental frequency during sustained phonation. Therefore, the CT muscle can potentially play a role in mitigating the RFF drop during offset. To examine this potential role, we increase the CT muscle activation

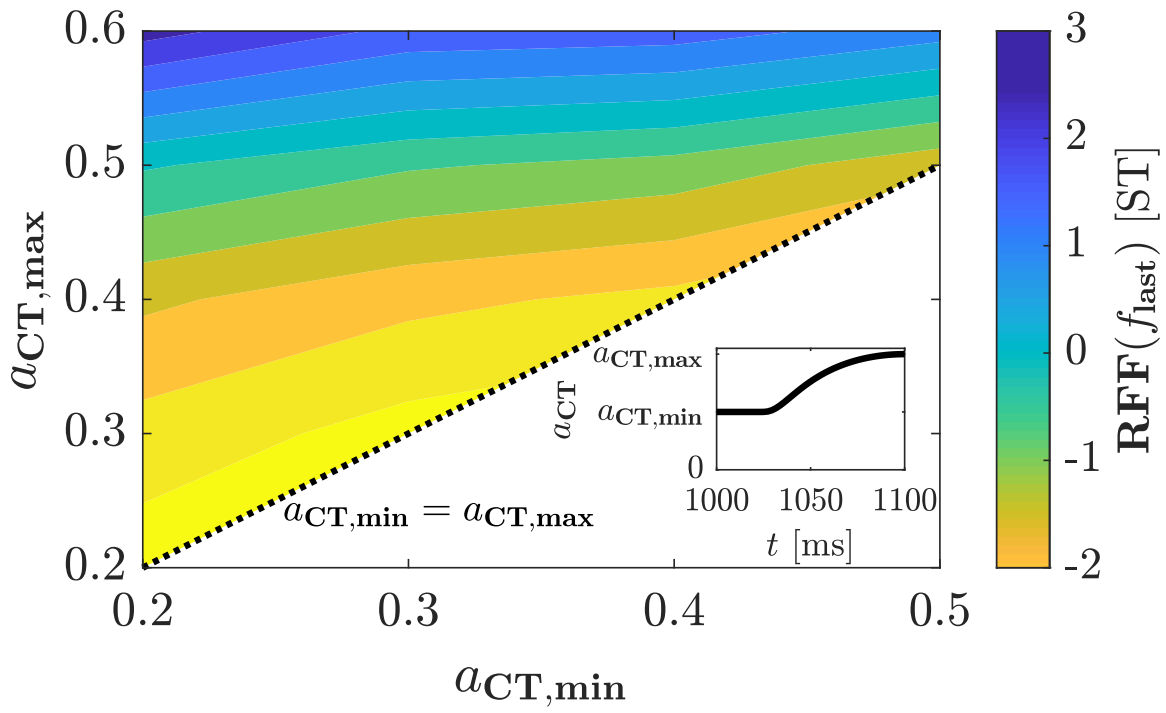


Figure 3.15: Cricothyroid muscle activation increase during abduction and its consequent effects on $\mathbf{RFF}(f_{\mathbf{last}})$.

during offset from a minimum value, $a_{\mathbf{CT,min}}$, to a maximum value, $a_{\mathbf{CT,max}}$, rather than keeping it at a constant value (see the inset of Figure 3.15). Figure 3.15 shows that increasing the CT muscle activation during offset counters, in general, the drop in RFF attributed to reduced collision forces. This supports previous numerical and experimental studies that showed an increase in the CT activation, or the stiffness of VFs, during offset in some cases [85, 59]. This may explain the relatively stable fundamental frequency during phonation offset observed in healthy speakers, should CT muscle activation increase during

offset in this population. Moreover, this potential role of the CT muscle may explain the sharper drop in RFF in the cases of aged and pathological speakers as CT activation in this population may be compromised [106].

3.2.8 Phonotraumatic vocal hyperfunction

PVH refers to the misuse of the VFs during phonation (e.g., excessive laryngeal muscle activation, high sub-glottal pressure, etc.) resulting in benign lesions, such as polyps and nodules [47]. It has been observed that speakers with PVH need to produce relatively high sub-glottal pressures during phonation to reach desired sound pressure levels, resulting in higher collision forces [71, 27].

In this subsection, we explore three different scenarios, associated with PVH, that produce higher collision forces and analyze the corresponding RFF behavior. The first scenario explores the impact of increasing the sub-glottal pressure. Here, we run offset simulations for various values of static sub-glottal pressure P_s and record the resulting $\mathbf{RFF}(f_{\mathbf{last}})$ and the maximum pre-offset collision force $F_{\mathbf{col,max}}$. Figure 3.16 shows that increasing the sub-glottal pressure leads to a sharper drop in RFF, which is correlated with higher pre-offset collision forces. We note, however, that the change in $\mathbf{RFF}(f_{\mathbf{last}})$ and $F_{\mathbf{col,max}}$ is relatively small over the 500 Pa range of P_s explored. We further note that the plateau in the RFF profile observed in Figure 3.16 is due to the choice of the last offset cycle; for slightly longer simulation times, the RFF profile is strictly decreasing.

The second scenario explores increasing the pre-offset LCA muscle activation values. We observed in the previous subsections how the LCA/PCA muscle activation is responsible for adducting/abducting the VFs and subsequently, altering the collision forces. Increasing pre-offset LCA muscle activation values then represents a crude model of laryngeal tension observed in some types of vocal hyperfunction [63]. Specifically, we employ the same LCA muscle activation waveform, but vary the initial value, $a_{\mathbf{LCA,max}}$, over the range of [0.45,0.53]. The final value in the LCA waveform is fixed at 0.3 for all cases, see the inset in Figure 3.17. Figure 3.17 presents $\mathbf{RFF}(f_{\mathbf{last}})$ and $F_{\mathbf{col,max}}$ as functions of $a_{\mathbf{LCA,max}}$. The figure shows how the increased pre-offset LCA muscle activation leads to a sharper drop in RFF, which, as observed previously, is correlated with higher pre-offset maximum collision forces. We note that for the highest value of $a_{\mathbf{LCA,max}}$ the maximum collision force actually decreases slightly, despite the continued drop in RFF, and that can be attributed, in part, to the complex dynamics of the model, which are beyond the scope of this work.

The last scenario we consider is increasing the cover mass, which corresponds to the formation of nodules or polyps [179]. Here, we study the effect of varying the lower cover

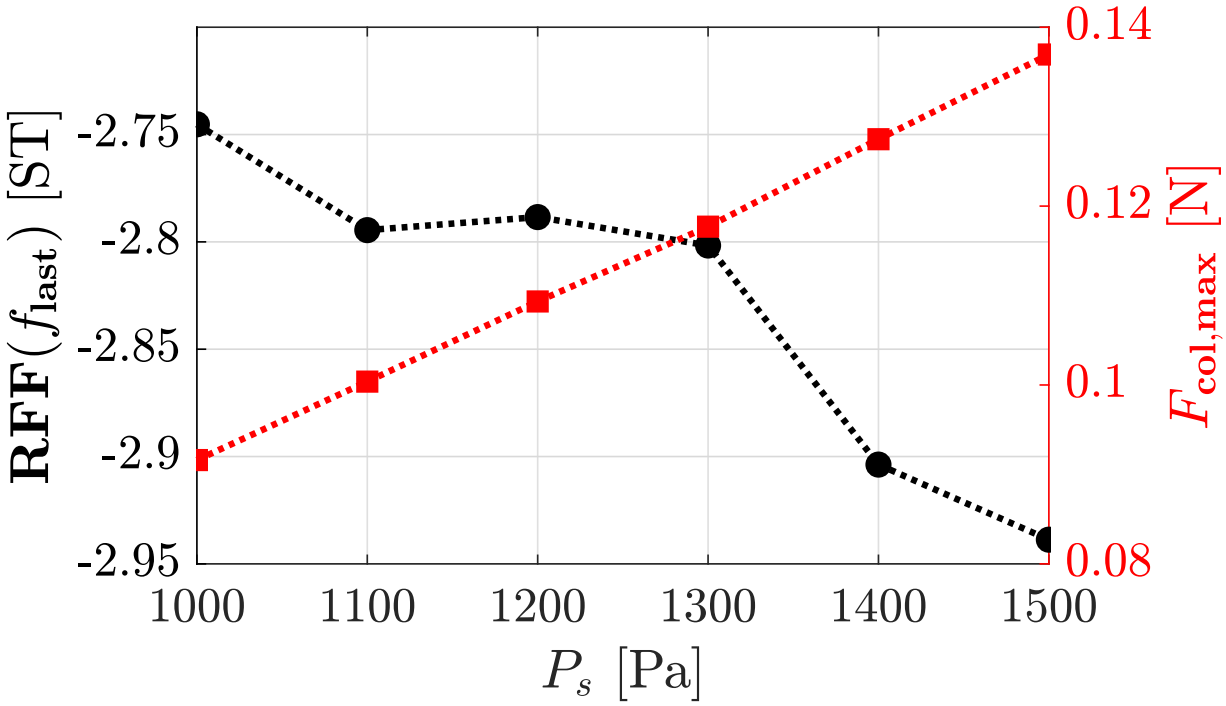


Figure 3.16: Effect of increasing sub-glottal pressure on $\mathbf{RFF}(f_{\mathbf{last}})$ and maximum collision force $F_{\mathbf{col,max}}$.

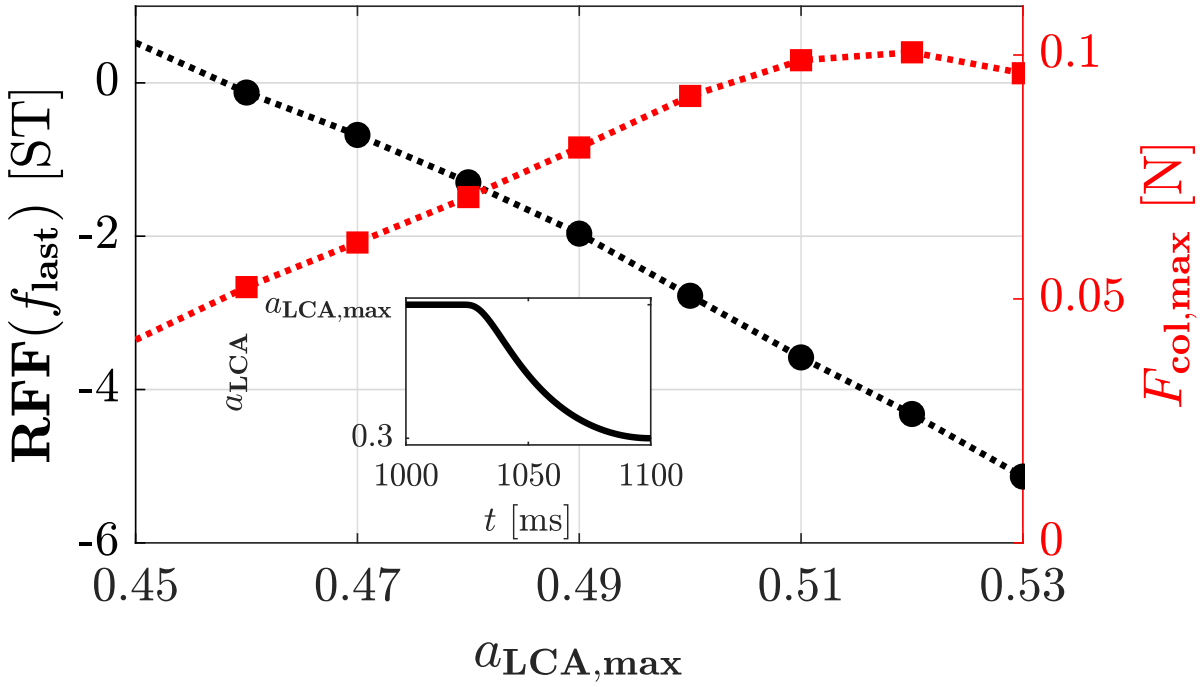


Figure 3.17: Lateral cricoarytenoid muscle activation profile with varying initial values $a_{\mathbf{LCA,max}}$ and its consequent effects on $\mathbf{RFF}(f_{\mathbf{last}})$ and maximum collision force $F_{\mathbf{col,max}}$.

mass, by adding a mass $\tilde{m}$ to m_1 (see Figure 3.6), on $\mathbf{RFF}(f_{\mathbf{last}})$. Figure 3.18 shows that

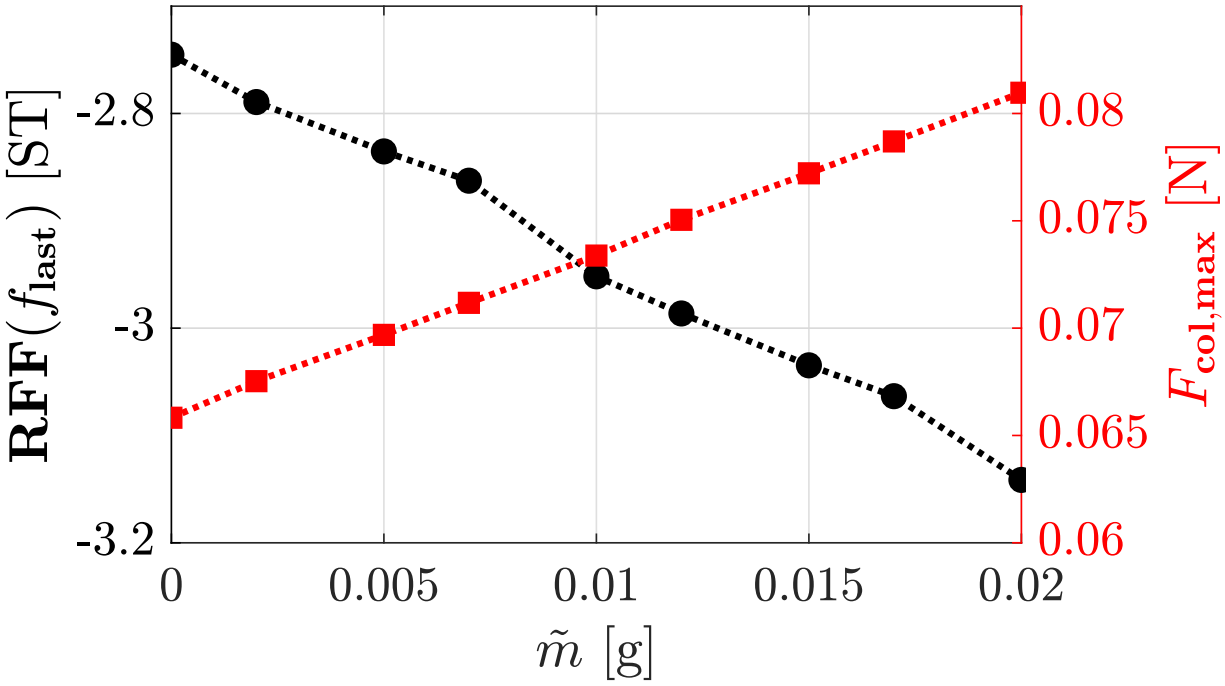


Figure 3.18: The effect of increasing the lower cover mass (by adding the mass $\tilde{m}$) on $\mathbf{RFF}(f_{\mathbf{last}})$ and $F_{\mathbf{col,max}}$.

increasing the lower cover mass leads to lower RFF values, which is again correlated to higher pre-offset collision forces. The increased forces can be attributed to the increased momentum of the VFs (especially prior to impact), which results in higher penetration into the collision plane and consequently higher collision forces. Figure 3.18 also indicates that the effect of increasing the cover mass on $\mathbf{RFF}(f_{\mathbf{last}})$ is insignificant, compared to, for example, the effect of pre-offset LCA muscle activation levels displayed in Figure 3.17, which agrees with clinical data showing that the surgical removal of polyps and nodules does not alter RFF significantly [136].

The three scenarios considered in this subsection show how the larger drops in RFF observed clinically for subjects with PVH can be potentially attributed, at least in part, to the increased pre-offset collision forces, which can be caused by increased sub-glottal pressure, pre-offset LCA muscle activation, or additional cover mass associated with an organic pathology.

3.3 Conclusion

In this chapter, we have introduced a theoretical and numerical framework that enabled understanding some of the laryngeal factors altering phonation frequency and RFF during phonation offset. This analysis was motivated by the promising application of RFF in assessing different voice pathologies. The main result of this work is that the clinically observed drop in RFF during offset is correlated with a decrease in collision forces as the VFs abduct. We demonstrate using a simple impact oscillator model the correlation between increased VF stiffness and sharper RFF drops observed in experimental studies. We further show that intra-subject variability in RFF may arise due to several factors, most notably the relative phase of the phonation cycle at which abduction initiates and the duration of the abduction period, which will vary from token to token and subject to subject. Moreover, we display the potential role of increasing the CT muscle activation during offset, which may be a compensatory mechanism adopted by normal speakers to stabilize frequency. Finally, we investigate possible scenarios connected to PVH showing that the clinically observed sharp RFF drops are associated with how high collision forces are prior to offset.

Our study employs several assumptions to simplify our analysis of phonation offset, which include: (a) the negligible role of laryngeal muscles, other than LCA and PCA, in abducting and adducting the VFs; (b) the constant nature of sub-glottal pressure during offset, whereas it has been speculated in the literature that the profile of the sub-glottal pressure during offset is time-variant [85]; and (c) the fixed supra-glottal geometry. The mentioned assumptions may limit the applicability of our findings, however the main hypothesis of this work on how RFF and collision forces are related is not affected by these assumptions.

Chapter 4

Exploring the Mechanics of Fundamental Frequency Variation During Phonation Onset

Phonation fundamental frequency and RFF during phonation onset exhibit temporal variations with characteristics that are dependent on the phonetic context and the nature of voice pathology if existent (see Section 2.5.3). As highlighted in Section 2.5.3, phonation onset is quite complex and studies of fundamental frequency variation during phonation onset do not provide a comprehensive understanding of its variation, which necessitates further extensive analysis.

This chapter extends the analysis in Chapter 3, where theoretical and numerical analyses are conducted to investigate the mechanisms underlying fundamental frequency patterns during phonation onset, and that in combination with the previous chapter, fulfills the first objective. We show, using the adopted modelling framework, that the temporal variation of fundamental frequency during phonation onset is influenced by the gradual increase in VF collision levels and the activation of laryngeal muscles (especially CT and TA muscles), where the differences in activation levels of laryngeal muscles may be responsible for the differences in RFF patterns between healthy speakers and speakers with VH.

The work presented in this chapter is based on a paper that has been published in the Biomechanics and Modelling in Mechanobiology journal under the title " Exploring the Mechanics of Fundamental Frequency Variation During Phonation Onset" [128], where Mohamed Ahmed Serry is the first author, and Cara E. Stepp and Sean D. Peterson are coauthors. M.A.S. and S.D.P. conceived of the study. M.A.S. developed the model. M.A.S.

generated preliminary results and figures. M.A.S. wrote the first draft of the article. All authors reviewed and approved the manuscript. This work was supported by the NIDCD of the NIH under awards P50DC015446 and R01DC015570. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. The authors thank Jonathan Deng for running finite element simulations to validate some of the modeling assumptions employed in this work and Dr. Matías Zañartu for insightful discussions on the intrinsic musculature of the larynx.

The organization of this chapter is as follows: in Section 4.1 we introduce the employed phonation models; the role of VF collision is discussed in Section 4.2; the influence of the cricothyroid and thyroarytenoid muscles during phonation onset is explored in Section 4.3; and Section 4.4 concludes the chapter.

4.1 Phonation models

In this section, we introduce the phonation models used in our analyses. The first is a hybrid phonation model that integrates the impact oscillator model introduced by [127] and a linearized version of the [155] single-mass model, which is used to explore the role of collision during phonation onset. The second is a body-cover reduced-order model [142] used to explore the role of muscle activation and corroborate findings from the hybrid model with a more physiologically relevant VF description.

Similar to Mergell et al. [94], it will be assumed, unless otherwise stated, that the neutral prephonatory gap between the VFs is fixed during onset. In the case of isolated vowels this assumption is supported by empirical data showing that VF oscillations are initiated from a fixed prephonatory neutral position [130]. In the case of vowels preceded by fricatives, onset has been found to start slightly before reaching the final prephonatory position [91, 114]. Moreover, for simplicity we neglect temporal variations in aerodynamic and acoustic parameters, such as the acoustic impedance at the mouth. Variations in such parameters are believed to play roles in altering fundamental frequency during transient periods of phonation [51]; however, their significance in comparison to VF contact and laryngeal muscle tension is the subject of some debate. Smith and Rob [132], for instance, found that onset fundamental frequency patterns of vowels preceded by fricatives and stop consonants are very similar despite their differing aerodynamic characteristics, implying a minor role of aerodynamics. On the contrary, the empirical data of relative fundamental frequency presented in [74] and [113] suggests that these factors may be prevalent. Herein, we focus on collision and muscle activation, leaving a comprehensive exploration of aerodynamics and acoustics for future work.

4.1.1 Hybrid phonation model

The hybrid model is shown schematically in Figure 4.1. It enables analysis of fluid-structure interactions during phonation onset, where the glottal flow is modeled using a linearized Bernoulli flow model, while incorporating the role of VF collision. The governing equations are

$$M\ddot{\xi} - \mathcal{B}_1\dot{\xi} + K\xi = 0,\ \xi(t) \geq -\delta, \tag{4.1a}$$

$$M\ddot{\xi} + \mathcal{B}_2\dot{\xi} + \mathcal{K}\xi = -k_{\text{col}}\delta,\ \xi(t) < -\delta, \tag{4.1b}$$

where $\xi(t)$ is the VF mass displacement from its neutral position, M is its mass, K is the tissue stiffness, k_{col} is collision stiffness, $\delta \geq 0$ is the neutral gap, and $\mathcal{K} = K + k_{\text{col}}$. The damping terms are given by $\mathcal{B}_1 = 2\tau P_L/(k_t\delta) - B$ and $\mathcal{B}_2 = B + b_{\text{col}}$, where B is the structural viscous damping coefficient, b_{col} is an additional damping coefficient incorporated during collision, P_L is the subglottal lung pressure, τ is a time delay term associated with the propagation of the mucosal wave on the medial surface of the VFs, and k_t is a pressure recovery term [155]. The mass, stiffness, and damping coefficients are given per unit area. The neutral gap, δ, serves as a proxy for degree of VF adduction, such that $\delta = 0$ corresponds to complete VF closure. This model neglects acoustic effects and assumes negligible supraglottal pressure; hence, P_L is correspondent to the transglottal pressure. It is assumed that the dynamics of the hybrid phonation model are oscillatory in both collision and non-collision regimes, that is,

$$\omega_1^2 := \frac{K}{M} - \frac{\mathcal{B}_1^2}{4M^2} > 0, \tag{4.2a}$$

$$\omega_2^2 := \frac{\mathcal{K}}{M} - \frac{\mathcal{B}_2^2}{4M^2} > 0, \tag{4.2b}$$

where $\omega_1,\ \omega_2 \geq 0$ denote the angular frequencies in the the non-collision and collision regimes, respectively.

The impact oscillator model of Serry et al. [127], referred to herein as the S21 model, can be recovered from Equation (4.1) by omitting the viscous forces (*i.e.*, by setting $\mathcal{B}_1 = \mathcal{B}_2 = 0$), resulting in

$$M\ddot{\xi} + K\xi = 0,\ \xi(t) \geq -\delta, \tag{4.3a}$$

$$M\ddot{\xi} + \mathcal{K}\xi = -k_{\text{col}}\delta,\ \xi(t) < -\delta. \tag{4.3b}$$

This model isolates the effects of collision and primitive parameters (*e.g.*, mass, stiffness, and neutral gap) on fundamental frequency, providing an abstract, yet useful, insight into the role of VF contact during real phonation scenarios.

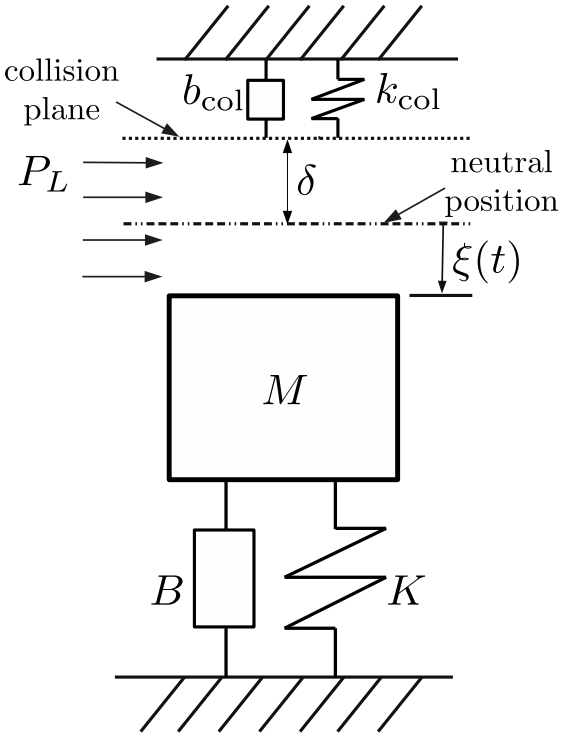


Figure 4.1: Schematic diagram of the hybrid phonation model.

The linearized version of the [155] model can be recovered from the hybrid phonation model by assuming collision-free oscillations, that is, $\xi(t) > -\delta$, yielding[1]

$$M\ddot{\xi} - \mathcal{B}_1\dot{\xi} + K\xi = 0. \tag{4.4}$$

Equation (4.4) provides useful insights into the fluid-structure interaction between the VFs and the glottal flow during phonation onset and, in particular, the role of aerodynamics (in the form of negative damping that results from linearizing the Bernoulli flow model) in initiating VF oscillations. The onset conditions predicted from Equation (4.4) (see Equation (4.6)) agree reasonably with experimental measurements from a physical model of the VF mucosa; in particular, phonation threshold pressure[2] is positively correlated with the neutral gap δ (for sufficiently large δ) and the VF viscous damping coefficient B [163].

4.1.2 Body-cover model

The reduced-order three-mass BCM [142] is employed to verify and extend the findings from the simpler, more analytically tractable, hybrid phonation model. This model, which embeds the essential physiological components of the VFs, consists of two cover masses and a body mass, all connected via springs and dampers to model the VFs viscoelastic tissues (see Figure B.1 in Appendix B.1). The model assumes the motion of the VFs to be symmetric about the medial plane; hence, only one of the folds is needed in the model construction. Collision of the opposing folds is modeled by activating additional nonlinear spring forces applied to the cover masses, where the spring forces are proportional to the degree of overlap of the cover masses with the medial (collision) plane, see Equations (6a) and (6b) in [142][3]. The model implements the muscle activation rules of [164] to control the primitive model variables via three dimensionless muscle activation parameters, a_{CT}, a_{TA}, and a_{LCA}, which account for the relative activation of the CT, TA, and LCA/PCA muscles, respectively. The neutral glottal gap in the BCM is modulated through activation of the LCA muscle, where the neutral glottal half width x_0 is given by $x_0 = 0.25L_0(1 - 2a_{\mathrm{LCA}})$, where L_0 is the resting VF length [164]. As seen from this relation, LCA activation is negatively correlated to the VF neutral gap, where increasing the activation of the LCA muscle leads to adducting the VFs.

[1] Equation (4.4) can be obtained by linearizing Equation (8) in [86] at the point of zero displacement (from the neutral position) and velocity.

[2] Phonation threshold pressure is the minimum subglottal pressure needed to initiate phonation [155]. It can also refer to the minimum subglottal pressure required to maintain VF oscillations [163].

[3] This is essentially the same collision model as for the hybrid phonation model. Both models neglect adhesive forces during collision, which can influence VF biomechanics [10].

Air flow through the glottis is modeled using a quasi-steady Bernoulli flow formulation with quasi-steady viscous correction for losses in the glottis [116, 84, 88]. The quasi-steady viscous model has shown good agreement with experimental observations of flow through larynx models [166]. We note herein that our glottal flow model is similar, but not identical, to those presented in [116, 88] as it incorporates flow separation and its formulation is suitable for modelling acoustic effects due to subglottal and supraglottal tracts. Viscous corrections are employed to account for non-negligible losses that occur during the initial stages of phonation when the flow speeds are relatively slow and during periods when the glottis is nearly closed [31]. See Appendix B.1 for further details on the employed flow model.

Acoustics are modeled using the WRA method [64, 78, 140]. Similar to [33] and [186], a subglottal tract area function is adapted from respiratory system measurements of human cadavers [178], covering only the trachea and bronchi. A supraglottal tract is also included, which is configured to simulate the /i/ vowel [150]. The BCM dynamics are driven by the lung pressure, P_l, input to the inferior end of the subglottal tract. To mitigate numerical instabilities in the WRA implementation, P_l is ramped up from zero to the desired value during phonation onset according to the relation $P_l(t) = P_{l,0}(1 - \mathrm{e}^{-t/\sigma})$, where $P_{l,0}$ is the steady-state lung pressure and $\sigma = 0.2$ ms. The settling time for the ramp is less than 1 ms. The system dynamics are solved using an explicit version of Newmark's method [104, 33] with a sampling frequency of 140 kHz. Initial conditions in all BCM simulations are identical, with zero velocity for all masses and unstretched model springs. As in [127], we consider the time-series of the glottal area, A_g, in our frequency analysis, where frequency is determined from the time duration between sequential signal peaks.

4.2 Relationship Between Collision and Fundamental Frequency

[127] demonstrated a direct correlation between VF collision and fundamental frequency, wherein transitioning from a VF oscillation regime with collision to one without collision during phonation offset results in a decrease in fundamental frequency due to the net reduction in "system stiffness". We posit that this mechanism is also a contributing factor underlying the temporal variation in fundamental frequency during phonation onset.

4.2.1 Insights from the S21 model

In this section, we summarize key findings from [127], which investigated phonation offset, and expand the analysis therein to explore phonation onset. We note that despite the symmetries between phonation offset (decaying VF oscillations) and phonation onset (rising VF oscillations), there exist some notable differences between the two phenomena, including phonation threshold pressure values [163], and aerodynamic characteristics (depending on the phonetic context) [81]. Herein, we aim to utilize the symmetries between the two phenomena to elucidate the role of VF collision in altering fundamental frequency during phonation onset.

From Equation (4.3)(a) (the S21 model without collision), we note that the natural frequency of the oscillator is $f_0 = \sqrt{K/M}/(2\pi)$. For convenience, we define the normalized neutral gap $\tilde{\delta} = \sqrt{K/(2E)}\delta$, where E is the energy of the VF system per unit area, which can loosely be considered as the energy originally imparted to the system via aerodynamics [4]. The system energy is constant in time in this case owing to the lack of viscous losses, as can be seen from Equation (4.3). We further define the stiffness ratio $\tilde{k} = K/\mathcal{K}$. The fundamental frequency of the model is [127]

$$f = \begin{cases} \frac{2f_0}{\frac{2}{\pi}\sqrt{\tilde{k}}\arctan\left(\sqrt{(\frac{1}{\tilde{\delta}^2}-1)/\tilde{k}}\right)+\frac{2}{\pi}\arcsin(\tilde{\delta})+1}, & \tilde{\delta} \le 1, \\ f_0, & \tilde{\delta} > 1, \end{cases} \tag{4.5}$$

for which the behavior depends on whether or not the system has sufficient energy (vibration amplitude) to cause collision. Note that when $\tilde{\delta} > 1$ (no collision), frequency is independent of the oscillator energy. Utilizing a quasi-steady assumption with E as a parameter, we explore the effect of varying the system energy on fundamental frequency. The role of energy term, E, becomes evident when collision occurs ($\tilde{\delta} \le 1$), wherein fundamental frequency increases as E increases, with an asymptotic value $2f_0/(\sqrt{\tilde{k}}+1)$, corresponding to oscillations at zero neutral gap ($\delta = 0$) (see Figure 4.2). The asymptotic behavior suggests that the collision-based mechanism is inefficient at changing frequency at high energy levels, as large energy increases result in modest gains in fundamental frequency.

In addition to system energy, collision is also modulated via VF adduction, which is embedded in the S21 model through the neutral gap, δ. Equation (4.5) shows that decreasing δ has a similar effect to increasing E (both lead to decreasing $\tilde{\delta}$). That is, for

[4] As can be seen from Equation (4.3), the S21 model does not incorporate aerodynamic effects. Stating that the system energy is imparted via aerodynamics is a crude assumption. In Section 4.2.2, we consider aerodynamic energy transfer during phonation onset more rigorously using the hybrid phonation model.

fixed system energy and stiffness, fundamental frequency can be increased purely through adduction. Similar to the energy rise, adduction only impacts fundamental frequency of the model when collision is present ($\tilde{\delta} \leq 1$). We note that the effect of increasing energy is mediated by the neutral gap, with a more muted response as the gap decreases. That is, the effectiveness of a rise in system energy at increasing frequency during phonation onset is dependent on the adduction level of the VFs. In reality the exact relation is naturally expected to be complex due, in part, to the geometry of the glottis and the high degrees of freedom of the VFs.

Finally, the S21 model provides useful insights into the role of VF stiffness during phonation onset. From Equation (4.3), we observe that K affects the dynamics of the VF system in both the collision and non-collision regimes, wherein increasing stiffness increases the (instantaneous) fundamental frequency. This indicates that changing VF stiffness during phonation onset through intrinsic muscle activation alters fundamental frequency even in the initial stage of onset in some phonetic contexts when VF oscillations are collision-free. Thus, we expect potentially competing factors of adduction, aerodynamic energy transfer, and laryngeal tension to influence fundamental frequency during phonation onset.

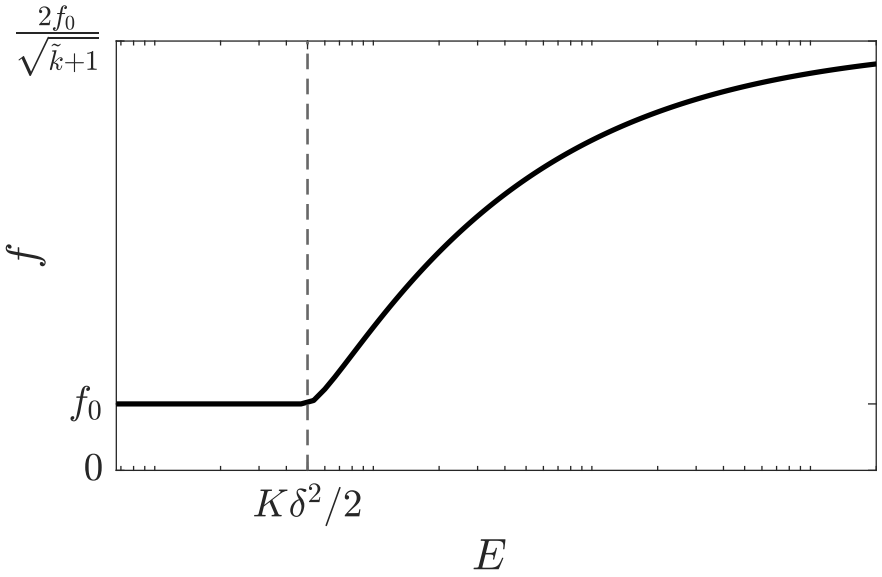


Figure 4.2: Fundamental frequency of the S21 model as a function of the system energy. The dashed vertical line indicates the energy level at which collision initiates.

4.2.2 Analysis using the hybrid phonation model

Analysis of the S21 model in the previous section relies on the quasi-steady assumption, wherein the dynamics of fluid-structure interaction during phonation onset and viscous friction losses are neglected. In this section, we consider an analytical treatment to the onset problem using the hybrid phonation model (see Section 4.1) and elucidate the dynamic nature of the collision-based mechanism. We consider the evolution of VF oscillations during onset while incorporating VF contact, which has been typically omitted in previous theoretical analyses of phonation onset [155, 190, 86].

Pre-collision, the hybrid model is equivalent to the linearized [155] model in Equation (4.4), which predicts VF oscillations with exponential growth when (see, for example, [155])

$$2\tau\frac{P_L}{k_t\delta} - B > 0. \tag{4.6}$$

However, realistic energy dynamics during phonation onset are complex and oscillatory due to several factors, including nonlinear fluid-structure interaction effects, VF collision, acoustics, and viscous losses. The primary energy transfer mechanism to the VF system is the temporal asymmetry of the average intraglottal pressure, where, loosely speaking, positive energy transfer from the glottal flow and energy dissipation to the flow take place when the VF configuration is convergent and divergent, respectively, with the total energy transferred from the flow being larger than that dissipated to the flow in order to sustain oscillations [151]. The hybrid model (Equation (4.1)) allows exploration of the general trends of the complex oscillatory energy dynamics beyond the initial onset of oscillations by incorporating simplified VF contact and aerodynamic effects.

We examine the energy evolution by considering the discrete system energy at the same phase in a sequence of oscillation cycles. Let $\tau_i,\ i = 0, 1, 2, \cdots$ be the time instances such that $\xi(\tau_i) = -\delta$ and $\dot{\xi}(\tau_i) < 0$, which correspond to the beginning of each collision. Let $\mathcal{V}_i = |\dot{\xi}(\tau_i)|$ be the oscillator velocity magnitude at time instance τ_i. The energy immediately prior to each collision (kinetic energy plus potential energy in the spring) is then $E(\tau_i) = M\mathcal{V}_i^2/2 + K\delta^2/2$. The velocity sequence $\{\mathcal{V}_i\}$ can be obtained approximately using the recurrence relation

$$\mathcal{V}_{i+1} = \mathcal{A}\mathcal{V}_i + \mathcal{W},\ \ i = 0, 1, 2, \cdots, \tag{4.7}$$

where the initial velocity $\mathcal{V}_0 > 0$ is given. The parameter $\mathcal{A}$ (a scaling term) is modulated by the energy losses and gains in the collision and collision-free regimes, respectively, and the parameter $\mathcal{W}$ (a drift term) is regulated by the neutral gap δ. Derivation of Equation (4.7) and the exact definitions of $\mathcal{A}$ and $\mathcal{W}$ are provided in Appendix B.2.

The dynamics of the recurrence relation given in Equation (4.7) exhibit various behaviors depending on the numerical values of $\mathcal{A}$ and $\mathcal{W}$ (*e.g.*, linear growth, exponential growth, and exponential decay). Herein, we are interested in cases where VF oscillations are bounded, thus corresponding to realistic phonation onset scenarios. On average, the aerodynamic energy transfer to the VF system is larger than viscous losses after phonation initiation, which induces VF oscillations of growing amplitude. The (average) difference between aerodynamic energy transfer to the VF system and viscous dissipation gradually decays over time until the difference becomes zero, which corresponds to steady-state VF oscillations of constant amplitude (that is, sustained phonation).

The case of bounded energy growth can be determined from Equation (4.7) when $\mathcal{W} \geq 0$ and $0 < \mathcal{A} < 1$, which is fulfilled when

$$0 < 2\tau \frac{P_L}{k_t \delta} - B < (B + b_{\text{col}})\sqrt{\tilde{k}}. \tag{4.8}$$

This corresponds to the onset condition given in Equation (4.6) under the additional constraint that the subglottal pressure is such that the damping ratio in the non-collision regime, $\mathcal{B}_1/(M\omega_1)$, is smaller than the damping ratio in the collision regime, $\mathcal{B}_2/(M\omega_2)$, to ensure VF oscillations of finite amplitude. As an example, if we set $k_t = 1.1$, $\delta = 10^{-3}\,\text{m}$, $\tau = 1.5 \times 10^{-3}\,\text{s}$, and $B = 2380\,\text{Pa}\cdot\text{s/m}$ (similar to values used in [155] and [84]) and additionally assume $b_{\text{col}} = 4B$ and $k_{\text{col}} = 3K$ (similar to assumptions in [135]), Equation (4.8) predicts that P_L should be within the approximate range $[875, 2180]$ Pa, in order to have VF oscillations of bounded amplitude[5]. In this case ($\mathcal{W} \geq 0$ and $0 < \mathcal{A} < 1$), Equation (4.7) can be rewritten as

$$\mathcal{V}_i = \frac{\mathcal{W}}{1-\mathcal{A}} - \mathcal{A}^i \left(\frac{\mathcal{W}}{1-\mathcal{A}} - \mathcal{V}_0 \right), \; i = 1, 2, \cdots \tag{4.9}$$

If

$$\mathcal{V}_0 < \frac{\mathcal{W}}{1-\mathcal{A}}, \tag{4.10}$$

that is, the kinetic energy of the VFs is initially low, then the sequence $\{\mathcal{V}_i\}$ is monotonically increasing with an asymptotic upper bound $\mathcal{V}_\infty = \mathcal{W}/(1-\mathcal{A})$ (see Figure 4.3). This shows that during phonation onset, the energy of the VF system increases gradually, on average, and achieves an asymptotic value where aerodynamic energy transfer equals to viscous losses, which is associated with sustained phonation.

[5]The additional constraint in Equation (4.8) is sensible as subglottal pressure during phonation is bounded by the physiological limitations of the vocal and respiratory systems. During normal speech, subglottal pressure values are within the approximate range of 200-800 Pa [192], whereas shouting can lead to subglottal pressures up to and beyond 10 kPa [72].

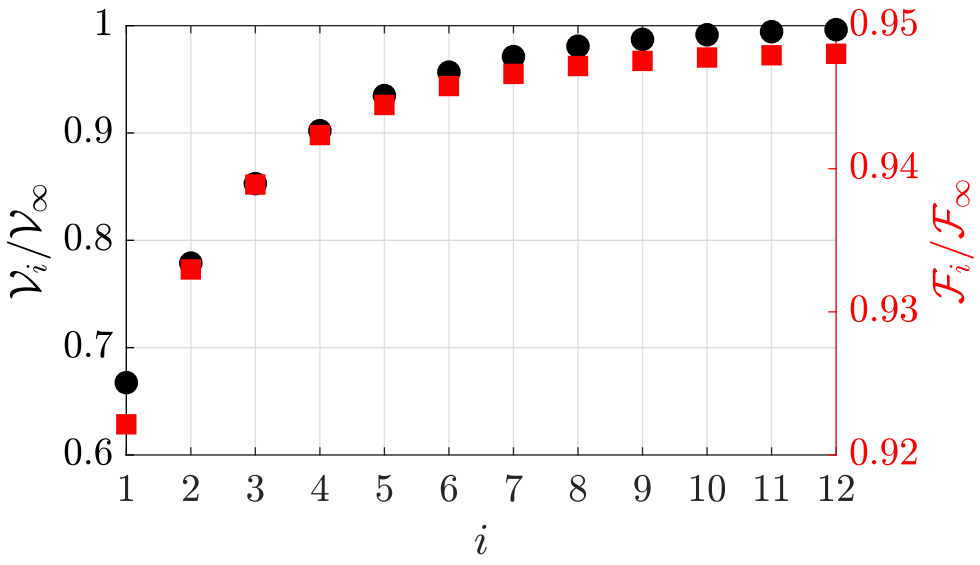


Figure 4.3: Illustrative example of the velocity (circles) and frequency (squares) sequences given in Equations (4.7) and (4.11), respectively, normalized with respect to the asymptotic upper bounds $\mathcal{V}_\infty$ and $\mathcal{F}_\infty$, respectively, with with fixed-in-time parameter values $M = \mathcal{B}_1 = K = 1$, $\mathcal{B}_2 = 2$, $\mathcal{K} = 3$, $\delta = 0.1$ and $\mathcal{V}_0 = 0.5\mathcal{W}/(1-\mathcal{A})$. Note that, for the given parameter values, the conditions in Equations (4.8), (4.10), and (4.12) are fulfilled, which explains the monotonically increasing and bounded behaviors.

The frequency of the i^{th} cycle, $\mathcal{F}_i$, is approximately given by

$$\mathcal{F}_i = \left[\frac{1}{\mathcal{F}_\infty} + \frac{\alpha_1}{\mathcal{V}_{i-1} + \beta} - \frac{\alpha_2}{\mathcal{V}_{i-1}} \right]^{-1}, \quad i = 1, 2, \cdots \tag{4.11}$$

where $\mathcal{F}_\infty = \left((1/\omega_1) + (1/\omega_2) \right)^{-1} / \pi$ (see Appendix B.3 for the derivation and the definitions of parameters α_1, α_2, and β). Note that if the sequence $\{\mathcal{V}_i\}$ is monotonically increasing, which is attained if the conditions in Equations (4.8) and (4.10) are satisfied, and that the initial velocity additionally satisfies the constraint[6]

$$\mathcal{V}_0 \geq \beta \frac{\alpha_2 + \sqrt{\alpha_1 \alpha_2}}{\alpha_1 - \alpha_2}, \tag{4.12}$$

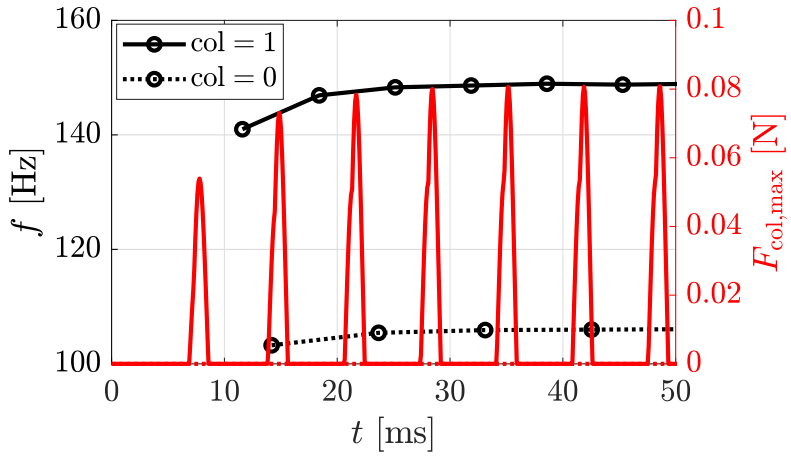
then the sequence $\{\mathcal{F}_i\}$ is guaranteed to be monotonically increasing (see Figure 4.3) with an asymptotic upper bound $\mathcal{F}_\infty$, which corresponds to the fundamental frequency in the case of zero neutral gap. In other words, the fundamental frequency exhibits a bounded increase during phonation onset due to the average increase in the kinetic energy of the VF system, which agrees in essence with the quasi-steady analysis in Section 4.2.1. This rising trend is also consistent with empirical studies of onset of initial and isolated vowels, see Figure 2.12. The implications of matching with empirical data are discussed in Section 4.2.4.

4.2.3 Numerical simulations with the body-cover model

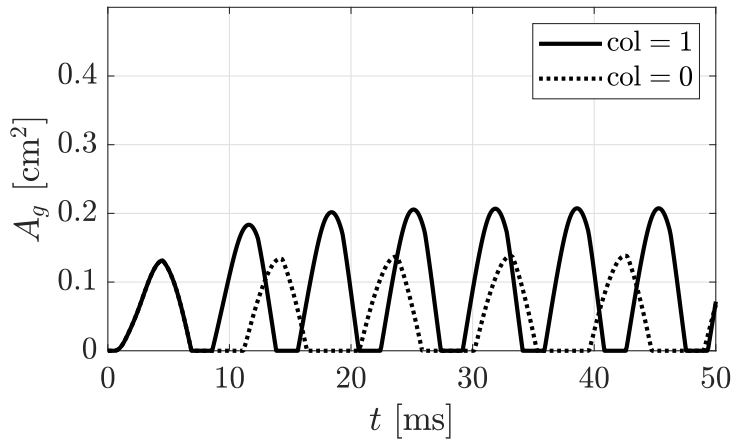
To ground the analysis from the simplified models in a more physiologically-relevant framework, we resort to simulations using the BCM.

First, we consider onset simulations with fixed steady-state lung pressure, $P_{l,0} = 800$ Pa, and muscle activation values corresponding to low/normal CT and TA activation levels and fully adducted VFs, where $a_{\text{CT}} = 0.2$, $a_{\text{TA}} = 0.2$, and $a_{\text{LCA}} = 0.5$. Figure 4.4 displays the fundamental frequency, maximum collision force (among the cover masses), and glottal area profiles of two exemplar cases during onset. In one case, the collision springs are activated, resulting in non-zero collision forces when VF contact occurs (col = 1), and in the other case, the collision springs are deactivated through the entire simulation period, resulting in zero collision forces (col = 0).

[6] Derivation of the lower bound is omitted for brevity. It can be obtained by imposing that the frequency function given in Equation (4.11) be increasing with respect to its velocity argument (*e.g.,* by setting the first derivative to be positive).



(a) Fundamental frequency (left), and maximum collision force (right) time series.



(b) Glottal area time series.

Figure 4.4: Two exemplar simulations of phonation onset using the body-cover model with fixed subglottal pressure, $P_{l,0}$ = 800 Pa, and muscle activation values a_{CT} = 0.2, a_{TA} = 0.2, and a_{LCA} = 0.5. Solid lines indicate the simulation with collision springs activated when contact occurs (col = 1), whereas dashed lines indicate the simulation with the collision springs deactivated, even when collision occurs (col = 0).

Figure 4.4(a: left axis) shows that there is a gradual increase in fundamental frequency for both cases, though the increase is greater for the col = 1 case. Figure 4.4(a: right axis) illustrates the increase in collision forces during onset for the col = 1 case, which is attributed to the increased energy of the VF system. It can be inferred from Figure 4.4(a) that the more rapid rise in fundamental frequency in the col = 1 case is correlated with the rise in collision forces, which is in agreement with our theoretical analysis in Section 4.2.2[7]. The rise in fundamental frequency in the case of deactivated collision springs highlights the complex nature of the process, wherein nonlinear stiffness and aerodynamic contributions can also influence fundamental frequency. Moreover, the relatively larger increase in fundamental frequency in the case of activated collision springs indicates that collision plays a significant role in increasing frequency during onset when all other controlling factors (*e.g.*, muscle activation) are fixed. Figure 4.4(b) shows that the amplitude of the glottal area waveform increases during onset also due to aerodynamic energy transfer, where a larger oscillation amplitude is noticed in the col = 1 case, which can be attributed to the (repulsive) contact forces during the contact periods.

To further highlight the influence of VF contact, we now consider an onset simulation with the same steady-state lung pressure, and CT and TA muscle activation values. However, we vary the LCA activation level from $a_{LCA} = 0.4$ to $a_{LCA} = 0.5$ over a period of 50 ms, which corresponds to the VFs being initially abducted then proceeding to the fully adducted state. This scenario simulates the glottal state during the onset of vowels preceded by voiceless consonants (see, for example, [25]). Figure 4.5 displays the fundamental frequency, maximum collision force, and glottal area time-series for the simulation (LCA activation and glottal area time series are shown in the inset). The figure shows that VF oscillations exhibit contact starting from $t \approx 30$ ms. Moreover, the figure shows that prior to the initial contact, oscillations exhibit variations in fundamental frequency, potentially due to nonlinear and aerodynamic effects as stated in the discussion of Figure 4.4. Furthermore, the figure displays that, starting from the initial VF contact instance, the oscillations exhibit a significant rise in fundamental frequency, which also correlates with the rise in collision forces, in agreement with the theoretical analysis in Sections 4.2.1 and 4.2.2.

4.2.4 Comments on relations to empirical observations

The increasing frequency resulting from progressively greater degrees of collision during phonation onset predicted by the models in this study aligns with empirical observations

[7]This correlation is also observed in other onset simulations with different fixed-in-time steady state subglottal pressure and muscle activation values.

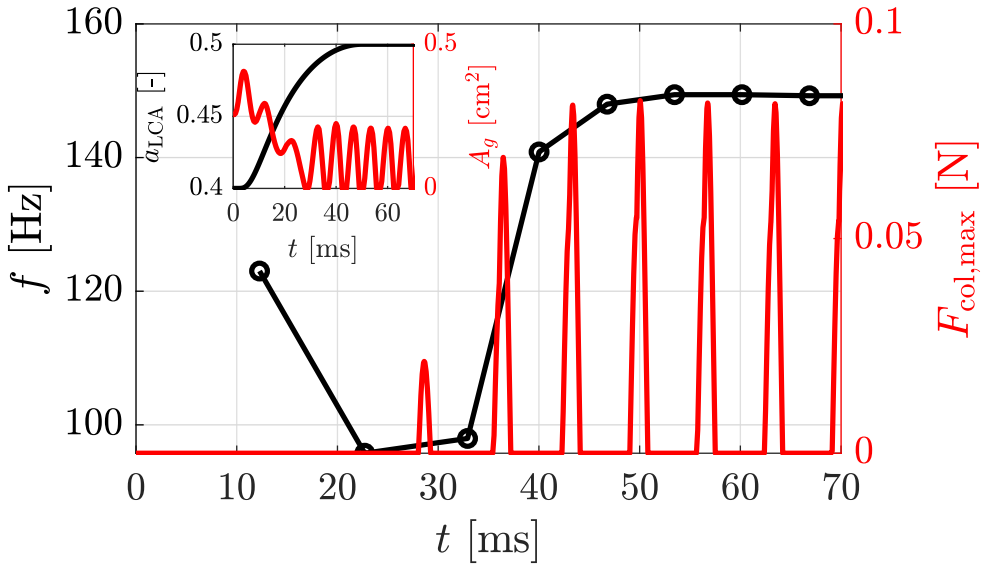


Figure 4.5: Frequency and maximum collision force versus time for an onset simulation with time-varying neutral glottal gap, where a_{LCA} varies from 0.4 to 0.5 over a 50 ms span (see the inset), $P_{l,0} = 800$ Pa, and $a_{\mathrm{CT}} = a_{\mathrm{TA}} = 0.2$. The inset of the figure also depicts the time-series of the glottal area waveform, A_g, showing its variation with changing lateral cricoarytenoid muscle activation.

for initial and isolated vowels [132, 97] (see Figure 2.12). This also agrees with some reported observations for vowels preceded by voiced consonants [51]. Whereas variations in laryngeal muscle tension and/or aerodynamics are often proposed to be the underlying factors governing the rise in frequency for these conditions [132], our study shows that these influences need not be present to generate the observed behavior, since the natural system dynamics tend to increase fundamental frequency during phonation onset. We emphasize that this does not mean these other factors are not playing a role during onset, only that they are not necessary to produce the observed frequency patterns.

As has been observed clinically (see Figure B.3), this gradually increasing effect of VF contact during phonation onset is also present in other phonetic contexts, including the onset of vowels preceded by voiceless consonants, implying the relevance of VF contact in various contexts. However, as shown in Figure 2.12, fundamental frequency tends to decrease during onset for a vowel preceded by a voiceless consonant, indicating that the collision-based rise in frequency is overshadowed by other factors. In Section 4.3, we show that laryngeal muscle activation can induce the observed decreasing trends of fundamental frequency during the onset of vowels preceded by voiceless consonants.

4.3 Muscle Tension and Frequency Regulation

In this section, we explore the influence of intrinsic laryngeal muscle tension and, in particular, the role of CT and TA muscles during phonation onset using the BCM. Intrinsic laryngeal muscles and their roles in phonation have been extensively investigated in several clinical [15, 17, 19, 18, 16] and numerical [34, 3, 101, 181, 182] studies, where it has been found that the CT and TA muscles are essential in regulating fundamental frequency. Increasing activation of the CT muscle has been found to increase phonation fundamental frequency [80, 17]. On the other hand, the role of the TA muscle in modulating fundamental frequency is more complex as its activation can either increase or decrease fundamental frequency, with some conflicting results in the literature (see [101]). Activation of the LCA and IA muscles have been found to be positively correlated with fundamental frequency [19], whereas PCA activation exhibits negative correlation with fundamental frequency [18].

To the best of our knowledge, there are few studies that have substantially investigated the temporal variations of laryngeal muscle activation during phonation onset and how these variations may underlie empirical observations of fundamental frequency (*e.g.*, [80]). In this study, we attempt to explore these temporal variations in order to elucidate some of the underlying mechanisms of phonation onset.

In real phonation scenarios, the intrinsic laryngeal muscles do not act in isolation and their effect on fundamental frequency depends on several factors, including the relative geometry and contraction levels of agonist/antagonist muscles [4]. To simplify our analysis in this section, we aim to isolate the effects of the CT and TA muscles and assume that the tension variation in other laryngeal muscles is negligible. In all simulations presented below, we set $P_{l,0}$ = 800 Pa and a_{LCA} = 0.5, which corresponds to fully adducted vocal folds.

4.3.1 Cricothyroid muscle

The CT muscle plays a crucial role in regulating fundamental frequency by elongating and tensioning the VFs [164, 134, 80, 6, 16]. Electromyography has shown that activation of the CT muscle is higher in phonetic contexts wherein vowels are preceded by voiceless consonants in comparison to vowels preceded by voiced consonants, which correlates with the empirically observed higher onset fundamental frequency in such conditions [80]. It has been speculated that higher VF tension, which correlates with higher activation of the CT muscle, is required to mitigate VF vibrations during the production of voiceless consonants and that the higher tension carries over into the adjacent vowel [51]. Here, we aim to investigate this hypothesis numerically by varying CT muscle activation while keeping the activation levels of the TA and LCA muscles fixed.

We begin with a quasi-steady analysis wherein CT activation is fixed in time. Figure 4.6 presents sustained phonation fundamental frequency as a function of CT activation for different TA activation levels. In all cases, sustained phonation fundamental frequency is positively correlated to the CT muscle activation level (for fixed TA activation level), in agreement with previous numerical and clinical studies studies [3, 164, 17]. There are slight fluctuations observed in the fundamental frequency curves for large a_{CT} values, which can be attributed, in part, to the non-linearity of the BCM. Assuming variations in the activation levels of other laryngeal muscles to be small, this suggests that, in contexts where vowels are preceded by voiceless consonants, CT muscle activation level may be decreasing in order to achieve the empirically observed decaying fundamental frequency patterns as seen in, for example, [136][8].

Figure 4.7 presents instances of the glottal area time series during phonation onset for monotonically decaying CT activation, with initial value $a_{\text{CT,i}}$ and final value $a_{\text{CT,f}}$. Specifically, initial activation level $a_{\text{CT,i}}$ is varied between 0.2 and 0.6 across simulations and final activation value is set to be 0.2. The transition between the two a_{CT} levels

[8]The extent of which laryngeal maneuvers, including the variation of the CT muscle activation, alter fundamental frequency is language-specific as highlighted in [30].

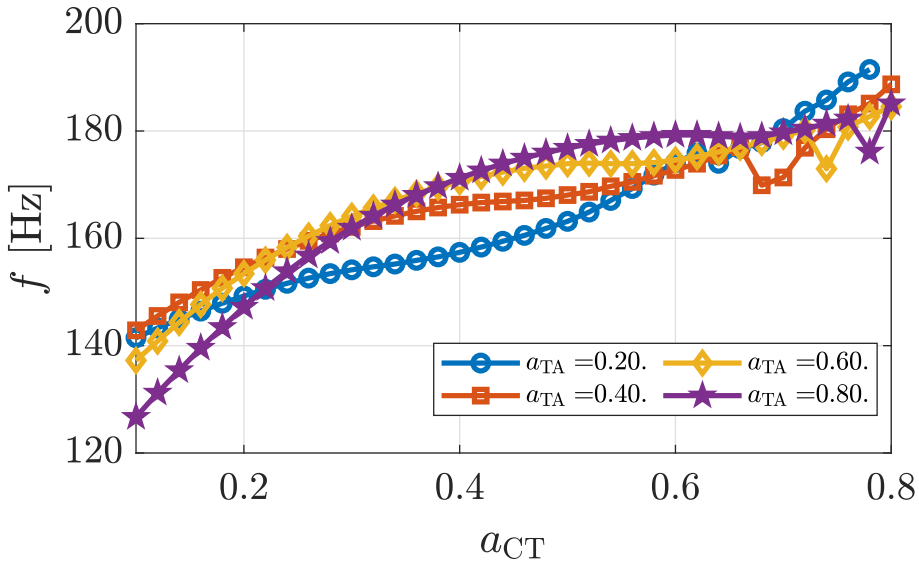


Figure 4.6: Sustained phonation fundamental frequency as a function of cricothyroid muscle activation for varying thyroarytenoid activation levels.

occurs over a duration of 50 ms (see the inset for the activation level temporal evolution), which is of the same order of magnitude as observed experimentally (see, for example, the electromyographic signals depicted in Figures 1-3 in [80]). The figure shows that the amplitude of vibration grows the most rapidly when there is no change in CT activation ($a_{\mathrm{CT,i}}$ = 0.2), with the rate of amplitude growth decreasing with increasing $a_{\mathrm{CT,i}}$. Higher CT activation levels result in stiffer folds, and thus higher frequency and generally lower amplitude, which relax as the activation level decreases in time. This is in agreement with the claim in [51] that the increased VF tension in phonetic contexts with vowels preceded by voiceless consonants is required to inhibit VF oscillations during the production of the voiceless consonant.

Figure 4.8 presents the temporal evolution of normalized fundamental frequency for the glottal area time-series shown in Figure 4.7, as well as for analogous cases with a_{TA} = 0.4. Decreasing CT activation levels during onset, with sufficiently large initial values, generally results in a decaying fundamental frequency profile, which matches empirical observations of vowels preceded by voiceless consonants, see Figure 2.12. Figure 4.8 displays non-monotonicity of the fundamental frequency profile in some cases, such as for a_{TA} = 0.4, $a_{\mathrm{CT,i}}$ = 0.4, wherein the frequency of the second cycle is larger than that of the first cycle despite the monotonic decay of CT activation levels. This has also been observed

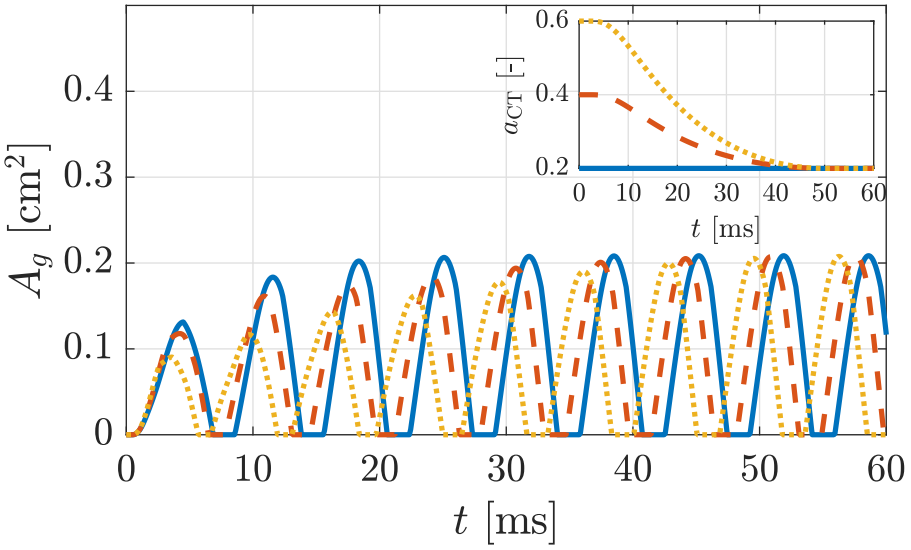


Figure 4.7: Time-series of glottal area for varying initial cricothyroid activation levels using the body-cover model at $a_{\mathrm{CT,f}} = 0.2$ and $a_{\mathrm{TA}} = 0.2$. Similar trends hold for $a_{\mathrm{TA}} = 0.4$ (not shown).

in empirical studies of onset fundamental frequency (see, for example, [80, 76], and Figure B.3 in Appendix B.4). This non-monotonic behavior can be attributed, in part, to the collision-based mechanism, which is dominant in some cases, such as when $a_{\mathrm{CT,i}} = 0.2$ (CT activation is constant-in-time). Collision onset causes fundamental frequency to increase, which opposes the effect of decreasing VF stiffness associated with the reduction in CT activation level. This demonstrates the complexity of phonation onset, where competing mechanisms are at play.

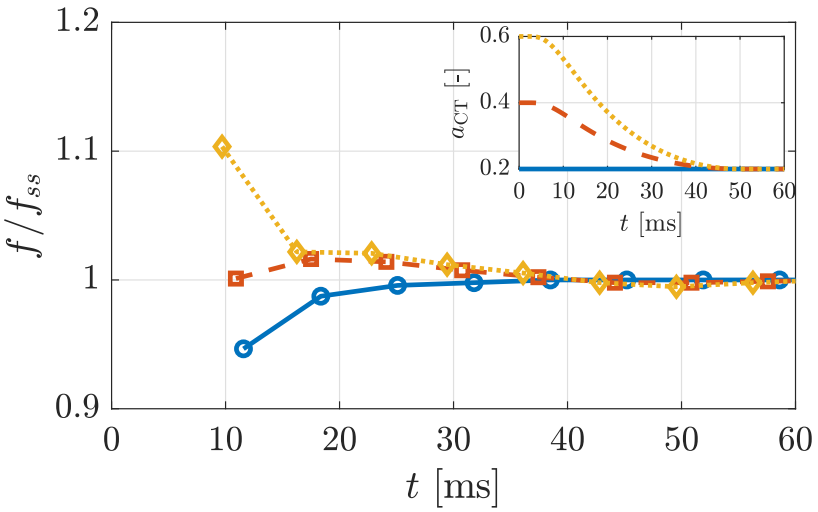
Similar fundamental frequency trends are observed when fixing initial CT activation and varying the final value, as shown in Figure 4.9. This figure, in combination with Figure 4.8, shows that when the drop in CT activation level is sufficiently large, the magnitude of the reduction in CT activation is correlated with the drop in (relative) fundamental frequency. This important finding will be used to explain some empirical observations in Section 4.3.3.

4.3.2 Thyroarytenoid muscle

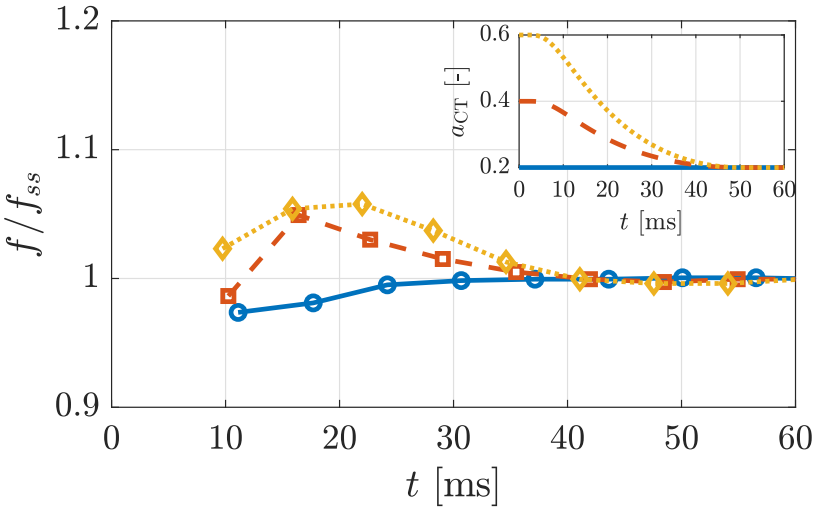
Activation of the TA muscle contributes to adducting and shortening the VFs [160, 16]. TA activation during phonation initiation has been found to begin when the VFs start adducting and carries over into sustained phonation [118]. Studies with human subjects have shown that the relation between TA activation and sustained phonation fundamental frequency is proportional when fundamental frequency values are low, whereas the relation is inverse at high frequency levels [162]. However, *in vitro* studies involving excised canine larynges [17] exhibit some deviations from human studies [162].

Figure 4.10 illustrates the relation between sustained phonation fundamental frequency and TA muscle activation for various CT activation levels. This figure shows that the relationship between steady-state fundamental frequency and TA activation level is relatively complex, with the influence of increasing TA depending on CT activation, in agreement with empirical data [162]. For example, when $a_{\mathrm{CT}} = 0.2$, corresponding to relatively low fundamental frequency, the relation is proportional when $a_{\mathrm{TA}} \in [0.1, 0.3]$, but it is inverse when $a_{\mathrm{CT}} = 0.6$. The steady state phonation results suggest that in scenarios where CT activation follows a decaying profile and TA activation is fixed, higher TA activation values can result in smaller differences between initial and final fundamental frequencies (see, for example, the decrease in spread between the curves in Figure 4.10 as a_{TA} increases from 0.1 to 0.3).

To test this, we perform onset simulations with decaying CT profiles similar to those shown in the inset of Figure 4.9 at various fixed TA activation values. TA activation



(a) $a_{\mathrm{TA}} = 0.20$.



(b) $a_{\mathrm{TA}} = 0.40$.

Figure 4.8: Time-series of fundamental frequency, normalized with respect to sustained phonation fundamental frequency f_{ss}, for varying initial cricothyroid activation levels using the body-cover model for (a) $a_{\mathrm{TA}} = 0.20$, and (b) $a_{\mathrm{TA}} = 0.40$. The time-series of cricothyroid activation are shown in the insets.

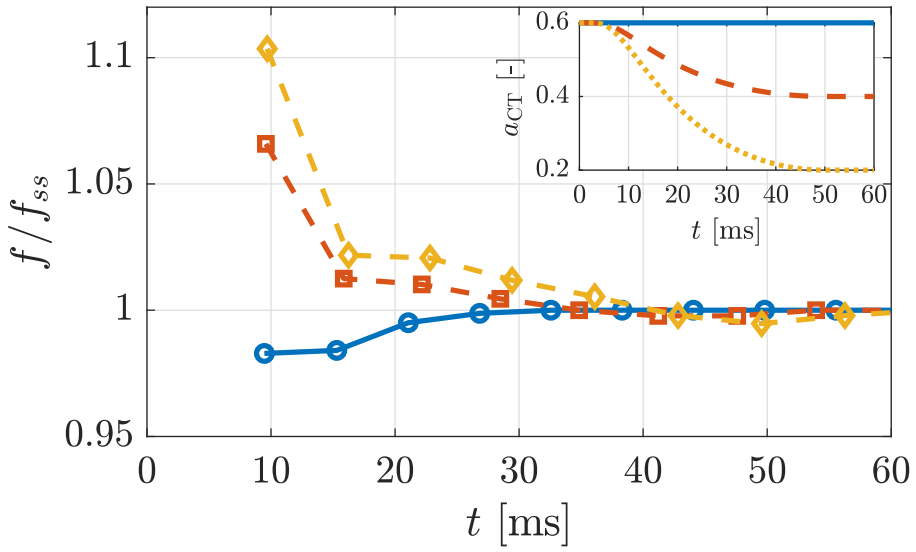


Figure 4.9: Time-series of fundamental frequency, normalized with respect to sustained phonation fundamental frequency f_{ss}, with a_{TA} = 0.2, $a_{\mathrm{CT,i}}$ = 0.6, and different final cricothyroid activation levels. The time-series of cricothyroid activation are shown in the inset.

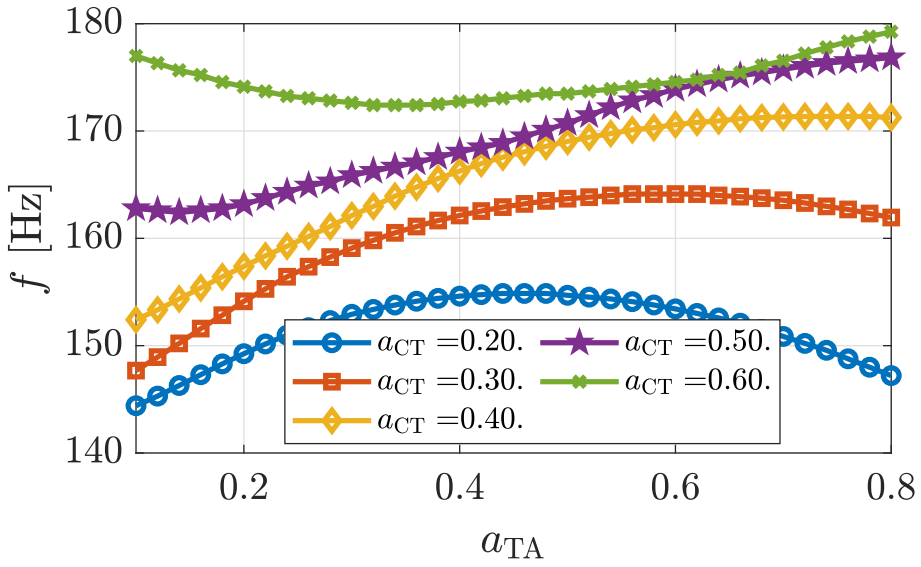


Figure 4.10: Fundamental frequency as a function of thyroarytenoid muscle activation level with fixed cricothyroid activation levels during steady phonation.

levels are selected such that states wherein CT and TA activation have agonistic and antagonistic influences on fundamental frequency in the steady state analysis are both represented (*i.e.*, the relations between a_{TA} and $a_{\mathrm{CT,i}}$ is inverse and the relation between a_{TA} and $a_{\mathrm{CT,f}}$ is proportional, according to our quasi-steady analysis). We record the fundamental frequency of the first cycle, $f_{\mathrm{c},1}$ (the inverse of the time difference between the first two detected peaks of the glottal area waveform) and normalize it with respect to the steady-state fundamental frequency during sustained phonation, f_{ss}. Figure 4.11 displays a contour plot of $f_{\mathrm{c},1}/f_{ss}$ as a function of $a_{\mathrm{CT,f}}$ and a_{TA}. The figure shows that increasing a_{TA} from 0.1 to 0.3 results in the initial normalized fundamental frequency decreasing for decaying CT profiles with $a_{\mathrm{CT,i}} = 0.6$ and $a_{\mathrm{CT,f}} \in [0.2, 0.4]$, in agreement with our quasi-steady analysis.

4.3.3 Comments on relations to empirical observations

As seen in Figure 2.12, empirical data from human studies indicate that fundamental frequency decays during the onset of vowels preceded by voiceless consonants, with higher onset relative fundamental frequency values in the case of adult speakers with healthy voices

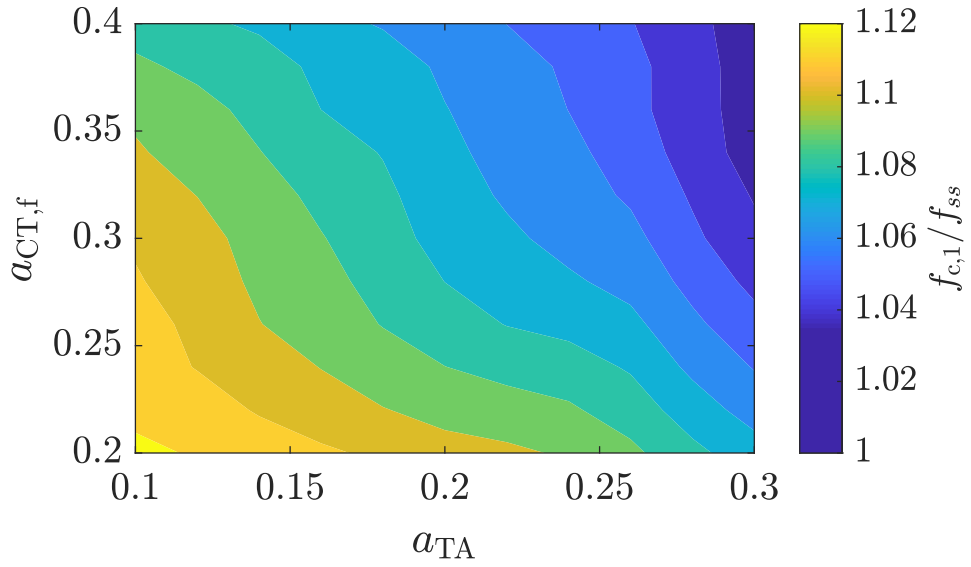


Figure 4.11: Contour plot of initial fundamental frequency normalized by steady state frequency, $f_{\mathrm{c},1}/f_{ss}$, as a function $a_{\mathrm{CT,f}}$ and a_{TA}, where $a_{\mathrm{CT,i}} = 0.6$.

in comparison with adult speakers with hyperfunctional voices [136][9], where speakers with hyperfunctional voices often exhibit excessive and/or imbalanced muscular forces [47, 48]. Figures 4.8 and 4.9 show that when the magnitude of the drop in CT activation during onset is sufficiently high there is a decrease in normalized fundamental frequency that is correlated to the magnitude of the reduction in CT activation. This suggests that CT activation levels may be a factor underlying the differences between healthy speakers and speakers with vocal hyperfunction, wherein healthy speakers potentially produce relatively larger variations in the CT activation levels during phonation onset. Interestingly, the non-monotonic behavior (initial rise followed by a fall) in fundamental frequency present in some cases in Figure 4.8 is also observed in empirical studies, see for example [80, 76], and Figure B.3 in Appendix B.4. Moreover, Figure 4.11 indicates that higher TA activation levels result in lower normalized fundamental frequency values during phonation onset, at least initially. Thus, TA activation level is another potential factor that may underlie clinical observations, where our results suggest that speakers with vocal hyperfunction may produce higher TA activation levels, which also may explain the empirically observed lower initial relative fundamental frequency values [136].

[9]Relative fundamental frequency is a normalized measure of fundamental frequency that is functionally similar to the frequency patterns presented in Sections 4.3.1 and 4.3.2

4.4 Conclusion

In this chapter, we aimed to uncover some of the potential underlying mechanisms driving the observed differences in onset fundamental frequency patterns in different phonetic contexts, see Figure 2.12, where we resorted to theoretical and numerical analyses of single- and multi-mass models, respectively (see Section 4.1). We found that the increasing degree of VF collision during onset, associated with the rise in vibration amplitude, and/or the decrease in the neutral glottal gap, naturally gives rise to an increase in system fundamental frequency (see Section 4.2). Such an increase in fundamental frequency is experimentally observed during the onset of isolated/initial vowels, and in some instances of a vowel preceded by a voiced consonant (see Section 4.2.4). In these cases, laryngeal muscle tension may still play a role, but it is not a prerequisite for the observed behavior.

On the other hand, our analysis suggested that muscle activation is necessary to produce the observed decay in fundamental frequency evident in vowels preceded by voiceless consonants, since the system dynamics with all control factors fixed produce the opposite trend (see Figure 4.8(a)). In particular, our analysis indicated that reduction in fundamental frequency is due, in part, to a concomitant decrease in cricothyroid muscle activation during onset (see Section 4.3.1). Interestingly, the competing mechanisms of muscle activation and collision can lead to a frequency pattern that initially rises then drops, which has been observed in experimental studies of onset fundamental frequency (see Section 4.3.1). The magnitude of the reduction in cricothyroid muscle activation was found to be a potential factor underlying the difference in relative fundamental frequency between healthy and hyperfunctional voices during the onset of vowels preceded by voiceless consonants (see Section 4.3.3). Furthermore, our investigation suggested that increased thyroarytenoid muscle activation mitigates the variation in relative fundamental frequency caused by a decrease in cricothyroid muscle activation, which may also contribute to the experimentally observed differences between hyperfunctional and normal phonation (see Sections 4.3.2 and 4.3.3).

Chapter 5

Modelling the influence of the extrinsic musculature on phonation

We saw in Section 2.4 the complexity of the anatomical structure of the extrinsic laryngeal muscles and their involvement in phonation. We also observed in Section 2.5.1 the relevance of these muscles in the case of patients with VH. As highlighted in Section 2.4, the findings in the literature regarding the role of extrinsic laryngeal muscles in phonation are conflicting, which necessitates further extensive research to elucidate the mechanisms of which extrinsic laryngeal muscles alter voice production in healthy speakers and patients with VH.

In this chapter, we explore, via means of simplified theoretical and numerical analyses of the **net** effects of extrinsic muscles, how the pulling action of extrinsic muscles exerted on the larynx affects the mechanics of phonation. We illustrate how the magnitude, direction, and location of the net force, corresponding to the activation of extrinsic laryngeal muscles, alters VF tension and phonation fundamental frequency. We combine the findings from the aforementioned simplified analysis with anatomical and clinical data to pinpoint potential roles of particular extrinsic muscles in normal and hyperfunctional phonation.

The analyses presented herein are adapted from a published research paper in the Biomechanics and Modelling in Mechanobiology journal under the title "Modeling the influence of the extrinsic musculature on phonation" [126], where Mohamed Ahmed Serry is the first author, and Gabriel A. Alzamendi, Matías Zañartu, and Sean D. Peterson are coauthors. M.A.S. and S.D.P. conceived of the study. M.A.S. and G.A. developed the model. M.A.S. generated preliminary results and figures. All authors analyzed and interpreted results. M.A.S. wrote the first draft of the article. All authors reviewed and

approved the manuscript. This work was supported by the NIDCD of the NIH under Award No. P50DC015446, and ANID BASAL FB0008. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Organization of this chapter is as follows: model development is discussed in Section 5.1; a simplified theoretical analysis of the adapted model is introduced in Section 5.2; numerical simulation results are presented in Section 5.3; the discussion of the results and their implications is presented in Section 5.4; and the salient conclusions are presented in Section 5.5.

5.1 Modelling

In an effort to capture the net effect of the extrinsic muscles on VF vibration, we propose the following modelling framework: we adapt the phonation model of Alzamendi et al. [4], which integrates the two-dimensional (2D) posturing model introduced by Titze and Hunter [160] and the triangular body-cover model (TBCM) developed by Galindo et al. [32]. As explained in [4], this model includes independent activation of all five intrinsic muscles. We utilize the acoustic, aerodynamic, and TBCM modules of [4] and introduce modifications to the posturing module to include the thyroid cartilage pulling effect due to extrinsic muscles. The model assumes the configuration of the VFs and the position of the arytenoid cartilages to be symmetric about the mid-coronal plane, and that the activation levels of the intrinsic laryngeal muscles are identical on the right and left sides. Therefore, only one VF, in addition to associated intrinsic and extrinsic muscles, is considered in our analysis.

According to the posturing model in [160], the VF strain, ϵ, is linearly decomposed into three components: strain due to VF adduction, ϵ_a; strain due to rotation of the CT joint, ϵ_r; and strain due translation of the CT joint, ϵ_t. That is,

$$\epsilon = \epsilon_a + \epsilon_t + \epsilon_r. \tag{5.1}$$

The adductory strain ϵ_a is obtained from the translation and rotation of the arytenoid cartilages. We refer the readers to [153] and [160] for detailed discussion and governing equations for ϵ_a.

In this work, we incorporate the net pulling effects from the neck muscles on the thyroid cartilage by modifying the rotational and translational stains, ϵ_r and ϵ_t . In our derivation, similar to [160, 153], the rotational and translational strains are assumed to be decoupled;

that is, the CT joint undergoes only rotation (translation) when estimating rotational (translational) strain. In the subsequent discussion, we derive static equations for the rotational and translational strains, as we are concerned with sustained phonation scenarios, where the mechanical and geometrical properties of the VFs in such scenarios are assumed constant over a sufficiently long time window. We note that the static equations can be easily extended to model dynamic changes in strain by adopting the derivations in [153, Chapter 3].

5.1.1 Rotational strain

Derivation of the rotational strain follows [161] and [153]. The strain ϵ_r is related to VF elongation due to CT joint rotation, ΔL_r, through the relation

$$\epsilon_r = \Delta L_r / L_0, \tag{5.2}$$

where L_0 denotes the resting VF length. CT joint rotation is modulated by the forces from the CT muscle, the TA muscle, and the vocal ligament, in addition to the net larynx pulling force from the extrinsic muscles (see Figure 5.1), and is given by the static moment equation

$$k_r \theta = r_{\mathrm{CT}} F_{\mathrm{CT}} - r_{\mathrm{p}} F_{\mathrm{p}} - r_{\mathrm{TA}} F_{\mathrm{TA}}, \tag{5.3}$$

where k_r is the rotational stiffness of the CT joint, θ is the rotational displacement of the CT joint, and F_{CT} and F_{TA} are the forces from the CT muscle and the TA muscle and vocal ligament. The variable F_{p} represents the net force on the thyroid cartilage due to the extrinsic muscles. The terms r_{CT}, r_{TA}, and r_{p} denote the moment arms corresponding to the respective forces. Elongation due to rotation, ΔL_r, can then be obtained by using a small angle approximation as

$$\Delta L_r \approx r_{\mathrm{TA}} \theta. \tag{5.4}$$

5.1.2 Translational strain

Similar to the rotational strain derivation, strain due to translation of the CT joint, ϵ_t, is estimated by

$$\epsilon_t = \Delta L_t / L_0, \tag{5.5}$$

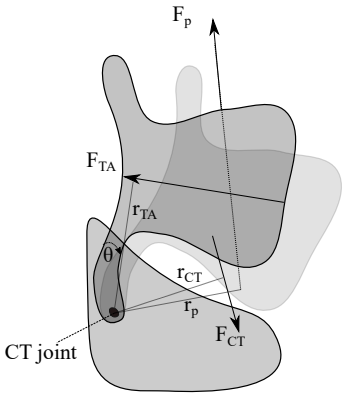


Figure 5.1: Free body diagram of the rotation of the CT joint.

where ΔL_t is the elongation of the VF due to translation of the CT joint parallel to the VF plane[1], which is given by the static equation

$$k_t \Delta L_t = \alpha_{\mathrm{CT}} F_{\mathrm{CT}} - F_{\mathrm{TA}} + \alpha_{\mathrm{p}} F_{\mathrm{p}}, \tag{5.6}$$

where k_t denotes the translational stiffness of the CT joint, and α_{CT} and α_{p} denote the direction cosines of the CT muscle and the resultant force due to extrinsic muscles, respectively (see Figure 5.2). As in [4], the forces F_{TA} and F_{CT} are related to VF strain; we refer interested readers to [4, 153] for detailed discussions.

5.1.3 Net extrinsic muscle force: magnitude, direction, and application location

In general, the force F_{p} and its direction cosine α_{p} and moment arm r_{p} are dependent on extrinsic muscle activation and the displacement of the thyroid cartilage with respect to the surrounding laryngeal structure. For simplicity, we set F_{p}, its direction ψ, defined with respect to a line perpendicular to the plane of the VFs (see Figure 5.3), and its

[1] We define the VF plane to be the plane parallel to the superior surfaces of the VFs.

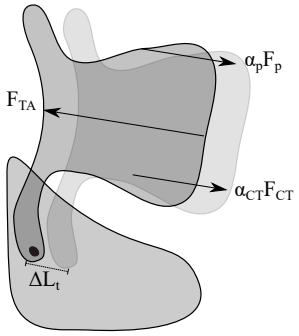


Figure 5.2: Free body diagram of the translation of the CT joint.

location of application κL_{thy} from the thyroid posterior margin, as prescribed parameters and investigate their effects on the mechanics of phonation. Herein, L_{thy} is the length of the lamina of the thyroid cartilage along the plane of the VFs and κ denotes the normalized location of the pulling force from the posterior margin. We assume that the direction ψ, once specified, remains unchanged during rotation of the CT joint. Moreover, we assume that F_{p} is applied at the superior end of the thyroid lamina. The distance from the CT joint (which is assumed to be located approximately at the tip of the inferior cornu) to the superior end of the thyroid cartilage lamina in the direction perpendicular to the VF plane is denoted by h_{thy}. Using trigonometry, the direction cosine and moment arm of F_{p} are given by

$$\alpha_{\text{p}} = \sin(\psi), \tag{5.7}$$

and

$$r_{\text{p}} = \kappa L_{\text{thy}} \cos(\psi) - h_{\text{thy}} \sin(\psi), \tag{5.8}$$

respectively. Note that the direction cosine α_{p} is a function of the angle ψ only, whereas the moment arm r_{p} is a function of ψ, κ, and the lengths L_{thy} and h_{thy}.

To the best of our knowledge, there exist no direct measurements of L_{thy} and h_{thy} in the literature as dimensions of the thyroid cartilage are typically collected after removing the VFs and surrounding tissues and cartilages; however, available laryngeal measurements

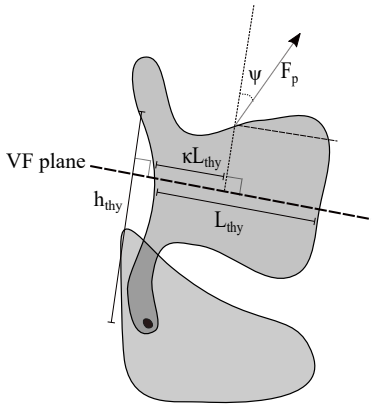


Figure 5.3: Schematic plot of the pulling force, F_{p}, and its direction and location.

can be used to obtain rough estimates of these lengths. Based on measurements of male cadaver thyroid cartilages from a European population [69], L_{thy} can be estimated as the projection of the superior width of the thyroid cartilage (denoted BE in [69]) on the medial plane, and h_{thy} can be estimated as the posterior height of the thyroid cartilage (denoted AD in [69]) minus the length of the superior horn (denoted AB in [69]). This yields values of $L_{\mathrm{thy}} \approx 32\,\mathrm{mm}$ and $h_{\mathrm{thy}} \approx 27\,\mathrm{mm}$ for male specimens; we adopt these values in this study. For female specimens, the values of L_{thy} and h_{thy}, based on the measurements in [69], are estimated to be approximately 22 mm for both lengths.

It has been found that the typical length of the TH muscle in a male cadaver subject is about 34 mm [168], and that, in swallowing maneuvers, the hyoid bone moves in the anterior-posterior direction about 1 cm [43]. Considering these measurements and assuming that at rest the hyoid bone is almost perpendicular to the plane of VFs, we estimate the angle ψ to be of the order of 0.1 rad in scenarios wherein the supra- and infra-hyoid muscles are jointly activated. In order to extensively investigate the influence of the thyroid cartilage pulling in the superior direction on phonation mechanics, we consider ψ values in the larger range of $[-1, 1]$ rad. On the other hand, to qualitatively assess the effect of pulling the thyroid cartilage in the inferior direction exerted by some neck muscles (e.g., ST), we additionally consider the ψ-range $[\pi - 1, \pi + 1]$ rad.

The cross-sectional area of the TH muscle in a male cadaver specimen is approximately 0.37 cm^2 [168]. By considering the average sarcomere length of dissected neck muscles of 2.6 μm and the method proposed in [168, Equations 7.2 and 7.3], the isometric maximum force in the TH muscle is estimated to be of the order of 10 N. This is of the same order of magnitude of estimates of isometric maximum forces for other neck muscles used in musculoskeletal modeling [100, Table 2]. As we expect that neck muscles will not be activated to their maximum capacity in phonation, we consider the range $[0, 1]$ N for the force magnitude F_{p} in this study.

As the moment arm r_{p} depends on the location of the pulling force and this location may change depending on activation levels of the various extrinsic muscles, we consider $\kappa \in [0, 1]$. That is, the location of the resultant pulling force from the extrinsic muscles may be anywhere along the length of the thyroid cartilage.

5.2 Net Extrinsic Muscle Effect: Analytical Insights

To gain insight into the influence of thyroid cartilage pulling on VF strain we define

$$\begin{aligned}\tilde{\epsilon}_{\text{p}} &= \left(-\frac{r_{\text{TA}}}{k_r L_0} r_{\text{p}} + \frac{1}{k_t L_0}\alpha_{\text{p}}\right) F_{\text{p}} \\ &= \frac{1}{L_0}\left(\left[\frac{r_{\text{TA}} h_{\text{thy}}}{k_r} + \frac{1}{k_t}\right]\sin(\psi) - \frac{r_{\text{TA}}\kappa L_{\text{thy}}}{k_r}\cos(\psi)\right) F_{\text{p}}\end{aligned} \tag{5.9}$$

to be the strain accumulated in the VFs due to F_{p}; (in other words, $\tilde{\epsilon}_{\text{p}}$ is the contribution of F_{p} to ϵ_r plus the contribution of F_{p} to ϵ_t). Assuming that the pulling effects from the neck muscles on VF strain do not alter the applied forces from the intrinsic muscles, and that the rotational and translational stiffnesses, k_r and k_t, are displacement-independent, $\tilde{\epsilon}_{\text{p}}$ positively correlates to VF strain, ϵ when activation of intrinsic laryngeal muscles is fixed[2]. In this section, we consider thyroid cartilage pulling in all directions (i.e., $\psi \in [-\pi/2, 3\pi/2]$ rad) while adopting the force magnitude range $[0, 1]$ N and κ in the range $[0, 1]$.

First, it can be deduced from Equation (5.9) that, when $\cos(\psi) \geq 0$ ($\psi \in [-\pi/2, \pi/2]$ rad; i.e., upward pulling), there is a negative correlation between $\tilde{\epsilon}_{\text{p}}$ and κ. This correlation becomes positive for $\psi \in [\pi/2, 3\pi/2]$ (downward pulling).

Next, let us consider the range of ordered pair values $(\psi, F_{\text{p}}) \in [-\pi/2, \pi/2]$ rad $\times [0, 1]$ N, which corresponds to thyroid cartilage pulling in the superior direction. To investigate

[2]These simplifying assumptions are not adopted in the full numerical phonation model used in the simulations presented in Section 5.3.

Table 5.1: Some laryngeal parameter values (rounded), based on measurements from canine and male cadaver specimens [153, 69].

Parameter	Value
r_{TA}	1.6 cm [153]
h_{thy}	27 mm (based on measurements from [69])
L_{thy}	32 mm (based on measurements from [69])
L_0	1.6 cm [153]
k_t	500 N/m [153] (zeroth order)[3]
k_r	0.05 N.m/rad [153] (zeroth order)

the sensitivity of $\tilde{\epsilon}$ to F_{p}, we compute the partial derivative of $\tilde{\epsilon}_{\mathrm{p}}$ with respect to F_{p}, $\partial_{F_{\mathrm{p}}}\tilde{\epsilon}_{\mathrm{p}}$, which yields

$$\partial_{F_{\mathrm{p}}}\tilde{\epsilon}_{\mathrm{p}} = \frac{1}{L_0}\left(\left[\frac{r_{\mathrm{TA}}h_{\mathrm{thy}}}{k_r} + \frac{1}{k_t}\right]\sin(\psi) - \frac{r_{\mathrm{TA}}\kappa L_{\mathrm{thy}}}{k_r}\cos(\psi)\right). \tag{5.10}$$

We see that the effect of the force magnitude F_{p} is direction-dependent, that is, it has a functional dependence on ψ. We can deduce from Equation (5.10) that for $\pi/2 \geq \psi \geq \psi_c$, where

$$\psi_c := \tan^{-1}\left(\frac{k_t r_{\mathrm{TA}}\kappa L_{\mathrm{thy}}}{k_t r_{\mathrm{TA}}h_{\mathrm{thy}} + k_r}\right), \tag{5.11}$$

changes in $\tilde{\epsilon}_{\mathrm{p}}$ are positively correlated to the force magnitude F_{p} ($\partial_{F_{\mathrm{p}}}\tilde{\epsilon}_{\mathrm{p}}$ is non-negative). On the other hand, if $-\pi/2 \leq \psi \leq \psi_c$, the correlation becomes negative ($\partial_{F_{\mathrm{p}}}\tilde{\epsilon}_{\mathrm{p}} \leq 0$). Note that the angle ψ_c is positively correlated with κ, implying that this critical angle increases as the pulling force location moves anteriorly.

Using clinical measurements of canine and human male cadaver models, see Table 5.1, and setting $\kappa = 0.5$, we estimate that

$$\frac{k_t r_{\mathrm{TA}}\kappa L_{\mathrm{thy}}}{k_t r_{\mathrm{TA}}h_{\mathrm{thy}} + k_r} \approx 0.51, \tag{5.12}$$

which, according to Equation (5.11), yields

$$\psi_c \approx 0.47 \text{ rad}. \tag{5.13}$$

Considering now the sensitivity of $\tilde{\epsilon}_{\mathrm{p}}$ to ψ, the gradient $\partial_{\psi}\tilde{\epsilon}_{\mathrm{p}}$ is given by

$$\partial_{\psi}\tilde{\epsilon}_{\mathrm{p}} = \frac{1}{L_0}\left(\left[\frac{r_{\mathrm{TA}}h_{\mathrm{thy}}}{k_r} + \frac{1}{k_t}\right]\cos(\psi) + \frac{r_{\mathrm{TA}}\kappa L_{\mathrm{thy}}}{k_r}\sin(\psi)\right)F_{\mathrm{p}}. \tag{5.14}$$

Equation (5.14) implies that for $\pi/2 \geq \psi \geq \bar{\psi}_c$, where

$$\bar{\psi}_c := -\tan^{-1}\left(\frac{k_t r_{\mathrm{TA}} h_{\mathrm{thy}} + k_r}{k_t r_{\mathrm{TA}} \kappa L_{\mathrm{thy}}}\right) = -\frac{\pi}{2} + \psi_c, \tag{5.15}$$

the strain $\tilde{\epsilon}_{\mathrm{p}}$ increases with ψ for fixed F_{p}; on the other hand, the strain decreases with ψ, for fixed F_p, when $-\pi/2 \leq \psi \leq \bar{\psi}_c$. Equation (5.15) indicates an increase in $\bar{\psi}_c$ when κ increases and that the difference between ψ_c and $\bar{\psi}_c$ is always $\pi/2$. The estimate in Equation (5.12), where κ is set to 0.5, implies that

$$\bar{\psi}_c \approx -1.1 \text{ rad}. \tag{5.16}$$

This indicates that for the range $\psi \in [-1, 1]$ rad, $\tilde{\epsilon}_{\mathrm{p}}$ is always positively correlated with ψ.

Due to the structure of Equation (5.9), the above findings can be easily extended to the range $(\psi, F_{\mathrm{p}}) \in [\pi/2, 3\pi/2] \times [0, 1]$, which corresponds to thyroid cartilage pulling with an inferior component. For inferior pulling, $\tilde{\epsilon}_{\mathrm{p}}$ is positively correlated to F_{p} in the ψ-range $[\pi/2, \psi_c + \pi]$ rad, whereas the correlation is negative for $\psi \in [\psi_c + \pi, 3\pi/2]$ rad. The strain $\tilde{\epsilon}_{\mathrm{p}}$ is positively correlated to ψ in the range $[\pi/2, \pi + \tilde{\psi}_c]$ rad and negatively correlated in the range $[\pi + \tilde{\psi}_c, 3\pi/2]$ rad.

The effects of F_{p} and ψ on $\tilde{\epsilon}_{\mathrm{p}}$ in the superior and inferior pulling ranges are depicted in Figure 5.4. This figure and the associated analysis show clearly that in the vicinity of the critical angles, ψ_c, $\psi_c + \pi$, $\bar{\psi}_c$, and $\bar{\psi}_c + \pi$, the influence of net neck muscle pulling can result in either an increase or decrease in VF strain. Moreover, the influence of the parameter κ on the critical angles ψ_c and $\bar{\psi}_c$ is depicted in Figure 5.5, which illustrates the positive correlation between κ and the aforementioned angles. In Section 5.3, we extensively investigate through numerical simulations the influence of the force F_{p}, its direction ψ, and its location parameterized by κ, on the VF posture and on different voice measures.

5.3 Numerical Simulations

In this section, we conduct numerical simulations using the modified TBCM (see Section 5.1), where the values of F_{p}, κ, and ψ are set in the ranges $[0, 1]$ N, $[0, 1]$, and $[-1, 1] \cup [\pi - 1, \pi + 1]$ rad, respectively. The focus herein is to investigate the influence of the pulling force parameters on the biomechanics of posturing and phonation.

The remaining parameters of the TBCM model are set as follows: we consider intrinsic muscle activation values in the ranges $a_{\mathrm{LCA}} = 0.5$, $a_{\mathrm{IA}} = 0.5$, $a_{\mathrm{PCA}} = 0$, $a_{\mathrm{CT}} \in [0.2, 0.6]$,

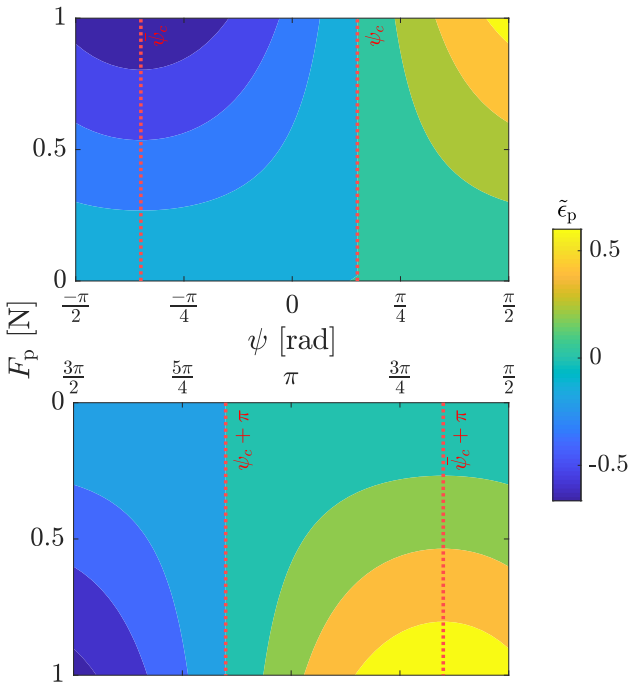


Figure 5.4: Contour plot of the strain $\tilde{\epsilon}_\mathrm{p}$ as a function of F_p and ψ from Equation (5.9): parameter values of Equation (5.9) are depicted in Table 5.1 and $\kappa = 0.5$. Scenarios for upward (superior) and downward (inferior) pulling forces are shown in the top and the bottom panels, respectively.

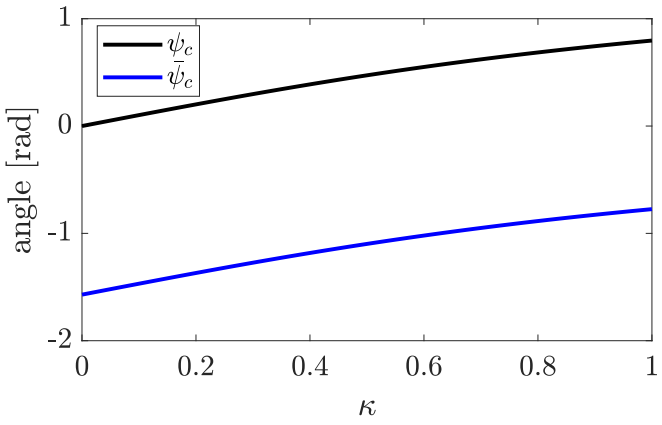


Figure 5.5: Plots of the critical angles ψ_c and $\bar{\psi}_c$ as functions of κ from Equations (5.11) and (5.15): adopted parameter values are listed in Table 5.1.

and $a_{\text{TA}} \in [0.2, 0.6]$, where a_{LCA}, a_{IA}, a_{PCA}, a_{CT}, and a_{TA} denote the non-dimensional normalized muscle activations for the LCA, IA, PCA, CT, and TA muscles, respectively. The above ranges correspond to phonation scenarios with low/normal activation levels of the CT and TA muscles, and fully (or near fully) adducted VFs (modal phonation). Moreover, we set the subglottal pressure P_s = 800 Pa, which is within the typical range during speech [192, 27]. The passive model parameters correspond to a male subject, and acoustics are modelled using the wave reflection analog [64, 140], where subglottal and supraglottal acoustic tracts are coupled with the TBCM. Similar to [33, 186], the subglottal tract area function is adapted from respiratory system measurements of human cadavers [178], covering only the trachea and bronchi. Moreover, a supra-glottal area function corresponding to a male tract configured to produce the American English open back unrounded vowel /*a*/ is adopted [141].

Each phonation simulation is run for 200 ms, with a sampling frequency of 44100 Hz. The first 100 ms of each simulation is discarded to remove transient effects. From the sustained oscillations portions, we compute the fundamental frequency, f_o, and the sound pressure level, SPL. Additionally, from the posturing computations, we estimate the VF strain, ϵ, and the glottal angle, θ_G, which is defined as the total included angle between the two VFs at the anterior attachment point.

Figure 5.6 depicts plots of vocal fold strain ϵ, sustained phonation fundamental frequency f_o, glottal angle θ_G, and sound pressure level SPL, as functions of the parameter κ, which specifies the location of the pulling force. The figure shows that, for the case

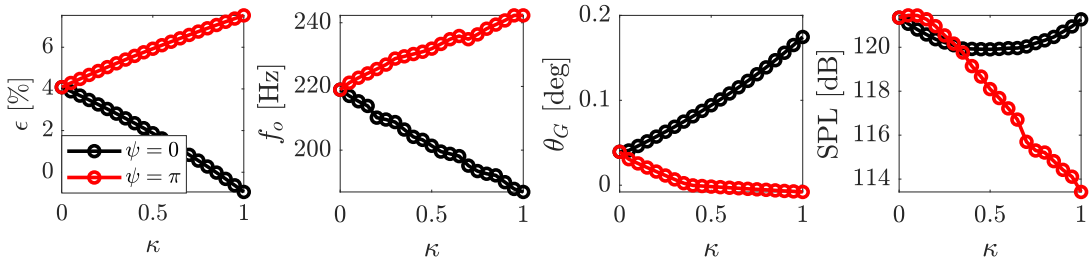


Figure 5.6: Vocal fold strain ϵ (first column), sustained phonation fundamental frequency f_o (second column), glottal angle θ_G (third column), and sound pressure level SPL (fourth column), as functions of κ, where $a_{\mathrm{CT}} = a_{\mathrm{TA}} = 0.4$ and $F_{\mathrm{p}} = 1$ N for ψ: $\psi = 0$ rad (black) and $\psi = \pi$ rad (red).

of upward pulling ($\psi = 0$), increasing κ, which corresponds to moving the pulling force location anteriorly along the thyroid cartilage length, leads to decreasing VF tension and fundamental frequency, in agreement with the analysis in Section 5.2, and slight variations in the glottal angle and SPL. Moreover, the figure indicates that, for the case of downward pulling ($\psi = \pi$), increasing κ leads to a general increase in VF tension and fundamental frequency, also in alignment with the analysis in Section 5.2, and insignificant variation in the glottal angle. However, for the case of downward pulling, a notable decrease in SPL is observed. We speculate that this decrease may be due to the fact that increased VF tension leads to VF oscillations of small amplitude, which may induce low-amplitude acoustic pressure.

Figure 5.7 displays contour plots of the VF strain, glottal angle, phonation fundamental frequency, and SPL as functions of F_{p} and ψ for various intrinsic muscle activation values and $\kappa = 0.5$. The first column shows that for all simulations when the angle ψ is in the approximate range [-1,0.5] rad the strain is negatively correlated to the force magnitude F_{p}, with the correlation becoming positive for the approximate range [0.5, 1] rad. This agrees reasonably with predictions from the theoretical analysis in Section 5.2. For the inferior thyroid cartilage pulling scenarios ($\psi \in [\pi - 1, \pi + 1]$ rad), Figure 5.7 shows that the VF strain increases with F_{p} in the approximate range $[\pi - 1, \pi + 0.4]$ rad. When the inferior pulling force has a sufficiently large posterior component (ψ is in the approximate range $[\pi + 0.4, \pi + 1]$ rad), VF strain decreases as F_{p} increases. In addition, the first column of Figure 5.7 shows that when ψ is in the range $[-1, 1]$ rad, VF strain is positively correlated to the angle ψ (for fixed F_{p}), which agrees with our theoretical predictions in Section 5.2. For the case of inferior pulling, ψ negatively correlates to the VF strain. These results indicate that pulling the thyroid cartilage superiorly, or superiorly and posteriorly, leads to

VF shortening. When superior pulling incorporates an anterior component with sufficiently large tilting angle, the VFs are elongated. In addition, the results indicate that inferior pulling of the thyroid cartilage typically lengthens the VFs, except in situations when the pulling has a sufficiently large posterior component, which causes the VFs to shorten and the strain to decrease.

Unsurprisingly, the second column of Figure 5.7 shows that fundamental frequency exhibits similar trends to VF strain, as fundamental frequency typically correlates with VF elongation (see, e.g., [187]). This implies that thyroid cartilage pulling in the superior or the superior-posterior directions causes a drop in fundamental frequency, which agrees with the analysis in [53]. Thyroid cartilage pulling in the superior-anterior (with sufficiently large anterior component), inferior, or inferior-anterior directions causes fundamental frequency to rise, which agrees qualitatively with the findings in [54, 53]. The third column of Figure 5.7 shows the influence of the pulling force F_{p} and its direction ψ on glottal angle. It can be seen that for the approximate ranges for ψ, $[-1, 0]$ rad and $[\pi+0.5, \pi+1]$ rad, increasing the force F_{p} leads to increasing the glottal angle; that is, the net extrinsic muscle action acts to abduct the VFs. In contrast, for the approximate ranges of ψ, $[0.5, 1]$ rad and $[\pi-1, \pi]$ rad, the glottal angle decreases with the force F_{p}, indicating an adducting action. We note, however, that for the considered ranges of F_{p} and ψ, changes in glottal angle are quite modest. The fourth column of Figure 5.7, which presents SPL, exhibits complex trends. However, in most of the considered muscle activation scenarios, the changes in SPL are within 4 dB[4]. Care should be taken when investigating the independence/dependence between intrinsic and extrinsic muscles. Despite the qualitative similarities of contour plots across different intrinsic muscle activation combinations in Figure 5.7, the significance of extrinsic musculature effects on different laryngeal and phonation variables depends upon the activation levels of the intrinsic muscles.

Figure 5.8 displays contour plots of VF strain as a function of F_{p} and ψ for various values of κ. These contour plots are similar in trends to those presented in Figure 5.7, where the aforementioned variable changes behaviour in vicinity of critical angles (recall the definitions of ψ_c and $\tilde{\psi}_c$ in Section 5.2). Such similarities have also been observed for the variables f_o, θ_G, and SPL (not shown). The figure illustrates how the critical angles corresponding to ψ_c and $\bar{\psi}_c$ increase as κ increases, which is in agreement with the theoretical analysis in Section 5.2.

[4]The case of inferior thyroid cartilage pulling with $a_{\mathrm{CT}} = a_{\mathrm{TA}} = 0.4$ exhibits changes in SPL within 8 dB.

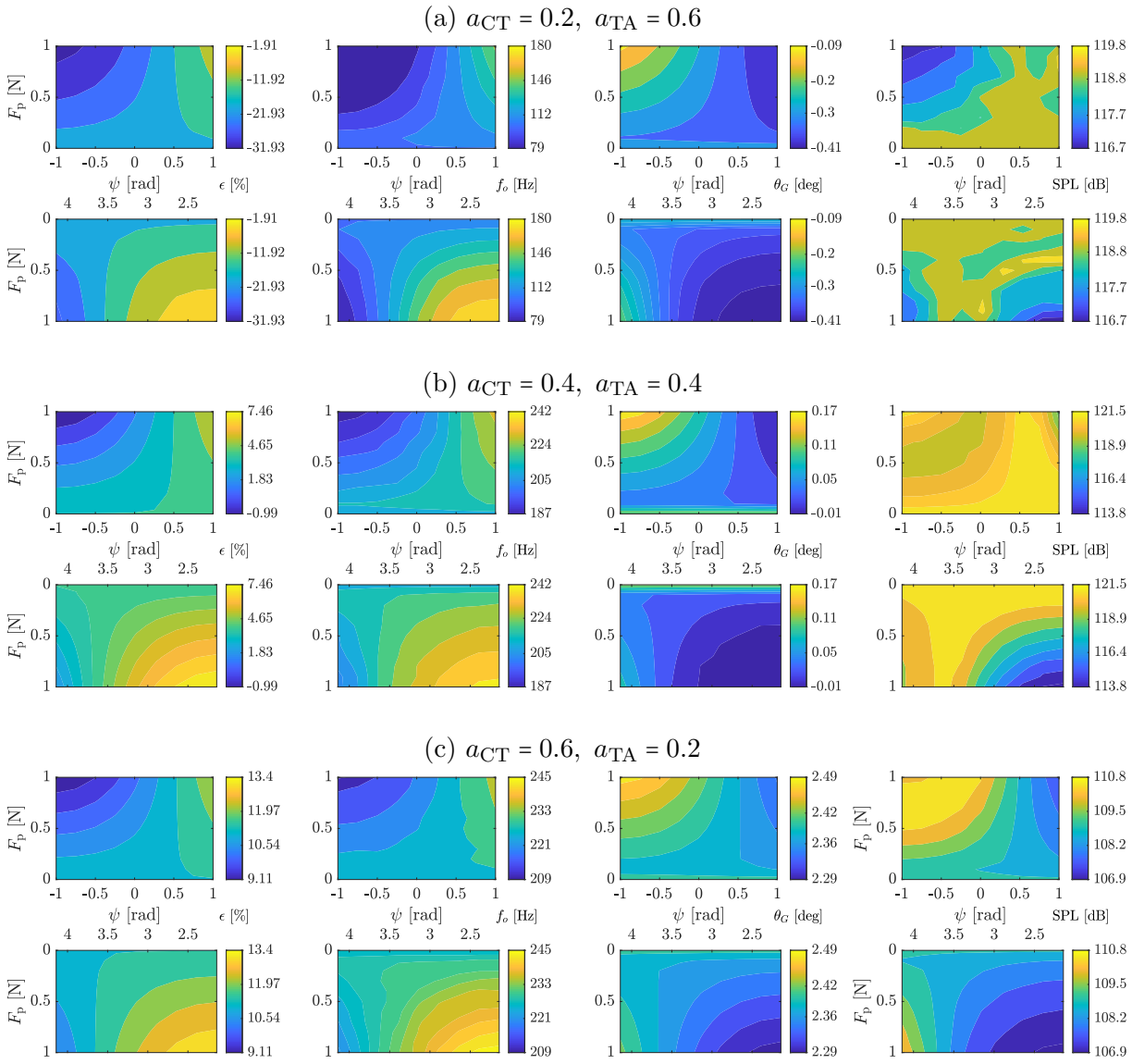


Figure 5.7: Contour plots of vocal fold strain ϵ (first column), sustained phonation fundamental frequency f_o (second column), glottal angle θ_G (third column), and sound pressure level SPL (fourth column), as functions of the force F_{p} and the angle ψ for different combinations of CT and TA activation levels with $\kappa = 0.5$: (a) $a_{\mathrm{CT}} = 0.2,\ a_{\mathrm{TA}} = 0.6$, (b) $a_{\mathrm{CT}} = 0.4,\ a_{\mathrm{TA}} = 0.4$, and (c) $a_{\mathrm{CT}} = 0.6,\ a_{\mathrm{TA}} = 0.2$.

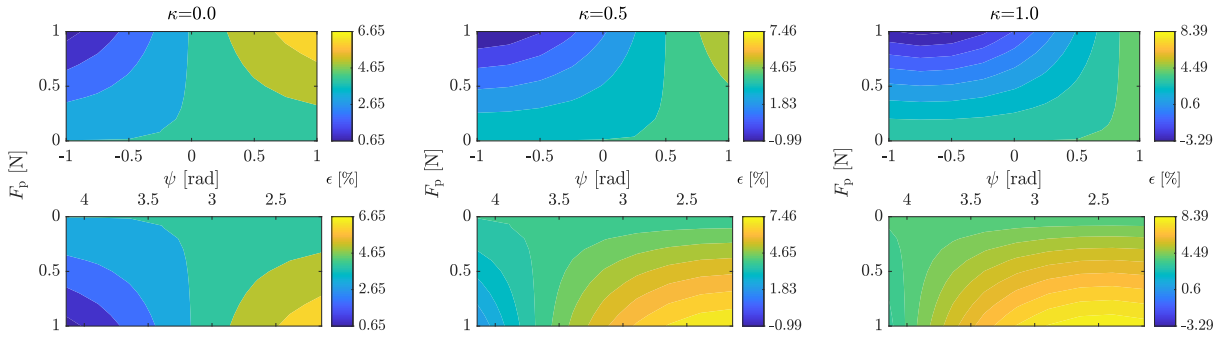


Figure 5.8: Contour plots of vocal fold strain ϵ as a function of the force F_{p} and the angle ψ for $a_{\mathrm{CT}} = 0.4$, $a_{\mathrm{TA}} = 0.4$ and different values of κ: $\kappa = 0$ (first column), $\kappa = 0.5$ (second column), and $\kappa = 1$ (third column).

5.4 Discussion

The results presented in Section 5.3 highlight how the magnitude, direction, and location of the net force on the thyroid cartilage due to the extrinsic laryngeal musculature influence glottal configuration and phonation characteristics, including fundamental frequency and SPL. Our analysis suggests that the suprahyoid muscles, which tend to pull the hyoid bone anteriorly (and thus pull the thyroid cartilage in the same direction), have the potential of increasing phonation fundamental frequency. Our analysis, in coordination with Figure 2.9, indicates that the GH muscle, which tends to move the hyoid bone anteriorly and superiorly [115], may act to raise fundamental frequency during phonation, agreeing with the analysis in [52]. Moreover, our analysis suggests that suprahyoid muscles that exert superior-posterior pulling (e.g., StyH [115]) on the thyroid cartilage may lower fundamental frequency, which agrees with the observations in [52] regarding the posterior motion of the hyoid bone and lowering f_o. In addition, our analysis shows that activation of extrinsic muscles that pull the thyroid cartilage superiorly (e.g., TH) may cause a decline in VF tension and f_o, which is in alignment with the analysis in [53]. Furthermore, our analysis indicates that neck muscles responsible of pulling the thyroid cartilage inferiorly (e.g., ST) may induce higher VF tension and consequently higher f_o, which agrees with the analysis in [53]. The potential roles of some extrinsic neck muscles, based on our analyses and anatomical data, are depicted in Figure 5.9. Finally, and most importantly, our analysis indicates that the effect of co-activated extrinsic laryngeal muscles depends upon the magnitude, direction, and location of the net force and subsequent net VF elongation, which all vary with different muscle co-activations, in alignment with the analysis in [174].

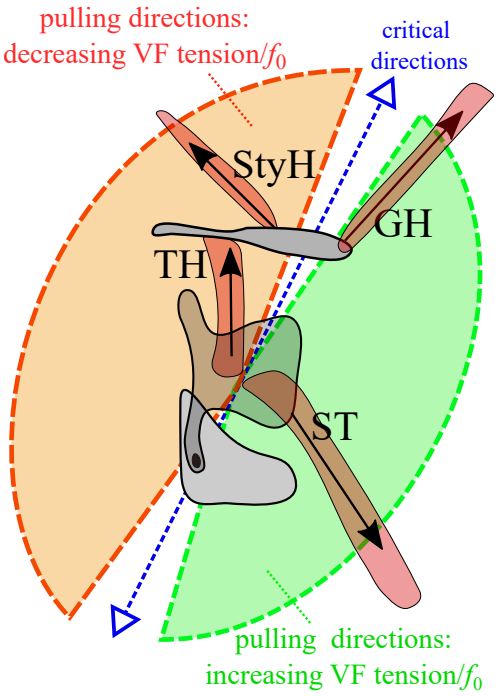


Figure 5.9: Potential roles of some extrinsic laryngeal muscles based on anatomical data and the analyses conducted in Sections 5.2 and 5.3.

This may explain, in part, inconsistencies appearing in the literature regarding the roles of extrinsic muscles in phonation. Given the experimental difficulties in assessing levels of muscle activity of human subjects, and in light of our findings on the dependence of VF strain on net force direction and application location, we postulate that inconclusive and contradictory findings from empirical studies may be associated with uncontrolled activation of additional strap muscles beyond those being studied in a given experimental campaign (see also the discussion in [174]). That is, in light of existing literature, our work suggests that to have a robust analysis and refined insight into the influence of neck muscles, the net effect of laryngeal muscles need be considered by investigating the larynx and hyoid bone movements, as in [54], and measuring the EMG activity of all extrinsic laryngeal muscles.

With regard to non-phonotraumatic hyperfunctional phonation, findings in the literature indicate excessive activation of neck muscles is common in patients with MTD [123]. These patients also typically suffer from abnormal phonation fundamental frequency, where some studies indicate that abnormality in f_o leans towards lower fundamental frequency [2], whereas recent studies indicate that MTD patients, especially female patients, exhibit higher f_o in comparison with control subjects [172]. In addition, clinical data reporting higher hyoid bone position in MTD subjects [82] imply relatively higher activation of suprahyoid muscles. These observations, in addition to the results obtained herein, suggest that for individuals with MTD with abnormally low f_o, the net pulling effect exerted on the thyroid cartilage by the suprahyoid muscles could potentially be in the superior or superior-posterior directions. This indicates that suprahyoid muscles contributing to pulling of the hyoid cartilage in these directions (e.g., StyH [115]) may exhibit higher activity in such patients. Besides the suprahyoid muscles, other extrinsic muscles that contribute to superior pulling of the thyroid cartilage, such as TH, may exhibit higher activity in MTD subjects, contributing to the clinically observed lower f_o. This partially agrees with the assessment in [5], who reported higher activation of the TH muscle in some MTD patients. In the case of MTD patients who exhibit abnormally high f_o [172], we hypothesize, based on the analysis in the current study, that suprahyoid muscles that are associated with superior-anterior pulling of the hyoid bone (e.g., the GH muscle [115]) may exhibit high activity in such patients. Finally, in the case of MTD subjects who suffer from tense neck muscles without abnormality in fundamental frequency, we speculate that an array of neck muscles are co-activated, such that their opposing effects neutralize each other (the net force is near zero), resulting in no significant changes in VF elongation.

There are a number of limitations in the present analysis that warrant consideration. The posturing model in this work considers only the net pulling effect of extrinsic muscles without incorporating the mechanics of individual muscles and connective tissue that will

influence hyoid bone movement. Furthermore, our model parameters are based on crude estimates of the pulling force direction and magnitude due to the paucity of relevant clinical measurements. Moreover, our modeling framework carries over the limitations of the phonation model of Alzemandi et al. [4], wherein for example, the direction cosines associated with intrinsic laryngeal muscles are assumed to be displacement-independent. Despite these limitations, the model is capable of capturing qualitatively some effects of extrinsic laryngeal muscles during phonation, in alignment with clinical observation. Still, the conjectures from the analyses above necessitate extensive clinical investigation into the involvement of neck muscles during phonation in vocally healthy subjects and patients with MTD.

5.5 Conclusion

In this chapter, we adapted a previously introduced muscle-controlled phonation model to incorporate the effects of extrinsic muscles on the mechanical and geometrical properties of the VFs, and subsequently applied the proposed model to elucidate the effects of larynx pulling on the mechanics of phonation. Theoretical analysis and simulations using the adapted model indicate that if the thyroid cartilage is pulled in the superior, superior-posterior, or inferior-posterior directions, VF tension tends to reduce and, subsequently, phonation fundamental frequency decreases. On the contrary, superior pulling with sufficiently large anterior component, inferior pulling, and inferior-anterior pulling, tend to raise the VF tension leading to fundamental frequency rise. Analysis also indicated a dependence of the VF tension level and fundamental frequency on the location of the net force application, where in the case of superior pulling, moving the force location anteriorly leads to a drop in tension and fundamental frequency, whereas a rise was found when pulling in the inferior direction. Comparison of the numerical results in this paper with reported clinical findings suggests potential involvements of particular neck muscles during phonation and the potential relevance of some neck muscles in patients with MTD.

Chapter 6

An Euler-Bernoulli-Type Beam Model of the Vocal Folds for Describing Curved and Incomplete Glottal Closure Patterns

From the discussion in Section 2.5.4, we observed how incomplete glottal closure can be associated with abnormal voice production and voice disorders (including VH) and how the manifestations of incomplete glottal closure exhibit a wide variety of glottal patterns (see Figure 2.13). We also noticed how some clinical/experimental and numerical studies explored the relations between laryngeal muscle activation and different incomplete glottal closure patterns; however, the physical mechanisms that induce the different glottal shapes remain poorly understood.

In this chapter, we show how different incomplete and curved glottal closure patterns can be induced by the layered structure of the VFs (see Section 2.2.1) and the reaction of this layered structure to the mechanical loads induced during different laryngeal maneuvers. To explore this interaction between the layered VF structure and mechanical loading induced during posturing, we propose an Euler-Bernoulli composite beam model of the VFs and integrate it with the muscle-controlled posturing model of Titze and Hunter [160] to gain refined physiological insights into how intrinsic laryngeal muscles may influence glottal geometry. Despite its simplicity, the model presented herein is capable of predicting the clinically observed glottal closure patterns depicted in Figure 2.13, where the model simplicity enables us to pinpoint potential mechanisms underlying the resulting glottal

shapes. These mechanisms are shown to be associated with the bending moments, induced inside the layers of the VFs during different posturing scenarios, which produce different convex and concave VF shapes.

The organization of this chapter is as follows: a detailed derivation of the composite beam VF model is introduced in Section 6.1, analysis of the derived model is conducted in Section 6.2, numerical simulations based on integrating the proposed beam model with the posturing model of Titze and Hunter [160] are presented in Section 6.3, the implications of the results of this work are discussed in Section 6.4, and the study is concluded in Section 6.5.

6.1 Model Development

The VFs are complex membrane-like structures, consisting of several superposed tissue layers with different mechanical and geometrical properties [153]. We hypothesize that this non-homogeneous structure underlies, in part, the different glottal shapes displayed in Figure 2.13. Herein, we propose a static Euler-Bernoulli-type composite beam model (see, e.g, [8]) for the VFs, with the different VF layers represented by strata in the beam. For simplicity, we assume symmetry with respect to the medial plane; hence we consider only one (the left) VF. A schematic representation of the composite beam model is given in Figure 6.1. The VF beam model consists of three layers: (1) the mucosa with depth d_{muc} and cross-sectional area A_{muc}, (2) the vocal ligament with depth d_{lig} and cross-sectional area A_{lig}, and (3) the thyroarytenoid (vocalis) muscle with depth d_{ta} and cross-sectional area A_{ta}. For the sake of compact presentation, we define the index set

$$\mathcal{I} = \{\text{muc}, \text{lig}, \text{ta}\}, \tag{6.1}$$

where muc, lig, and ta refer to the mucosa, ligament, and TA muscle tissue, respectively. Herein, we assume that each layer has a rectangular cross-section, and that the thicknesses of all layers are identical, and denoted by b; thus layer depth can be computed as A_i/b, $i \in \mathcal{I}$.

Our modelling framework assumes that VF deformation consists of (a) potentially large longitudinal stretching/compression with uniform strain, and (b) slight bending due to the induced moments inside the VF. Let L_0 denote the resting VF length and L denote the VF length after longitudinal deformation due to the associated nominal uniform strain $\bar{\varepsilon}$; that is,

$$L = (1 + \bar{\varepsilon})L_0. \tag{6.2}$$

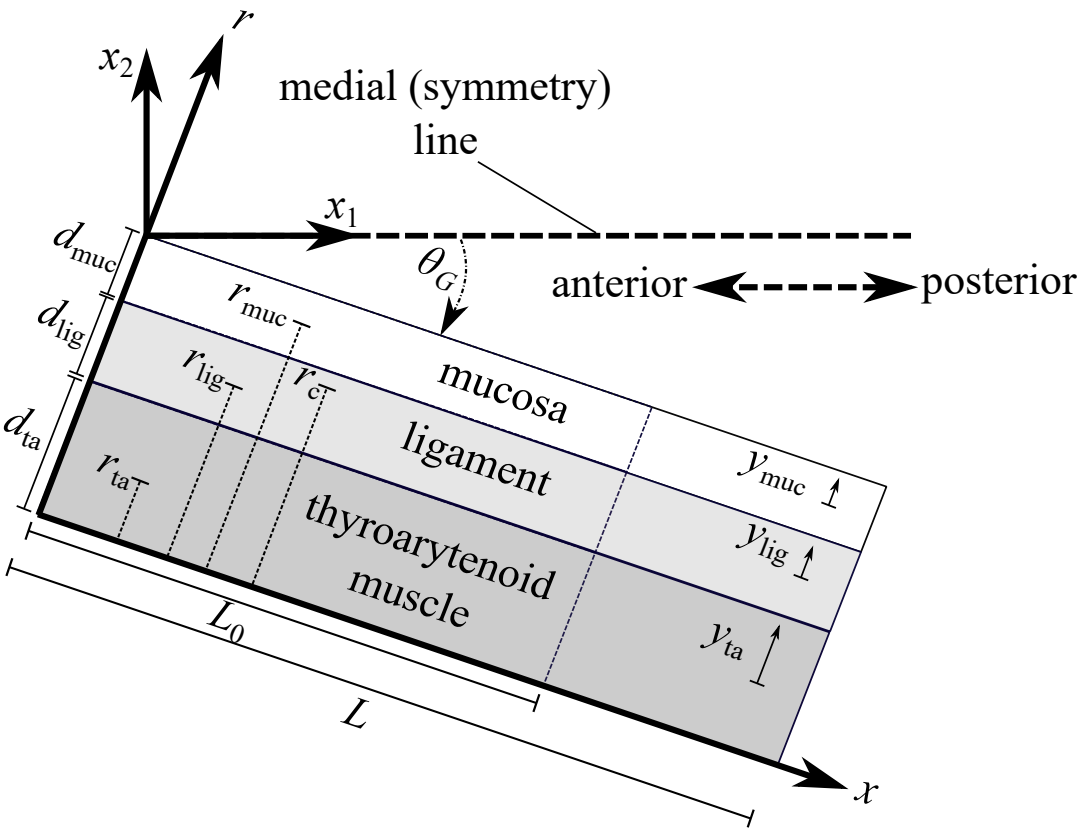


Figure 6.1: Schematic diagram (superior view) of the vocal fold composite beam model.

We assume that the nominal strain is known a priori, e.g., from the two-dimensional posturing model of Titze and Hunter [160], which incorporates the mechanics of the arytenoid cartilages and cricothyroid joints, and relates them to VF strain, see the discussion in Section 6.3.

Let $x \in [0, L]$ denote the position along the deformed VF configuration (after applying strain $\bar{\varepsilon}$) relative to the anterior VF margin, and r denote the depth position along the axis perpendicular to the VF axis relative to the base of the TA muscle (see Figure 6.1). Consider a plane VF cross-section at position x, and let y_{muc} denote the relative position along the r-axis with respect to the geometrical center of the mucosa (i.e., $y_{\text{muc}} = 0$ corresponds to the geometrical center of the mucousal cross-section). Similarly, let y_{lig} and y_{ta} be the counterparts of y_{muc} for the ligament and TA muscle, respectively (see Figure 6.1). Note that the range of y_i is $[-d_i/2, d_i/2],\ i \in \mathcal{I}$.

Let $w(x)$ denote the transverse deflection of the beam (in the r-direction). Moreover, let $u_i(x, y_i),\ i \in \mathcal{I}$, denote the longitudinal displacement associated with the i^{th} VF layer, where longitudinal displacements are with respect to the deformed VF configuration under $\bar{\varepsilon}$. In addition, let $\bar{u}_i(x) = u_i(x, y_i = 0),\ i \in \mathcal{I}$, denote the longitudinal displacement at the center of the i^{th} layer. Under Euler-Bernoulli beam theory (see, e.g., [8]), the longitudinal displacement functions can be written as

$$u_i = \bar{u}_i - y_i w',\ i \in \mathcal{I}, \tag{6.3}$$

where the prime symbol denotes differentiation with respect to x. Continuity of displacement field necessitates that $u_{\text{muc}}(x, -d_{\text{muc}}/2) = u_{\text{lig}}(x, d_{\text{lig}}/2)$ and $u_{\text{lig}}(x, -d_{\text{lig}}/2) = u_{\text{ta}}(x, d_{\text{ta}}/2)$ for all $x \in [0, L]$, which yields the conditions

$$\begin{aligned} \bar{u}_{\text{lig}} &= \bar{u}_{\text{muc}} + \frac{1}{2}(d_{\text{lig}} + d_{\text{muc}})w', \\ \bar{u}_{\text{ta}} &= \bar{u}_{\text{muc}} + \frac{1}{2}(d_{\text{ta}} + 2d_{\text{lig}} + d_{\text{muc}})w'. \end{aligned} \tag{6.4}$$

Given longitudinal displacement in the i^{th} layer with respect to the deformed configuration under $\bar{\varepsilon}$, the total strain in that layer is given by [1]

$$\varepsilon_i = \bar{\varepsilon} + (1 + \bar{\varepsilon})u_i',\ i \in \mathcal{I}. \tag{6.5}$$

[1] Consider an infinitesimal line element $\mathrm{d}x_0$ that experiences a composition of two deformations: the first is longitudinal deformation with associated uniform normal strain ε_0, and the second is due to a longitudinal displacement field u (with respect to the configuration after applying the strain ε_0). The length of the line element after applying strain ε_0 is $\mathrm{d}x_1 = (1 + \varepsilon_0)\mathrm{d}x_0$ ($\mathrm{d}x_0 = \mathrm{d}x_1/(1 + \varepsilon_0)$) and the length after applying the displacement field is $\mathrm{d}x_2 = (1 + \mathrm{d}u/\mathrm{d}x_1)\mathrm{d}x_1$. Therefore, the total strain due to the combination of the two deformations is $\varepsilon = (\mathrm{d}x_2 - \mathrm{d}x_0)/\mathrm{d}x_0 = \varepsilon_0 + (1 + \varepsilon_0)\mathrm{d}u/\mathrm{d}x_1$.

Substituting Equation (6.3) into Equation (6.5) results in

$$\varepsilon_i = \bar{\varepsilon} + (1 + \bar{\varepsilon})\,(\bar{u}_i' - y_i w'')\,,\ i \in \mathcal{I}. \tag{6.6}$$

The stress field is estimated from strain and, in the case of the TA muscle layer, TA muscle activation a_{ta}, which is a non-dimensional parameter, ranging between 0 and 1, that corresponds to the activation level in the TA muscle, with 0 indicating a completely flaccid muscle and 1 being maximum contraction. Herein, we utilize local linearization about the nominal strain $\bar{\varepsilon}$. That is, the stress functions in the VF layers, $\sigma_i,\ i \in \mathcal{I}$, are given by the approximate relations

$$\sigma_i = \sigma_{i,0} + E_i(\varepsilon_i - \bar{\varepsilon}),\ i \in \mathcal{I}, \tag{6.7}$$

where

$$\sigma_{j,0} = \bar{\sigma}_j(\bar{\varepsilon}),\ E_j = \mathrm{d}\bar{\sigma}_j(\bar{\varepsilon})/\mathrm{d}\varepsilon,\ j \in \{\mathrm{muc}, \mathrm{lig}\},$$

$$\sigma_{\mathrm{ta},0} = \bar{\sigma}_{\mathrm{ta}}(\bar{\varepsilon}, \mathsf{a}_{\mathrm{ta}}),\ E_{\mathrm{ta}} = \frac{\mathrm{d}\bar{\sigma}_{\mathrm{ta}}(\bar{\varepsilon}, \mathsf{a}_{\mathrm{ta}})}{\mathrm{d}\varepsilon},$$

and $\bar{\sigma}_i,\ i \in \mathcal{I}$, denotes the nonlinear stress function associated with the i^{th} layer.

The VF tissues exhibit a highly nonlinear hysteretic viscoelatic behaviour [193, 13, 187]. Several models have been proposed to capture such complex behaviours, including the modified Kelvin Model [153], nonlinear time-dependent parallel network models [187], and structurally-based nonlinear constitutive models [193]. Herein, and for simplicity, we assume the constitutive stress-strain relations associated with the nonlinear stresses to be of exponential type (see [56]) with symmetry about zero strain. In particular,

$$\bar{\sigma}_j = \mathrm{sign}(\varepsilon) m_j (\mathrm{e}^{|n_j \varepsilon|} - 1),\ j \in \{\mathrm{muc}, \mathrm{lig}\},$$

(sign is the signum function) and in the case of the TA muscle, we include stress induced by muscle activation, resulting in

$$\bar{\sigma}_{\mathrm{ta}} = \mathrm{sign}(\varepsilon) m_{\mathrm{ta}} (\mathrm{e}^{|n_{\mathrm{ta}} \varepsilon|} - 1) + \mathsf{a}_{\mathrm{ta}} \sigma_{\mathrm{a,max}},$$

where $m_i, n_i,\ i \in \mathcal{I}$, are parameters of the constitutive relations, and $\sigma_{\mathrm{a,max}}$ is the maximum active stress in the TA muscle. Symmetric stress-strain relations are employed herein to account for compressive forces developed in the VF, which have been often dismissed in previous studies of VF biomechanics. The numerical values of the constitutive relation parameters adopted in this study are listed in Table 6.1.

The normal forces in the VF layers are computed as $N_i = b \int_{-d_i/2}^{d_i/2} \sigma_i \mathrm{d}y_i,\ i \in \mathcal{I}$; substituting Equations (6.7) and (6.6) in yields

Table 6.1: Numerical values of the geometrical and mechanical properties for every layer in the composite vocal fold model: muc (mucosa), lig (ligament), and ta (thyroarytenoid). Cross-sectional areas are adopted from [153], whereas the parameters m_i, n_i, and $\sigma_{\mathrm{a,max}}$ are tuned to match experimental stress-strain curves from cadaver and canine models presented in [153, Figure 2.17, p. 88] based on human and canine samples.

parameter \ layer	muc	lig	ta
A_i [mm^2]	5	6.1	40.9
m_i [kPa]	1.5	2	1
n_i[-]	7	10	8
$\sigma_{\mathrm{a,max}}$ [kPa]	-	-	105

$$N_i = F_{i,0} + (1+\bar{\varepsilon})E_i A_i \bar{u}_i', \ i \in \mathcal{I}, \tag{6.8}$$

where

$$F_{i,0} = A_i \sigma_{i,0}, \ i \in \mathcal{I}, \tag{6.9}$$

denote the nominal normal forces generated by each layer. The total internal normal force is then

$$N = \sum_{i \in \mathcal{I}} N_i. \tag{6.10}$$

The moment about the center of the i^{th} layer due to the stress developed in that layer is given by $M_i = -b\int_{-d_i/2}^{d_i/2} y_i \sigma_i \mathrm{d}y_i, \ i \in \mathcal{I}$. Substituting Equations (6.7) and (6.6) into this formula gives

$$M_i = (1+\bar{\varepsilon})E_i I_i w'', \ i \in \mathcal{I}, \tag{6.11}$$

where $I_i = b\int_{-d_i/2}^{-d_i/2} y_i^2 \mathrm{d}y_i, \ i \in \mathcal{I}$, denotes the area moment of inertia of the i^{th} layer.

Let the positions of the geometric centers of the VF layers along the r-axis be denoted $r_i, \ i \in \mathcal{I}$; that is,

$$\begin{aligned} r_{\mathrm{muc}} &= d_{\mathrm{ta}} + d_{\mathrm{lig}} + \frac{d_{\mathrm{muc}}}{2}, \\ r_{\mathrm{lig}} &= d_{\mathrm{ta}} + \frac{d_{\mathrm{lig}}}{2}, \\ r_{\mathrm{ta}} &= \frac{d_{\mathrm{ta}}}{2}, \end{aligned} \tag{6.12}$$

see Figure 6.1. Consider a cross-section at longitudinal position x, and consider an arbitrary point on the cross-section located at a vertical position $r = r_c$ (see Figure 6.1). The moment at r_c, denoted M_c, is given by

$$\begin{aligned} M_c &= \sum_{i\in\mathcal{I}} M_i + (r_c - r_i)N_i \\ &= (1+\bar{\varepsilon})\Big(\sum_{i\in\mathcal{I}} E_i I_i\Big)w'' \\ &\quad + (1+\bar{\varepsilon})\sum_{i\in\mathcal{I}}(r_c - r_i)A_i E_i \bar{u}_i' \\ &\quad + \sum_{i\in\mathcal{I}}(r_c - r_i)F_{i,0}. \end{aligned} \tag{6.13}$$

Consider an infinitesimal element of longitudinal infinitesimal length $\mathrm{d}x$ with left edge at position x (see Figure 6.2), and let $V(x)$ and $q(x)$ denote the shear force and distributed load per unit length, respectively. The force and moment balances on the infinitesimal element yield

$$N' = 0, \tag{6.14}$$

$$V' - q = 0, \tag{6.15}$$

$$M' + V = 0, \tag{6.16}$$

where second order and higher terms are omitted.

From Equation (6.14), we deduce that the total normal force N is constant through the VF length. By the assumption that the VF undergoes compression/elongation with associated strain $\bar{\varepsilon}$ (see Equation (6.2)), the force N should be equal to the force that results in that strain, which is the sum of nominal forces $F_{i,0}$, $i \in \mathcal{I}$, (in Equation (6.9)). That is,

$$N = \sum_{i\in\mathcal{I}} F_{i,0}. \tag{6.17}$$

Therefore, by substituting Equations (6.8) and (6.10) into Equation (6.17), we have $\sum_{i\in\mathcal{I}}(1+\bar{\varepsilon})E_i A_i \bar{u}_i' = 0$, implying

$$\sum_{i\in\mathcal{I}} E_i A_i \bar{u}_i' = 0. \tag{6.18}$$

As we are interested in transverse deflection, we aim to obtain a balance equation solely

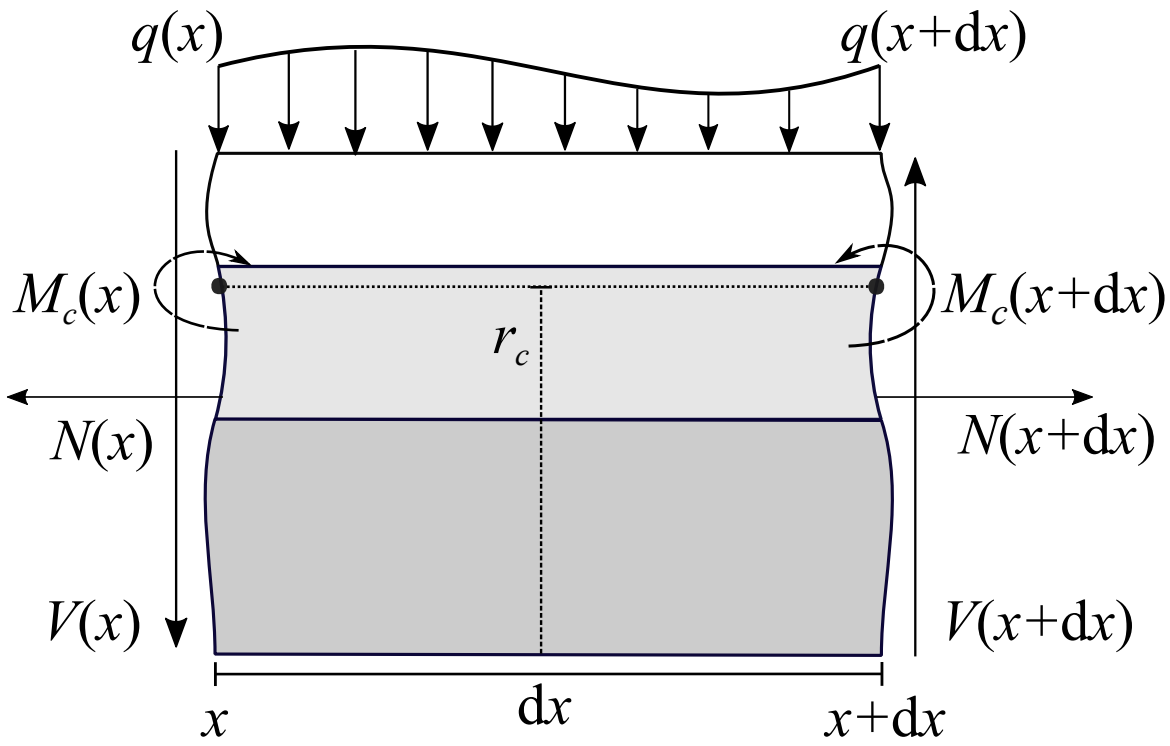


Figure 6.2: Free-body-diagram of an infinitesimal element of the composite beam vocal fold model.

in terms of w. For convenience, we define

$$\begin{aligned} l_{\text{lig}} &= \frac{1}{2}(d_{\text{lig}} + d_{\text{muc}}) \\ l_{\text{ta}} &= \frac{1}{2}(d_{\text{ta}} + 2d_{\text{lig}} + d_{\text{muc}}), \\ l_{\text{muc}} &= \frac{l_{\text{lig}}E_{\text{lig}}A_{\text{lig}} + l_{\text{ta}}E_{\text{ta}}A_{\text{ta}}}{E_{\text{muc}}A_{\text{muc}} + E_{\text{lig}}A_{\text{lig}} + E_{\text{ta}}A_{\text{ta}}}, \end{aligned} \tag{6.19}$$

and

$$\begin{aligned} \alpha_{\text{muc}} &= -l_{\text{muc}}, \\ \alpha_{\text{lig}} &= l_{\text{lig}} - l_{\text{muc}}, \\ \alpha_{\text{ta}} &= l_{\text{ta}} - l_{\text{muc}}. \end{aligned} \tag{6.20}$$

From the continuity condition in Equation (6.4) and the zero force condition in Equation (6.18), it can be deduced that the displacement functions $\bar{u}_i,\ i \in \mathcal{I}$, satisfy the relations

$$\bar{u}_i' = \alpha_i w'',\ i \in \mathcal{I}. \tag{6.21}$$

Substituting Equation (6.21) into Equation (6.13), we obtain

$$M_c = \mu_c w'' + M_{c,0}, \tag{6.22}$$

where

$$\mu_c = (1 + \bar{\varepsilon}) \sum_{i \in \mathcal{I}} E_i I_i + (r_c - r_i) A_i E_i \alpha_i \tag{6.23}$$

is the composite bending stiffness and

$$M_{c,0} = \sum_{i \in \mathcal{I}} (r_c - r_i) F_{i,0} \tag{6.24}$$

is the nominal moment at $r = r_c$ due to the nominal normal forces. Note that the bending stiffness is strain-dependent, which can be of importance in posturing scenarios with large VF strains.

Combining Equations (6.15) and (6.16) results in

$$M_c'' + q = 0,$$

which, in terms of the deflection w, where Equation (6.22) is substituted in, is

$$\mu_c w'''' + q = 0. \tag{6.25}$$

The distributed load q is assumed to be associated with VF contact, being proportional to the VF transverse motion overlap beyond the medial plane. That is,

$$q = K_{\text{col}}(w - x\tan(\theta_G))\mathbf{H}(w - x\tan(\theta_G)), \tag{6.26}$$

where K_{col} is a stiffness coefficient associated with VF contact, θ_G is the clockwise angle between the medial plane and the deformed VF configuration under the strain $\bar{\varepsilon}$ (see Figure 6.1), and $\mathbf{H}$ is the Heaviside function.

In this work, we assume zero transverse deflection at the anterior and posterior ends of the VF. That is,

$$w(0) = w(L) = 0. \tag{6.27}$$

Moreover, we assume zero moment at the posterior VF margin, i.e.,

$$M_c(L) = 0. \tag{6.28}$$

Furthermore, we assume a reactive moment at the anterior VF margin that is proportional to the rotational displacement with respect to a rest VF angle θ_0. The total angle at the anterior margin between the medial plane and the VF is approximately given by $\theta_G - w'(0)$. Consequently, the moment boundary condition at the anterior margin is given by

$$M_c(0) = -K_r(\theta_G - w'(0) - \theta_0), \tag{6.29}$$

where K_r is a rotational stiffness coefficient. Like the nominal strain $\bar{\varepsilon}$, it is assumed that the angle θ_G is known a priori. Finally, we assume that r_c corresponds to the geometrical center of the ligament, that is $r_c = r_{\text{lig}}$. This assumption, in addition to the boundary condition given in Equation (6.28), implies that the total normal force N is positioned at the geometrical center of the ligament. This can be deduced from the fact that $M_c(L) = (r_c - r_{\text{N}})N$, where r_{N} denotes the r-position of the total normal force N (i.e., force centroid).

6.2 Analytical Solution for a Special Case

To gain simple yet useful insights into how internal moments inside the VF beam model affect its curvature, we consider the scenario of zero contact forces (i.e., $q(x) = 0$) and assume $|w'(0)| \ll |\theta_G|$, which reduces the boundary condition given in Equation (6.29) to

$$M_c(0) = -K_r(\theta_G - \theta_0). \tag{6.30}$$

Equation (6.25), with the boundary conditions given in Equations (6.27), (6.30), and (6.28), and the definition of M_c in Equation (6.22), can be solved analytically, where the curvature of the VF beam model, w'', is given explicitly by

$$w'' = -\frac{M_{c,0}}{\mu_c} + \left(1 - \frac{x}{L}\right)\frac{K_r}{\mu_c}(\theta_0 - \theta_G). \tag{6.31}$$

Note that positive w'' implies convex VF geometry, whereas negative curvature implies concave geometry. Recalling the definition of $M_{c,0}$ given in Equation (6.24) and implementing the assumption that $r_c = r_{\text{lig}}$ result in

$$\begin{aligned} -M_{c,0} &= -\frac{d_{\text{ta}} + d_{\text{lig}}}{2}F_{\text{ta},0} + \frac{d_{\text{muc}} + d_{\text{lig}}}{2}F_{\text{muc},0} \\ &= \tilde{M}_{\text{ta}} + \tilde{M}_{\text{muc}}, \end{aligned}$$

where

$$\tilde{M}_{\text{ta}} = -\frac{d_{\text{ta}} + d_{\text{lig}}}{2}F_{\text{ta},0}, \tag{6.32}$$

and

$$\tilde{M}_{\text{muc}} = \frac{d_{\text{muc}} + d_{\text{lig}}}{2}F_{\text{muc},0}. \tag{6.33}$$

By additionally defining

$$\tilde{M}_r = K_r(\theta_0 - \theta_G), \tag{6.34}$$

Equation (6.31) can be rewritten as

$$w'' = \frac{1}{\mu_c}\left[\tilde{M}_{\text{ta}} + \tilde{M}_{\text{muc}} + \left(1 - \frac{x}{L}\right)\tilde{M}_r\right]. \tag{6.35}$$

In the following discussion, we assume that bending stiffness μ_c is always positive (μ_c according to Equation (6.23) changes with the elongation or compression of the VF beam model). First, let us analyse abstractly the effects of the moment terms $\tilde{M}_{\text{ta}}$, $\tilde{M}_{\text{muc}}$, and $\tilde{M}_r$ on the VF curvature. We can observe from Equation (6.35) that the effect of the reactive moment $\tilde{M}_r$ on the VF curvature decays linearly with a maximum effect (in magnitude) at $x = 0$ and zero effect at $x = L$, whereas the moment terms $\tilde{M}_{\text{ta}}$ and $\tilde{M}_{\text{muc}}$ have spatially invariant (i.e., constant) effects on w''. We can see that the curvature is positively correlated with $\tilde{M}_r$ and $\tilde{M}_{\text{ta}} + \tilde{M}_{\text{muc}}$, wherein the case $\tilde{M}_r > 0$ and $\tilde{M}_{\text{ta}} + \tilde{M}_{\text{muc}} > 0$ implies positive curvature (i.e., convex VF geometry), whereas the case $\tilde{M}_r < 0$ and $\tilde{M}_{\text{ta}} + \tilde{M}_{\text{muc}} < 0$ implies negative curvature (i.e., concave VF geometry). Considering the fact that the effect of the anterior reactive moment $\tilde{M}_r$ is decaying linearly along the VF length and that the

nominal moments induced by the VF layers, $\tilde{M}_{\text{ta}}$ and $\tilde{M}_{\text{muc}}$, are spatially invariant, we can deduce an interesting scenario for which the second order derivative changes its sign from positive to negative. In particular, in the case

$$\tilde{M}_r > -(\tilde{M}_{\text{ta}} + \tilde{M}_{\text{muc}}) > 0, \tag{6.36}$$

we can see that w'' is positive on $[0, x_{\text{cr}})$, where

$$x_{\text{cr}} = L\left(1 + \frac{\tilde{M}_{\text{ta}} + \tilde{M}_{\text{muc}}}{\tilde{M}_r}\right),$$

and negative for $x \in (x_{\text{cr}}, L]$, implying a change from convexity to concavity. The conditions on the internal moments and resulting VF shapes from the analysis above are summarized in Table 6.2.

Table 6.2: Conditions on the moments applied to the vocal fold composite model based on Equation (6.35), and the resulting vocal fold shapes: $\backsim$ (convex) and $\sim$ (concave).

condition	shape
$\tilde{M}_{\text{muc}} + \tilde{M}_{\text{ta}}$ and $\tilde{M}_r > 0$	$\backsim$
$\tilde{M}_{\text{muc}} + \tilde{M}_{\text{ta}}$ and $\tilde{M}_r < 0$	$\sim$
$\tilde{M}_r > -(\tilde{M}_{\text{muc}} + \tilde{M}_{\text{ta}}) > 0$	$\backsim\!\sim$

We note that convex VF geometry is a major characteristic of the bowed VF pattern. Moreover, the concave VF geometry can be associated with posterior glottal opening. Furthermore, the transition along the VF length from convexity to concavity resembles the VF shapes in the glottal hourglass pattern (see Figure 2.13).

Now, let us relate the findings listed in Table 6.2 to physiological posturing scenarios. The term $\tilde{M}_r$, as seen from its definition in Equation (6.34), is related to VF adduction and abduction, wherein $\tilde{M}_r > 0$ corresponds to VF adduction ($\theta_G < \theta_0$) and $\tilde{M}_r < 0$ corresponds to VF abduction ($\theta_G > \theta_0$).

The term $\tilde{M}_{\text{muc}} + \tilde{M}_{\text{ta}}$, which is defined according to Equations (6.32) and (6.33), is determined by the reactive moments developed in the VF layers, especially the mucosa and TA muscle, during VF tensioning or compression. Note that, based on the area measurements (Table 6.1) and the assumption of uniform thickness b , the moment arm of the TA muscle, $(d_{\text{ta}} + d_{\text{lig}})/2$, is larger than that of the mucosa, $(d_{\text{muc}} + d_{\text{lig}})/2$. In the case of VF compression, the compressive forces in the TA muscle are typically larger in magnitude than that in the mucosa (i.e., $F_{\text{ta},0} \ll F_{\text{muc},0} < 0$); hence, the moment induced

Table 6.3: Conditions on the moments applied to the vocal fold composite beam model and corresponding physiological posturing scenarios

condition	physiological scenario
$\tilde{M}_r > 0$	VF adduction
$\tilde{M}_r < 0$	VF abduction
$\tilde{M}_{\text{muc}} + \tilde{M}_{\text{ta}} > 0$	VF compression
$\tilde{M}_{\text{muc}} + \tilde{M}_{\text{ta}} < 0$	TA muscle activation

by the TA muscle, $\tilde{M}_{\text{ta}}$, is positive and predominant making $\tilde{M}_{\text{ta}} + \tilde{M}_{\text{muc}} > 0$. On the other hand, when the VF is tensioned due to activating the TA muscle, the force $F_{\text{ta},0}$ is positive and predominant and, consequently, the term $\tilde{M}_{\text{ta}}$ is negative (see Equation (6.32)) and predominant. This scenario results in $\tilde{M}_{\text{ta}} + \tilde{M}_{\text{muc}} < 0$. These relations between the moment terms and corresponding laryngeal posturing scenarios are summarized in Table 6.3.

By combining the analyses above (see Tables 6.2 and 6.3) we can deduce the following. The bowed shape with convex VF geometry can be due to (a) positive reactive moment at the anterior margin ($\tilde{M}_r > 0$) during VF adduction, and/or (b) internal moments during VF compression, wherein $\tilde{M}_{\text{ta}} + \tilde{M}_{\text{muc}} > 0$. Moreover, the concave VF shape arising in the case of posterior glottal opening can be due to (a) negative reactive moment at the anterior margin ($\tilde{M}_r < 0$) during VF abduction, and/or (b) sufficiently large activation of the TA muscle, wherein $\tilde{M}_{\text{ta}} + \tilde{M}_{\text{muc}} < 0$. The hourglass shape, with transitioning from convexity to concavity, may necessitate a tuned laryngeal maneuver that involves sufficient TA activation ($\tilde{M}_{\text{ta}} + \tilde{M}_{\text{muc}} < 0$) and sufficient VF adduction ($\tilde{M}_r > 0$) such that $\tilde{M}_r + \tilde{M}_{\text{ta}} + \tilde{M}_{\text{muc}} > 0$.

6.3 Simulations

In this section we further investigate VF curvature and incomplete glottal closure by incorporating the VF posturing modeling introduced by Titze and Hunter [160]. In particular, we adopt the implementation of Alzamendi et al. [4]. The posture model relates activation of the five intrinsic muscles to the prephonatory glottal configuration, and in particular, the rotational and linear displacements of the cricothyroid joints and the arytenoid cartilages, where the VFs and intrinsic muscles are modelled as spring-like elements. From the aforementioned displacements, the VF nominal strain $\bar{\varepsilon}$ and glottal angle θ_G are estimated. Similar to the muscle activation parameter a_{ta} embedded in the VF beam model,

Table 6.4: Numerical values of some of the vocal fold beam model parameters.

parameter	numerical value
L_0 [mm]	15
b [mm]	5
θ_0 [rad]	0.2540
K_{col} [N/m^2]	2×10^8
K_r [N.m]	0.05

the posture model relies on five normalized muscle activation parameters, a_{ta}, a_{ct}, a_{lca}, a_{ia}, and a_{pca}, which correspond to the TA, CT, LCA, IA, and PCA muscles, respectively. In this study, we assume that the muscle activation parameter a_{ta} embedded in the VF beam model is identical in value to the muscle activation parameter a_{ta} embedded in the posturing model ($\text{a}_{\text{ta}} = a_{\text{ta}}$).

It is important to mention that the constitutive relations embedded in the VF beam model are different from those in the posture model implementation adopted from [4]. The focus of the current study is to replicate the VF static configurations, wherein we employ experimental stress-strain data based on human and canine samples [153, Figure 2.17, p. 88] to prescribe the mechanical behaviors of the tissues. The focus of the posture model in [4] instead focused on physiologically accurate posturing and phonation outputs, which required adhoc tuning of some posturing model parameters. Prior experimental and numerical studies typically suffer from significant variability in the reported numerical values of biomechanical parameters (see, e.g., [153, 160, 56, 112]), and in some cases, numerical values are missing, which further motivates the adhoc tuning approach adopted in [3].

The posturing model in [160, 4] is dynamic due to inertial and viscous effects. In this study, and as we are interested in static posturing scenarios, given the numerical values of the five intrinsic muscle activation parameters, the posture model is run until the VF strain and glottal angle reach steady-state and such values are input to the composite beam model. Once the VF strain and glottal angle parameters are fed into the beam model, Equation (6.25), supplemented with Equations (6.26) and (6.13) and the boundary conditions in Equations (6.27), (6.28), and (6.29), is solved numerically. The aforementioned equations and boundary conditions are discretized by means of finite difference. In our simulations, the angle θ_0 is set as the glottal angle from the posturing simulations when all laryngeal muscles are inactive. The numerical values adopted in this study for the remaining VF beam model parameters are listed in Table 6.4.

To clearly illustrate the glottal geometries resulting from the simulations, a coordinate

system (x_1, x_2) with origin at the anterior margin of the VFs is established. The x_1-axis is aligned along the medial plane pointing in the posterior direction and the x_2-axis is perpendicular to the medial plane pointing to the right, relative to the human body frame (see Figure 6.1). In all figures presented below, the configurations of the VFs with respect to the (x_1, x_2) coordinate system are plotted, where we utilize the model symmetry to produce the opposing VF shape.

The numerical investigation in this section aims at exploring several laryngeal maneuvers and how they influence the VF geometry[2]. In particular, and motivated by previous clinical, experimental, and numerical findings [182, 99, 15, 183], we consider laryngeal maneuvers associated with adductory (TA, LCA, and IA) and abductory (PCA) muscles as they have been found to play major roles in inducing curved glottal geometries. The remaining laryngeal muscle, namely the CT muscle, has been found to play a major role in regulating phonation fundamental frequency by stretching the VFs and that is outside the scope of this study. Herein, we compare some simulation results with findings from previous clinical, experimental, and high-fidelity numerical studies to verify the accuracy of the proposed VF beam model. Moreover, we attempt to elucidate potential mechanisms underlying the curved VF geometries observed clinically (see Figure 2.13).

First, we investigate the effects of increasing coactivation of the LCA and IA muscles, which are responsible for adducting the VFs [4], while the TA muscle and other intrinsic muscles are inactive. Figure 6.3 exhibits the glottal shapes and the induced moments corresponding to simulations wherein LCA and IA muscle activation levels are increased simultaneously. Figure 6.3(left) shows that coactivation of the LCA and IA muscles leads to posterior glottal closure with a remaining mid-membranous gap. This convex VF shape matches previous clinical and numerical findings [15, 183]. Figure 6.3(right) shows that the VF convexity is due to the predominance of the reactive moments at the anterior VF margin ($\tilde{M}_r$ is positive and relatively large), which arises due to VF adduction ($\theta_G < \theta_0$), and that agrees with the theoretical predictions in Section 6.2.

Next, we explore the effects of isolated activation of the TA muscle, which is responsible for both adducting and shortening the VFs. Figure 6.4 exhibits the glottal shapes and induced moments corresponding to simulations wherein TA muscle activation levels are increased, while all other intrinsic muscles are inactive. Figure 6.4(left) shows that isolated activation of the TA muscle leads to anterior and mid-membranous glottal closure with remaining posterior opening. The resulting concave VF shapes are in agreement with

[2]The end points (anterior and posterior margins) of the VFs resulting from the proposed beam model are identical to those established by the posturing model of [160]. Accounting for internal bending moments results in deviation of the VF shape from the linear prescription (with angle θ_G) of [160].

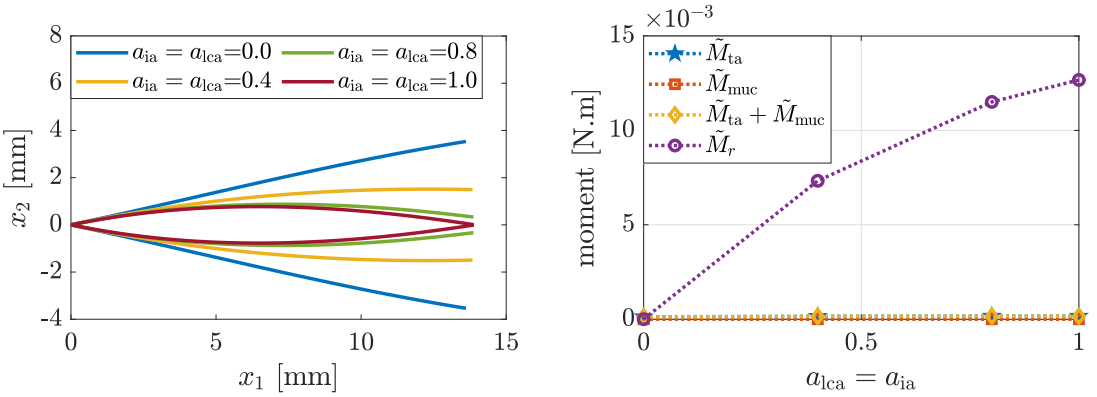


Figure 6.3: (Left) glottal profile and (right) induced moments for increasing lateral cricoarytenoid and interarytenoid muscle activation levels and other intrinsic muscles being inactive.

previous experimental and numerical investigations [15, 183]. Figure 6.4(right) shows that the concavity is primarily determined by the internal moments induced by the TA muscle activation ($\tilde{M}_{ta}$ is negative and relatively large in magnitude), which is in alignment with the analysis in Section 6.2.

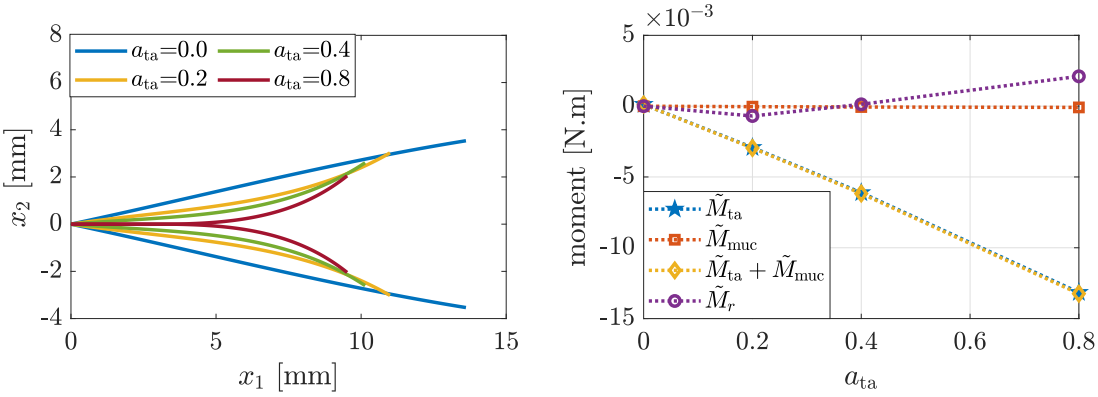


Figure 6.4: (Left) glottal profile and (right) induced moments for increasing thyroarytenoid activation levels and other intrinsic muscles being inactive.

In an effort to explore the mechanics of the hourglass glottal shape, we explore the glottal shape associated with increasing activation of the TA muscle while the LCA and IA are kept at constant non-zero activation levels. Simulating such maneuvers is encouraged by the findings from the theoretical analysis in Section 6.2. Figure 6.5 displays the glottal

shapes and induced moments associated with slight increasing activation of the TA muscle, while the LCA and IA are kept at constant levels ($a_{\text{lca}} = a_{\text{ta}} = 0.6$). Figure 6.5(left) shows that in the case of zero TA activation, the glottal shape is bowed with slight, but not full, posterior adduction. As TA activation is increased, a medial bulge is observed whereas anteriorly the glottal geometry is still convex, resulting in an overall hourglass shape. Figure 6.5(right) displays how the internal moments $\tilde{M}_{\text{ta}}$ and $\tilde{M}_{\text{muc}}$ and the reactive moment at the anterior margin $\tilde{M}_r$ satisfy the condition of Equation (6.36). This aligns with the analysis in Section 6.2 and suggests that the hourglass glottal shape necessitates involvement of reactive moments at the anterior VF margin (associated with VF adduction) and internal moments induced inside the VF layers (primarily the TA muscle).

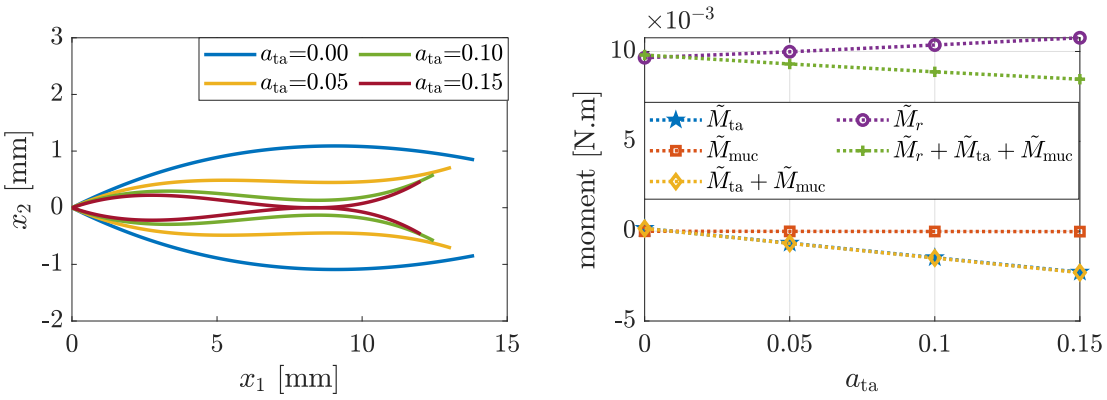


Figure 6.5: (Left) glottal shapes, and (right) induced moments for increasing thyroarytenoid muscle activation levels and other intrinsic muscle activation levels being at $a_{\text{lca}} = a_{\text{ta}} = 0.6$ and $a_{\text{ct}} = a_{\text{pca}} = 0$.

As seen from the previous simulations, dismissing the TA muscle or the LCA and IA muscles can not produce full glottal closure, therefore, in the next set of simulations, we investigate the effects of coactivating all the adductory muscles. Figure 6.6 depicts the glottal patterns associated with increasing coactivation of all adductory muscles. It is observed that (almost) full glottal closure is attained when all adductors are coactivated ($a_{\text{ia}} = a_{\text{lca}} = 0.45$, $a_{\text{ta}} = 0.7$), which is in alignment with previous experimental and numerical investigations [15, 183].

Finally, we explore the effects of the PCA muscle, a primary VF abductor, on the glottal geometry, where we consider simulations motivated by the clinical observations highlighted in [99]. Figure 6.7 displays glottal patterns associated with increasing PCA activation where adductory muscles are kept at activation levels associated with near full

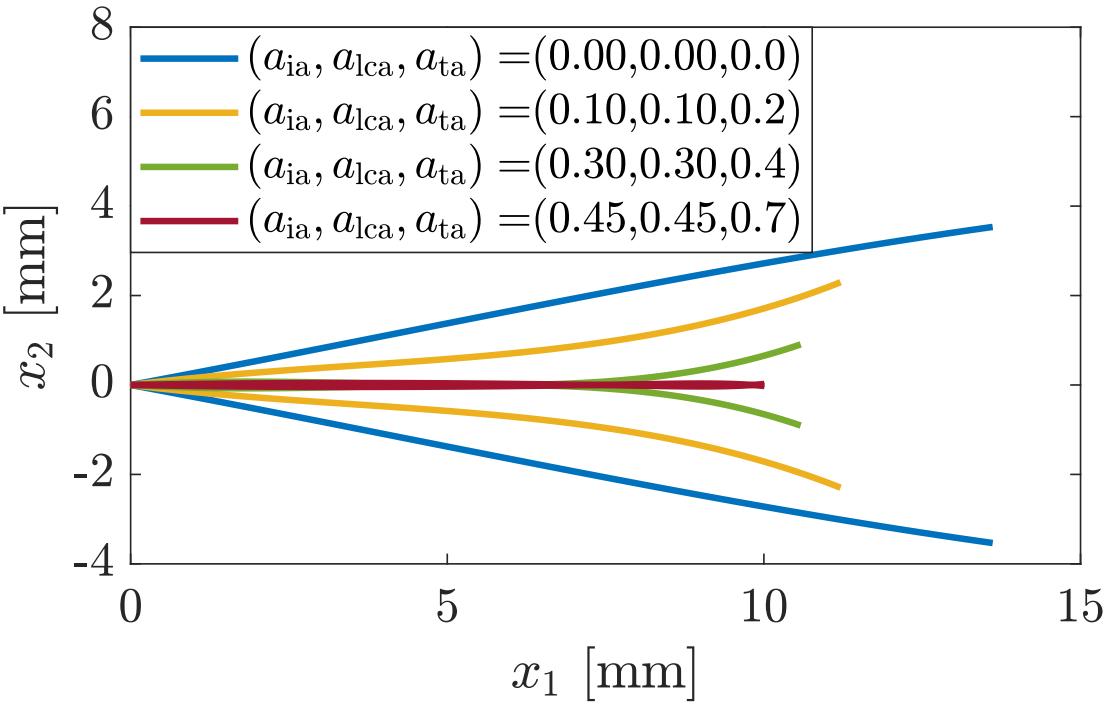


Figure 6.6: Glottal shapes for increasing coactivation levels of all adductory muscles, and the posterior cricoarytenoid and cricothyroid muscles being inactive.

closure. The figure displays that increasing PCA activation leads to posterior opening, while the VFs are sustaining concave shapes similar to those presented when TA alone is activated (see Figure 6.4). This shows that PCA activation in such laryngeal maneuvers acts to neutralize the posterior adductory effects of the LCA and IA muscles, which further supports the clinical observations highlighted in [99].

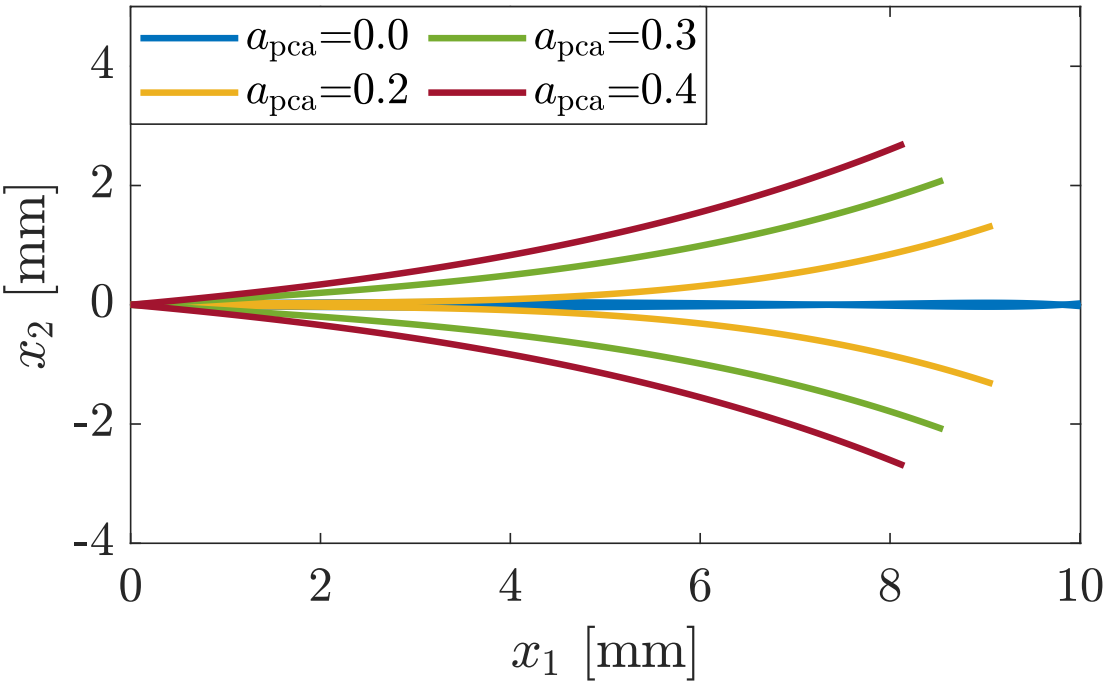


Figure 6.7: Glottal shapes for increasing activation of the posterior cricoarytenoid muscle and the activation levels of the other muscles being at $(a_{\text{lca}}, a_{\text{ia}}, a_{\text{ta}}) = (0.45, 0.45, 0.7)$ and $a_{\text{ct}} = 0$.

6.4 Discussion

The results of Sections 6.2 and 6.3 highlight potential mechanisms underlying different patterns of incomplete glottal closure. In particular, results indicate that bowed VF shapes result, in part, from low or null activation of the TA muscle in combination with coactivation of the LCA and IA muscles. The predominant mechanism in this case is the anterior

reactive moments that resist bringing the VFs together during adduction. This pattern can also arise in the case of low TA muscle activation and VF compression, as suggested by the analysis in Section 6.2, where in this case, the internal moment induced by the TA muscle tissue ($\tilde{M}_{\text{ta}}$ is positive and predominantly large), making it the driving factor. This scenario (bowing due to VF compression) can potentially take place when extrinsic laryngeal muscles are excessively activated, especially those associated with VF compressing, such as the thyrohyoid muscle [53].

In addition, our analysis suggests that posterior glottal opening with combined VF concavity results from high activation of the TA muscle and low or null activation of the LCA and IA muscles.The driving mechanism herein is the internal moment induced by the TA muscle activation $\tilde{M}_{\text{ta}}$, which is negative and predominantly large in magnitude in this case. A similar glottal pattern also occurs when all adductory muscles are activated in addition to the activation of the PCA muscle. This supports the hypothesis of [99], regarding the excessive activation of the PCA muscle in patients with MTD.

Finally, our analysis suggests that the hourglass glottal shape may emerge from laryngeal maneuvers that involve, for example, moderate coactivation of all adductory muscles, where both anterior reactive moment and internal moment due to TA muscle activation are at play opposing each other.

As highlighted in Chapter 2, incomplete glottal closure can be linked to voice disorders, including MTD, which is characterised by excessive, imbalanced, and poorly regulated activity of the intrinsic and extrinsic muscles [48]. MTD patients typically suffer from incomplete glottal closure, exhibiting various patterns including those depicted in Figure 2.13 [99, 105]. Our analysis in this paper suggests two potential mechanisms underlying bowed VFs in MTD patients (1) the TA muscle is not properly activated (possibly due to muscle activation imbalance), and (2) excessive activation of neck muscles, leading to VF compression. In the case of MTD subjects with posterior glottal opening and concave VF geometry, our model suggests that there may be either (1) inefficient use of the LCA and IA muscles, whereas the TA muscle is activated sufficiently (in comparison to normal posturing scenarios), or (2) excessive activation of all adducotry and abductory muscles, where the PCA muscle activation cancels out the adductory effects of the LCA and IA muscles, in agreement with [99].

A number of simplifying assumptions are embedded in the presented model, including (1) negligible shear deformation, (2) negligible elastic forces from the connective tissues attached to the TA muscle, (3) negligible motions in the superior-inferior direction, (4) neglecting potential bending effects from the vocal ligament by setting $r_c = r_{\text{lig}}$, (5) zero moments at the posterior ends of the VFs, (6) small transverse VF deflections, and (7) one-

way coupling between the VF beam model and the posturing model, where any contact forces emerging due to the VF curvature do not alter the mechanics of the laryngeal cartilages.

Assumption (1) is a consequence of the adopted Euler-Bernoulli framework. Note that with the uniform thickness assumption, and considering the model dimensions given in Tables 6.1 and 6.4, the total VF depth, $\sum_{i\in\mathcal{I}} d_i$, is approximately 10 mm whereas the resting VF length is 15 mm; hence, the depth and length dimensions are quite comparable. For such cases (thick beams), Timoshenko beam theory [152] is typically adopted to account for shear stresses[3]. As the goal of the current work is to construct a simple analytically-tractable model that predicts qualitatively the curved glottal configurations observed clinically, we adhered to the Euler-Bernoulli framework, leaving derivations of more complex models to future work. We posit that assumption (2) is reasonable as the elastic forces from the connective tissues are passive, mostly only restricting the extent to which the VF deflects. Moreover, assumption (3) is suitable as the majority of the VF motion during posturing occurs medially and/or laterally (see, e.g., the findings of [15]). Regarding assumption (4), the ligament is stiffer than other VF layers and it geometrically forms the intermediate VF layer, making it the 'chassis' of the VF layered structure; hence setting $r_c = r_{\text{lig}}$, which indicates that the total normal force in the VF is positioned at the center of the ligament (see the end of Section 6.1), is sensible. Assumptions (5)-(7), in addition to other assumptions such as the rectangular geometries of the VF layers, are introduced to primarily simplify our analysis; hence, further investigation is needed to verify the validity of such assumptions in different posturing scenarios, and refine them when needed.

Despite these simplifying assumptions, our modelling framework is capable of predicting some of the glottal patterns observed in previous clinical and high-fidelity numerical studies, which is encouraging. Still, the speculations and potential explanations provided in this work need further extensive investigation into the biomechanics of VF posturing in both healthy subjects and patients with MTD.

[3]It is worth noting that the two theories (Euler-Bernoulli and Timoshenko) do coincide for a uniform homogeneous simply-supported linear beam with specified moments at the end points and zero distributive load, regardless of the beam thickness or mechanical properties (in Section 5.2, we studied a similar simply-supported case with zero distributive load). These two theories, when compared, tend to produce qualitatively, but not necessarily quantitatively, similar predictions (see, e.g., [9]).

6.5 Conclusion

In this chapter, we introduced a simple one-dimensional Euler-Bernouilli-type composite beam VF model to understand the mechanisms underlying glottal configuration and incomplete glottal closure. The model, despite its simplicity, was capable of predicting some of the glottal configurations observed clinically or replicated using high-fidelity modeling. Our analysis highlighted how the different patterns of incomplete glottal closure can arise naturally due to the layered VF structure and the associated induced moments. We coupled the proposed beam model with the posturing model of [160] to gain physiologically relevant insights into the role of laryngeal muscle activation. Our analysis showed that bowed VF shape can arise due to activation of the LCA and IA muscles without incorporating the TA muscle during adduction or, as predicted theoretically in Section 6.2, due to VF compression. On the other hand, isolated activation of the TA muscle results in medial bulging and posterior glottal opening. Posterior opening can also occur due to activating all adductors in addition to activating the PCA muscle. Moreover, our analysis suggested that the hourglass glottal shape can arise from specific laryngeal maneuvers involving the adductory laryngeal muscles. These results provided potential explanations and conjectures regarding the posturing mechanics of patients with MTD.

Chapter 7

Closing Remarks

In this thesis, we investigated, by means of theoretical analyses and numerical simulations, different aspects of laryngeal biomechanics and phonation and attempted to reveal some of the underlying mechanisms associated with vocal hyperfunction. In particular, we accomplished the following:

1. We explored the effects of laryngeal (muscle activation) and biomechanical (VF contact) factors on phonation offset in Chapter 3, showing that the clinically observed drop in RFF is potentially due to the reduction of VF collision levels associated with increasing the gap between the folds (i.e., abduction). We also showed the drop rate of RFF can be modulated by the activation of the CT muscle, which we conjectured to be a mechanism underlying the differences between healthy and hyperfunctional speakers. Furthermore, we showed how certain manifestations of PVH (increased VF mass, subglottal pressure, and pre-offset neutral gap) affect the drop rate of RFF during phonation offset.

2. In Chapter 4, we extended the analysis of the previous chapter to cover phonation onset. We showed that, when all laryngeal and aerodynamic parameters are fixed in time, fundamental frequency tends to rise due to the rise in VF collision forces, and that matches with the clinical observations of the onset of initial or isolated vowels and, in some cases, vowels preceded by voiced consonants. On the other hand, we showed through numerical simulations that the decline in fundamental frequency in the case of onset of vowels preceded by voiceless consonants requires involvement of laryngeal muscles (especially CT and TA muscles) and illustrated how laryngeal factors can underlie the differences between healthy speakers and patients with vocal

hyperfunction. The analyses conducted in Chapters 3 and 4 address the first objective of this thesis.

3. We developed in Chapter 5 a phonation model that incorporates the net effect of extrinsic laryngeal muscles and investigated the influence of the net effect on posturing mechanics and sustained phonation. More specifically, we showed that the pulling effect from the extrinsic laryngeal muscles on VF tension and phonation fundamental frequency depends on the pulling magnitude, direction, and location. Guided by clinical findings, we used our analysis of the pulling effect to pinpoint potential roles of specific extrinsic muscles in modulating fundamental frequency. Moreover, based on sparse clinical data, we deduced some implications regarding the potential roles of extrinsic laryngeal muscles in hyperfunctional phonation. The investigation conducted in Chapter 5 addresses the second objective of this thesis.

4. We derived in Chapter 6 a simple yet insightful Euler-Bernoulli-type beam vocal fold model, and incorporated this model with the posturing model of Titze and Hunter [160] to explore the influences of different laryngeal maneuvers. This model, despite is simplicity, is capable of predicting some of the curved and incomplete glottal closure patterns observed clinically in some hyperfunctional patients [99, 105]. We conducted theoretical analysis on this model showing how the layered VF structure and induced bending moments during posturing are potential factors underlying these glottal configurations. We then suggested some potential laryngeal mechanisms that may be at play in hyperfunctional speakers. Our analysis of incomplete and curved glottal closure patterns in Chapter 6 fulfills the third objective of this thesis.

7.1 Future Work

We suggest the following extensions and refinements to the investigations conducted in this thesis.

(1) The analyses of transient phonation in Chapters 3 and 4 can be refined by (a) implementing more sophisticated and physiologically relevant muscle activation rules in phonation simulations (e.g., the model in[4]) and (b) considering the the time–varying nature of subglottal pressure and supraglottal vocal tract during transient phonation.

(2) The work in Chapter 5 on the influence of neck muscles can be refined, where the influence of individual neck muscles can be analysed by adopting more refined posturing

models that incorporate separately controlled extrinsic muscles. For that purpose, models used in investigations of neck stabilization, swallowing, and phonatory postures (e.g., [43, 100, 98]) may be adapted. In addition, the analysis can be extended to investigate the influence of tracheal pulling on the mechanics of phonation, which was highlighted, e.g., in [174].

(3) There are various possible refinements to the VF beam model presented in Chapter 6. For example, the infinitesimal transverse deflection assumption can be slightly weakened, where large deformations may be accounted for. In addition, a Timoshenko-type theory [152] can be incorporated to account for shear deformation. Moreover, two-way coupling between the VF beam model and the posturing model of Titze and Hunter [160] can be incorporated to account for potential effects that curved VF geometries may exert on the mechanics of laryngeal cartilages. Furthermore, the VF beam model can be integrated with numerical phonation models, to study how curved glottal geometries may influence the vocal function and the tissue-flow-acoustics interaction in the glottis during phonation in patients with vocal hyperfunction.

Finally, it is important to utilize the predictions and conjectures presented in this thesis to guide future experimental and clinical investigations of phonation and vocal hyperfunction. For example, clinical and experimental investigations can be conducted to further explore: (a) the roles of intrinsic laryngeal muscles, especially CT and TA, in modulating fundamental frequency during transient phonation, (b) the effects of coactivating supra- and infra-hyoid muscles on phonation fundamental frequency in both healthy subjects and patients with VH, and (c) the relation between the posturing of the laryngeal cartilages and the curved configurations of the VFs. It is worth noting that our analysis in Chapter 3 has, in fact, guided a recent clinical study by Groll et al. [38]. This study showed the relevance of VF contact in modulating RFF during phonation offset, verifying partially our theoretical and numerical investigations in Chapter 3.

References

[1] Fariborz Alipour, David A Berry, and Ingo R Titze. A finite-element model of vocal-fold vibration. *The Journal of the Acoustical Society of America*, 108(6):3003–3012, 2000.

[2] Kenneth W Altman, Cory Atkinson, and Cathy Lazarus. Current and emerging concepts in muscle tension dysphonia: a 30-month review. *Journal of Voice*, 19(2):261–267, 2005.

[3] Gabriel A Alzamendi, Rodrigo Manríquez, Paul J Hadwin, Jonathan J Deng, Sean D Peterson, Byron D Erath, Daryush D Mehta, Robert E Hillman, and Matías Zañartu. Bayesian estimation of vocal function measures using laryngeal high-speed videoendoscopy and glottal airflow estimates: An in vivo case study. *The Journal of the Acoustical Society of America*, 147(5):EL434–EL439, 2020.

[4] Gabriel A Alzamendi, Sean D Peterson, Byron D Erath, Robert E Hillman, and Matías Zañartu. Triangular body-cover model of the vocal folds with coordinated activation of the five intrinsic laryngeal muscles. *The Journal of the Acoustical Society of America*, 151(1):17–30, 2022.

[5] Thana Angsuwarangsee and Murray Morrison. Extrinsic laryngeal muscular tension in patients with voice disorders. *Journal of Voice*, 16(3):333–343, 2002.

[6] James E Atkinson. Correlation analysis of the physiological factors controlling fundamental voice frequency. *The Journal of the Acoustical Society of America*, 63(1):211–222, 1978.

[7] Ronald C Auvenshine and Nathan J Pettit. The hyoid bone: an overview. *CRANIO®*, 2018.

[8] Olivier Andre Bauchau and James I Craig. *Structural analysis: with applications to aerospace structures*, volume 163. Springer Science & Business Media, 2009.

[9] André Teófilo Beck and Cláudio RA da Silva Jr. Timoshenko versus euler beam theory: Pitfalls of a deterministic approach. *Structural Safety*, 33(1):19–25, 2011.

[10] Pinaki Bhattacharya and Thomas Siegmund. The role of glottal surface adhesion on vocal folds biomechanics. *Biomechanics and modeling in mechanobiology*, 14(2):283–295, 2015.

[11] Neil Bhattacharyya. The prevalence of voice problems among adults in the united states. *The Laryngoscope*, 124(10):2359–2362, 2014.

[12] Heiko Braak and Eva Braak. Pathoanatomy of parkinson's disease. *Journal of neurology*, 247(2):II3–II10, 2000.

[13] Roger W Chan, Min Fu, Lindsay Young, and Neeraj Tirunagari. Relative contributions of collagen and elastin to elasticity of the vocal fold under tension. *Annals of biomedical engineering*, 35(8):1471–1483, 2007.

[14] Ye Chen, Zheng Li, Siyuan Chang, Bernard Rousseau, and Haoxiang Luo. A reduced-order flow model for vocal fold vibration: From idealized to subject-specific models. *Journal of Fluids and Structures*, 94:102940, 2020.

[15] Dinesh K Chhetri and Juergen Neubauer. Differential roles for the thyroarytenoid and lateral cricoarytenoid muscles in phonation. *The Laryngoscope*, 125(12):2772–2777, 2015.

[16] Dinesh K Chhetri, Juergen Neubauer, and David A Berry. Neuromuscular control of fundamental frequency and glottal posture at phonation onset. *The Journal of the Acoustical Society of America*, 131(2):1401–1412, 2012.

[17] Dinesh K Chhetri, Juergen Neubauer, Elazar Sofer, and David A Berry. Influence and interactions of laryngeal adductors and cricothyroid muscles on fundamental frequency and glottal posture control. *The Journal of the Acoustical Society of America*, 135(4):2052–2064, 2014.

[18] Hong-Shik Choi, Gerald S Berke, Ming Ye, and Jody Kreiman. Function of the posterior cricoarytenoid muscle in phonation: in vivo laryngeal model. *Otolaryngology—Head and Neck Surgery*, 109(6):1043–1051, 1993.

[19] Hong Shik Choi, Ming Ye, and Gerald S Berke. Function of the interarytenoid (ia) muscle in phonation: in vivo laryngeal model. *Yonsei Medical Journal*, 36(1):58–67, 1995.

[20] Hong-Shik Choi, Ming Ye, Gerald S Berke, and Jody Kreiman. Function of the thyroarytenoid muscle in a canine laryngeal model. *Annals of Otology, Rhinology & Laryngology*, 102(10):769–776, 1993.

[21] Frederic L Darley, Arnold Elvin Aronson, and Joe Robert Brown. *Motor speech disorders*. WB Saunders Company, 1975.

[22] Gifford Z Decker and Scott L Thomson. Computational simulations of vocal fold vibration: Bernoulli versus navier–stokes. *Journal of Voice*, 21(3):273–284, 2007.

[23] Philippe H Dejonckere and Malte Kob. Pathogenesis of vocal fold nodules: new insights from a modelling approach. *Folia Phoniatrica et Logopaedica*, 61(3):171–179, 2009.

[24] Jonathan J Deng, Paul J Hadwin, and Sean D Peterson. The effect of high-speed videoendoscopy configuration on reduced-order model parameter estimates by bayesian inference. *The Journal of the Acoustical Society of America*, 146(2):1492–1502, 2019.

[25] Manuel Diaz-Cadiz, Victoria S McKenna, Jennifer M Vojtech, and Cara E Stepp. Adductory vocal fold kinematic trajectories during conventional versus high-speed videoendoscopy. *Journal of Speech, Language, and Hearing Research*, 62(6):1685–1706, 2019.

[26] Byron D Erath, Matias Zanartu, Kelley C Stewart, Michael W Plesniak, David E Sommer, and Sean D Peterson. A review of lumped-element models of voiced speech. *Speech Communication*, 55(5):667–690, 2013.

[27] Víctor M Espinoza, Matías Zañartu, Jarrad H Van Stan, Daryush D Mehta, and Robert E Hillman. Glottal aerodynamic measures in women with phonotraumatic and nonphonotraumatic vocal hyperfunction. *Journal of Speech, Language, and Hearing Research*, 60(8):2159–2169, 2017.

[28] João Aragão Ximenes Filho, Paulo Hilário Saldiva Do Nascimento, Domingos Hiroshi Tsuji, and Luiz Ubirajara Sennes. Histologic changes in human vocal folds correlated with aging: a histomorphometric study. *Annals of Otology, Rhinology & Laryngology*, 112(10):894–898, 2003.

[29] James L Flanagan. *Speech analysis: Synthesis and perception*. Springer-Verlag, 1972.

[30] Alexander L Francis, Valter Ciocca, Virginia Ka Man Wong, and Jess Ka Lam Chan. Is fundamental frequency a cue to aspiration in initial stops? *The Journal of the Acoustical Society of America*, 120(5):2884–2895, 2006.

[31] Lewis P Fulcher, Ronald C Scherer, and Travis Powell. Viscous effects in a static physical model of the uniform glottis. *The Journal of the Acoustical Society of America*, 134(2):1253–1260, 2013.

[32] Gabriel E Galindo, Sean D Peterson, Byron D Erath, Christian Castro, Robert E Hillman, and Matías Zañartu. Modeling the pathophysiology of phonotraumatic vocal hyperfunction with a triangular glottal model of the vocal folds. *Journal of Speech, Language, and Hearing Research*, 60(9):2452–2471, 2017.

[33] Gabriel E Galindo, Matías Zañartu, and Juan I Yuz. A discrete-time model for the vocal folds. In *IEEE EMBS International Student Conference*, pages 74–77. Citeseer, 2014.

[34] Biao Geng, Mohammadreza Movahhedi, Qian Xue, and Xudong Zheng. Vocal fold vibration mode changes due to cricothyroid and thyroarytenoid muscle interaction in a three-dimensional model of the canine larynx. *The Journal of the Acoustical Society of America*, 150(2):1176–1187, 2021.

[35] A Gervasio, I Mujahed, A Biasio, and S Alessi. Ultrasound anatomy of the neck: The infrahyoid region. *Journal of ultrasound*, 13(3):85–89, 2010.

[36] Alexander M Goberman and Michael Blomgren. Fundamental frequency change during offset and onset of voicing in individuals with parkinson disease. *Journal of Voice*, 22(2):178–191, 2008.

[37] IF Grace and RA Ibrahim. Modelling and analysis of ship roll oscillations interacting with stationary icebergs. *Proceedings of the Institution of Mechanical Engineers, Part C: Journal of Mechanical Engineering Science*, 222(10):1873–1884, 2008.

[38] Matti D Groll, Sean D Peterson, Matías Zañartu, Jennifer M Vojtech, and Cara E Stepp. Empirical evaluation of the role of vocal fold collision on relative fundamental frequency in voicing offset. *Journal of Voice*, 2022.

[39] Paul J Hadwin, Gabriel E Galindo, Kyle J Daun, Matías Zañartu, Byron D Erath, Edson Cataldo, and Sean D Peterson. Non-stationary bayesian estimation of parameters from a body cover model of the vocal folds. *The Journal of the Acoustical Society of America*, 139(5):2683–2696, 2016.

[40] Paul J Hadwin and Sean D Peterson. An extended kalman filter approach to non-stationary bayesian estimation of reduced-order vocal fold model parameters. *The Journal of the Acoustical Society of America*, 141(4):2909–2920, 2017.

[41] David G Hanson, Bruce R Gerratt, and Paul H Ward. Cinegraphic observations of laryngeal function in parkinson's disease. *The Laryngoscope*, 94(3):348–353, 1984.

[42] Helen M Hanson. Effects of obstruent consonants on fundamental frequency at vowel onset in english. *The Journal of the Acoustical Society of America*, 125(1):425–441, 2009.

[43] Takuya Hashimoto, Mariko Urabe, Foo Chee-Sheng, Atsuko Murakoshi, Takahiro Kikuchi, Yukihiro Michiwaki, and Takuji Koike. Development of a musculoskeletal model of hyolaryngeal elements for understanding pharyngeal swallowing mechanics. *Applied Sciences*, 10(18):6276, 2020.

[44] Elizabeth S Heller Murray, Yu-An S Lien, Jarrad H Van Stan, Daryush D Mehta, Robert E Hillman, J Pieter Noordzij, and Cara E Stepp. Relative fundamental frequency distinguishes between phonotraumatic and non-phonotraumatic vocal hyperfunction. *Journal of Speech, Language, and Hearing Research*, 60(6):1507–1515, 2017.

[45] Elizabeth S Heller Murray, Roxanne K Segina, Geralyn Harvey Woodnorth, and Cara E Stepp. Relative fundamental frequency in children with and without vocal fold nodules. *Journal of Speech, Language, and Hearing Research*, 63(2):361–371, 2020.

[46] Patricia Henríquez, Jesús B Alonso, Miguel A Ferrer, Carlos M Travieso, Juan I Godino-Llorente, and Fernando Díaz-de María. Characterization of healthy and pathological voice through measures based on nonlinear dynamics. *IEEE transactions on audio, speech, and language processing*, 17(6):1186–1195, 2009.

[47] Robert E Hillman, Eva B Holmberg, Joseph S Perkell, Michael Walsh, and Charles Vaughan. Objective assessment of vocal hyperfunction: An experimental framework and initial results. *Journal of Speech, Language, and Hearing Research*, 32(2):373–392, 1989.

[48] Robert E Hillman, Cara E Stepp, Jarrad H Van Stan, Matías Zañartu, and Daryush D Mehta. An updated theoretical framework for vocal hyperfunction. *American Journal of Speech-Language Pathology*, pages 1–7, 2020.

[49] Minoru Hirano. Morphological structure of the vocal cord as a vibrator and its variations. *Folia phoniatrica et logopaedica*, 26(2):89–94, 1974.

[50] Minoru Hirano. Phonosurgery: basic and clinical investigations. *Otol (Fukuoka)*, 21:239–242, 1975.

[51] Jean-Marie Hombert, John J Ohala, and William G Ewan. Phonetic explanations for the development of tones. *Language*, 55(1):37–58, 1979.

[52] Kiyoshi Honda, Hiroyuki Hirai, Shinobu Masaki, and Yasuhiro Shimada. Role of vertical larynx movement and cervical lordosis in f0 control. *Language and Speech*, 42(4):401–411, 1999.

[53] K H Hong, M Ye, Y M Kim, K F Kevorkian, and G S Berke. The role of strap muscles in phonation: in vivo canine laryngeal model. *Journal of Voice*, 32, 1997.

[54] Yong Tae Hong, Ki Hwan Hong, Je-Pyo Jun, and Pyung Han Hwang. The effects of dynamic laryngeal movements on pitch control. *American Journal of Otolaryngology*, 36(5):660–665, 2015.

[55] MS Howe and RS McGowan. On the single-mass model of the vocal folds. *Fluid dynamics research*, 42(1):015001, 2010.

[56] Eric J Hunter and Ingo R Titze. Refinements in modeling the passive properties of laryngeal soft tissue. *Journal of Applied Physiology*, 103(1):206–219, 2007.

[57] James Ing, Ekaterina Pavlovskaia, Marian Wiercigroch, and Soumitro Banerjee. Experimental study of impact oscillator with one-sided elastic constraint. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 366(1866):679–705, 2008.

[58] Kenzo Ishizaka and James L Flanagan. Synthesis of voiced sounds from a two-mass model of the vocal cords. *Bell System Technical Journal*, 51(6):1233–1268, 1972.

[59] Sanyukta Jaiswal. *Cricothyroid muscle activity at voicing transitions*. PhD thesis, University of Iowa, 2011.

[60] Jack J Jiang, Yu Zhang, Julia MacCallum, Alicia Sprecher, and Liang Zhou. Objective acoustic analysis of pathological voices from patients with vocal nodules and polyps. *Folia Phoniatrica et Logopaedica*, 61(6):342–349, 2009.

[61] Jack J Jiang, Yu Zhang, and Clancy McGilligan. Chaos in voice, from modeling to measurement. *Journal of Voice*, 20(1):2–17, 2006.

[62] Jack J Jiang, Yu Zhang, and Jennifer Stern. Modeling of chaotic vibrations in symmetric vocal folds. *The Journal of the Acoustical Society of America*, 110(4):2120–2128, 2001.

[63] Jack Jiaqi Jiang and Ingo R Titze. Measurement of vocal fold intraglottal pressure and impact stress. *Journal of Voice*, 8(2):132–144, 1994.

[64] J L Kelly and C C Lochbaum. Speech synthesis. In *Proceedings of the Fourth International Congress on Acoustics*, 1962.

[65] Brian Quinn Kettlewell. The influence of intraglottal vortices upon the dynamics of the vocal folds. Master's thesis, University of Waterloo, 2015.

[66] Sid Khosla, Shanmugam Muruguppan, Ephraim Gutmark, and Ronald Scherer. Vortical flow field during phonation in an excised canine larynx model. *Annals of Otology, Rhinology & Laryngology*, 116(3):217–228, 2007.

[67] James P Kirby and D Robert Ladd. Effects of obstruent voicing on vowel f 0: Evidence from "true voicing" languages. *The Journal of the Acoustical Society of America*, 140(4):2400–2411, 2016.

[68] A Kirzinger and U Jürgens. Role of extralaryngeal muscles in phonation of subhuman primates. *Journal of comparative physiology. A, Sensory, neural, and behavioral physiology*, 175(2):215–222, 1994.

[69] Tanja Kovač, Branka Popović, Ksenija Marjanović, Vjekoslav Wertheimer, Marko Kovačević, Vasilije Nikolić, Ana Jo-Osvatić, and Radivoje Radić. Morphometric characteristics of thyroid cartilage in people of eastern croatia. *Collegium antropologicum*, 34(3):1069–1073, 2010.

[70] Michael H Krane and Timothy Wei. Theoretical assessment of unsteady aerodynamic effects in phonation. *The Journal of the Acoustical Society of America*, 120(3):1578–1588, 2006.

[71] Jeff Kuo, Eva B Holmberg, and Robert E Hillman. Discriminating speakers with vocal nodules using aerodynamic and acoustic features. In *1999 IEEE International Conference on Acoustics, Speech, and Signal Processing. Proceedings. ICASSP99 (Cat. No. 99CH36258)*, volume 1, pages 77–80. IEEE, 1999.

[72] Aude Lagier, Thierry Legou, Camille Galant, Benoit Amy de La Bretèque, Yohann Meynadier, and Antoine Giovanni. The shouted voice: A pilot study of laryngeal physiology under extreme aerodynamic pressure. *Logopedics Phoniatrics Vocology*, 42(4):141–145, 2017.

[73] M Drew LaMar, Yingyong Qi, and Jack Xin. Modeling vocal fold motion with a hydrodynamic semicontinuum model. *The Journal of the Acoustical Society of America*, 114(1):455–464, 2003.

[74] Yu-An S Lien, Caitlin I Gattuccio, and Cara E Stepp. Effects of phonetic context on relative fundamental frequency. *Journal of Speech, Language, and Hearing Research*, 57(4):1259–1267, 2014.

[75] Yu An S Lien, Elizabeth S Heller Murray, Carolyn R Calabrese, Carolyn M Michener, Jarrad H Van Stan, Daryush D Mehta, Robert E Hillman, J Pieter Noordzij, and Cara E Stepp. Validation of an algorithm for semi-automated estimation of voice relative fundamental frequency. *Annals of Otology, Rhinology & Laryngology*, 126(10):712–716, 2017.

[76] Yu-An S Lien, Carolyn M Michener, Tanya L Eadie, and Cara E Stepp. Individual monitoring of vocal effort with relative fundamental frequency: Relationships with aerodynamics and listener perception. *Journal of Speech, Language, and Hearing Research*, 58(3):566–575, 2015.

[77] YuAn S Lien, Carolyn R Calabrese, Carolyn M Michener, Elizabeth Heller Murray, Jarrad H Van Stan, Daryush D Mehta, Robert E Hillman, J Pieter Noordzij, and Cara E Stepp. Voice relative fundamental frequency via neck-skin acceleration in individuals with voice disorders. *Journal of Speech, Language, and Hearing Research*, 58(5):1482–1487, 2015.

[78] Johan Liljencrants. *Speech synthesis with a reflection-type line analog*. PhD thesis, Royal Institute of Technology, 1985.

[79] Anders Löfqvist, Thomas Baer, Nancy S McGarr, and Robin Seider Story. The cricothyroid muscle in voicing control. *The Journal of the Acoustical Society of America*, 85(3):1314–1321, 1989.

[80] Anders Löfqvist, Thomas Baer, Nancy S McGarr, and Robin Seider Story. The cricothyroid muscle in voicing control. *The Journal of the Acoustical Society of America*, 85(3):1314–1321, 1989.

[81] Anders Löfqvist, Laura L Koenig, and Richard S McGowan. Vocal tract aerodynamics in/aca/utterances: Measurements. *Speech Communication*, 16(1):49–66, 1995.

[82] Soren Y Lowell, Richard T Kelley, Raymond H Colton, Patrick B Smith, and Joel E Portnoy. Position of the hyoid and larynx in people with muscle tension dysphonia. *The laryngoscope*, 122(2):370–377, 2012.

[83] Jorge C Lucero. The minimum lung pressure to sustain vocal fold oscillation. *The Journal of the Acoustical Society of America*, 98(2):779–784, 1995.

[84] Jorge C Lucero. Relation between the phonation threshold pressure and the prephonatory glottal width in a rectangular glottis. *The Journal of the Acoustical Society of America*, 100(4):2551–2554, 1996.

[85] Jorge C Lucero and Laura L Koenig. Phonation thresholds as a function of laryngeal size in a two-mass model of the vocal folds. *The Journal of the Acoustical Society of America*, 118(5):2798–2801, 2005.

[86] Jorge C Lucero and Laura L Koenig. On the relation between the phonation threshold lung pressure and the oscillation frequency of the vocal folds. *The Journal of the Acoustical Society of America*, 121(6):3280–3283, 2007.

[87] Jorge C Lucero, Kélem G Lourenço, Nicolas Hermant, Annemie Van Hirtum, and Xavier Pelorson. Effect of source–tract acoustical coupling on the oscillation onset of the vocal folds. *The Journal of the Acoustical Society of America*, 132(1):403–411, 2012.

[88] Jorge C Lucero and Jean Schoentgen. Smoothness of an equation for the glottal flow rate versus the glottal area. *The Journal of the Acoustical Society of America*, 137(5):2970–2973, 2015.

[89] Paul Luizard and Xavier Pelorson. Threshold of oscillation of a vocal fold replica with unilateral surface growths. *The Journal of the Acoustical Society of America*, 141(5):3050–3058, 2017.

[90] Ted Mau, Joseph Muhlestein, Sean Callahan, Kent T Weinheimer, and Roger W Chan. Phonation threshold pressure and flow in excised human larynges. *The Laryngoscope*, 121(8):1743–1751, 2011.

[91] Victoria S McKenna, Elizabeth S Heller Murray, Yu-An S Lien, and Cara E Stepp. The relationship between relative fundamental frequency and a kinematic estimate

of laryngeal stiffness in healthy adults. *Journal of Speech, Language, and Hearing Research*, 59(6):1283–1294, 2016.

[92] Daryush D Mehta, Jarrad H Van Stan, Matías Zañartu, Marzyeh Ghassemi, John V Guttag, Víctor M Espinoza, Juan P Cortés, Harold A Cheyne, and Robert E Hillman. Using ambulatory voice monitoring to investigate common voice disorders: Research update. *Frontiers in bioengineering and biotechnology*, 3:155, 2015.

[93] Daryush D Mehta, Matias Zanartu, Shengran W Feng, Harold A Cheyne II, and Robert E Hillman. Mobile voice health monitoring using a wearable accelerometer sensor and a smartphone platform. *IEEE Transactions on Biomedical Engineering*, 59(11):3090–3096, 2012.

[94] Patrick Mergell, Hanspeter Herzel, Thomas Wittenberg, Monika Tigges, and Ulrich Eysholdt. Phonation onset: vocal fold modeling and high-speed glottography. *The Journal of the Acoustical Society of America*, 104(1):464–470, 1998.

[95] Mihai Mihaescu, Sid M Khosla, Shanmugam Murugappan, and Ephraim J Gutmark. Unsteady laryngeal airflow simulations of the intra-glottal vortical structures. *The Journal of the Acoustical Society of America*, 127(1):435–444, 2010.

[96] Rajat Mittal, Byron D Erath, and Michael W Plesniak. Fluid dynamics of human phonation and speech. *Annual Review of Fluid Mechanics*, 45:437–467, 2013.

[97] Burkhart Mohr. Intrinsic variations in the speech signal. *Phonetica*, 23(2):65–93, 1971.

[98] Scott Reid Moisik and Bryan Gick. The quantal larynx: The stable regions of laryngeal biomechanics and implications for speech production. *Journal of Speech, Language, and Hearing Research*, 60(3):540–560, 2017.

[99] Murray D Morrison and Linda A Rammage. Muscle misuse voice disorders: description and classification. *Acta oto-laryngologica*, 113(3):428–434, 1993.

[100] Jonathan D Mortensen, Anita N Vasavada, and Andrew S Merryweather. The inclusion of hyoid muscles improve moment generating capacity and dynamic simulations in musculoskeletal models of the head and neck. *PloS one*, 13(6):e0199912, 2018.

[101] Mohammadreza Movahhedi, Biao Geng, Qian Xue, and Xudong Zheng. Effects of cricothyroid and thyroarytenoid interaction on voice control: Muscle activity, vocal fold biomechanics, flow, and acoustics. *The Journal of the Acoustical Society of America*, 150(1):29–42, 2021.

[102] Jack Mullen, David M Howard, and Damian T Murphy. Real-time dynamic articulations in the 2-d waveguide mesh vocal tract model. *IEEE Transactions on Audio, Speech, and Language Processing*, 15(2):577–585, 2007.

[103] Mark Nagurka and Shuguang Huang. A mass-spring-damper model of a bouncing ball. In *Proceedings of the 2004 American Control Conference*, volume 1, pages 499–504. IEEE, 2004.

[104] Nathan M Newmark. A method of computation for structural dynamics. *Journal of the Engineering Mechanics Division*, 85(3):67–94, 1959.

[105] Duong Duy Nguyen, Dianna T Kenny, Ninh Duy Tran, and Jonathan R Livesey. Muscle tension dysphonia in vietnamese female teachers. *Journal of Voice*, 23(2):195–208, 2009.

[106] Naoya Nishida, Aki Taguchi, Kazumi Motoyoshi, Masamitsu Hyodo, Kiyofumi Gyo, and Junzo Desaki. Age-related changes in rat intrinsic laryngeal muscles: analysis of muscle fibers, muscle fiber proteins, and subneural apparatuses. *European Archives of Oto-Rhino-Laryngology*, 270(3):975–984, 2013.

[107] Arne B Nordmark et al. Non-periodic motion caused by grazing incidence in an impact oscillator. *Journal of Sound and Vibration*, 145(2):279–297, 1991.

[108] Richard Ogden. *Introduction to English Phonetics*. Edinburgh University Press, 2017.

[109] Ralph N Ohde. Fundamental frequency as an acoustic correlate of stop consonant voicing. *The Journal of the Acoustical Society of America*, 75(1):224–230, 1984.

[110] Rory O'Keeffe, Seyed Yahya Shirazi, Sarmad Mehrdad, Tyler Crosby, Aaron M Johnson, and S Farokh Atashzar. Perilaryngeal-cranial functional muscle network differentiates vocal tasks: A multi-channel semg approach. *IEEE Transactions on Biomedical Engineering*, 2022.

[111] Rory O'Keeffe, Yahya Shirazi, Sarmad Mehrdad, Tyler Crosby, Aaron M Johnson, and S Farokh Atashzar. Perilaryngeal functional muscle network in patients with vocal hyperfunction-a case study. *bioRxiv*, pages 2023–01, 2023.

[112] Anil Palaparthi, Simeon Smith, and Ingo R Titze. Mapping thyroarytenoid and cricothyroid activations to postural and acoustic features in a fiber-gel model of the vocal folds. *Applied Sciences*, 9(21):4671, 2019.

[113] Yeonggwang Park, Feng Wang, Manuel Díaz-Cádiz, Jennifer M Vojtech, Matti D Groll, and Cara E Stepp. Vocal fold kinematics and relative fundamental frequency as a function of obstruent type and speaker age. *The Journal of the Acoustical Society of America*, 149(4):2189–2199, 2021.

[114] Rita R Patel, Karen Forrest, and Drew Hedges. Relationship between acoustic voice onset and offset and selected instances of oscillatory onset and offset in young healthy men and women. *Journal of Voice*, 31(3):389.e9–389.e17, 2017.

[115] William G Pearson, Susan E Langmore, and Ann C Zumwalt. Evaluating the structural properties of suprahyoid muscles and their potential for moving the hyoid. *Dysphagia*, 26(4):345–351, 2011.

[116] Xavier Pelorson, Avraham Hirschberg, RR Van Hassel, APJ Wijnands, and Yves Auregan. Theoretical and experimental study of quasisteady-flow separation within the glottis during phonation. application to a modified two-mass model. *The Journal of the Acoustical Society of America*, 96(6):3416–3431, 1994.

[117] Pranati Pillutla, Neha K Reddy, Patrick Schlegel, Zhaoyan Zhang, and Dinesh K Chhetri. Control of pre-phonatory glottal shape by intrinsic laryngeal muscles. *The Laryngoscope*, 2022.

[118] Christopher J Poletto, Laura P Verdun, Robert Strominger, and Christy L Ludlow. Correspondence between laryngeal vocal fold movement and muscle activity during speech and nonspeech gestures. *Journal of Applied Physiology*, 97(3):858–866, 2004.

[119] Ali Rajaei, Ebrahim Barzegar Bafrooei, Fariba Mojiri, and Mohammad Hussein Nilforoush. The occurrence of laryngeal penetration and aspiration in patients with glottal closure insufficiency. *International Scholarly Research Notices*, 2014, 2014.

[120] Martin Rothenberg. A new inverse-filtering technique for deriving the glottal air flow waveform during voicing. *The Journal of the Acoustical Society of America*, 53(6):1632–1645, 1973.

[121] Bernard Roubeau, Claude Chevrie-Muller, and Jean Lacau Saint Guily. Electromyographic activity of strap and cricothyroid muscles in pitch change. *Acta oto-laryngologica*, 117(3):459–464, 1997.

[122] Nelson Roy. Functional dysphonia. *Current Opinion in Otolaryngology & Head and Neck Surgery*, 11(3):144–148, 2003.

[123] Nelson Roy. Assessment and treatment of musculoskeletal tension in hyperfunctional voice disorders. *International Journal of Speech-Language Pathology*, 10(4):195–209, 2008.

[124] Nelson Roy, Ray M Merrill, Steven D Gray, and Elaine M Smith. Voice disorders in the general population: prevalence, risk factors, and occupational impact. *The Laryngoscope*, 115(11):1988–1995, 2005.

[125] M Senator. Existence and stability of periodic motions of a harmonically forced impacting system. *The Journal of the Acoustical Society of America*, 47(5B):1390–1397, 1970.

[126] Mohamed A Serry, Gabriel A Alzamendi, Matías Zañartu, and Sean D Peterson. Modeling the influence of the extrinsic musculature on phonation. *Biomechanics and Modeling in Mechanobiology*, pages 1–14, 2023.

[127] Mohamed A Serry, Cara E Stepp, and Sean D Peterson. Physics of phonation offset: Towards understanding relative fundamental frequency observations. *The Journal of the Acoustical Society of America*, 149(5):3654–3664, 2021.

[128] Mohamed A Serry, Cara E Stepp, and Sean D Peterson. Exploring the mechanics of fundamental frequency variation during phonation onset. *Biomechanics and Modeling in Mechanobiology*, pages 1–18, 2022.

[129] Steven Wayne Shaw and PJ Holmes. A periodically forced piecewise linear oscillator. *Journal of sound and vibration*, 90(1):129–155, 1983.

[130] Travis L Shiba and Dinesh K Chhetri. Dynamics of phonatory posturing at phonation onset. *The Laryngoscope*, 126(8):1837–1843, 2016.

[131] Petr Šidlof, Olivier Doaré, Olivier Cadot, and Antoine Chaigne. Measurement of flow separation in a human vocal folds model. *Experiments in Fluids*, 51(1):123–136, 2011.

[132] Allan B Smith and Michael P Robb. Factors underlying short-term fundamental frequency variation during vocal onset and offset. *Speech, Language and Hearing*, 16(4):208–214, 2013.

[133] Maria Södersten, Stellan Hertegård, and Britta Hammarberg. Glottal closure, transglottal airflow, and voice quality in healthy middle-aged women. *Journal of Voice*, 9(2):182–197, 1995.

[134] B. Sonesson. Vocal fold kinesiology. In Sten Grillner, B. Lindblom, J. Lubker, and A. Persson, editors, *Speech Motor Control*, Wenner–Gren Center International Symposium Series, pages 113–117. Pergamon, 1982.

[135] Ina Steinecke and Hanspeter Herzel. Bifurcations in an asymmetric vocal-fold model. *The Journal of the Acoustical Society of America*, 97(3):1874–1884, 1995.

[136] Cara E Stepp, Robert E Hillman, and James T Heaton. The impact of vocal hyperfunction on relative fundamental frequency during voicing offset and onset. *Journal of Speech, Language, and Hearing Research*, 53(5):1220–1226, 2010.

[137] Cara E Stepp, Rosemary A Lester-Smith, Defne Abur, Ayoub Daliri, J Pieter Noordzij, and Ashling A Lupiani. Evidence for auditory-motor impairment in individuals with hyperfunctional voice disorders. *Journal of Speech, Language, and Hearing Research*, 60(6):1545–1550, 2017.

[138] Cara E. Stepp, Gabrielle R. Merchant, James T. Heaton, and Robert E. Hillman. Effects of voice therapy on relative fundamental frequency during voicing offset and onset in patients with vocal hyperfunction. *Journal of Speech, Language, and Hearing Research*, 54(5):1260–1266, 2011.

[139] Brad H Story. An overview of the physiology, physics and modeling of the sound source for vowels. *Acoustical Science and Technology*, 23(4):195–206, 2002.

[140] Brad H Story. 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, 2005.

[141] Brad H Story. Comparison of magnetic resonance imaging-based vocal tract area functions obtained from the same speaker in 1994 and 2002. *The Journal of the Acoustical Society of America*, 123(1):327–335, 2008.

[142] Brad H Story and Ingo R Titze. Voice simulation with a body-cover model of the vocal folds. *The Journal of the Acoustical Society of America*, 97(2):1249–1260, 1995.

[143] Brad H Story, Ingo R Titze, and Eric A Hoffman. Vocal tract area functions from magnetic resonance imaging. *The Journal of the Acoustical Society of America*, 100(1):537–554, 1996.

[144] Brad Hudson Story. *Physiologically-based speech simulation using an enhanced wave-reflection model of the vocal tract.* PhD thesis, University of Iowa, 1995.

[145] Hans Werner Strube. The meaning of the kelly–lochbaum acoustic-tube model. *The Journal of the Acoustical Society of America*, 108(4):1850–1855, 2000.

[146] Hans Werner Strube. Are conical segments useful for vocal-tract simulation?(l). *The Journal of the Acoustical Society of America*, 114(6):3028–3031, 2003.

[147] Juan Suárez-Quintanilla, Alejandro Fernández Cabrera, and Sandeep Sharma. Anatomy, head and neck, larynx. *StatPearls [Internet]*, 2020.

[148] Reiki Sukekawa and Ichizoh Itoh. Anatomical study of the human omohyoid muscle: regarding intermediate morphologies between normal and anomalous morphologies of the superior belly. *Anatomical science international*, 81(2):107–114, 2006.

[149] Lucian Sulica and Alison Behrman. Management of benign vocal fold lesions: a survey of current opinion and practice. *Annals of Otology, Rhinology & Laryngology*, 112(10):827–833, 2003.

[150] Hironori Takemoto, Kiyoshi Honda, Shinobu Masaki, Yasuhiro Shimada, and Ichiro Fujimoto. Measurement of temporal changes in vocal tract area function from 3D cine-MRI data. *The Journal of the Acoustical Society of America*, 119(2):1037–1049, 2006.

[151] Scott L Thomson, Luc Mongeau, and Steven H Frankel. Aerodynamic transfer of energy to the vocal folds. *The Journal of the Acoustical Society of America*, 118(3):1689–1700, 2005.

[152] Prof. S.P. Timoshenko. Lxvi. on the correction for shear of the differential equation for transverse vibrations of prismatic bars. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, 41(245):744–746, 1921.

[153] Ingo Titze and Fariborz Alipour. *The Myoelastic-Aerodynamic Theory of Phonation*. 01 2006.

[154] Ingo R Titze. Parameterization of the glottal area, glottal flow, and vocal fold contact area. *The Journal of the Acoustical Society of America*, 75(2):570–580, 1984.

[155] Ingo R Titze. The physics of small-amplitude oscillation of the vocal folds. *The Journal of the Acoustical Society of America*, 83(4):1536–1552, 1988.

[156] Ingo R Titze. On the relation between subglottal pressure and fundamental frequency in phonation. *The Journal of the Acoustical Society of America*, 85(2):901–906, 1989.

[157] Ingo R Titze. Physiologic and acoustic differences between male and female voices. *The Journal of the Acoustical Society of America*, 85(4):1699–1707, 1989.

[158] Ingo R Titze. Regulating glottal airflow in phonation: Application of the maximum power transfer theorem to a low dimensional phonation model. *The Journal of the Acoustical Society of America*, 111(1):367–376, 2002.

[159] Ingo R Titze. Nonlinear source–filter coupling in phonation: Theory. *The Journal of the Acoustical Society of America*, 123(4):1902–1915, 2008.

[160] Ingo R Titze and Eric J Hunter. A two-dimensional biomechanical model of vocal fold posturing. *The Journal of the Acoustical Society of America*, 121(4):2254–2260, 2007.

[161] Ingo R Titze, Jiaqi Jiang, and David G Drucker. Preliminaries to the body-cover theory of pitch control. *Journal of Voice*, 1(4):314–319, 1988.

[162] Ingo R Titze, Erich S Luschei, and Minoru Hirano. Role of the thyroarytenoid muscle in regulation of fundamental frequency. *Journal of Voice*, 3(3):213–224, 1989.

[163] Ingo R Titze, Sheila S Schmidt, and Michael R Titze. Phonation threshold pressure in a physical model of the vocal fold mucosa. *The Journal of the Acoustical Society of America*, 97(5):3080–3084, 1995.

[164] Ingo R Titze and Brad H Story. Rules for controlling low-dimensional vocal fold models with muscle activation. *The Journal of the Acoustical Society of America*, 112(3):1064–1076, 2002.

[165] Tomáš Vampola, Jaromír Horáček, and Ivo Klepáček. Computer simulation of mucosal waves on vibrating human vocal folds. *Biocybernetics and biomedical engineering*, 36(3):451–465, 2016.

[166] JW Van den Berg, JT Zantema, and P Doornenbal Jr. On the air resistance and the bernoulli effect of the human larynx. *The Journal of the Acoustical Society of America*, 29(5):626–631, 1957.

[167] Kees van den Doel and Uri M Ascher. Real-time numerical solution of webster's equation on a nonuniform grid. *IEEE transactions on audio, speech, and language processing*, 16(6):1163–1172, 2008.

[168] Chris Alan Van Ee. *Tensile properties of the human muscular and ligamentous cervical spine*. Duke University, 2000.

[169] Evelyne Van Houtte, Kristiane Van Lierde, and Sofie Claeys. Pathophysiology and treatment of muscle tension dysphonia: a review of the current knowledge. *Journal of Voice*, 25(2):202–207, 2011.

[170] Jarrad H Van Stan, Daryush D Mehta, Andrew J Ortiz, James A Burns, Katherine L Marks, Laura E Toles, Tara Stadelman-Cohen, Carol Krusemark, Jason Muise, Tiffiny Hron, Steven M Zeitels, Annie B Fox, and Robert E Hillman. Changes in a daily phonotrauma index after laryngeal surgery and voice therapy: Implications for the role of daily voice use in the etiology and pathophysiology of phonotraumatic vocal hyperfunction. *Journal of Speech, Language, and Hearing Research*, 63(12):3934–3944, 2020.

[171] Jarrad H Van Stan, Daryush D Mehta, Andrew J Ortiz, James A Burns, Laura E Toles, Katherine L Marks, Mark Vangel, Tiffiny Hron, Steven Zeitels, and Robert E Hillman. Differences in weeklong ambulatory vocal behavior between female patients with phonotraumatic lesions and matched controls. *Journal of Speech, Language, and Hearing Research*, 63(2):372–384, 2020.

[172] Jarrad H Van Stan, Andrew J Ortiz, Juan P Cortes, Katherine L Marks, Laura E Toles, Daryush D Mehta, James A Burns, Tiffiny Hron, Tara Stadelman-Cohen, Carol Krusemark, et al. Differences in daily voice use measures between female patients with nonphonotraumatic vocal hyperfunction and matched controls. *Journal of Speech, Language, and Hearing Research*, 64(5):1457–1470, 2021.

[173] Katherine Verdolini, Young Min, Ingo R Titze, Jon Lemke, Kice Brown, Miriam van Mersbergen, Jack Jiang, and Kim Fisher. Biological mechanisms underlying voice changes due to dehydration. *Journal of Speech, Language, and Hearing Research*, 2002.

[174] Erkki Vilkman, Aatto Sonninen, Perti Hurme, and Pentti Körkkö. External laryngeal frame function in voice production revisited: a review. *Journal of Voice*, 10(1):78–92, 1996.

[175] Jennifer M Vojtech, Roxanne K Segina, Daniel P Buckley, Katharine R Kolin, Monique C Tardif, J Pieter Noordzij, and Cara E Stepp. Refining algorithmic estimation of relative fundamental frequency: Accounting for sample characteristics and fundamental frequency estimation method. *The Journal of the Acoustical Society of America*, 146(5):3184–3202, 2019.

[176] Ben C Watson. Fundamental frequency during phonetically governed devoicing in normal young and aged speakers. *The Journal of the Acoustical Society of America*, 103(6):3642–3647, 1998.

[177] Hasini R Weerathunge, Gabriel A Alzamendi, Gabriel J Cler, Frank H Guenther, Cara E Stepp, and Matías Zañartu. Ladiva: A neurocomputational model providing laryngeal motor control for speech acquisition and production. *PLOS Computational Biology*, 18(6):e1010159, 2022.

[178] Ewald R Weibel, André Frédérick Cournand, and Dickinson W Richards. *Morphometry of the human lung*, volume 1. Springer, 1963.

[179] Darrell Wong, Mabo R Ito, Neil B Cox, and Ingo R Titze. 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, 1991.

[180] Qian Xue, Xudong Zheng, Rajat Mittal, and Steven Bielamowicz. Subject-specific computational modeling of human phonation. *The Journal of the Acoustical Society of America*, 135(3):1445–1456, 2014.

[181] Jun Yin and Zhaoyan Zhang. 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, 2013.

[182] Jun Yin and Zhaoyan Zhang. Interaction between the thyroarytenoid and lateral cricoarytenoid muscles in the control of vocal fold adduction and eigenfrequencies. *Journal of biomechanical engineering*, 136(11):111006, 2014.

[183] Jun Yin and Zhaoyan Zhang. Laryngeal muscular control of vocal fold posturing: Numerical modeling and experimental validation. *The Journal of the Acoustical Society of America*, 140(3):EL280–EL284, 2016.

[184] Matías Zañartu. Influence of acoustic loading on the flowinduced oscillations of single mass models of the human larynx. Master's thesis, Purdue University, 2006.

[185] Matías Zañartu. *Acoustic coupling in phonation and its effect on inverse filtering of oral airflow and neck surface acceleration*. PhD thesis, Citeseer, 2010.

[186] Matías Zañartu, Gabriel E Galindo, Byron D Erath, Sean D Peterson, George R Wodicka, and Robert E Hillman. Modeling the effects of a posterior glottal opening

on vocal fold dynamics with implications for vocal hyperfunction. *The Journal of the Acoustical Society of America*, 136(6):3262–3271, 2014.

[187] Kai Zhang, Thomas Siegmund, and Roger W Chan. A constitutive model of the human vocal fold cover for fundamental frequency regulation. *The Journal of the Acoustical Society of America*, 119(2):1050–1062, 2006.

[188] Kai Zhang, Thomas Siegmund, and Roger W Chan. Modeling of the transient responses of the vocal fold lamina propria. *Journal of the mechanical behavior of biomedical materials*, 2(1):93–104, 2009.

[189] Yu Zhang and Jack J Jiang. Chaotic vibrations of a vocal fold model with a unilateral polyp. *The Journal of the Acoustical Society of America*, 115(3):1266–1269, 2004.

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

[191] Zhaoyan Zhang. Mechanics of human voice production and control. *The Journal of the Acoustical Society of America*, 140(4):2614–2635, 2016.

[192] Zhaoyan Zhang. Respiratory laryngeal coordination in airflow conservation and reduction of respiratory effort of phonation. *Journal of Voice*, 30(6):760–e7, 2016.

[193] Zhaoyan Zhang. Structural constitutive modeling of the anisotropic mechanical properties of human vocal fold lamina propria. *The Journal of the Acoustical Society of America*, 145(6):EL476–EL482, 2019.

[194] Zhaoyan Zhang. Vocal fold contact pressure in a three-dimensional body-cover phonation model. *The Journal of the Acoustical Society of America*, 146(1):256–265, 2019.

APPENDICES

Appendix A

Physics of Phonation Offset

Important note: The analysis in this appendix is related to the investigation in Chapter 3. Variables with no definition herein have been previously defined in Chapter 3.

A.1 Derivation of the Frequency Formula

Let ϕ denote the solution to Equation (3.1), with $\phi(t_0) = 0$, $\dot{\phi}(t_0) = v_0 = \sqrt{2E_0/M}$, in the collision-free interval $[t_0, t_2]$, see Figure 3.2. Then, ϕ is given explicitly by $\phi(t) = v_0\sqrt{M/K}\sin(\sqrt{K/M}(t - t_0))$, $t \in [t_0, t_2]$. The term $t_f = t_1 - t_0$ is obtained by solving the algebraic equation $\phi(t_f + t_0) = 0$ which results in $t_f = \pi\sqrt{M/K}$. Moreover, the term $t_\delta = t_2 - t_1$ can be obtained by solving the algebraic equation $\phi(t_\delta + t_0) = \delta$, which has the explicit solution $t_\delta = \sqrt{M/K}\arcsin(\tilde{\delta})$. To obtain an analytic expression of the contact time t_c, let $\eta(t) = -(\xi(t) + \delta)$ be the displacement beyond the collision plane; then Equation (3.2) can be rewritten as

$$M\ddot{\eta} + (K + k_{\mathbf{col}})\eta = -K\delta,\ \eta(t) > 0. \tag{A.1}$$

Deriving an analytical expression of t_c requires solving Equation (A.1) over the interval $[t_2, t_3]$ in the collision regime. The velocity prior to the collision, namely $v_{\mathbf{col}}$ can be obtained from the energy equation $2E_0 = K\delta^2 + Mv_{\mathbf{col}}^2$ resulting in $v_{\mathbf{col}} = \sqrt{(2E_0 - K\delta^2)/M}$. Now, let ψ be the solution to Equation (A.1), over the interval $[t_2, t_3]$ with $\psi(t_2) = 0$ and $\dot{\psi}(t_2) = v_{\mathbf{col}}$, which is given explicitly as $\psi(t) = \tilde{k}\delta(C\sin(\omega(t - t_2)) + \cos(\omega(t - t_2)) - 1)$, $t \in [t_2, t_3]$, where $\omega = \sqrt{(K + k_{\mathbf{col}})/M}$ and $C = \sqrt{(1/\tilde{\delta}^2 - 1)/\tilde{k}}$. Direct substitution of

ψ into Equation (A.1) demonstrates that it is indeed a solution to the ODE. Then, t_c is obtained by solving the algebraic equation $\psi(t_c + t_2) = 0$, which has the analytic solution $t_c = 2\arctan(C)/\omega$.

A.2 Fundamental Frequency Bounds

The upper bound follows easily from the fact that $\arctan(x), \arcsin(x) \geq 0$ for all $x \geq 0$. The lower bound is obtained as follows. Note that for all $y \in (0, 1]$ and all $x \in [0, \infty)$, $y\arctan(x/y) \leq \arctan(x)$. This inequality can be obtained by comparing the integrands of $\arctan(x) = \int_0^x 1/(1 + t^2)\,\mathrm{d}t$ and $y\arctan(x/y) = \int_0^x 1/(1 + (t/y)^2)\,\mathrm{d}t$. Using this inequality in the expression of f results in the inequality $f \geq 2f_n/([2/\pi]\arctan\sqrt{1/\tilde{\delta}^2 - 1} + [2/\pi]\arcsin(\tilde{\delta})+1) = g$. Moreover, using the identity $\arctan\left(\sqrt{1-x^2}/x\right) = \pi/2-\arcsin(x)$, $x > 0$, in the expression of g simplifies it to be $g = f_n$.

A.3 Maximum Collision Force

As the collision forces in the collision regime are linearly elastic, we need to find the maximum displacement $\eta_{\mathbf{max}}$ beyond the collision plane. To do so, we resort to the energy equation (deduced from Equation (A.1) by integrating both sides with respect to η) $M\dot{\eta}(\tau_1)^2/2+(K+k_{\mathbf{col}})\eta(\tau_1)^2/2+K\delta\eta(\tau_1) = M\dot{\eta}(\tau_2)^2/2+(K+k_{\mathbf{col}})\eta(\tau_2)^2/2+K\delta\eta(\tau_2)$, $\tau_1, \tau_2 \in [t_2, t_3]$. Let $\tau_1 = t_2$ (beginning of collision) and τ_2 be the time instance corresponding to the maximum displacement $\eta_{\mathbf{max}}$. Note that at τ_1, $\eta(\tau_1) = 0$, and $\dot{\eta}(\tau_1) = v_{\mathbf{col}} = \sqrt{(2E_0 - K\delta^2)/M}$ and at τ_2, $\eta(\tau_2) = \eta_{\mathbf{max}}$ and $\dot{\eta}(\tau_2) = 0$. Consequently, we have $(K+k_{\mathbf{col}})\eta_{\mathbf{max}}^2+2K\delta\eta_{\mathbf{max}}-Mv_{\mathbf{col}}^2 = 0$, which results in $\eta_{\mathbf{max}} = \tilde{k}\delta(-1 + \sqrt{1 + (1/\tilde{\delta}^2 - 1)/\tilde{k}})$. Substituting the formula of $\eta_{\mathbf{max}}$ into the formula $F_{\mathbf{col,max}} = k_{\mathbf{col}}\eta_{\mathbf{max}}$ and rearranging results in Equation (3.4).

Appendix B

Fundamental Frequency Variation during Phonation Onset

Important note: The analysis in this appendix is related to the investigation in Chapter 4. Variables with no definition herein have been previously defined in Chapter 4.

B.1 Quasi-Steady Viscous Glottal Flow Model

Herein, we derive a viscous glottal flow model similar to that presented in [88], but that accounts for the convergent/divergent configurations arising in the body-cover VF model. A schematic diagram of the body-cover model [142] is shown in Figure B.1, where y_1, y_2, and y_b denote the displacements of the inferior cover mass m_1, superior cover mass m_2, and body mass m_b, respectively. The thicknesses of m_1 and m_2 are given by T_1 and T_2, respectively, and $T = T_1 + T_2$ yields the total VF thickness. The glottal areas associated with the displacements of the inferior and superior masses are given by $a_1 = 2L\max\{y_1, 0\}$ and $a_2 = 2L\max\{y_2, 0\}$, respectively, where L is the VF length. The subglottal and supraglottal areas are A_s and A_e, respectively, with associated pressures P_s and P_e.

The x-axis is aligned with the streamwise direction with $x = 0$ located at the inferior margin of m_1, the junction between the lower and upper masses is located at $x = T_1$, and $x = T$ corresponds to the superior margin of m_2. The density of air is ρ, the viscosity is μ, and the associated speed of sound is c_a. We assume the volumetric flow rate, q, to be quasi-steady. We assume Poiseuille flow throughout the glottis when the flow attachment criterion $0 < a_2 \leq 1.2a_1$ is satisfied [116, 131, 158]. In the case of flow separation, wherein

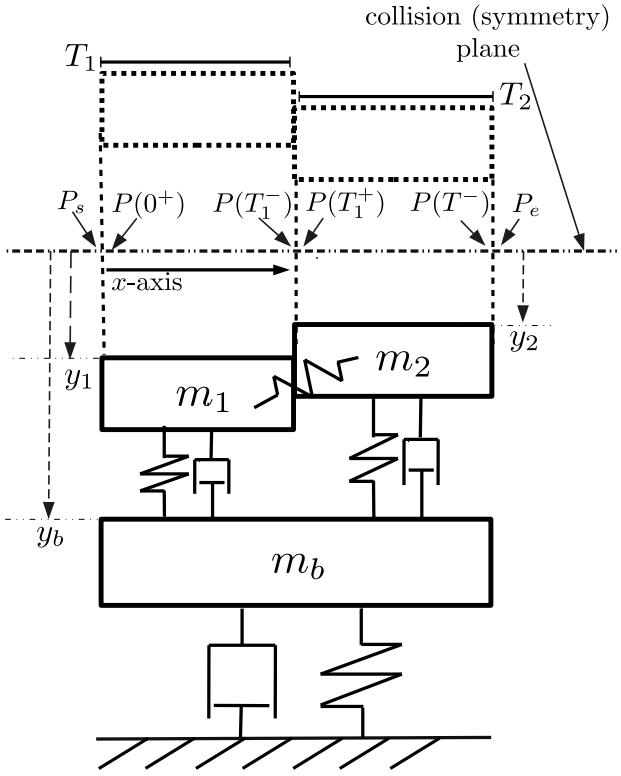


Figure B.1: Schematic of the body-cover model with locations of important pressure values identified.

$a_2 > 1.2a_1 > 0$, Poiseuille flow is assumed over the inferior mass only and the supraglotal pressure ($P = P_e$) is applied over the superior mass. When the masses are in contact, the loading conditions described by [142] are employed. In particular, if $a_1 > 0$ and $a_2 = 0$, then $P = P_s$ and $P = 0$ over the lower mass and upper mass, respectively. Moreover, if $a_1 = 0$ and $a_2 > 0$, then $P = 0$ and $P = P_e$ over the lower mass and upper mass, respectively. Finally, if $a_1 = a_2 = 0$, then $P = 0$ over the lower and upper masses. We note that in our derivation of the flow model, it is assumed that the glottal flow does not revert (i.e., $P_s - P_e \geq 0$).

Considering the attached flow case, $0 < a_2 \leq 1.2a_1$, let $P(0^+)$ denote the pressure at the inferior margin of the inferior mass, that is, $P(0^+) = \lim_{x \to 0^+} P(x)$ (see Figure B.1). We assume Bernoulli flow between the subglottal region and the entry to the glottis, with

$A_s \gg a_1$, resulting in the equation

$$P(0^+) + \frac{1}{2}\rho\left(\frac{q}{a_1}\right)^2 = P_s. \tag{B.1}$$

Moreover, we assume Poiseuille flow along mass m_1 ($x \in (0, T_1)$), yielding

$$P(0^+) = P(x) + x\frac{12\mu L^2 q}{a_1^3}, \tag{B.2}$$

and, taking the limit as x approaches T_1 from the left side, yields

$$P(0^+) = P(T_1^-) + T_1\frac{12\mu L^2 q}{a_1^3}. \tag{B.3}$$

Assuming a Bernoulli flow through the junction between the cover masses at $x = T_1$ yields

$$P(T_1^-) + \frac{1}{2}\rho\left(\frac{q}{a_1}\right)^2 = P(T_1^+) + \frac{1}{2}\rho\left(\frac{q}{a_2}\right)^2. \tag{B.4}$$

Assuming a Poiseuille flow again along mass m_2 ($x \in (T_1, T)$) results in

$$P(T_1^+) = P(x) + (x - T_1)\frac{12\mu L^2 q}{a_2^3},\ x \in (T_1, T). \tag{B.5}$$

Taking the limit as x approaches T from the left side yields

$$P(T_1^+) = P(T^-) + T_2\frac{12\mu L^2 q}{a_2^3}. \tag{B.6}$$

Finally, at the superior margin of the VFs, we assume flow separation with no pressure recovery, which gives

$$P(T^-) = P_e. \tag{B.7}$$

Combining Equations (B.1), (B.3), (B.4), (B.6), and (B.7) results in a quadratic equation for the flow rate given by

$$\frac{\rho}{2a_2^2}q^2 + 12\mu L^2\left(\frac{T_1}{a_1^3} + \frac{T_2}{a_2^3}\right)q - (P_s - P_e) = 0, \tag{B.8}$$

the solution of which is

$$q = \frac{2a_2^3\delta_p}{\gamma + \sqrt{\gamma^2 + 2\rho a_2^4\delta_p}}, \tag{B.9}$$

where $\delta_p = P_s - P_e$ and $\gamma = 12\mu L^2 \left(T_1(a_2/a_1)^3 + T_2\right)$. Equation (B.9) can then be used to determine the pressures applied to each cover mass through substitution back into the Bernoulli/Poiseuille flow equations from which they were obtained. The aerodynamic forces over the lower and upper masses, F_l and F_u, are then computed as

$$F_l = L\int_0^{T_1} P(x)dx = LT_1P(0^+) - L\frac{T_1^2}{2}\frac{12\mu L^2 q}{a_1^3} \tag{B.10}$$

and

$$F_u = L\int_{T_1}^{T} P(x)dx = LT_2P(T_1^+) - L\frac{T_2^2}{2}\frac{12\mu L^2 q}{a_2^3}. \tag{B.11}$$

For the detached flow case, $a_2 > 1.2a_1 > 0$, we assume Bernoulli flow from the subglottal region to the inferior margin of the inferior mass, viscous flow over the inferior mass, and flow separation at the mass junction ($x = T_1$) leading to zero pressure recovery and uniform pressure P_e over the upper mass. Following a derivation similar to that of Equation (B.8) yields a quadratic equation for flow as

$$\frac{\rho}{2a_1^2}q^2 + 12\mu L^2\frac{T_1}{a_1^3}q - (P_s - P_e) = 0, \tag{B.12}$$

which has the solution

$$q = \frac{2a_1^3\delta_p}{\bar{\gamma} + \sqrt{\gamma^2 + 2\rho a_1^4\delta_p}}, \tag{B.13}$$

where $\bar{\gamma} = 12\mu L^2 T_1$. From this, we can obtain the aerodynamic force on the inferior mass, which is given by Equation B.10, and the superior mass, which is given by

$$F_u = L\int_{T_1}^{T} P(x)dx = LT_2P_e. \tag{B.14}$$

When acoustics are modelled using the wave reflection analog, flow rate has to be given in terms of incident pressure waves. The subglottal pressure P_s can be written as $P_s = P_s^+ + P_s^-$, where P_s^+ denotes the incident subglottal pressure and P_s^- denotes the outward travelling subglottal pressure. Similarly, the supraglottal pressure P_e can be written as $P_e = P_e^+ + P_e^-$, where P_e^+ denotes the outward travelling supraglottal pressure and P_e^- denotes the incident supraglottal pressure. Continuity then yields $P_s^- = P_s^+ - \rho c_a q/A_s$ and $P_e^+ = P_e^- + \rho c_a q/A_e$. Consequently, the transglottal pressure $P_s - P_e$ can be written as

$$P_s - P_e = 2(P_s^+ - P_e^-) - \frac{\rho c_a}{A^*}q, \tag{B.15}$$

where $A^* = A_s A_e/(A_s + A_e)$. By plugging Equation (B.15) into Equation (B.8) and solving for q, we obtain, for fully attached flow,

$$q = \frac{4a_2^3\tilde{\delta}_p}{\chi + \sqrt{\chi^2 + 4\rho a_2^4\tilde{\delta}_p}}, \tag{B.16}$$

where $\tilde{\delta}_p = P_s^+ - P_e^-$ and $\chi = 12\mu L^2\left(T_1(a_2/a_1)^3 + T_2\right) + \rho c_a a_2^3/A_*$. Similarly, for the detached flow, we obtain

$$q = \frac{4a_1^3\tilde{\delta}p}{\bar{\chi} + \sqrt{\bar{\chi}^2 + 4\rho a_1^4\tilde{\delta}_p}}, \tag{B.17}$$

where $\bar{\chi} = 12\mu L^2 T_1 + \rho c_a a_1^3/A^*$. Finally, if one of the cover masses is in contact, such that $a_1 = 0$ or $a_2 = 0$, then $q = 0$.

B.2 Derivation of the Approximate Velocity Recurrence Relation

Herein, we derive the approximate recurrence relation in Equation (4.7) for the hybrid model. Let $\eta : [0, \infty) \to \mathbb{R}$ be an oscillatory solution to Equation (4.1) with oscillations of possibly varying amplitude and frequency. Consider an oscillation period $\mathcal{I} = [t_0, t_2]$ depicted in Figure B.2, where it is assumed that $\eta(t_0) = -\delta$ and that the initial velocity $\dot{\eta}(t_0) = v_0 < 0$. To obtain the discrete system, we seek velocities v_1 and v_2 at time instances t_1 and t_2, corresponding to the first and second times after t_0 where the displacement $\eta(t) = -\delta$, respectively.

During the interval $\mathcal{I}_1 = [t_0, t_1]$, $\eta(t) \leq -\delta$ and the dynamics satisfy Equation (4.1b). For convenience, define $\phi(\cdot) = -(\eta(\cdot) + \delta)$, then on the interval $\mathcal{I}_1$, ϕ satisfies $M\ddot{\phi} + \mathcal{B}_2\dot{\phi} + \mathcal{K}\phi = c$, where $c = -K\delta$. Note that $\phi(t_0) = 0$ and $\bar{v}_0 = \dot{\phi}(t_0) = -v_0$. Therefore, ϕ is given explicitly as

$$\begin{aligned}\phi(t) = &\frac{\bar{v}_0 - \frac{\mathcal{B}_2 c}{2M\mathcal{K}}}{\omega_2} e^{-\frac{\mathcal{B}_2}{2M}(t-t_0)} \sin(\omega_2(t-t_0)) \\ &- \frac{c}{\mathcal{K}} e^{-\frac{\mathcal{B}_2}{2M}(t-t_0)} \cos(\omega_2(t-t_0)) + \frac{c}{\mathcal{K}}, \ t \in \mathcal{I}_1,\end{aligned}$$

and $\dot{\phi}(t) = \exp(-\mathcal{B}_2(t-t_0)/(2M))(\lambda \sin(\omega_2(t-t_0)) + \bar{v}_0 \cos(\omega_2(t-t_0)))$, where $\lambda = \omega_2 c/\mathcal{K} - \mathcal{B}_2\left[\bar{v}_0 - \mathcal{B}_2 c/(2M\mathcal{K})\right]/(2M\omega_2)$. In the case $c = 0$, $t_1 = t_0 + \pi/\omega_2$.

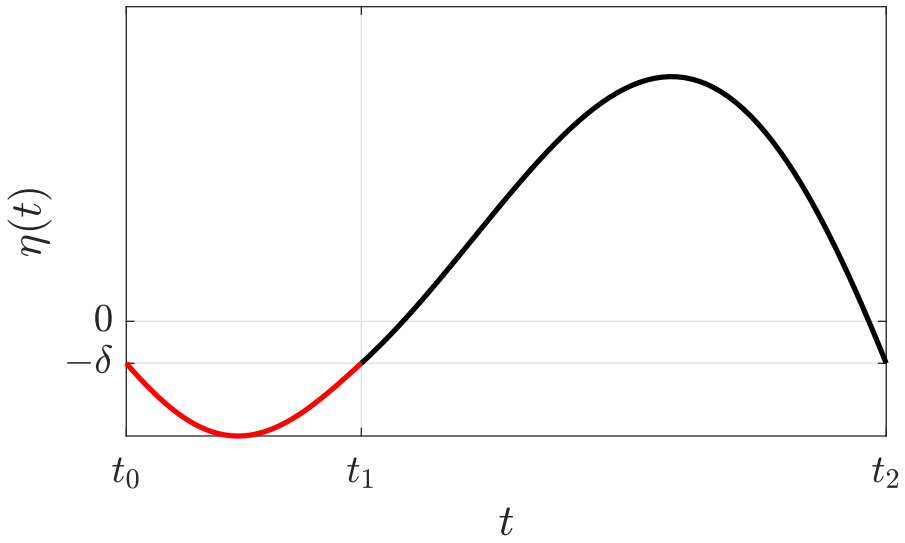


Figure B.2: One oscillation period of the hybrid phonation model.

If we assume c to be sufficiently small, utilizing a small neutral gap approximation, then t_1 can be obtained approximately as follows: let us write $t_1, \phi, \dot{\phi}$ as functions of c ($t_1 = t_1(c)$, $\phi = \phi(t, c)$, $\dot{\phi} = \dot{\phi}(t, c)$). Then, by implicitly differentiating the equation $\phi(t_1(c), c) = 0$ with respect to c and solving for $\mathrm{d}t_1/\mathrm{d}c$, we get $t_1' := \mathrm{d}t_1/\mathrm{d}c = -D_2\phi(t_1(c), c)/\dot{\phi}(t_1(c), c)$, where $D_i f$ denotes the partial derivative of a multivariable function f with respect to its i^{th} argument. By direct substitution, and using the fact that $t_1(0) = t_0 + \pi/\omega_2$, we get that $t_1'(0) = [1 + \exp(\mathcal{B}_2\pi/(2M\omega_2))]/(\mathcal{K}\bar{v}_0)$. Then, using first order Taylor expansion, we get $t_1(c) \approx t_1(0) + t_1'(0)c$, which results in the approximate formula $t_1 = t_0 + \pi/\omega_2 + \bar{t}$, where

$$\bar{t} = c\frac{1 + \mathrm{e}^{\frac{\mathcal{B}_2\pi}{2M\omega_2}}}{\mathcal{K}\bar{v}_0}. \tag{B.18}$$

By plugging the approximate formula of t_1 into the formula for $\dot{\phi}$ and using the trigonometric identities $\sin(\pi + x) = -\sin(x)$, $\cos(\pi + x) = -\cos(x)$, we obtain $\dot{\phi}(t_1) \approx \exp(-\mathcal{B}_2(\pi/\omega_2 + \bar{t})/(2M))(-\lambda\sin(\omega_2\bar{t}) - \bar{v}_0\cos(\omega_2\bar{t}))$. Note that, by definition, $\bar{t}$ is proportional to δ. By further utilizing the small neutral gap assumption, we implement the approximations $\exp(-\mathcal{B}_2\bar{t}/(2M)) \approx 1$, $\sin(\omega_2\bar{t}) \approx \omega_2\bar{t}$, and $\cos(\omega_2\bar{t}) \approx 1$ in the above expression, in addi-

tion to the definition of λ, which yields

$$\dot{\phi}(t_1) \approx -\left(-c\frac{\mathcal{B}_2}{2M}\frac{1+\mathrm{e}^{-\frac{\mathcal{B}_2\pi}{2M\omega_2}}}{\mathcal{K}\bar{v}_0}+\mathrm{e}^{-\frac{\mathcal{B}_2\pi}{2M\omega_2}}\right)\bar{v}_0 - \omega_2 c^2\frac{1+\mathrm{e}^{-\frac{\mathcal{B}_2\pi}{2M\omega_2}}}{\mathcal{K}\bar{v}_0}\left[\frac{\omega_2}{\mathcal{K}}+\frac{\mathcal{B}_2}{2M}\frac{\frac{\mathcal{B}_2}{2M\mathcal{K}}}{\omega_2}\right].$$

By neglecting the second order term (proportional to c^2), we obtain the final approximate formula $\dot{\phi}(t_1) \approx -\exp(-\mathcal{B}_2\pi/(2M\omega_2))\bar{v}_0+c\mathcal{B}_2[1+\exp(-\mathcal{B}_2\pi/(2M\omega_2))]/(2M\mathcal{K})$, and in terms of v_1 and v_0, where $c=-K\delta$ is substituted back, we get the approximate relation

$$v_1 = -\mathrm{e}^{-\frac{\mathcal{B}_2\pi}{2M\omega_2}}v_0+\tilde{k}\delta\frac{\mathcal{B}_2}{2M}\left(1+\mathrm{e}^{-\frac{\mathcal{B}_2\pi}{2M\omega_2}}\right). \tag{B.19}$$

By applying a similar analysis on the interval $\mathcal{I}_2=[t_1,t_2]$, where the small neutral gap approximation is utilized [1], we obtain the approximate formulas $t_2=t_1+\pi/\omega_1+\tilde{t}$, where

$$\tilde{t}=\delta\frac{1+\mathrm{e}^{-\frac{\mathcal{B}_1\pi}{2M\omega_1}}}{v_1}. \tag{B.20}$$

and

$$v_2=-\mathrm{e}^{\frac{\mathcal{B}_1\pi}{2M\omega_1}}v_1-\delta\frac{\mathcal{B}_1}{2M}(1+\mathrm{e}^{\frac{\mathcal{B}_1\pi}{2M\omega_1}}). \tag{B.21}$$

Combining Equations (B.19) and (B.21) yields $|v_2|=\mathcal{A}|v_0|+\mathcal{W}$, where

$$\mathcal{A}=\mathrm{e}^{\frac{\pi}{2M}\left[\frac{\mathcal{B}_1}{\omega_1}-\frac{\mathcal{B}_2}{\omega_2}\right]},$$
$$\mathcal{W}=\frac{\delta}{2M}\left[\tilde{k}\mathcal{B}_2\left(\mathcal{A}+\mathrm{e}^{\frac{\mathcal{B}_1\pi}{2M\omega_1}}\right)+\mathcal{B}_1\left(1+\mathrm{e}^{\frac{\mathcal{B}_1\pi}{2M\omega_1}}\right)\right].$$

By repeating the approximate solution procedure recursively over all oscillation periods, where collision is assumed to occur in each period, we obtain Equation (4.7).

[1] The damping term $\mathcal{B}_1$ grows unboundedly as δ approaches zero ($\mathcal{B}_1=2\tau P_L/(k_t\delta)-B$). Therefore, we additionally assume in the subsequent derivation that the subglottal lung pressure P_L is sufficiently small (depending on the value of δ), in order to have the damping term $\mathcal{B}_1$ bounded.

B.3 Derivation of the approximate frequency relation

The frequency of the oscillation period given in Figure B.2 is $\mathcal{F} = [t_2 - t_0]^{-1} = [(t_2 - t_1) + (t_1 - t_0)]^{-1}$. From the analysis in Appendix B.2, we have $t_2 - t_1 = \pi/\omega_1 + \tilde{t}$, and $t_1 - t_0 = \pi/\omega_2 + \bar{t}$, where $\bar{t}$ and $\tilde{t}$ are given in Equations (B.18) and (B.20), respectively. Consequently, we obtain $\mathcal{F} = [1/\mathcal{F}_\infty + \bar{t} + \tilde{t}]^{-1}$. By substituting Equation (B.19) into Equation (B.20), $\mathcal{F}$ can be written in terms of $|v_0|$ as:

$$\mathcal{F} = \left[\frac{1}{\mathcal{F}_\infty} + \frac{\alpha_1}{|v_0| + \beta} - \frac{\alpha_2}{|v_0|} \right]^{-1},$$

where

$$\begin{aligned}
\alpha_1 &= \delta \left(e^{\frac{\mathcal{B}_2 \pi}{2M\omega_2}} + \frac{1}{\mathcal{A}} \right), \\
\alpha_2 &= \tilde{k}\delta \left(e^{\frac{\mathcal{B}_2 \pi}{2M\omega_2}} + 1 \right), \\
\beta &= \tilde{k}\delta \frac{\mathcal{B}_2}{2M} \left(e^{\frac{\mathcal{B}_2 \pi}{2M\omega_2}} + 1 \right).
\end{aligned}$$

By repeating the above derivation over each oscillation period, we obtain Equation (4.11).

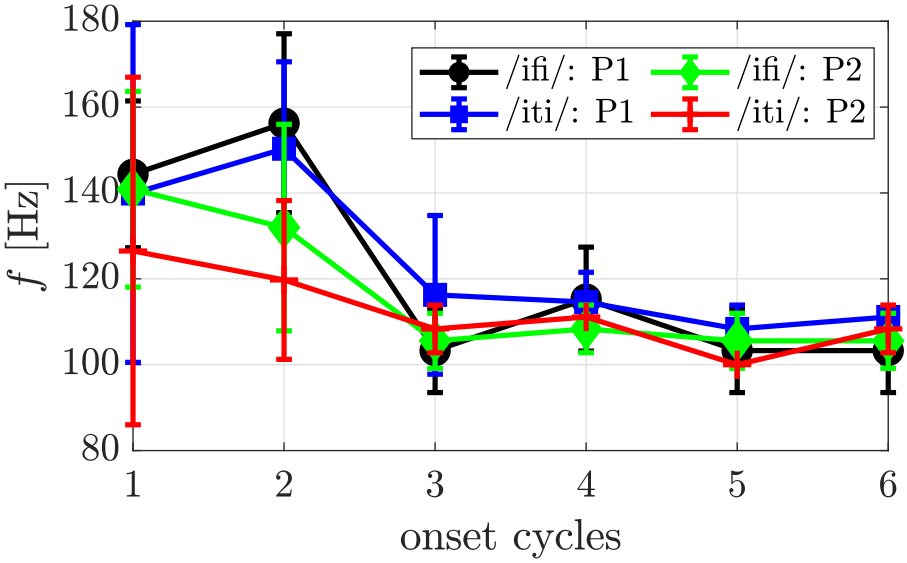
B.4 Empirical Data and Observations

Figure B.3(a) presents fundamental frequency data from two healthy male participants during the onset portion of repeated /ifi/ and /iti/ utterances[2]. For these utterances, the VFs transition from an abducted to an adducted state prior to and during the onset of the second vowel. Each data point corresponds to an average over four utterances, with error bars showing one standard deviation.

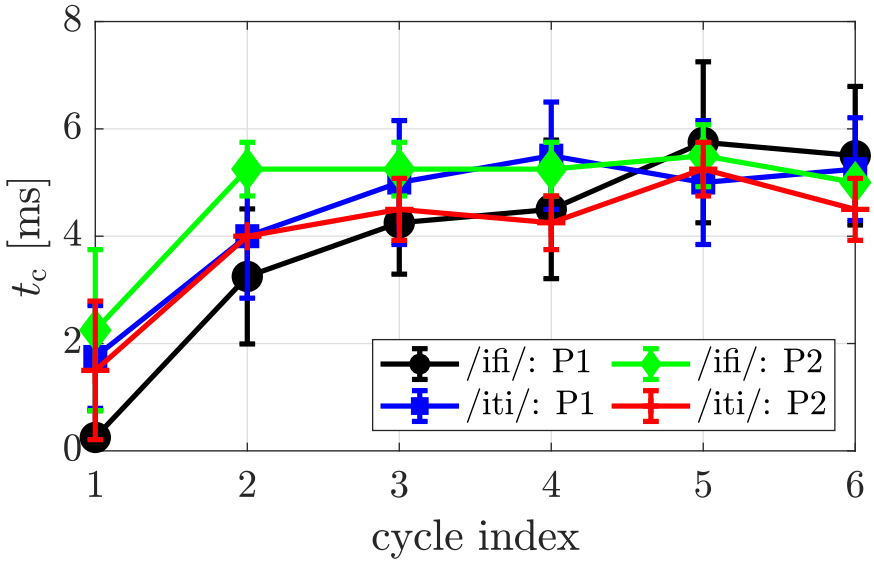
Figure B.3(b) presents the duration of contact for the first several oscillation cycles incorporating collision. Contact time is estimated from the glottal angle, θ_G, extracted from endoscopic video recordings during the utterances. The contact time per cycle, t_c, corresponds to the interval when $\theta_G = 0$. The figure shows that, in general, VF contact time

[2]Clinical data provided by the STEPP Lab, Boston University; these data were utilized in several previous studies on the kinematics of VFs [91, 25, 113].

per cycle increases gradually during phonation onset for both participants. Based upon the analysis in Section 4.2.1, the increasing role of collision in the dynamics should result in an increase in fundamental frequency. As per the discussion in Section 4.3.1, decaying CT activation leads to a decrease in fundamental frequency. Comparing empirical data in Figure B.3(a) with the simulation results in Figure 4.8, we note qualitative similarities between the frequency patterns of participant 1 (P1) and the numerical case with $a_{\mathrm{TA}} = 0.4$ and $a_{\mathrm{CT,i}} = 0.6$, and the frequency patterns of participant 2 (P2) and the numerical case with $a_{\mathrm{TA}} = 0.2$ and $a_{\mathrm{CT,i}} = 0.6$. That is, the modeling exercise is successful in qualitatively replicating several experimentally observed patterns during phonation onset.



(a) Fundamental frequency variation during phonation onset.



(b) Contact time per cycle during phonation onset.

Figure B.3: (a) Fundamental frequency and (b) contact duration during the onset portion of /ifi/ and /iti/ utterances collected from two participants (P1 and P2): points indicate the average across 4 utterances and error bars indicate one standard deviation.