REDUCED-ORDER VOCAL FOLD MODELS IN IN-VIVO DATA FOR THE ASSESSMENT OF UNDERLYING PHYSIOLOGICAL PARAMETERS AND MECHANISMS IN VOCAL HYPERFUNCTION

BY

Jesús Alberto Parra Peña, B.S.c.

A dissertation submitted

in partial fulfillment of the requirements

for the degree

Doctor of Philosophy

in Electronic Engineering

Universidad Técnica Federico Santa María

Valparaíso, Chile

October 2024

"Reduced-order vocal fold models in in-vivo data for the assessment of underlying physiological parameters and mechanisms in vocal hyperfunction," a dissertation prepared by Jesús Alberto Parra Peña in partial fulfillment of the requirements for the degree, Doctor of Philosophy, has been approved and accepted by the following:

Matías Zañartu
Thesis Advisor

Alejandro J. Weinstein
Chair of the Examining Committee

Date

Committee in charge:

Dr. Matias Zañartu

Dr. Alejandro J. Weinstein, Chair

Dr. Sean D. Peterson

Dr. Jarrad H. Van Stan

DEDICATION

I would like to dedicate this thesis, first and foremost, to my family, for all their love and support throughout this process. Every step I have taken has been thanks to their guidance and care.

I also dedicate this work to my friends and colleagues in Chile, for being there and making me feel at home at every moment. You have allowed me to make this country my second home.

To everyone at USM, AC3E, and VPLab for being my academic home throughout this journey. You have not only fostered my intellectual growth but also contributed to my personal development.

ACKNOWLEDGMENTS

I am profoundly grateful to my advisor, Professor Matías Zañartu, whose guidance and vast knowledge in speech, acoustic modeling, and signal processing were pivotal throughout my doctoral studies. His patience, continuous guidance, and support not only made this thesis possible but also transformed my doctoral journey at UTFSM into a truly memorable experience. Thank you, Professor Zañartu, for providing me with the opportunity and confidence to develop this research and for instilling in me the rigor and passion that define your own work. Your enthusiasm and constant encouragement have left an indelible mark on my career and personal growth.

I extend my thanks to the members of my thesis committee: Professors Dr. Alejandro Weinstein, Dr. Sean Peterson, and Dr. Jarrad Van Stan, for their valuable insights and feedback, which have greatly enriched this work.

I am grateful to my fellow members of the Voice Production Lab at UTFSM—Dr. Gabriel Alzamendi, Dr. Emiro Ibarra, Dr. Juan P. Cortes, Josue Martinez, Nicolas Quinteros, Carlos Calvache, Rocio Ortega, and Christian Castro—for their friendship and support throughout this journey. Your companionship and the fruitful discussions we had influenced many aspects of my work, and your friendship has been essential during this entire process.

I gratefully acknowledge the financial support provided by the scholarships

from ANID Beca Doctorado Nacional 21202490, ANID Basal FB0008, the UTFSM PIIC programs N° 042/2021 and N° 077/2022, the Beca de Término de Tesis, the AC3E and the National Institutes of Health (NIH) National Institute on Deafness and Other Communication Disorders grant P50 DC015446.

On a personal note, I extend my gratitude to the AC3E family for the wonderful moments we have shared. It has been an incredible place to develop my work and to engage with colleagues and friends. I feel fortunate to have been part of such a great family.

ABSTRACT

REDUCED-ORDER VOCAL FOLD MODELS IN IN-VIVO DATA FOR THE ASSESSMENT OF UNDERLYING PHYSIOLOGICAL PARAMETERS AND MECHANISMS IN VOCAL HYPERFUNCTION

Jesús Alberto Parra Peña, BSc.

Doctor of Philosophy

Universidad Técnica Federico Santa María

Valparaíso, Chile, 2024

Dr. Matías Zañartu Salas, Chair

This thesis focuses on the mathematical modeling of vocal hyperfunction using physiologically-inspired vocal fold models, with the goal of improving both laboratory and ambulatory assessments of vocal pathologies. The motivation lies in the need for accurate models capable of capturing the biomechanical and physiological characteristics of normal and pathological vocal fold conditions. This research develops and validates vocal fold models that represent critical features of vocal hyperfunction, including asymmetrical oscillations and the presence of vocal fold lesions such as nodules, while also integrating vocal fold models with

ambulatory data to estimate physiological parameters like subglottal pressure and muscle activation.

The first part of this thesis involves the integration an existing vocal fold models with ambulatory data using a sampling and filtering approach. The study successfully replicated ambulatory distributions of physiological parameters and identified significant statistical differences between control subjects and patients with phonotraumatic vocal hyperfunction. These results demonstrate the potential of vocal fold models to improve the understanding of vocal hyperfunction by estimating subglottal pressure and laryngeal muscle activation in real-world settings. This foundational work serves as the basis for future efforts to refine vocal fold models for broader clinical applications.

The second stage of this research focuses on the development of an asymmetrical vocal fold model to represent non-phonotraumatic vocal hyperfunction . By introducing an asymmetry factor in muscle activation, the model captures the biomechanical irregularities associated with this vocal hyperfunction, such as imbalanced muscle control and its effects on vocal fold oscillation. Initial validation of the model against clinical data, including high-speed video recordings, shows promising results, making it a powerful tool for understanding the mechanics of asymmetric vocal fold pathologies. This work paves the way for extending the model to other vocal disorders, such as unilateral vocal fold paralysis. In the next stage, a model was developed to characterize the effects of nodules on vocal fold

mechanics, focusing on how these lesions alter phonation thresholds. The fixed-mass approach adopted in the nodule model effectively represents the in-silico and in-vivo characteristics of phonotraumatic vocal hyperfunction.

Using the pathological models, this thesis explores the acoustic compensation mechanisms in vocal hyperfunction, revealing key differences between phonotraumatic and non-phonotraumatic cases. While both conditions present elevated subglottal pressure, only phonotraumatic hyperfunction is characterized by increased adduction and a higher collision-to-subglottal pressure ratio, reinforcing the likelihood of trauma in such cases.

Looking ahead, future work will focus on refining these models with more complex muscle activation patterns and dynamic controls, incorporating additional physiological features, and validating the models with clinical data such as high-speed video or intramuscular electromyography.

This thesis makes three key contributions: first, it introduces a novel approach to integrating vocal fold models with ambulatory data, enabling the estimation of physiological parameters in real-world contexts. Second, it develops two models to capture the distinct features of vocal hyperfunction, including asymmetry in non-phonotraumatic vocal hyperfunction and the presence of nodules in phonotraumatic cases. Third, it presents an analysis of compensatory mechanisms, revealing that while both types of hyperfunction involve elevated subglottal pressure, only phonotraumatic hyperfunction is characterized by increased adduction

and higher aerodynamic forces, providing a clear distinction between these two vocal pathologies.

Contents

List of Tables

List of Figures

ABBREVIATIONS

AA	**A**mplitude **A**symmetry
ACC	**Acc**eleration sensor or signal
CAJ	**C**rico**A**rytenoid Junction
CPP	**C**epral **P**ick Prominence
CT	**C**rico**t**hyroid muscle
CTJ	**C**rico**T**hyroid Junction
DKG	**D**igital **K**ymo**g**ram
DPI	**D**aily **P**honotrauma **I**ndex
EMG	**E**lectro**m**yo**g**raphy
f_o	fundamental frequency
GAW	**G**lottal **A**rea **W**ave
GVV	**G**lottal **V**olume **V**elocity
$H_1 - H_2$	**H**armonic **1** to **H**armonic **2** ratio
HSV	**H**igh **S**peed **V**ideo
IA	**I**nterarytenoid muscle
IBIF	**I**mpedance **B**ased **I**nverse **F**iltering
KS	**K**olmogorov-**S**mirnov
LCA	**L**ateral **C**rico**a**rytenoid muscle
LIG	**Lig**ament

MAP	**M**uscle **A**ctivation **P**lot
MFDR	**M**aximum **F**low **D**eclination **R**ate
MRI	**M**agnetic **R**esonance **I**maging
MUC	**Muc**osa
NPVH	**N**on-**P**honotraumatic **V**ocal **H**yperfunction
NSAM	**N**eck-**S**urface **A**cceleration **M**agnitude
OQ	**O**pen **Q**uotient
PA	**P**hase **A**symmetry
PCA	**P**osterior **C**rico**a**rytenoid muscle
PGO	**P**osterior **G**lottal **O**pening
PTP	**P**honation **T**reshold **P**ressure
PVH	**P**honotraumatic **V**ocal **H**yperfunction
RFF	**R**elative **F**undamental **F**requency
SD	**S**tandard **D**eviation
SPL	**S**ound **P**ressure **L**evel
TA	**T**hyro**a**rytenoid muscle
TBCM	**T**riangular **B**ody **C**over **M**odel
VE	**V**isco-**E**lastic
VF	**V**ocal **F**old
VH	**V**ocal **H**yperfunction

VP **V**ocal **P**rocess

WRA **W**ave **R**eflaction **A**nalog

Chapter 1
Introduction

1.1 Motivation

Laryngeal voice disorders have a significant impact on daily life, affecting communication, social interactions, work performance, and psychological well-being [1]. These disorders are prevalent globally, affecting approximately 7.7% of the adult population annually [2], and about 30% of adults in the United States have experienced voice-related issues at some point in their lives [3]. In particular, professionals who rely heavily on their voice, such as teachers, singers, and actors, are especially vulnerable to these conditions. For example, in Chile, around 75% of school teachers report experiencing voice problems [4], making it a critical occupational health issue.

Among the most common voice disorders is vocal hyperfunction (VH) [5, 6]. VH is generally classified into two types: phonotraumatic (PVH) and non-phonotraumatic (NPVH). PVH is associated with trauma to the vocal folds (VFs), often resulting in nodules or polyps [7], while NPVH involves excessive tension

in the laryngeal muscles without visible trauma to the VFs [8, 9]. VH involves adaptive mechanisms where trauma or imbalances in laryngeal tension lead to compensatory overexertion, reinforcing the disorder [6]. While increased subglottal pressure (P_s) in patients with VH can be observed in short-term recordings [10, 11], long-term trends, especially in ambulatory settings, are harder to assess due to the challenges of continuous monitoring [12].

Ambulatory monitoring is essential for tracking how a patient's vocal condition changes during daily activities and therapy. Recent tools, such as vocal dose measurements [13, 14] and regressor models [15, 16, 17, 18, 19], have proven effective in differentiating VH patients from controls by measuring vocal load and its fluctuations over time. These tools primarily rely on measures like sound pressure level (SPL), harmonic differences ($H_1 - H_2$), fundamental frequency (f_o), and cepstral peak prominence (CPP). However, these parameters are not the underlying physiological variables that define vocal function, such as laryngeal muscle activity and subglottal pressure.

Measuring these intrinsic physiological variables presents significant challenges. P_s can be assessed using invasive methods such as tracheal punctures [20, 21, 22] or transoral pressure transducers [23, 24], but these techniques are often impractical in clinical settings. Indirect methods, including intraoral pressure sensors [25, 26, 27] and regression models based on aerodynamic data [28, 29, 30, 31, 32, 33], offer less invasive alternatives and have shown promise even in ambulatory applications

[34].

Similarly, intrinsic laryngeal muscle activity is critical for understanding VH but is difficult to measure directly [35]. In-vitro studies, such as those by Chhetri et al. [36, 37], have provided valuable insights into how intrinsic laryngeal muscles exhibit hyperfunction in conditions like laryngeal dystonia, while hypofunction is observed in conditions such as paresis or paralysis [38]. Despite this, direct muscle activity measurements often require invasive electromyography [39], limiting their widespread use.

In response to these limitations, mathematical models of phonation provide a powerful tool for estimating these hard-to-measure variables. By using non-invasive data sources, such as acoustic signals or accelerometer recordings, these models can estimate physiological variables like P_s and laryngeal muscle activity [33, 40, 41, 42]. This approach offers a non-invasive way to gain insights into the mechanisms of VH, advancing diagnosis and management in real-world or ambulatory settings.

To model VF dynamics, a lumped-element approach is selected due to its balance between simplicity and biological relevance [43, 44]. While more complex models, such as those based on finite elements, offer detailed representations of VF tissue behavior, they involve a large number of parameters that need to be defined or optimized, making them cumbersome for practical applications [45, 46]. Lumped-element models, on the other hand, provide a more straightforward al-

ternative by requiring fewer parameters and offering a direct way to interpret changes or additions to the models [47]. This simplicity allows for easier exploration of how specific modifications affect VF dynamics, making them well-suited for studying VF disorders and refining model assumptions.

Furthermore, to effectively characterize VH, it is necessary to extend these models to account for the specific features of vocal hyperfunction. This includes incorporating elements that reflect the distinguishing characteristics of PVH and NPVH, such as VF trauma [48, 49] or asymmetry between the left and right VFs [50, 51]. Existing approaches in the literature provide a foundation [52, 53, 54, 55, 56, 57, 49], but additional work is required to refine these models for inclusion in an integrated framework that links clinical, laboratory, and ambulatory data.

Models that better represent the role of muscle activation or the impact of pathophysiological elements in VH will allow us to begin studying the role of underlying variables in the compensation mechanisms for VH and to understand the differences between PVH and NPVH. By incorporating these refinements, we can more accurately capture how VF behavior changes in response to the distinct pathologies, providing insight into the compensatory adjustments made by the vocal system in each case.

In summary, this thesis will focus on two key objectives: integrating VF models with ambulatory and clinical data to provide a more comprehensive understanding of phonation, and extending these models to improve the representation of VH in

a compensation environment.

1.2 Aims and hypotheses

1.2.1 General aim

To develop and validate frameworks based on low order models of voice production for the representation and analysis of vocal hyperfunction in ambulatory and laboratory settings.

1.2.2 Specific aims

1. To integrate an existing low-order vocal fold model and ambulatory accelerometer distributions for the estimation of underlying parameters in ambulatory conditions of PVH patients.

2. To develop low-order vocal fold models for describing phonotraumatic and nonphonotraumatic vocal hyperfunction to describe their various compensatory mechanisms using both in-vivo and in-silico data.

1.2.3 Hypotheses

- **H1:** *Specific acoustic and aerodynamic signal features are indicative of distinct physiological patterns between pathological subjects and controls, reflect-*

ing underlying statistical differences in intrinsic laryngeal muscle activation and subglottal pressure during long-term ambulatory monitoring.

Previous research has shown statistical differences in ambulatory acoustic and aerodynamic feature distributions between normal and hyperfunctional populations [15, 19, 16, 17, 18, 58]. Based on these findings, we hypothesize that specific acoustic and aerodynamic features, captured through long-term monitoring, will be indicative of underlying physiological variations, particularly in laryngeal muscle activation and subglottal pressure. Using mathematical VF models, we aim to replicate SPL and $H_1 - H_2$ signals and observe significant differences in model inputs, such as laryngeal tension and subglottal pressure. This hypothesis will be tested using paired t-tests to compare patient-control groups.

H2: *Intensity and pitch compensation mechanisms differ between phonotraumatic and non-phonotraumatic vocal hyperfunction, reflected in distinct collision pressure profiles and muscle activation patterns.*

Previous work has shown that both PVH and NPVH exhibit increased subglottal pressure compared to normal phonation [10, 11]. We hypothesize that this occurs through distinct compensatory muscle activation strategies. In PVH, incomplete glottal closure due to lesions leads to elevated subglottal pressure and increased posterior muscle activity, resulting in higher

collision pressure. In contrast, NPVH also presents increased subglottal pressure, but asymmetrical muscle activity produces lower collision pressure compared to PVH, potentially explaining why nodules form in PVH but not in NPVH. These compensatory differences are expected to become evident in the models under sustained vowel conditions, helping to differentiate the pathophysiological mechanisms between the two types of hyperfunction.

1.3 Methodology overview

The first aim of this thesis is to integrate an existing low-order VF model with ambulatory accelerometer data to estimate underlying parameters in patients with PVH. This process begins by transforming traditional in-laboratory simulations of the Triangular Body-Cover Model (TBCM), which primarily focus on sustained vowel phonations with fixed vocal tract configurations, into simulations that more accurately reflect real-world ambulatory conditions. To achieve this, phoneme usage frequencies observed in conversational speech will be incorporated into the simulations. By using phonetically balanced texts to calculate the probability of vowel usage, we ensure that the simulations represent realistic speech patterns. This approach allows for a more accurate representation of phonation in an ambulatory setting, aligning the model more closely with the conditions patients experience in daily life.

Next, to effectively link model parameters to real-world ambulatory features, a methodology will be developed to filter the model simulations using experimental ambulatory data. This involves weighting the model outputs based on phoneme usage criteria to generate input-output vectors that correspond to the distributions of ambulatory features observed in real-world scenarios. By employing a Monte Carlo sampling approach, a dense and homogeneous representation of the input variables of the VF model will be achieved, providing a comprehensive mapping of the model's input-output space. This will allow for meaningful comparisons between the predictions generated by the model and actual ambulatory measurements, thereby improving the robustness and accuracy of the model.

Finally, to describe changes in pathology conditions during ambulatory monitoring, the resulting model-feature distributions will undergo rigorous statistical analysis, similar to methodologies used in VH classification studies. For example, the Daily Phonotrauma Index (DPI) will be used as a benchmark. Central tendency and higher-order statistical measures will be applied to the model distributions, followed by statistical tests such as Wilcoxon tests and t-tests, with effect size being quantified using Cohen's distance. The ultimate goal of this analysis is to identify key model parameters, including muscular activation and subglottal pressure, that can statistically differentiate between control subjects and PVH patients. Additionally, the study will explore how these parameters behave as patients transition between control and pathological states, with clusters being

defined based on DPI values.

Moving to the second aim, which focuses on developing low-order VF models for describing PVH and NPVH, this study aims to characterize the compensatory mechanisms using both in-vivo and in-silico data. The first step in this process involves extending the existing TBCM to represent VF asymmetry, a key characteristic of NPVH. By allowing independent control over the activation of each VF's musculature, the model simulates imbalances in VF posture and mechanical properties. This approach enhances the ability to study how asymmetric muscle activation affects phonation, offering a more precise understanding of the biomechanical irregularities found in NPVH.

For PVH, the TBCM is further modified to include the presence of vocal nodules, which are modeled as fixed masses with specific viscoelastic properties. These modifications alter the collision dynamics of one VF cover layer and replicate the impact that nodules have on vocal function, based on existing data regarding the properties of nodular tissue. This approach helps to analyze how nodule parameters affects phonation.

The outputs from these models will be validated by comparing them with established modeling approaches and experimental or clinical data. For NPVH, the model's asymmetric configurations will be compared with previous studies where mechanical properties were directly altered. By analyzing simulated kymogram signals and comparing them with clinical data, we will fine-tune the model to

accurately reflect the effects of muscle imbalance. For PVH, the validation will focus on replicating known trends from previous studies regarding nodule size and stiffness, while also simulating phonation threshold pressures to better understand the role of vocal nodule parameters.

Additionally, these models will be used to explore how compensation occurs in terms of acoustic parameters like intensity (SPL) and pitch (f_o). Through the application of the inverse Jacobian method, we will simulate how the pathological models restore the initial acoustic target. This optimization process will allow us to characterize the compensatory mechanisms for both PVH and NPVH, assessing key aerodynamic parameters such as collision pressure and Maximum Flow Declination Rate (MFDR). This analysis will provide deeper insights into how these compensations impact vocal efficiency and aerodynamic behavior in different pathological contexts.

1.4 Significance and contributions

Understanding the compensatory mechanisms that contribute to the onset and progression of vocal VH, remains an important area of study [6]. This thesis seeks to address this challenge by developing models and methodologies that capture the physiological dynamics of VH in both clinical and ambulatory contexts.

Firstly, this work integrates an existing VF model with ambulatory data to

simulate and analyze vocal behavior over voice monitoring. By applying the base TBCM to ambulatory accelerometer distributions, this thesis seeks to estimate how physiological compensations manifest in daily life, providing insights into long-term vocal function in patients with PVH. This approach extends the application of VF models beyond controlled laboratory settings, allowing for a more comprehensive understanding of VH in dynamic, real-world environments.

In the next phase, the TBCM was extended to include pathological modifications, such as VF nodules and muscle activation imbalances, to better reflect the clinical characteristics of PVH and NPVH. These modifications, validated and calibrated using both in-vivo and in-silico data, offer an accurate representation of the behaviors typically seen in patients with VH. By introducing these pathological elements into the model, this thesis enhances our ability to simulate how changes in pathology severity, such as nodule size and stiffness, muscle imbalance, or DPI value, impact key clinical metrics, including Phonation Threshold Pressure (PTP), larygeal muscle activation, f_o, SPL and others.

Additionally, these models, built to represent the physiological aspects of the pathology, allow for a detailed examination of the underlying parameters or input conditions for voice production. This enables the study of how maintaining consistent acoustic outputs, despite the presence of pathology, reflects in critical values such as collision pressure, offering new insights into the compensatory strategies employed by the vocal system and its effect in VF condition.

In conclusion, this thesis advances VF modeling for ambulatory applications and extends these models to better represent pathological conditions, contributing to a deeper understanding of the compensatory mechanisms in VH. The most significant contribution lies in presenting pathological VF models with a muscle control system that governs both mechanical properties and posture, offering a comprehensive tool for studying compensatory behaviors in VH.

1.5 Publications

This thesis is supported by a set of publications. Related publications made during the research period are also included.

1.5.1 Journals papers

1. **Parra J.A.**, Ibarra E.J., Calvache C., Van Stan J.H., Hillman R.E., and Zañartu M. (2025) *"Estimating the pathophysiology of phonotraumatic vocal hyperfunction using ambulatory data and a computational model"*. Accepted for publication at Journal of Speech, Language, and Hearing Research.

2. **Parra, J. A.**, Calvache, C., Alzamendi, G. A., Ibarra, E. J., Soláque, L., Peterson, S. D., and Zañartu, M. (2024). *"Asymmetric triangular body-cover model of the vocal folds with bilateral intrinsic muscle activation"*. The Journal of the Acoustical Society of America, 156(2), 939–953. DOI:

10.1121/10.0028164.

3. Ibarra, E. J., **Parra, J. A.**, Alzamendi, G. A., Cortés, J. P., Espinoza, V. M., Mehta, D. D., Hillman, R. E., and Zañartu, M. (2021). *"Estimation of Subglottal Pressure, Vocal Fold Collision Pressure, and Intrinsic Laryngeal Muscle Activation From Neck-Surface Vibration Using a Neural Network Framework and a Voice Production Model"*. Frontiers in physiology, 12, 732244. DOI: 10.3389/fphys.2021.732244.

1.5.2 Journals papers in review

1. Sepúlveda J., **Parra J.A.**, Ibarra E.J., Araya M., De La cuadra P., and Zañartu M. *"Enhanced prediction of model-based features in neck surface acceleration signals using a probabilistic bayesian neural network with uncertainty estimation"*. Submitted for publication at IEEE Transactions on Audio, Speech and Language Processing.

1.5.3 Conferences

1. Sorolla C., **Parra J.A.**, Ibarra E.J., Alzamendi G.A., and Zañartu M. *"Forward mapping estimation of speech-motor control using machine learning techniques"*. 13th International Conference on Voice Physiology and Biomechanics, July, 22-26 2024. Erlangen, Germany.

2. **Parra J.A.**, Ibarra E.J., Calvache C., Van Stan J.H., Hillman R.E., and Zañartu M. *"Inferring physiological parameters in phonotraumatic vocal hyperfunction: Insights"*. 15th Congress of the PanEuropean Voice Conference, PEVOC 2024, September, 4-7, 2024. Santander, Spain.

3. **Parra J. A.**, Calvache C., and Zañartu M. *"An asymmetric triangular body-cover model with unbalanced intrinsic muscle activation"*. 15th International Conference on Advances in Quantitative Laryngology, Voice and Speech Research, AQL 2023, March 30 - April 1, 2023. Phoenix, US.

4. Calvache C., **Parra J.A.**, Soláque L., and Zañartu M. *"Describing non-phonotraumatic vocal hyperfunction with an asymmetric triangular body-cover model with unbalanced intrinsic muscle activation"*. 15th International Conference on Advances in Quantitative Laryngology, Voice and Speech Research, AQL 2023, March 30 - April 1, 2023. Phoenix, US.

5. **Parra J.A.**, Alzamendi G.A., and Zañartu M. *"Nodule approximation in a triangular shaped vocal fold model"*. 52nd Annual Symposium: Care of the Professional Voice, May 31 – June 4, 2023. Philadelphia, US.

6. **Parra J.A.**, Calvache C., and Zañartu M. *"Assessing phoneme distribution for speech modeling"*. Proceedings of 18th International Symposium on Medical Information Processing and Analysis, SPIAIM 2022, November, 9-11, 2022. Valparaiso, Chile. DOI: 10.1117/12.2670042.

7. **Parra J.A.**, Ibarra E.J., Calvache C., and Zañartu M. *"Estudio de la activación muscular desbalanceada mediante el uso de modelos de pliegues vocales"*. 78° Congreso Chileno de Otorrinolaringología, ORL 2022, November, 9-11, 2022. Viña del Mar, Chile.

8. **Parra J.A.**, *"Modelos oscilantes de los pliegues vocales: Una perspectiva física a un problema clínico"*. I Congreso de la Sociedad Chilena de Fonoaudiología and IV Congreso de Fonoaudiólogos Investigadores, October, 21-22, 2022. Valparaiso, Chile.

9. Ibarra E.J., **Parra J.A.**, Alzamendi G.A., Cortés J.P., Espinoza V.M., and Zañartu M. *"Método basado en inteligencia artificial para la estimación de la función vocal usando un sensor de aceleración"*. 77° Congreso Chileno de Otorrinolaringología, ORL 2021, November, 9-12, 2021. Online.

10. **Parra J.A.**, Ibarra E.J., Alzamendi G.A., Cortés J.P., and Zañartu M. *"Descubriendo parámetros físicos subyacentes a partir de las distribuciones diarias de características acústicas mediante simulaciones monte carlo de un modelo de producción de voz de baja dimensión"*. 77° Congreso Chileno de Otorrinolaringología, ORL 2021, November, 9-12, 2021. Online.

11. Ibarra E.J., **Parra J.A.**, Alzamendi G.A., Cortés J.P., Espinoza V.M., and Zañartu M. *"A Machine Learning Framework for Estimating Subglottal*

Pressure during Running Speech from Glottal Airflow measures". 14th International Conference on Advances in Quantitative Laryngology, Voice and Speech Research, AQL 2021, June, 7-10, 2021. Online.

12. **Parra J.A.**, Ibarra E.J., Alzamendi G.A., Cortés J.P., and Zañartu M. *"Discovering Underlying Physical Parameters From Daily Phonotrauma Index Distributions using Monte Carlo Simulations of a Low-Dimensional Voice Production Model"*. 14th International Conference on Advances in Quantitative Laryngology, Voice and Speech Research, AQL 2021, June, 7-10, 2021. Online.

1.6 Document structure

The chapters in this thesis are organized as follows: The chapter 2 provides a background including the state of the art, a general description of voice production models, the source-filter theory of phonation, the VH statement, and the foundational concepts of ambulatory sensing and monitoring. The chapter 3 explores the connections between ambulatory monitoring and VF models for the PVH population. It discusses the underlying parameters that distinguish controls from patients and their clinical implications. The chapter 4 and chapter 5 offers insights into the pathological models developed in this research, introducing technical and anatomical modifications to enhance model representativity of

VH conditions. These chapters also discuss the results, effects of pathological parameters, and comparisons with other approaches and experimental data. The chapter 6 summarizes the contributions of the thesis and outlines future research directions.

Chapter 2
Background

This chapter outlines the theoretical and methodological foundations for modeling vocal hyperfunction, focusing on phonation models and their application to vocal function estimation. The chapter reviews key models, including those of the vocal tract, glottal airflow, vocal folds, and laryngeal muscle activation. It also discusses the biomechanical modeling of VH and its monitoring in ambulatory settings. Lastly, the two central components of this work are introduced: the Triangular Body-Cover Model for vocal fold dynamics and the Daily Phonotrauma Index for tracking PVH in real-world conditions.

2.1 Phonation models

Phonation is the physiological process through which voice is produced. This complex process involves the generation of sound through the interaction of subglottal airflow with the VFs, creating a sound wave that propagates through the vocal tract. These interactions are detailed in what is known as the myoelastic theory of phonation [59], where the composition of the VFs plays a key role in the

generation of this sound wave.

From an engineering and systems perspective, this process can be expressed in terms of a source-filter system [60], where there is a primary source - the lungs, which produce a airflow. This airflow, when interacting with the folds, generates the glottal airflow wave that propagates and is modulated/filtered by the vocal tract. It is important to clarify that this is a simplified way of describing phonation. Other considerations can include third-type interactions between flow, tract, and folds [61, 62].

Within this system, we identify three key elements in the interpretation of the vocal process: the vocal tract, glottal airflow, and VFs, as illustrated in Figure 2.1. We will now delve into a more detailed description of these elements and the ways in which they have been modeled.

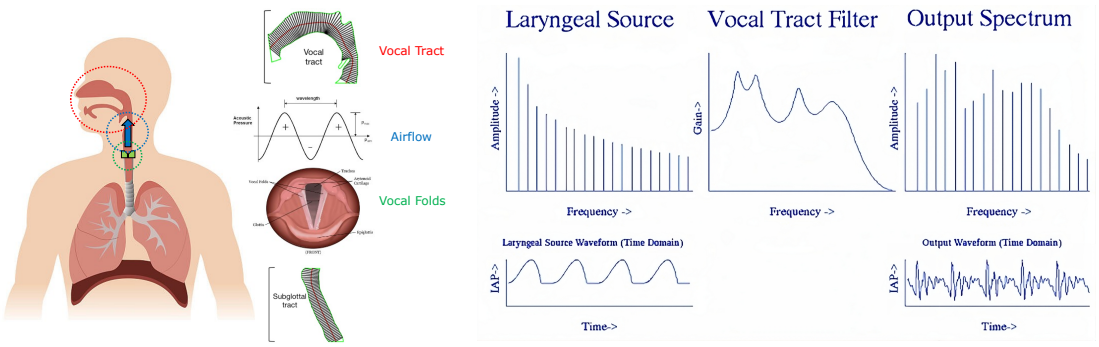

Figure 2.1: Key elements in voice production: (A) vocal folds, glottal airflow and vocal tracts. (B) Scheme of source-filter theory.

2.1.1 Vocal tract models

The vocal tract is the part of the phonatory apparatus that extends from the larynx to the lips and nose. It is a flexible cavity that connects the VFs with the outside environment, through which sound propagates. As with any cavity in acoustics, it functions as a filter that modifies the characteristics of the sound produced by the vibration of the VFs. These modifications are achieved through variations in the shape and size of the vocal tract, facilitated by the movement of the articulators - regions comprising the tongue, lips, soft palate, hard palate, uvula and pharynx [60].

In terms of a system or filter, the configuration of the vocal tract determines the resonances, more commonly known as formants, which are essential for the quality and intelligibility of the produced sound [60, 63].

Modeling the vocal tract has been a significant area of interest in the fields of phonetics, linguistics, bioengineering, and acoustics for many decades. The earliest models of the vocal tract were physical and mechanical. Gunnar Fant and other researchers developed acoustic models based on the source-filter theory [60]. These models treated the vocal tract as a series of coupled tubes with varying lengths and diameters, simplifying the complexity of the actual tract [60]. Since then, the study of the tract's dimensions using Magnetic Resonance Imaging (MRI) imaging has begun, allowing for the characterization of lengths and areas

[64].

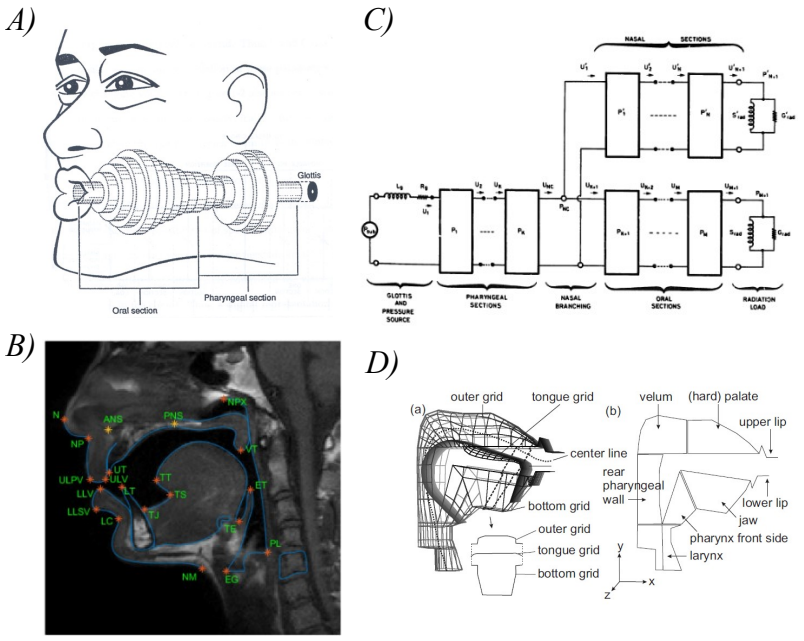

Figure 2.2: Vocal tract modelingexamples: (A) Vocal tract as cylindrical areas [65]. (B) MRI image for area computation [66]. (C) vocal tract model using electro-acoustic analog [67]. (D) vocal tract model from sectional movement [68].

Digital models began to replace the simpler physical or propagation models. Finite element models enabled more precise and detailed simulations of the vocal tract and its interaction with sound [69].

Modern imaging technology, such as MRI and computed tomography, has facilitated the creation of three-dimensional models of the vocal tract [64, 70, 68]. These models can simulate not only the static anatomy but also the dynamic changes during articulation. Over the years, vocal tract modeling has evolved from simplistic representations to highly sophisticated models that can accurately

capture the complexity and dynamics of human speech.

For the purposes of this thesis, we have chosen to use a vocal tract model as a waveguide, through which sound waves propagate forwards and backwards. This idea draws inspiration from the works of Maeda [67], who treated the vocal tract as a transmission line using the electroacoustic analog [71, 67]. The version used here does not model the loss elements of the tract as reactances and impedances but as exponential losses, and is detailed in the thesis of Story[43], where he introduces the methodology we use, called the Wave Reflection Analog (WRA).

In the WRA approach, the vocal tract is modeled as a series of concatenated cylindrical sections, each representing a specific part of the tract. This method effectively captures the acoustical properties and resonances of the vocal tract, allowing for a detailed analysis of how sound is filtered and modulated as it travels through this complex system. This modeling technique is particularly beneficial in understanding the nuances of vocal tract dynamics in different phonatory conditions, providing valuable insights for both normal and pathological speech analysis.

Let's delve into some detail here. The vocal tract is considered as a succession of cylindrical sections of varying cross-sectional areas. The pressure or airflow waves traveling between two sections, at the interface, they must satisfy the continuity of both airflow and pressure, in the n-th junction in Figure 2.3:

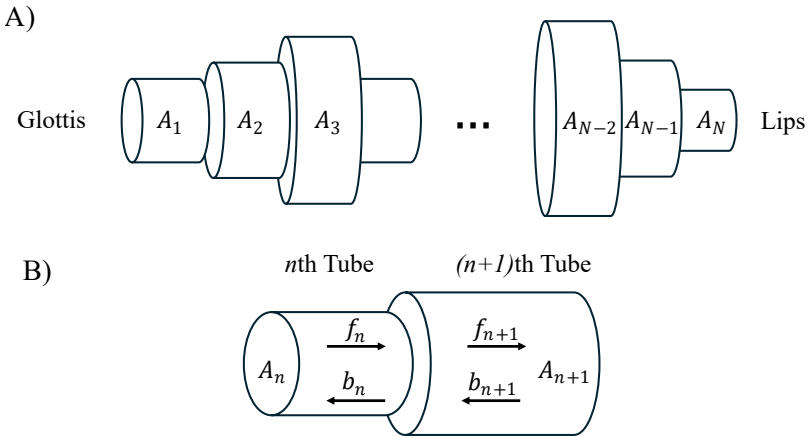

Figure 2.3: Scheme of vocal tract area model: (A) Tract areas. (B) Scheme of forward and backward flow. Adapted from Sahoo [72]

$$f_n + b_n = f_{n+1} + b_{n+1}, \tag{2.1a}$$

$$\frac{1}{Z_n}(f_n - b_n) = \frac{1}{Z_{n+1}}(f_{n+1} - b_{n+1}). \tag{2.1b}$$

Here, it was possible to connect airflow and pressure through the acoustic impedance of the medium which is defined as: $Z_n = \frac{\rho c}{A_n}$, with ρ the air density, c sound air velocity and A_n the cross-sectional area of the n-th cylindrical section . The emergence of the cross-sectional area of the tract section is noteworthy, indicating the influence of its geometry. If we resolve the set of equations in terms of the inputs and outputs at the interface:

$$b_n = r_n f_n + (1 - r_n) b_{n+1}, \tag{2.2a}$$

$$f_{n+1} = (1 + r_n) f_n - r_n b_{n+1}, \tag{2.2b}$$

with

$$r_n = \frac{A_n - A_{n+1}}{A_n + A_{n+1}}. \tag{2.3}$$

Where r_n is the reflection coefficient of the n-th interface. This shows that it is the geometry that defines how much of the wave is transmitted or reflected. Additionally, losses in this propagation are introduced through an exponential factor, derived from the solution of the wave equation [73, 74]. This factor accounts for losses due to the medium's viscosity or energy absorption by the walls, among others [74, 75]. For example, the attenuation of the forward wave to travel in th n-th section is:

$$f_n(z + l) = e^{-\beta l} f_n(z) \tag{2.4}$$

Where β is the attenuation factor and l the section length. Another way to model these losses is by maintaining them as circuit elements to be considered in each section [60, 74]. Additionally, the output of the voice wave to the exterior is modeled with piston-type impedance elements [60, 43]. This is the summarize of how we will consider the propagation of both the subglottal and supraglottal tract.

2.1.2 Glottal airflow

The glottal flow wave that propagates through the tract is due to the interaction of subglottal and supraglottal with the VFs [76]. Figure 2.4 shows the steps in the glottic cycle and how the self-sustained oscillation of the VFs is generated, involving subglottal pressure and the tension of them.

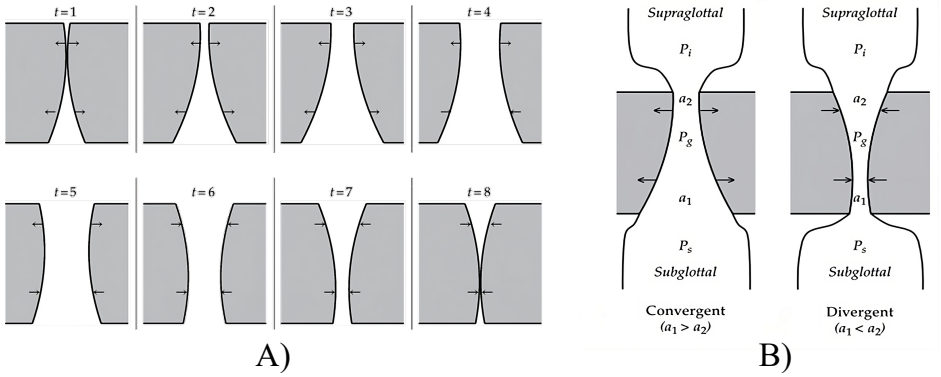

Figure 2.4: Glottal cycle diagram: (A) Glottal cycle in time instance. (B) Converge and divergence VF position. Adapted from Story [76]

To quickly describe the glottic cycle: In the convergent stage (from t=1 to t=4), the subglottal pressure breaks the barrier of the VFs and displaces them due to the force exerted by this pressure. Additionally, the opening that is created begins to allow air flow, which, according to Bernoulli's law, produces a decrease in pressure between the folds. As the folds move and the pressure decreases, they tend to return to their initial position due to the elastic component of the phenomenon (t4). In the following phase (from t=5 to t=8), the force from

the pressure does not separate the folds but seeks to bring them together, along with the elastic action of the folds. Once closed and due to the inertia of their movement, the cycle restarts.

$$P = P_i + (P_S - P_i)\left(1 - \frac{a_1^2}{a_2^2}\right). \tag{2.5}$$

Equation 2.5 shows the ideal change in pressure in the middle of the folds due to the differences in upper and lower areas and pressures. With P_i is the supraglottal and P_S subglottal pressure.

$$u_g = \frac{a_g c}{k_t}\left\{-\frac{a_g c}{A^*} + \left[\left(\frac{a_g}{A^*}\right)^2 + \frac{4k_t}{c^2\rho}(P_S - P_i)\right]^{1/2}\right\}. \tag{2.6}$$

Equation 2.6 shows how glottal airflow changes in terms of glottal area $a_g = \min(a_1, a_2)$ and the pressure differences [65]. Where A^* corresponds to the effective area between the upper and lower parts of the VFs, and k_t is an empirical transglottal pressure coefficient [77].

This is a simplification of how aerodynamic energy is transformed into sound. Over the years, adjustments have been made to this type of model, introducing factors for loss in the conversion of flow to pressure, regimes where the flow is laminar or turbulent, among other types of interactions [78, 79, 80]. For the models presented in this thesis, the solution for glottal flow taken from the work of Lucero [81] is used.

2.1.3 Vocal folds models and their parameters

The VFs can be conceptualized as an oscillatory membrane that interacts with subglottal pressure, generating the glottal wave essential for voice production. The structure of the VFs has been meticulously studied from histological and physiological perspectives [82, 83], leading to the establishment of the body-cover theory as proposed by [59]. This theory delineates the VF composition into models of five, three, and two layers, as summarized in Table 2.1.

Table 2.1: Comparison of the common VFs division [65].

Layers	**Five**	**Three**	**Two**
Ephithelium	Ephithelium	Mucosa	Cover
Lamina Propia	Superficial Layer		
	Intermediate Layer	Ligament	
Muscle	Deep Layer		Body
	Muscle	Muscle	

Each layer possesses distinct mechanical properties, contributing to the VFs density, stiffness, and other viscoelastic characteristics. Furthermore, the VFs exhibit a unique geometric arrangement, including the vocal process, which grants them their characteristic zipper or V shape [84]. This shape is primarily deter-

mined by the positioning of the arytenoids among other structures.

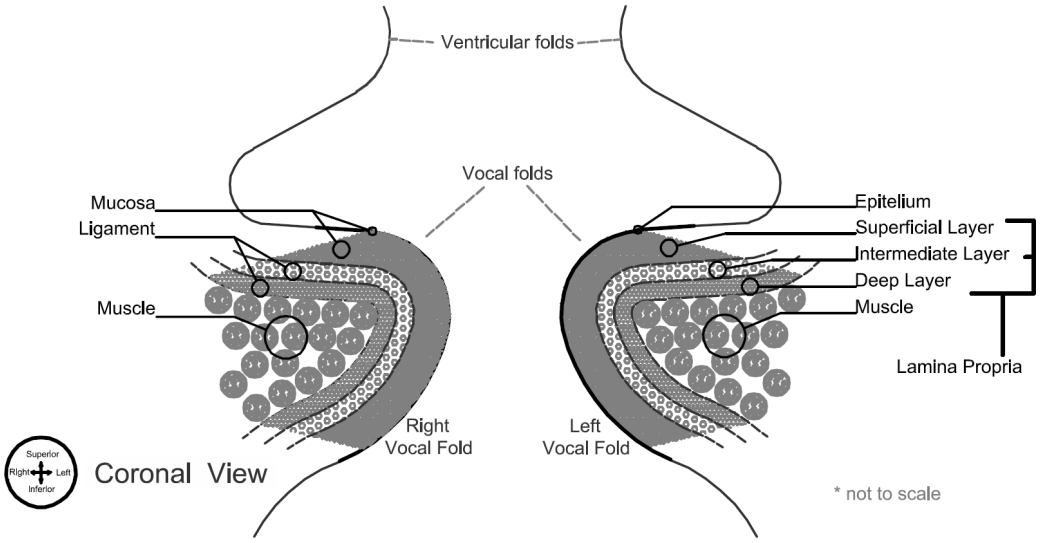

Figure 2.5: VFs layered structure. Adapted from Titze [65]

In modeling the VFs, various types of models have been developed, which can be categorized based on how they describe the VFs. Here, we classify them into two major groups: continuous models and lumped element models, as discussed in revious works [45, 47, 46, 85].

This classification provides a foundation for understanding the different approaches used to represent the complex mechanics of VF vibration and the resulting voice production.

Continuous models of VFs are marked by detailed geometric and mechanical descriptions. Typically, the geometry of the VF is defined by a mesh that outlines its shape. For resolving the dynamics of the VFs, these models often employ techniques with a high number of degrees of freedom, such as finite differences or finite

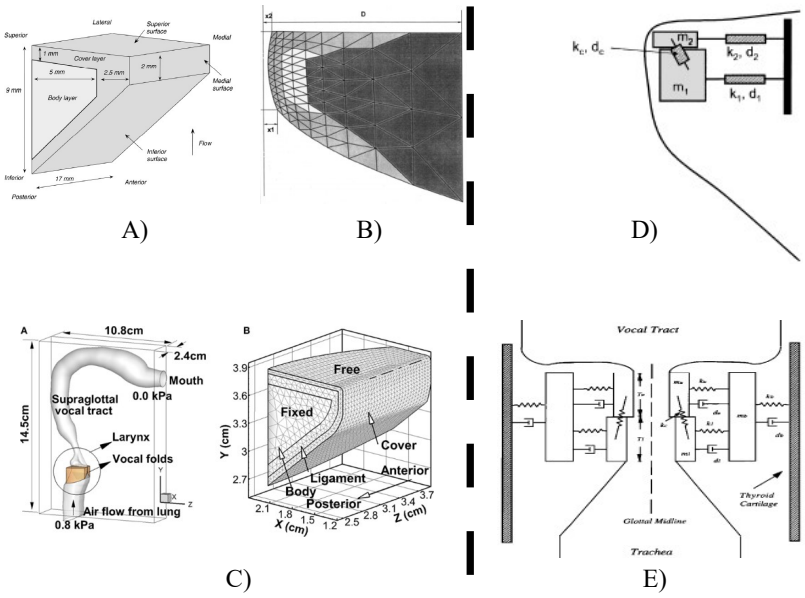

Figure 2.6: Examples of VF models: (left) continuos models: (A) Zhang et al. [86], (B) Alipour et al. [87] and (C) Jiang et al. [88]. (right) lumped element models: (D) Story [89] and (E) Story and Titze [90].

elements. In these methods, each node of the mesh is governed by an equation of motion accounting for forces or energy, allowing for a detailed description of the mechanical interaction between the fold and the airflow. However, these models generally have a high computational cost due to their complexity and the number of degrees of freedom involved.

On the other hand, lumped element models approximate VFs as a system of coupled oscillators, condensing the mechanical information of the folds into mass, elasticity, and viscosity components of the oscillator. These models offer a more direct interpretability of parameters and facilitate large-scale exploration of their

operational domains, which can be cumbersome for continuous models.

Both groups of models have their advantages, whether in terms of degrees of freedom and execution speed or in providing a detailed description of the phenomenon and geometry. Lumped element models are an excellent choice when a versatile, quick model is required for running across a wide range of parameters. In contrast, continuous models are ideal for detailed descriptions of the composition and interaction of various elements in the representation of VFs.

Every model of VFs requires parameter characterization, usually involving geometric descriptions like length and thickness, as well as the mechanical properties of the different layers comprising the folds. These values are often captured from measurements on excised larynges, such as the Young's modulus, Poisson's ratio, and compressibility modulus for mechanical components and parameters like fold length, thickness, and arytenoid distance for geometry [91, 92, 93, 94, 59].

In the spirit of this thesis, as defined in the motivation section 1.1, a better understanding of the elements present in lumped-element models is necessary. To illustrate this, the body-cover model [90] is used as a basis (see Figure 2.7 left). Here, the VF is defined as three masses, in a representation of two layers: body and cover as mention in Table 2.1.

The general form of the equation of motion for one of these masses is:

$$m\ddot{x} = F_{\text{elas}} + F_{\text{visc}} + F_{\text{aero}} + F_{\text{col}}, \tag{2.7}$$

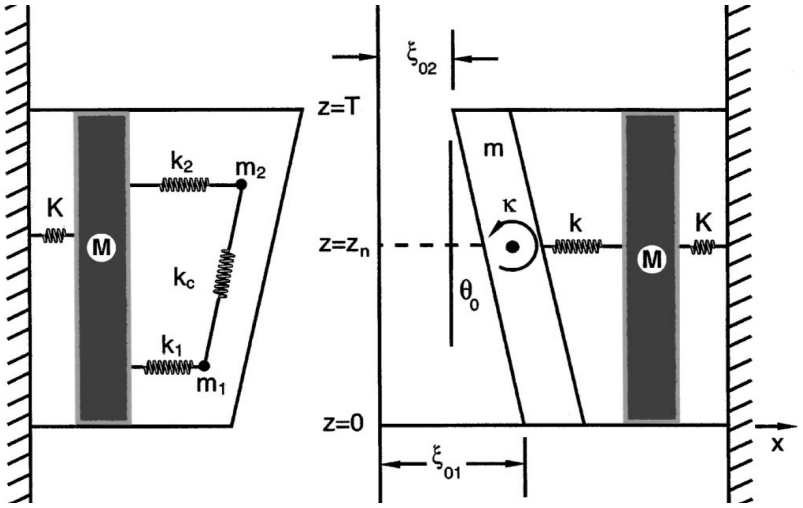

Figure 2.7: Scheme of lumped-element body-cover model: (left) three-mass and (right) bar-plate in frontal cross section. Adapted from [95]

where the forces include elastic, viscous, aerodynamic, and collision forces. In this equation of motion, the elastic and viscous forces are modeled as a critically damped harmonic oscillator of the mass-spring type. The collision and aerodynamic forces only apply to the masses of the cover.

This section will describe how to calculate the mass and elasticity (spring) components of each mass, based on the geometry and tension of the layers. Let m_1 and m_2 describe the mass of cover, and M the mass of the body, their values are defined by:

$$m_1 = \rho L T D_c \frac{Z_n}{T}, \tag{2.8a}$$

$$m_2 = \rho L T D_c \left(1 - \frac{Z_n}{T}\right), \tag{2.8b}$$

$$M = \rho L T D_b. \tag{2.8c}$$

where L, T, D_c and D_b represent the length, thickness, and depth of the cover and body layers, and Z_n is the point that divides the upper and lower component in the cover. For the elastic component, i.e., the springs, the stress (σ) and shear modulus (μ) properties of the two layers representing the VFs are used. Using the representation presented in [95], the springs values are defined:

$$k_1 = 2\mu_c \left(\frac{LT}{D_c}\right) \frac{Z_n}{T} + \pi^2 \sigma_c \frac{D}{L} z_n, \tag{2.9a}$$

$$k_2 = 2\mu_c \left(\frac{LT}{D_c}\right) \left(1 - \frac{Z_n}{T}\right) + \pi^2 \sigma_c \frac{D}{L} T \left(1 - \frac{Z_n}{T}\right), \tag{2.9b}$$

$$k_c = \left\{ \frac{1}{2} \mu_c \frac{L D_c}{T} \left[\frac{1}{3} - \frac{Z_n}{T}\left(1 - \frac{Z_n}{T}\right)\right]^{-1} - 2\mu_c \frac{LT}{D_c} \right\} \frac{Z_n}{T} \left(1 - \frac{Z_n}{T}\right), \tag{2.9c}$$

$$K = 2\mu_b \left(\frac{LT}{D_b}\right) + \pi^2 \sigma_c \frac{D}{L} T. \tag{2.9d}$$

Where k_1 and k_2 are the spring that connect the cover with body block, k_c is the coouppling srping for the two cover blocks and K represent the spring component of the body block.

This method of calculating the properties of the three-mass model provides a static view of the mechanical and geometric components of the VFs, yet these

are dynamic properties. Both mechanical and geometrical properties of the VFs change dynamically due to the action of the laryngeal musculature.

2.1.4 Laryngeal muscles activation models

Intrinsic laryngeal muscles

The dynamic properties of the VFs are largely influenced by the action of the intrinsic laryngeal muscles, which play a crucial role in modulating the mechanical properties and shape of the VFs, thereby affecting their vibration and the resulting sound production [96]. These muscles include the Lateral Cricoarytenoid (LCA), Interarytenoid (IA), Posterior Cricoarytenoid (PCA), Cricothyroid (CT), and Thyroarytenoid (TA), each of which plays a specific role in VF function, as illustrated in Figure 2.8 and Figure 2.9

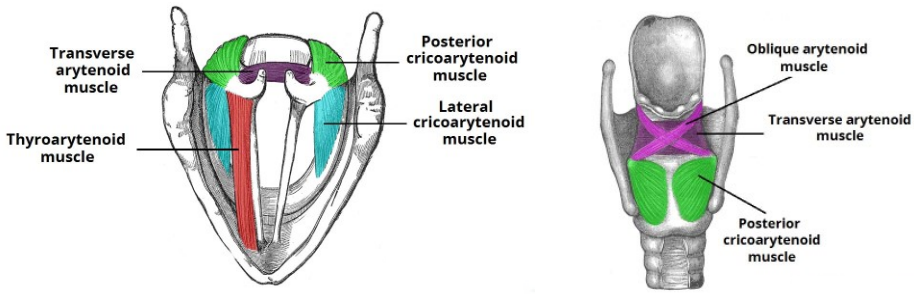

Figure 2.8: Intrinsic laryngeal muscles.

The LCA muscles are primarily responsible for adducting the VFs, pulling them together. This action is critical for initiating and sustaining phonation by

closing the glottal gap, allowing for sound production. Similarly, the IA muscles aid in adduction by pulling the arytenoid cartilages towards each other, assisting in closing the posterior portion of the glottal gap.

In contrast, the PCA is the only muscle involved in abducting the VFs, moving them apart. This action is essential for breathing, as it opens the glottal gap, enabling airflow through the vocal folds.

The CT and TA muscles primarily control the tension and length of the VFs. The CT muscle functions to tense and elongate the VFs, thereby increasing the pitch of the voice. Conversely, the TA muscle, which includes the vocalis muscle, acts to relax and shorten the VFs, resulting in a lower pitch. Through their antagonistic relationship, the CT and TA muscles regulate the natural oscillation frequency of the VFs, as well as their passive and active stiffness, and contribute to shaping the vocal process.

In summary, the LCA and IA muscles are primarily responsible for VF adduction, while the PCA handles abduction. Together, these muscles manage the position and configuration of the arytenoid cartilages at the back of the glottis. Meanwhile, the CT and TA muscles modulate the length and stiffness of the VFs, influencing the frequency of vocal fold oscillation and the overall quality of phonation.

To develop a VF muscle control model, it is essential to characterize the vocal folds in terms of both their mechanical properties and posture. The mechanical

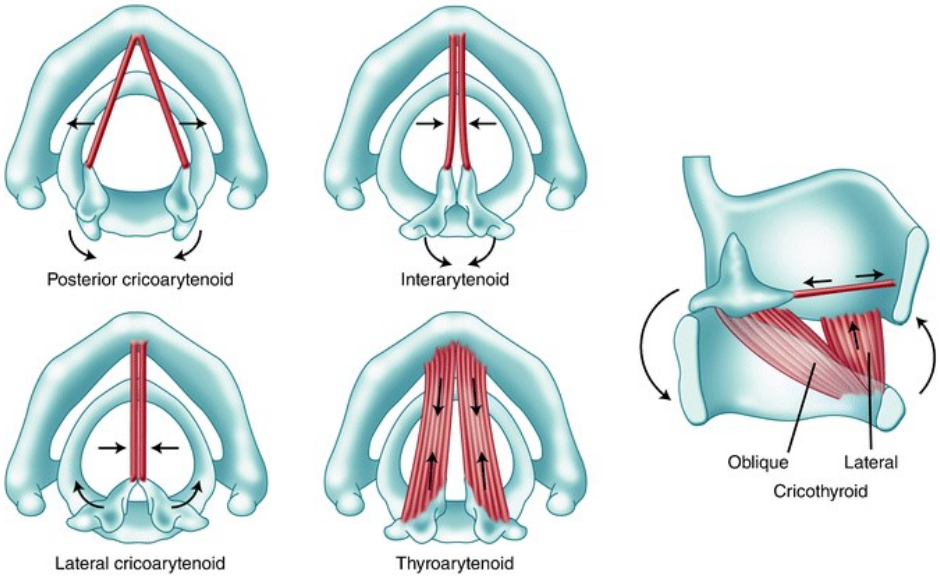

Figure 2.9: Representation of laryngeal muscles action. Adapted from [96]

aspect explores the relationship between strain (ϵ) and stress (σ), a key concept in tissue mechanics, especially for soft tissues like the VFs.

Lumped element models provide a useful approach for illustrating how laryngeal muscle activity influences the geometry and mechanics of the VFs. In this context, the vocal folds are modeled using a system of three masses representing the body and cover of the folds. As depicted in Figure 2.7, this approach allows for a detailed characterization of the upper and lower cover layers through distinct masses and springs, which better reflects the mechanical complexity of the VFs compared to a single mass model for the body.

Intrinsic laryngeal muscles action from [95]

Before delving into a complex representation of this phenomenon, in [95] the authors provide a straightforward approach to calculate parameters such as elongation, thickness, and elastic constants in a simplified representation of the VFs controlled by the action of intrinsic laryngeal muscles in lumped element models.

The discussion starts by defining a crucial element in this and future developments, the muscle activation level (a). This level can be understood as the normalized contraction force of the muscle from 0 to 1, where 0 is the muscle at rest and 1 is the muscle fully contracted, or at its maximum voluntary contraction [97].

With this definition, along with the action of the muscles in the glottis, the geometrical properties of the VF: the elongation (ϵ) and length (L) are defined as:

$$\epsilon = G\left(Ra_{\mathrm{CT}} - a_{\mathrm{TA}}\right) - Ha_{\mathrm{LC}}, \tag{2.10}$$

$$L = L_0(1 + \epsilon). \tag{2.11}$$

Where G, R and H are empirical constants to match the muscle activation action with excised larynx experiment, and L_0 is the rest VF length [95]. At the same time, the other geometrical dimension for the blocks that represent Vfs, the

thickness (T) of the fold, depth (D) of the body and cover layers, and the nodal point (Z_n) for upper and lower distribution of the VFs are defined as:

$$T = \frac{T_0}{1 + 0.8\epsilon}, \tag{2.12}$$

$$\begin{aligned} D_b &= \frac{a_{\mathrm{TA}} D_{\mathrm{mus}} + 0.5 D_{\mathrm{lig}}}{1 + 0.2\epsilon}, & (2.13a) \\ D_c &= \frac{D_{\mathrm{muc}} + 0.5 D_{\mathrm{lig}}}{1 + 0.2\epsilon}, & (2.13b) \end{aligned}$$

$$Zn = (1 + a_{\mathrm{TA}})\, T/3. \tag{2.14}$$

Here, the activation of the muscles responsible for elongating and relaxing the fold, CT and TA, is observed, in addition to the presence of a component due to adduction $a_{\mathrm{LC}} = a_{\mathrm{LCA}} - a_{\mathrm{PCA}}$. The constants accompanying these equations are empirically determined by adjusting experiments with dogs excised larynges [95].

Furthermore, in this work, elongation is linked to stress through: passive components, dependent on elongation; and active components, dependent on muscle activation.

$$\begin{aligned} \sigma_b &= \left[0.5\sigma_{\mathrm{lig}} D_{\mathrm{lig}} + \sigma_{\mathrm{mus}} D_{\mathrm{mus}}\right] D_b, & (2.15a) \\ \sigma_c &= \left[0.5\sigma_{\mathrm{lig}} D_{\mathrm{lig}} + \sigma_{\mathrm{muc}} D_{\mathrm{muc}}\right] D_c, & (2.15b) \\ \sigma_{\mathrm{mus}} &= a_{\mathrm{TA}} \sigma_{am} \max\left[0, 1 - b(\epsilon - \epsilon_m)^2\right] + \sigma_p. & (2.15c) \end{aligned}$$

$$\sigma_p = \begin{cases} 0 & \text{for } \epsilon \leq \epsilon_1, \\ -\frac{\sigma_0}{\sigma_1}(\epsilon - \epsilon_1) & \text{for } \epsilon_1 \leq \epsilon \leq \epsilon_2, \\ -\frac{\sigma_0}{\sigma_1}(\epsilon - \epsilon_1) + \sigma_2 \left[e^{C(\epsilon - \epsilon_2)} - C(\epsilon - \epsilon_2) - 1\right] & \text{for } \epsilon > \epsilon_2. \end{cases} \tag{2.16}$$

The value for Equation 2.16 constants is in Table I of [95].

In terms of geometry, this approach define two rules for ξ_{02} and ξ_{01}, the adduction level of upper and lower block, respectively:

$$\xi_{02} = 0.25L_0(1 - 2.0a_{LC}), \tag{2.17}$$

$$\xi_{01} = \xi_{02} + T(0.05 - 0.15a_{TA}). \tag{2.18}$$

In this example, muscle activation defines both stress and strain, with strain determining stress. However, for this model, there is not inverse connection, meaning that strain is not affected by the total stress of the muscles; something that we know occurs in soft materials.

A more complex representation of this relationship for intrinsic muscles can be achieved through a 1D Kelvin model for the muscles [98, 44], a differential equation inspired in a spring-damping representation of the muscle elongation; and we will also associate the elongation of the VF, not to an empirical equation with the muscle activation, but to the displacement of the cricoarytenoid (CAJ) and cricothyroid (CTJ) junctions.

Intrinsic laryngeal muscles action in elongation and posture, [98]

This approach to the stress-strain relationship involves seven elements: the five intrinsic muscles plus the ligament and mucosa: $\mathcal{I} = \{\text{LCA, IA, PCA, CT, TA, LIG, MUC}\}$. Each element is modeled using a Kelvin model where total stress (σ) depends on strain (ϵ), and both active (σ_a) and passive (σ_p) stress components. This results in a set of two coupled equations for passive stress:

$$t_s\dot{\sigma} + \sigma = \sigma_a + \sigma_p(\epsilon) + E(\epsilon)t_p\dot{\epsilon}, \tag{2.19}$$

and active stress

$$t_a\dot{\sigma_a} + \sigma_a = a\sigma_m f(\epsilon)g(\dot{\epsilon}). \tag{2.20}$$

Where t_s, t_p and t_a are constant; E, f and g are functions dependent on ϵ; σ_p is defined by Equation 2.16 and a is the activation of each intrinsic muscle. The explicit forms and values of constants can be found in Section 2.2 of [97], noting that the LIG and MUC components lack an active stress component (σ_a), being non-contractile tissue.

To solve the Equation 2.19 and Equation 2.20, the strain of each muscle and the VF elongation is defined by the displacement of the CAJ and CTJ. A diagram of these joints and their rotational and translational displacement is seen in Figure 2.10. The dynamics for both joints is defined by the nexts motion equation:

The displacement in x (ξ) and y (ψ), and the rotation (θ) of the CAJ, is given

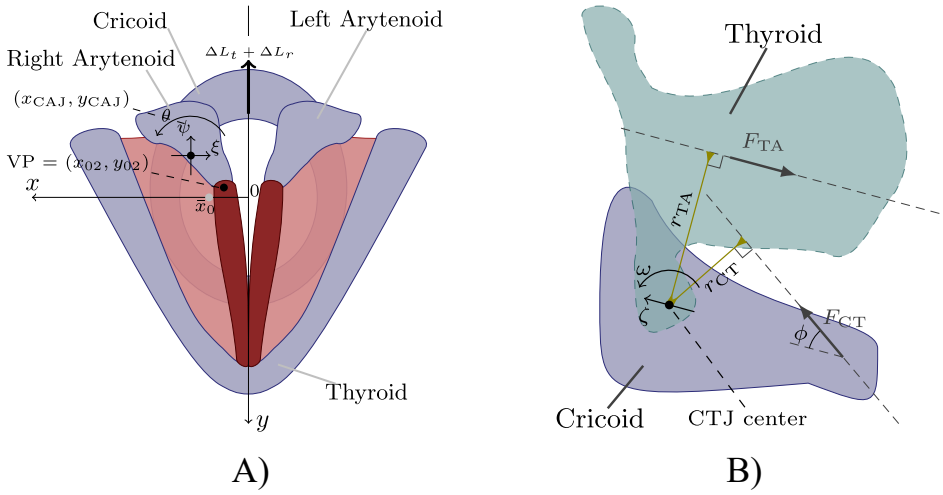

Figure 2.10: Main laryngeal structures involved in the prephonatory posturing. (A) Cricoarytenoid accommodation at the glottal plane; and (B) Projection of effective cricothyroid accommodation on the glottal plane. Glottal geometry and VF adjustment are controlled via the relative accommodation of major laryngeal cartilages. Figures adapted from [44]. CAJ: cricoarytenoid joint, CTJ: cricothyroid joint.

by:

$$M_{\text{ac}}\,\ddot{\xi} + d_x\dot{\xi} + k_x\xi = \sum_{i\in\mathcal{I}} \delta_i F_i, \tag{2.21a}$$

$$M_{\text{ac}}\,\ddot{\psi} + d_y\dot{\psi} + k_y\psi = \sum_{i\in\mathcal{I}} \beta_i F_i, \tag{2.21b}$$

$$I_{\text{ac}}\,\ddot{\theta} + \delta\dot{\theta} + \kappa\theta = \sum_{i\in\mathcal{I}} \gamma_i F_i. \tag{2.21c}$$

Where F_i represents the force exerted by each muscle: $\mathcal{I} = \{\text{LCA, IA, PCA, CT, TA, LIG, MUC}\}$, which can be calculated as $F_i = \sigma_i \times A_i$, with A_i the cross-sectional area of the muscle. The δ, β, γ are the directional cosines for the forces. The constants are presented in Table I of [44].

For CTJ, the translational (ζ) and rotational (ω) movement, is also defined by:

$$M_{\mathrm{cc}}\ddot{\zeta} + k_t\left(t_t\dot{\zeta} + \zeta\right) = \left[\cos\phi\, F_{\mathrm{CT}} - (F_{\mathrm{TA}} + F_{\mathrm{LIG}} + F_{\mathrm{MUC}})\right], \quad (2.22a)$$

$$I_{\mathrm{cc}}\ddot{\omega} + k_r\left(t_r\dot{\omega} + \omega\right) = \left[r_{\mathrm{CT}}F_{\mathrm{CT}} - r_{\mathrm{TA}}\left(F_{\mathrm{TA}} + F_{\mathrm{LIG}} + F_{\mathrm{MUC}}\right)\right], \quad (2.22b)$$

Where M and I are the traslational mass and the inertial moment; k_t and k_r are the translational and rotational stiffnesses, t_t and t_r are the time constants for the translational and rotational viscous damping, r_{TA} and r_{CT} are the moment arms for the TA and CT torques relative to CTJ center, and ϕ is the angle between the line of action of F_{CT} and the translation axis of the cricoid cartilage, respectively. The value of the constants is detailed in Chapter 3 of [97] and Table I of [44].

Solving the differential equations for ξ, ψ, ζ, θ and ω allows defining the elongation effect of the CAJ and CTJ, resulting in:

$$\Delta L_t = \zeta, \quad (2.23a)$$

$$\Delta L_r = r_{\mathrm{TA}}\,\omega, \quad (2.23b)$$

$$\Delta L_a = -\left[y_{CAJ}(1-\cos\theta) - (x_{CAJ} - \bar{x}_{02})\sin\theta + \psi\right]. \quad (2.23c)$$

And also the aryteniod position in x-y plane, that define the vocal process:

$$x_{02} = x_{CAJ} - (x_{CAJ} - \bar{x}_0)\cos\theta + y_{CAJ}\sin\theta + \xi, \tag{2.24a}$$

$$y_{02} = y_{CAJ}(1 - \cos\theta) - (x_{CAJ} - \bar{x}_0)\sin\theta + \psi - (\Delta L_t + \Delta L_r). \tag{2.24b}$$

And additional element, the x-displacement of the lower portion of the VF is computed following the idea of [95]:

$$x_{01} = T \cdot tan(0.001) + x_{02}. \tag{2.25}$$

The total elongation or strain of the VF can then be expressed as:

$$\epsilon_{VF} = \frac{1}{L_0}\left(\Delta L_a + \Delta L_t + \Delta L_r\right), \tag{2.26}$$

where L_0 is the initial VF length. Equation 2.26 dynamically couples the adduction and elongation procedures [95, 98]. The VF length is $L_g = (1 + \epsilon)\, L_0$.

Thus characterizing the elongation of each of the 7 muscle components as:

$$\epsilon_{\mathrm{CT}} = \frac{L_0}{L_{CT}}\left(r_{CT}\omega + \zeta/cos(\phi)\right), \tag{2.27a}$$

$$\epsilon_{\mathrm{TA}} = \epsilon_{\mathrm{LIG}} = \epsilon_{\mathrm{MUC}} = \frac{1}{L_0}\left(\Delta L_a + \Delta L_t + \Delta L_r\right), \tag{2.27b}$$

$$\epsilon_{\mathrm{LCA}} = -\frac{\gamma_{LCA}\theta}{L_{LCA}}, \tag{2.27c}$$

$$\epsilon_{\mathrm{IA}} = -\frac{\gamma_{IA}\theta}{L_{IA}}, \tag{2.27d}$$

$$\epsilon_{\mathrm{PCA}} = -\frac{\gamma_{PCA}\theta}{L_{PCA}}, \tag{2.27e}$$

$$\tag{2.27f}$$

$L_{\mathcal{I}}$ denotes the initial length of the $\mathcal{I}$ muscle. This reveals how the CT muscle strain is defined by the extension due to CTJ; for the TA muscle, Ligament, and Mucosa, the strain is the elongation of the VF, and for posterior muscles LCA, IA and PCA the strain is the rotational contribution in the CAJ displacement.

For a straightforward implementation, a presentation as an algorithm for solving the components of strain and stress of the VFs, given a variable input of muscular activation, would be, following [98]:

1. Define the forces $F_{\mathcal{I}}$ and active stresses $\sigma_{\mathcal{I}}$ (Equation 2.19 and Equation 2.20) for all the muscles as fourth-order Runge-Kutta dependent variables. There are a total of 12 variables; however, the LIG and MUC do not have active stresses.

2. Additionally, define ξ, $\dot{\xi}$, ψ, $\dot{\psi}$, θ, $\dot{\theta}$, ζ, $\dot{\zeta}$, ω, $\dot{\omega}$ as Runge-Kutta dependent variables, summing up to 10 variables, using Equation 2.21 and Equation 2.22.

3. The derivatives of all the 22 variables are further defined as independent variables, culminating in a total of 44 first-order differential equations that need simultaneous solving.

4. Muscle activations ($a_{\mathcal{I}}$) are regarded as inputs, which can be either constant or time-varying.

5. Within the Runge-Kutta loop, compute strains ($\epsilon_{\mathcal{I}}$, Equation 2.27), strain rates ($\dot{\epsilon_{\mathcal{I}}}$), active and passive forces of all muscles, passive forces of the ligament and mucosa, and the first-order derivatives as stipulated in step three.

6. Obtain the time-varying 44 variables.

2.1.5 Triangular body-cover model: TBCM

Having established the role of laryngeal musculature in controlling both the mechanical properties and posture of the VFs, the next step is to introduce a model that integrates these elements into a comprehensive framework. This model allows for the representation of the body-cover structure of the VFs while incorporating muscle control over both their mechanical characteristics and laryngeal posture.

The TBCM [42, 44], a symmetrical low-order model of the VFs, incorporates tissue-fluid-acoustic interactions at the glottis, enabling simulations of sustained vowels and time-varying glottal gestures.

Building upon prior efforts in low-order modeling [95], VF posturing [98], body-cover VF models [90], and the triangular shape of the glottis [42, 84], the TBCM consists of paired three-mass body-cover systems (see Figure 2.12) and a muscle-controlled model of all five intrinsic laryngeal muscles: CT, TA, LCA, IA, PCA. These muscles allow for dynamic control of the prephonatory posture (i.e., vocal

process accommodation) and the VF configuration (i.e., viscoelastic VF properties) during phonation. Each intrinsic muscle is treated using a modified Kelvin model for the stress-strain response in the tissues, considering both the passive and active stress components [97]. Following the model described in subsubsection 2.1.4. The model offers flexibility by allowing normalized (beween 0 and 1) individual actuation levels for each muscle, namely, $(a_{\mathrm{LCA}}, a_{\mathrm{IA}}, a_{\mathrm{PCA}}, a_{\mathrm{CT}}, a_{\mathrm{TA}})$. Moreover, this model allows for the incorporation of pulmonary pressure (P_L) as an adjustable input parameter, working in tandem with the muscle activation profiles.

General scheme of TBCM in the phonation model:

The TBCM model operates within a broader phonation model, capturing essential elements of vocal fold dynamics and the interactions between physiological systems during voice production. Figure 2.11 provides an overview of how the TBCM is embedded in a global phonatory framework, including key elements such as the subglottal tract, laryngeal muscle control, and the vocal fold model itself, along with their respective inputs and outputs. Those components have been detailed in the preceding sections of this chapter.

The **subglottal tract** plays a vital role in propagating lung pressure (P_L) to the glottis, where the subglottal pressure (P_S) is distributed along the tract. This pressure propagation is modeled using WRA [43], where the subglottal tract

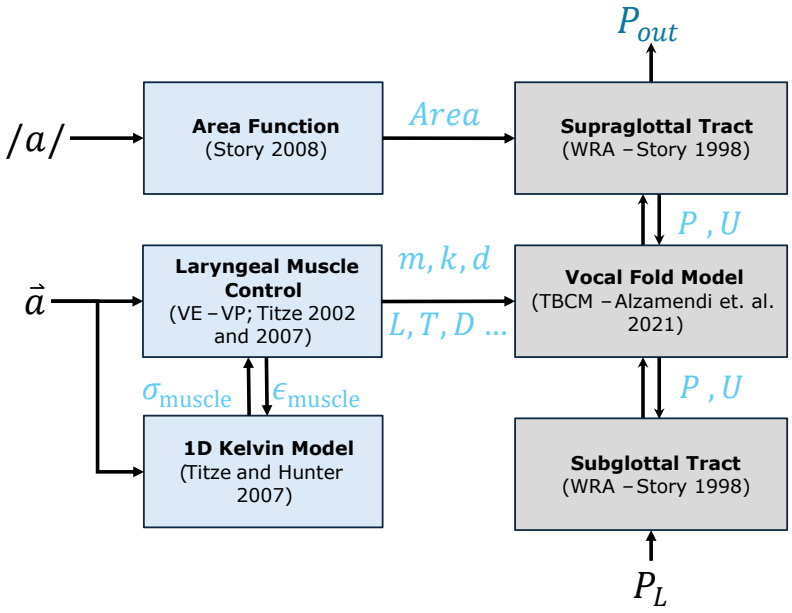

Figure 2.11: Scheme full phonation model TBCM.

is represented as an area function. This principle is briefly introduced in Equation 2.2. The input to this component is lung pressure, while the output is the calculated subglottal pressure and airflow along various points in the tract.

Next, the **laryngeal muscle control** system is tasked with defining the VF posture and viscoelastic (VE) properties that are essential for VF oscillation. The posture, primarily driven by the positioning of the vocal process (VP), is influenced by muscle activations [98]. Additionally, the mechanical properties of the VFs are characterized by their mass and spring dynamics [95]. The inputs for this section include muscle activation vectors and stress/strain (σ/ϵ) equations, which are outputs of the earlier described 1D model that links muscle activation with stress. The outputs are the physical displacement of the arytenoids, posterior

glottal opening (PGO), and the resulting geometric and mechanical properties of the VFs.

The **vocal fold model** itself is represented through a system of three coupled mass-spring oscillators, where each mass corresponds to a section of the vocal fold layers (body and cover) [90, 42, 44]. The mechanical parameters derived from muscle control and the subglottal/supraglottal pressures (P_i, P_S) serve as inputs to this system [81]. The resulting outputs are critical variables such as the forces exerted on the VFs, the glottal area, and the resulting glottal airflow (U_t).

In this broader context, the **supraglottal tract** manages the airflow and pressure from the glottis to the output at the mouth. It uses the same WRA [43] approach as the subglottal tract, with reflection coefficients calculated based on the area of each section of the vocal tract. The input to this system is the glottal airflow (U_t), and the output includes both the supraglottal pressure (P_i) and the pressure radiated at the mouth (P_{out}).

Finally, the **area function** plays an essential role in defining the geometry of the vocal tract, with a dictionary of cylindrical section areas for various vowels [64]. This input enables the vocal tract to be dynamically adjusted for different phonatory gestures. The vocal tract geometry, in turn, affects how pressure and airflow propagate through the system.

By integrating all of these elements, the TBCM model offers a comprehensive framework for simulating and understanding the intricate processes that occur

during phonation.

TBCM motion equation:

In TBCM, the aerodynamic forces over the VF cover layer are computed according to [99], while elastic, damping, and collision forces over the body-cover masses follow [42]. Glottal airflow is computed from the acoustic driving pressures on the glottal area, as [97, 100, 81]. The acoustic wave propagation is simulated using the wave reflection analog scheme, modeling subglottal and supraglottal tracts as a discrete concatenation of acoustic cylinders with variable cross-sectional areas [73], similar to Equation 2.2.

TBCM stands out as a particularly effective model for describing VH due to its comprehensive approach in simulating the VF dynamics. By integrating the effects of PGO and allowing for both posterior and membranous glottal gaps within a single framework, TBCM facilitates a more nuanced analysis of VF posturing and the resulting phonatory behavior.

The coupled equations for model the dynamics of the three masses in TBCM are:

$$F_u = m_u\,\ddot{x}_u = F_{k,\,u} + F_{d,\,u} - F_{kc} + F_{e,\,u} + F_{Col,\,u}, \tag{2.28a}$$

$$F_l = m_l\,\ddot{x}_l = F_{k,\,l} + F_{d,\,l} - F_{kc} + F_{e,\,l} + F_{Col,\,l}, \tag{2.28b}$$

$$F_b = m_b\,\ddot{x}_b = F_{k,\,b} + F_{d,\,b} - \left[F_{k,\,u} + F_{d,\,u} + F_{k,\,l} + F_{d,\,l}\right], \tag{2.28c}$$

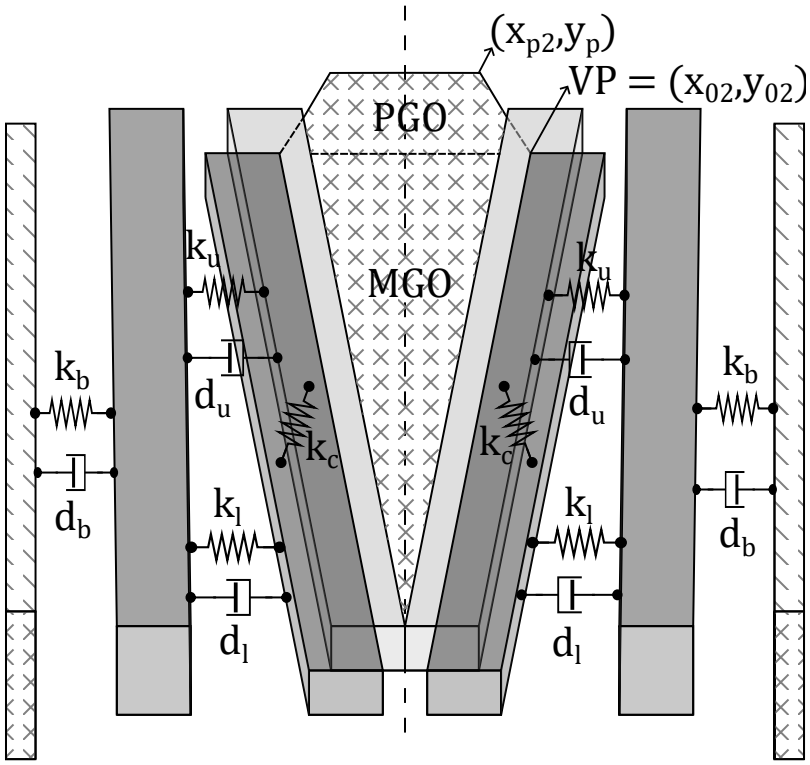

Figure 2.12: Scheme of TBCM for the VFs.

where m is mass, x is the medial-lateral displacement measured in the middle of the VF, and F is the force component for each block. Furthermore, force subscripts k, d, e, and kc represent the mechanical forces produced by the springs, dampers, flow pressures, and elastic coupling between the upper and lower masses, respectively. An additional spring force, F_{Col}, is introduced during VF collision to capture the effects of impact between opposite upper/lower cover masses.

The definitions of the forces in Equation 2.28 are similar to the forces originally expressed in [90], the elastic forces modeled as non-linear spring force. Defining a

general displacement around equilibrium position:

$$\Delta x_i = (x_i - x_{i,0}), \tag{2.29}$$

where i can be u for upper, l for lower and b body block. the elastic forces are modeled as:

$$F_{k,u} = -k_u \left\{ [\Delta x_u - \Delta x_b] + \eta_u \left[\Delta x_u - \Delta x_b \right]^3 \right\}, \tag{2.30a}$$

$$F_{k,l} = -k_l \left\{ [\Delta x_l - \Delta x_b] + \eta_l \left[\Delta x_l - \Delta x_b \right]^3 \right\}, \tag{2.30b}$$

$$F_{k,b} = -k_b \left[\Delta x_b - \eta_b \Delta x_b^3 \right], \tag{2.30c}$$

$$F_{kc} = -k_c \left[\Delta x_l - \Delta x_u \right]. \tag{2.30d}$$

where the k_s are linear spring constants, the η_s are nonlinear spring constants, and the $x_{s,0}$ are the rest positions of the respective masses, with $s \in \{u, l, b\}$. Rest positions for the upper and lower blocks are $x_{u,0} = 0.5\, x_{02}$, $x_{l,0} = 0.5\, x_{01}$ pursuant to VF posturing, whereas $x_{b,0} = 3$ mm [42]. For all these parameters, a set of controlling rules are used to assign a determinate value [95], computed using Equation 2.9.

In similar manner, the collision force is modeled as spring force. This force needs to consider the triangular shape of the VF definer by the vocal process. To compute the total collision force for lower or upper block, the collision fraction is defined:

$$\alpha_i = L_{Col,\,i}/L_g, \tag{2.31}$$

where the subscript $i \in \{u,\, l\}$, and α_i is the proportions of block length ($L_{i,\,\mathrm{col}}$) undergoing collision at the given time, with $0.0 \leq \alpha_i \leq 1.0$. The collision for is modeled as the integral for the differential force, with the force as a non-linear Hooke's law.

$$\begin{aligned} F_{Col,\,u} &= -k_{Col,\,u} \int_0^{L_{Col,\,u}} x_u(y) + \eta_{Col,\,u} x_u^3(y) dy, & \text{(2.32a)} \\ F_{Col,\,l} &= -k_{Col,\,l} \int_0^{L_{Col,\,l}} x_l(y) + \eta_{Col,\,l} x_l^3(y) dy, & \text{(2.32b)} \end{aligned}$$

$$\begin{aligned} x_u(y) &= x_u - \frac{x_{02}}{2} + x_{02}\frac{y}{L_g} = c_u + b_u \frac{y}{L_g}, & \text{(2.33a)} \\ x_l(y) &= x_l - \frac{x_{01}}{2} + x_{01}\frac{y}{L_g} = c_l + b_l \frac{y}{L_g}, & \text{(2.33b)} \end{aligned}$$

The integrals are solved considering the triangular shape given by the position of the VFs (c_i , b_i). Obtaining a simplified form, we have:

$$\begin{aligned} F_{Col,\,u} &= -\frac{k_{Col,\,u}}{4}\alpha_u \left(2c_u + b_u\alpha_u\right)\left[2 + \eta_{Col,\,u}\left(2c_u^2 + 2b_u c_u \alpha_u + b_u^2 \alpha_u^2\right)\right], & \text{(2.34a)} \\ F_{Col,\,l} &= -\frac{k_{Col,\,l}}{4}\alpha_l \left(2c_l + b_l\alpha_l\right)\left[2 + \eta_{Col,\,l}\left(2c_l^2 + 2b_l c_l \alpha_l + b_l^2 \alpha_l^2\right)\right]. & \text{(2.34b)} \end{aligned}$$

Damping forces are modeled as a typical proportional velocity force in a harmonic oscillator:

$$F_{d,u} = -2\nu_u\sqrt{m_u k_u}\,(\dot{x}_u - \dot{x}_b)\,, \tag{2.35a}$$

$$F_{d,l} = -2\nu_l\sqrt{m_l k_l}\,(\dot{x}_l - \dot{x}_b)\,, \tag{2.35b}$$

$$F_{d,b} = -2\nu_b\sqrt{m_b k_b}\dot{x}_b. \tag{2.35c}$$

where ν_s is the damping ratio, with $s \in \{u, l, b\}$. For $\nu_s = 1$, the damping force is exactly that of a critically damped oscillator.

The aerodynamic force F_e is the force exerted by the intraglottal flow. To calculate it, it's necessary to determine the glottal area defined by the two blocks composing the cover. Considering VF posture and symmetry concerning the mid-sagittal plane, the glottal area defined by the upper and lower cover blocks can be expressed as follows:

$$A_u = 2\max\left\{0,\, x_u - \frac{x_{02}}{2}\right\} + 2(1-\alpha_u)^2 L_g \frac{x_{02}}{2}, \tag{2.36a}$$

$$A_l = 2\max\left\{0,\, x_u - \frac{x_{02}}{2}\right\} + 2(1-\alpha_u)^2 L_g \frac{x_{02}}{2}, \tag{2.36b}$$

The first term of the equation represents the rectangular area generated when the anterior part of the block is not in collision, which becomes zero during collision; meanwhile, the second term represents the triangular area that is present at all times and is weighted by $1-\alpha$, which is the fraction not in collision. the 2 factor correspond to the VF symmetry.

The aerodynamic force will use the pressure per area expression, considering the cross-sectional area of each block that is not in collision. Thus, the expression for the pressures in the blocks when one of them is fully colliding, i.e., when $A_g = \min\{A_u,\, A_l\}$, means $A_g = 0$, is:

$$P_u = \begin{cases} 0 & \text{for } A_u \leq 0 \text{ and } A_l \leq 0, \\ 0 & \text{for } A_u \leq 0 \text{ and } A_l > 0, , \\ P_i & \text{for } A_u > 0 \text{ and } A_l \leq 0, \end{cases} \tag{2.37a}$$

$$P_l = \begin{cases} 0 & \text{for } A_u \leq 0 \text{ and } A_l \leq 0, \\ P_S & \text{for } A_u \leq 0 \text{ and } A_l > 0, , \\ 0 & \text{for } A_u > 0 \text{ and } A_l \leq 0, \end{cases} \tag{2.37b}$$

Meanwhile, when the glottis is open, the Equation 2.5 is utilized:

$$P_u = P_i, \tag{2.38a}$$

$$P_l = P = P_i + (P_S - P_i)\left(1 - \frac{A_g^2}{A_l^2}\right), \tag{2.38b}$$

with the pressure expressions, the force exerted on each block is calculated as:

$$F_{e,\,u} = P_u(1 - \alpha_u)L_g T_u, \tag{2.39a}$$

$$F_{e,\,l} = P_l(1 - \alpha_u)L_g T_l. \tag{2.39b}$$

Another way to represent the pressure in each block is presented in [99], where the maximum power transfer theorem is applied.

With the representation of forces acting on the blocks, it becomes possible to characterize the time-varying glottal area. The missing element in this VF model is the construction of the glottal flow, defined as the difference between subglottal and supraglottal pressure. This flow is influenced by the glottal area and the posterior glottal opening, which is defined by the displacement of the arytenoids. The PGO is quantified as the area of the trapezoid formed at the posterior aspect, as depicted in Figure 2.12. The equation for PGO is as follows:

$$A_{\text{PGO}} = \max\{0,\, \min\{(x_{p1} + x_{01}),\, (x_{p2} + x_{02})\}(y_{02} - y_p)\}. \tag{2.40}$$

Regarding the glottal flow, it can be calculated as presented in subsection 2.1.2 or through the smoothed representation provided by [81], which is expressed by:

$$u_g = \frac{2A_T^3\delta_p}{\left(\frac{\rho c A_T^3}{A^*} + \gamma_m + \sqrt{\left(\frac{\rho c A_T^3}{A^*} + \gamma_m\right)^2 + 2k_t\rho A_T^4|\delta_p|}\right)}. \tag{2.41}$$

Where A_T is the sum of the membranous component and the contribution of the PGO.

$$A_T = A_g + A_{\text{PGO}}, \tag{2.42}$$

additionally,with A^* the effective area computed as:

$$A^* = \left(\frac{1}{A_i} + \frac{1}{A_s}\right)^{-1}, \tag{2.43}$$

with A_i and A_s the area of the first supraglottal and subglottal section respectively, δ_p is the pressure difference $\delta_p = P_S - P_i$, γ_m is the viscosity air contribution and k_t the coefficiented introduced in Equation 2.6.

2.2 Vocal hyperfunction

The modeling component of this thesis is designed to characterize VH, making it essential to first define what VH entails and to review the key developments in this field.

VH encapsulates a range of conditions characterized by an overexertion or imbalance in the musculature surrounding the larynx during the act of phonation [101]. This overexertion can manifest through benign VF pathologies, such as nodules, or through forms of dysphonia that do not coincide with any visible VF abnormalities, commonly referred to as primary muscle tension dysphonia [102].

Fundamentally, VH is a consequence of vocal misuse or abuse, where the phonatory system is subjected to excessive strain due to unbalanced or excessive muscular forces. This strain often arises from an individual's attempt to compensate for perceived deficiencies in vocal output, such as insufficient volume, which can initiate a harmful cycle. Increased muscle and aerodynamic forces exerted to augment vocal projection can lead to vocal strain, tissue damage, and a progressive decline in vocal quality, necessitating even greater effort for voice pro-

duction. VH is broadly categorized based on its etiology into two distinct types: PVH and NPVH.

PVH is identified by visible signs of damage on the VFs contact surfaces, such as nodules or polyps [12]. This damage often stems from daily or habitual vocal behaviors that exert excessive strain on the VFs. Such behaviors may include speaking at an intensity that is too high, using a pitch that is unsuitable for the individual's voice, engaging in extended periods of speech without adequate rest [5, 6, 103, 104, 105]. The diagnosis of PVH typically requires a laryngoscopic examination to directly observe the VFs for any lesions or changes in tissue [106]. Further assessment of voice quality, along with a detailed patient history, assists in identifying vocal behaviors that may contribute to the disorder. The primary treatment for PVH involves voice therapy, aimed at modifying detrimental vocal habits to alleviate strain on the VFs. In cases where lesions have developed, surgical removal may be deemed necessary. Post-surgical voice therapy is critical in altering the vocal habits that led to the disorder, thus helping to prevent its recurrence.

On the other hand, NPVH, alternatively known as primary muscle tension dysphonia or functional dysphonia, represents a prevalent and diverse category of voice disorders frequently addressed by laryngologists and speech-language pathologists [107, 108, 109]. Characterized by a range of habitual and chronic vocal symptoms encountered in daily activities—such as dysphonia [110, 111],

increased vocal effort or fatigue [112], and anterior neck discomfort or excessive muscle activity. NPVH manifests without evident phonotraumatic damage or other structural or neurological conditions that might impair phonation. Diagnosis of NPVH is thorough, incorporating auditory-perceptual analysis, acoustic measurements, and laryngoscopic examinations to exclude the presence of phonotraumatic lesions, with a focus on detecting misuse or dysfunction of laryngeal muscles. The cornerstone of NPVH treatment is voice therapy, which emphasizes techniques aimed at minimizing laryngeal tension, enhancing VF vibration, and promoting efficient voice production.

Research into VH reveals two distinct patterns of biomechanical behavior, distinguishing PVH from NPVH. Individuals with PVH exhibit significantly higher subglottal pressures and airflow rates during phonation, indicative of more intense VF vibrations and faster closure rates necessary to produce standard voice levels [5, 6, 10, 11]. This biomechanical profile suggests a heightened risk of VF damage, potentially leading to persistent lesions and dysphonia. The concept of a "vicious cycle" in PVH, as proposed by clinicians, illustrates the condition's exacerbation due to increasing hyperfunctional efforts to counter VF trauma [42, 100]. Computational models corroborate this, showing that compensatory measures, such as augmented subglottal pressure, can indeed achieve desired vocal outcomes but at the cost of elevated VF collision forces, thereby escalating the risk of phonotrauma. Contrastingly, NPVH is marked by inefficient phonation also characterized by ele-

vated subglottal pressures and incomplete glottal closure—resulting in dysphonia with a lower risk of tissue damage [5, 6, 10, 11]. Studies utilizing ambulatory voice monitoring have further delineated PVH, identifying distinct vocal usage patterns that underscore the biomechanical divergence between PVH and NPVH [18, 16, 113].

2.2.1 Modeling biomechanical elements in VH

Following our discussion on VH, we now explore biomechanical modeling approaches for PVH and NPVH. For PVH, the focus is on modeling the VF tissue trauma. This element can explain the need for elevated subglottal pressure, and the resultant high VF collision forces, effect to increased VF mass and incomplete closure, as identified in [10, 5, 6]. The nodules and polyps have been modeled in previous approaches as a swelling in the VFs characterized by different mechanical properties: density, size and elastic properties as Young modulus and Poisson ratio [92, 93, 94]. Table 2.2 present an overview of parameters range in different experimental, clinical and modeling studies.

From the perspective of lumped element for example, previous works introduce the nodule or polyp as a change in mass and spring values of multi-mass oscillator model and change in the VF collision instance [49] or as a new oscillator that change the VF shape and collision [56, 48, 119]. In continuous model, the trauma is defined as a different material adding to the VF, changes que geometry and

Table 2.2: VF parameters for nodule implementation.

Feature	**Abbr.**	**Range**	**Unit**	**Reference**
Nodule Length	L_n	1 to 3	mm	[7, 52, 114]
VF Young Modulus	E_{VF}	1 to 40	kPa	[91, 92, 115, 94, 52, 59, 88, 116, 117]
Nodule Young Modulus	E_n	10 to 500	kPa	[52, 53, 88, 114]
Shear Modulus	μ_{VF}	0.5 to 1	kPa	[97, 92, 93, 59, 118, 116]
Poisson Ratio	ν_{VF}	0.3	–	[92, 115, 52, 53, 97, 88, 118]

producing new oscillation patrons [88, 52, 53, 115, 57]. These approaches are studied the effect of nodule parameters (size, stiffness, position) on PTP and show how the presence of the trauma increase this [114, 119]. This type of studies can be a guideline to represent the increase in subglottal pressure in PVH and understand the compensatory mechanism.

For NPVH, the focus is on modeling the increase in VF tension, and how to represent an imbalance in laryngeal muscle activity, and asymmetry in the vocal process. These elements can explain dysphonia and incomplete closure without a trauma. The asymmetry in the VF has been modeled from mechanical and geometrical differentiation of the proprieties of left and right VFs. Those approaches do not consider the muscle activity as the different element that control and produce the asymmetry. Many approaches utilize lumped element models where an

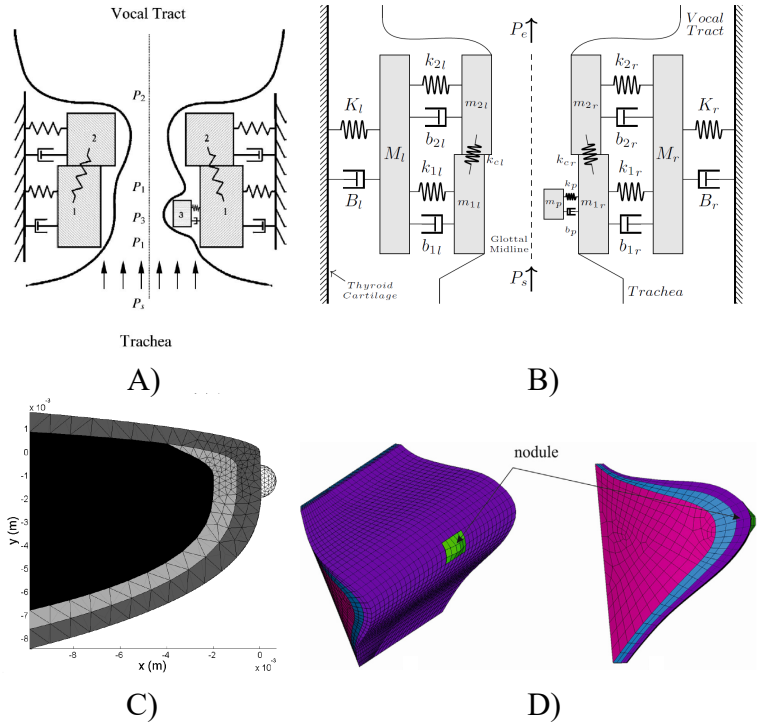

Figure 2.13: Examples of Nodules or PVH modeling approaches: (A) Zhang and Jiang [48], (B) Santos et al.[56], (C) Greiss et al. [53] and (D) Vampola et al. [115].

asymmetry factor q alters the ratio of left/right mass and spring component of the VF [50, 51, 122, 123, 120]. Continuous and synthetic models often directly modify the geometry or mechanical properties of the VFs [87, 54, 55, 124, 125]. Use a parameter that control the level of left/right asymmetry can help to understand how evolve the dysphonia. Include asymmetry factor in the muscle activity, that is the control element in the VF properties have the potential to represent the muscle imbalance in NPVH.

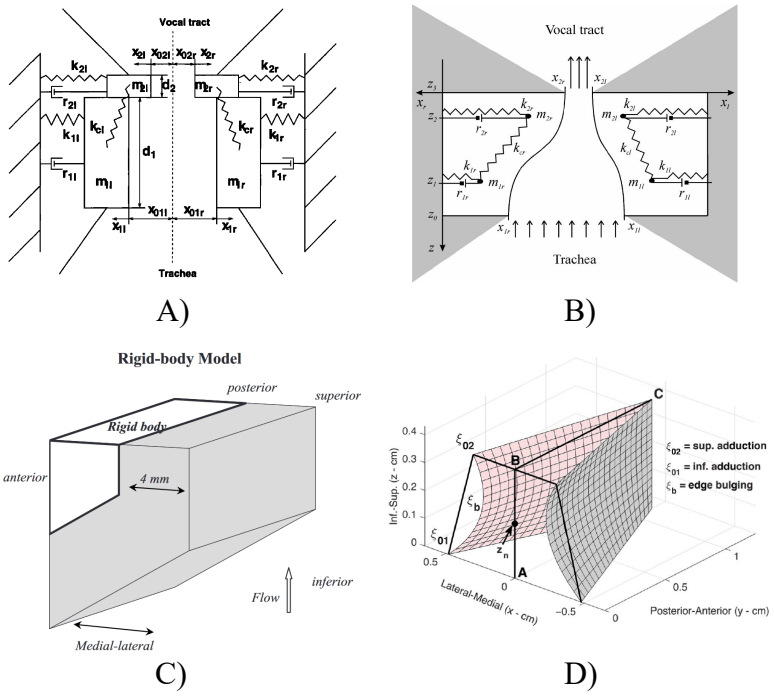

Figure 2.14: Examples of asymmetry modeling approaches: (A) Steinecke and Herzel [50], (B) Dresel et al. [120], (C) Zhang and Hieu Luu [121] and (D) Samlan et al. [54].

2.2.2 Ambulatory monitoring in VH

We have explored both direct (models of nodules and polyps) and indirect (models of asymmetry) approaches to identify elements present in the characterization of this vocal condition. Due to the significant daily behavioral component of voice disorders in VH, ambulatory monitoring is considered an ideal tool for examining voice use and providing insight into the characterization of voice behavior during daily activities [126, 14, 127]. Monitoring devices typically employ sensors placed on the neck, such as miniaturized acelerometers (ACC), to detect

neck skin vibration and, thus, non-invasively monitor phonation [13, 14, 128].

Monitoring efforts have historically concentrated on vocal intensity, fundamental frequency (f_o), and vocal use (vocal dose) as primary metrics for study [129]. Data collection and analysis have often utilized statistical perspectives, leveraging means, deviations, ranges, and other descriptors to manage the extensive volume of recordings collected throughout the day [130, 129, 131, 132, 128].

Research has leveraged the potential of ACC to estimate vital voice-related parameters, such as subglottal pressure and power dissipated by VFs, among others. These measurements have demonstrated a strong correlation with glottal flow, highlighting the ACC value in non-invasively assessing vocal function [133, 134]. Moreover, findings show that incorporating additional characteristics, such as spectral slope and RMS amplitude, can refine vocal function estimation [29, 15]. The transformative approach of impedance based inverse filtering (IBIF) [135] has further enabled the conversion of ACC signals into their glottal flow analogs, providing a more nuanced understanding of vocal mechanics.

In VH, aerodynamic quantities estimated from the ACC have proven differential between pathological and control populations, underscoring the tool diagnostic potential [10, 11]. Recent efforts have focused on developing classification and monitoring tools that are both straightforward and interpretable, facilitating their use in clinical settings [15, 136, 137].

These advancements have implications for post-therapy and surgery follow-

up, enabling the development of feedback protocols that empower users to better control their voice [136]. However, a significant challenge remains in ambulatory representation of the underlying mechanisms that produce aerodynamic characteristic differences, linking them to vocal function, subglottal pressure, and muscle activation.

2.2.3 Monitoring tools in VH

Accurate and continuous monitoring is crucial for understanding the dynamics of VH in real-world environments. Ambulatory monitoring tools provide valuable insights into the day-to-day vocal behaviors of individuals, helping to detect patterns associated with PVH and NPVH. Among these tools, the Daily Phonotrauma Index (DPI) stands out as a key metric for assessing the logit of VF trauma and classifying patients based on their vocal behaviors.

The DPI, a logistic regression for PVH and control using weekly ambulatory data, has emerged as a pivotal tool in classifying patients with PVH against control subjects by evaluating the potential for VF trauma. This study elaborates on two DPI versions—one utilizing SPL skewness [19] and the other employing Neck-Surface Acceleration Magnitude (NSAM) skewness [16], each combined with $H_1 - H_2$ standard deviation of ACC signal to reflect phonatory forces and VF closure dynamics.

In the initial version [19], the study involved 180 adult women, half diagnosed

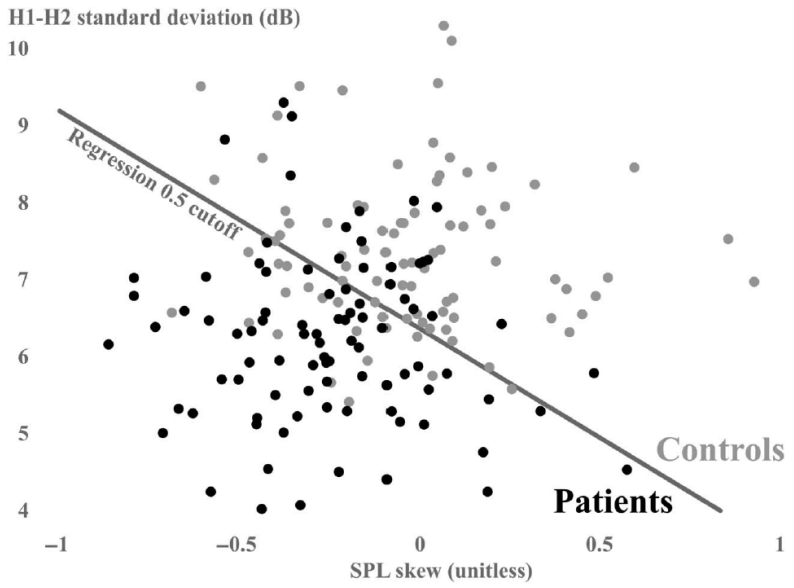

Figure 2.15: Scatter plot of DPI with SPL, input space: the standard deviation of $H_1 - H_2$ on the y-axis against the skewness of SPL on the x-axis. Patients with VH are represented in black, while matched controls are in gray. Each dot represents the weekly distribution parameters of a single patient. The logistic regression cutoff is depicted as a gray diagonal line. Adapted from [19].

with VF nodules or polyps and half as age, gender and occupation-matched controls without voice disorders. The detailed diagnosis process included comprehensive evaluations by a laryngologist or speech-language pathologist. Ambulatory voice data were collected using a neck-placed sensor (ACC) to unobtrusively monitor phonation [134]. The device recorded neck-surface acceleration, providing a weeklong summary of voice use including SPL, f_o, CPP, $H_1 - H_2$, among others. The study employed a subject specific linear regression to transform NSAM

into estimated SPL values [134]. This transformation allowed for the calculation of vocal dose measures and the identification of voiced and nonvoiced segments. Statistical analyses included paired t-tests and Wilcoxon signed-ranks tests to compare weekly voice use summary statistics between patients and controls. A logistic regression model was then constructed to classify participants based on SPL skewness and $H_1 - H_2$ variability, demonstrating the model effectiveness in distinguishing between PVH patients and matched controls. Figure 2.15 shows Logistic regression space reported in [19] The line represents the cutoff for classification.

To avoid the linear regression in the SPL calculation, the formal version of DPI uses NSAM skewness instead SPL skewness [16]. For the propose of this thesis, we use the SPL estimation data, because the current VF models do not have ACC as available signal but have sound pressure and SPL as possible output features. In terms of $H_1 - H_2$, the DPI uses the one calculated from the ACC signal. On the other hand, the VF models available to us have not yet considered this sensor. This would imply a drawback when trying to calculate DPI from VF models. However, there is a linear relationship between the $H_1 - H_2$ of ACC and the $H_1 - H_2$ of glottal flow [138], and additionally using the IBIF methodology allows us to calculate the equivalent $H_1 - H_2$ through the transformed glottal flow.

Therefore, we will use as equivalent to DPI in PVH population, the one constructed from the SPL adjusted by subject, through the linear calibration with

NSAM; and the $H_1 - H_2$ calculated from IBIF. The explicit expression for the DPI value reported in [16]:

$$DPI = \frac{1}{1 + e^{-(-4.195\text{NSAM}_{skew} - 1.254H_1 - H_{2std} + 6.729)}} \tag{2.44}$$

This value can be interpreted as the probability of be PVH, and is used to classify PVH (high DPI) or control (low DPI).

This regressor has been used to evaluate and monitor the evolution of patients during therapy and surgery [16, 17, 136]. along with being incorporated in ambulatory feedback schemes [58].

In addition to the DPI for PVH, a similar approach has been developed to monitor and classify patients with NPVH [18]. This non-linear logistic regression model was constructed using one week of ambulatory voice data from a population of 36 female patients with NPVH and 36 vocally healthy matched controls. A subset of 11 patients with NPVH was also monitored after voice therapy, allowing for the evaluation of changes in vocal behavior post-treatment. As in the case of PVH, the collection of ambulatory data focused on NSAM, f_o, CPP, and $H_1 - H_2$.

The final non-linear logistic regression model used two key voice parameters—CPP mean and $H_1 - H_2$ mode—to differentiate between NPVH patients and controls. The parameter space is presented in figure 2.16.

These features reflect important aspects of vocal fold closure and phonatory efficiency: the reduced CPP in NPVH patients indicates less periodic voice sig-

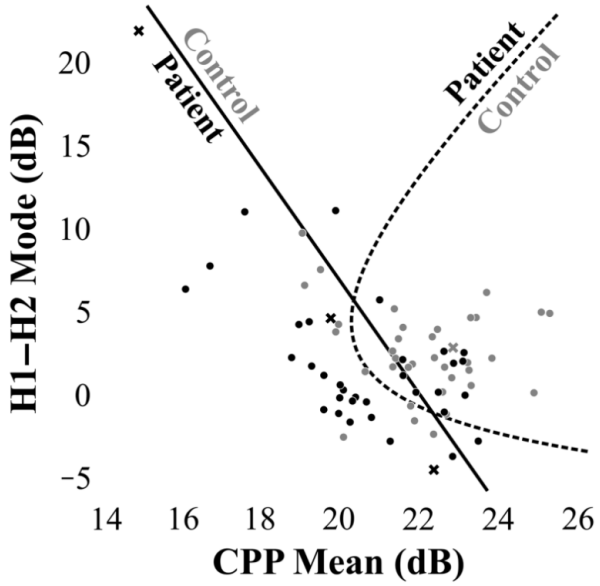

Figure 2.16: Scatter plot of NPVH regressor input space: the mode of $H_1 - H_2$ on the y-axis against the mean of CPP on the x-axis. Patients with NPVH are represented as black crosses, while matched controls are represented as gray circles. Each point represents the weekly distribution parameters of a single patient. The linear logistic regression cutoff is depicted as a black diagonal line, and the non-linear logistic regression cutoff is depicted as a dashed black line. Adapted from [18].

nals, while the less positive $H_1 - H_2$ skewness points to a less abrupt closure of the vocal folds.

Although the ambulatory component of this thesis focuses primarily on PVH and its associated DPI, the NPVH regression model is briefly introduced here for completeness. It serves to highlight how similar methodologies can be applied to other forms of vocal hyperfunction. Both models share similarities in

data collection and feature extraction, yet they target distinct pathophysiological mechanisms, providing complementary insights into vocal fold behavior across different hyperfunctional conditions.

2.3 Chapter conclusions

This chapter provides the theoretical foundation for phonation modeling, identifying VFs, glottal flow, and vocal tract models as critical components for understanding voice dynamics [60, 139]. Among the various modeling approaches, lumped-element models stand out due to their efficiency and physiological relevance in representing VF behavior [47, 45]. These models allow for accurate phonation simulation with reduced computational cost, which is essential for exploring latent variables like muscle activation that are difficult to measure directly in clinical settings [95, 98, 57].

In the study of vocal hyperfunction, significant progress has been made in modeling VF pathologies, such as nodules, yet there remains no consensus on the precise mechanical properties of swelling in PVH [53, 7, 52, 49, 48, 56]. Moreover, while models that incorporate VF asymmetries have been proposed for NPVH, these models tend to focus on biomechanical parameters rather than control variables like muscle activation, leaving open questions about how these asymmetries are directly linked to dysphonia [54, 55, 51, 50].

The TBCM emerges as a robust and versatile model for phonation due to its detailed representation of the five intrinsic muscles, allowing for accurate simulations of VF dynamics [44, 42]. This model has been applied to characterize specific data in Kalman filter schemes, fitting flow and glottal area signals in individual subjects [40], and to generalize patterns across larger populations, serving as a foundation for regressors of vocal characteristics using data from multiple subjects [33, 140]. Its capacity to capture complex physiological dynamics with low computational demands makes it an ideal choice for studying muscle activation and its role in vocal hyperfunction.

In terms of ambulatory monitoring, the DPI has proven to be an effective tool for classifying and tracking patients with PVH in clinical settings [19, 16, 17, 58]. However, further research is needed to determine how ambulatory-monitored variables, such as muscle activation or subglottal pressure, can accurately reflect underlying physiological mechanisms [34], and whether these differences can be correlated with estimates from VF models.

The following chapters will address these unresolved questions, investigating how ambulatory data can enhance the understanding of vocal function and proposing modifications to the existing models to better capture the observed physiological phenomena.

Chapter 3

Ambulatory analysis using an existing vocal fold model

This chapter explores how the TBCM can be integrated with ambulatory data for the estimation of underlying physiological parameters, aligning with the first aim of this research. In the previous chapter, the DPI, an ambulatory monitoring tool for PVH, was described. While its applicability in tracking the effects of therapy and surgery is clear, it remains to be explored how this regressor, specifically its feature space, relates to the physiological parameters that control voice production. Building on this, a sampling-filtering scheme is introduced in this research, which combines features obtained from VF model simulations with the ambulatory features that comprise the DPI. This approach proposes a Monte Carlo sampling method for the VF model simulation, a phonetically balanced rule for selecting the vocal tract configuration, and an inverse mapping procedure to mimic ambulatory distributions using model simulations. Additionally, a statistical analysis is conducted to examine whether the statistical differences observed in the DPI feature space persist in the space of physiological variables. In this

context, this chapter presents the description of model simulations, ambulatory recordings, the inverse mapping procedure, and the statistical analysis in both a patient-control paradigm and based on changes in DPI. The first version of the method and preliminary results were presented at AQL 2021, and the current version of the method and results detailed in this chapter have been submitted for publication in a peer-reviewed journal (JSLHR).

3.1 Connect models and ambulatory data

Historically, monitoring has focused on the characterization of vocal intensity, fundamental frequency, and vocal use (vocal dose) due to the association between PVH and excessive effort, typically resulting in elevated pitch and intensity [6]. In addition, data collection and reporting are often performed from a statistical perspective, such as means, deviations, ranges, and other quantities that describe the distribution of data, thanks to the volume of recordings that can be collected throughout the day [130, 129, 131, 132, 128].

In the context of voice disorder classification, an approach named the DPI has been adopted to assess significant differences in daily vocal behavior between PVH patients and matched controls [58, 16, 19, 18]. This index has been proven to capture vocal use patterns that distinguish PVH patients from controls, including characteristics such as SPL [19], NSAM [16] and variability in $H_1 - H_2$ [15, 19].

The question that emerges when employing this tracking tool is how the reported differences in aerodynamic measures, such as sound pressure level and harmonic variability, correlate with the underlying physiological mechanisms, such as laryngeal tension and subglottal pressure. To facilitate the interpretation of the DPI as a statistical classifier, it becomes crucial to establish a connection between aerodynamic characteristics obtained through ambulatory monitoring and the analysis of the physical processes inherent in normal and disordered voice production. Several studies in the literature have estimated subglottal pressure using different types of regressions or empirical formulas involving aerodynamic quantities [141, 33, 29, 133, 28, 142, 32], with the subglottal pressure value measured in the /pa/ gesture serving as ground truth [143]. The ACC has been used as a sensor in several of these estimation works [144, 33, 29, 30, 145, 146], showing its value as an instrument for obtaining vocal function. Recent work has taken this sensor to the ambulatory regime, taking estimates of subglottal pressure distributions using all-day recording [34], by linear regressions of subglottal pressure using parameters acquired from the ACC and calibrated in the laboratory, thereby generating estimates with ambulatory data.

In this study, we propose to bridge the gap between ambulatory aerodynamic distributions and their physiological counterparts by means of mathematical models of voice production. These models allow for the simulation and study of interactions between anatomical, biomechanical [46], and acoustic variables involved

in phonation [47, 45]. For this purpose, a fully interactive implementation of the TBCM [44] is utilized a voice production model for simulating different phonatory gestures that can be related to typical and hyperfunctional [44, 42, 100]. In addition, this model has also been successfully contrasted with clinical data in the context of the estimation of laryngeal muscle activations and subglottal pressure using glottal area and glottal flow observations [40, 33], thus making it an ideal model for the purpose of this study.

This chapter has two aims. First, we will replicate patient and healthy control ambulatory distributions underlying the DPI ($H_1 - H_2$ and SPL) with a TBCM to estimate differences in muscle activation levels and subglottal pressure. It is hypothesized that subglottal pressure, muscle activation levels, and adduction/ glottal closure will be higher in the patients than controls. Second, we will replicate five ambulatory distributions along the severity continuum of the DPI (as a probability). We expect the differences in muscle activation and subglottal pressure, between patients and controls from the first aim, to increase or decrease along with the DPI. To validate these hypotheses, we will statistically compare the resulting distributions and features between controls and patients with paired t-tests and Cohen's d values [147]. This study represents a synergistic approach, integrating ambulatory data from the PVH population, mathematical models of phonation (using the TBCM), and statistical analysis tools (using DPI) to identify the underlying elements of vocal hyperfunction in daily life.

3.2 Inverse mapping scheme

The proposed methodology for linking ambulatory data distributions with the speech production model is illustrated in Figure 3.1. This figure encapsulates the convergence of the three pivotal elements required to generate subglottal pressure and muscle activation distributions based on SPL and $H_1 - H_2$ distributions. These foundational components, in more detail, encompass: the ambulatory data, the simulations from the model, and the set of guidelines devised to mimic different vowels in the ambulatory environment using the model, i.e., the inverse mapping scheme.

3.2.1 Experimental data for ambulatory distributions of PVH and control subjects

The ambulatory database employed in this study is obtained and described in detail in [19]. This dataset included recordings from 180 adult women—90 of whom had been diagnosed with vocal fold nodules or polyps, and another 90 who acted as matched control subjects. The selection process for pairing patients and controls was meticulous, with careful consideration of age, gender, and occupation, alongside the prerequisite of having no recorded history of voice disorders. The governing institutional review board from the Massachusetts General Hospital

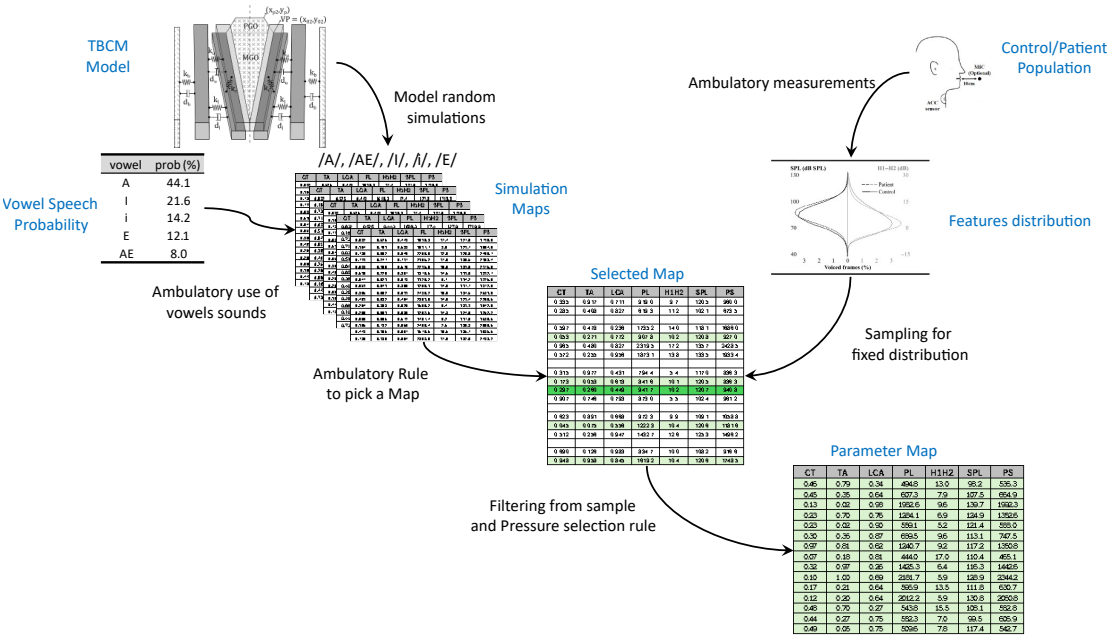

Figure 3.1: Diagram of the Inverse Mapping Scheme. The right track delineates ambulatory data and distributions from both PVH subjects and controls. The left track showcases data from TBCM simulations. At the center, the figure emphasizes the filtering of simulated data, grounded in established distributions.

approved all experimental aspects related to the use of human subjects for this study in the IRB protocol 2011P002376.

Over the course of a week, each participant was recorded using the Voice Health Monitor, as described in [134]. This device features a lightweight accelerometer (Model BU-27135, Knowles Electronics), placed on the anterior portion of the neck—just below the larynx and above the sternal notch—to capture phonation through the vibrations of the neck surface. In addition, a handheld microphone (H1 Handy Recorder, Zoom Corporation) was utilized for calibration purposes

once a day. For the ambulatory voice monitoring, we derived one-week summary statistics which included the mean, variability, skewness, and kurtosis of the recorded data. The selected voice features were SPL, f_o, and H_1-H_2. To convert the accelerometer signal into a glottal flow signal and subsequently quantify the aerodynamic features, we implemented the IBIF algorithm as described in [135] for each subject. The SPL signal was calculated using a subject specific linear regression to transform NSAM into estimated SPL values [134]. To generate the SPL and H_1-H_2 distributions, we processed the ambulatory data for each participant. This process involved the compilation of the first four statistical moments: mean, variance, skewness, and kurtosis. We utilized the 'pearsonnd' function in MatLab® to produce samples of size N that embodied these characteristics.

3.2.2 Simulations from phonation models

The ambulatory ACC data is known to be influenced by the subglottal and supraglottal systems [135]. To account for these complex interactions between tissue, airflow, and acoustic waves above and below the vocal folds we integrated a vocal tract model that simulates the propagation of acoustic waves in the time domain through both the subglottal and supraglottal systems using a wave reflection analog scheme [100]. A schematic representation of the TBCM is shown in the top left corner of Figure 3.1.

The model simulates sustained phonation segments of 0.5s duration with only

one vowel (i.e., supraglottal vocal tract) and vocal register (i.e., TBCM) per trial. However, the ambulatory data consists of spontaneous speech in daily life (i.e., many different vowels) and multiple vocal registers. Thus, to achieve different vowels and multiple modal registers (variations in pitch, loudness, and phonatory quality), the model uses input time-varying parameters such as the supraglottal tract cross-sectional area for a given vowel, normalized muscle activation levels (a_{CT}, a_{TA}, a_{LCA} and a_{PCA}), and lung pressure (P_L). A truncated Taylor-series approximation is implemented to simultaneously solve the differential equations of motion for the three masses that compose the body-cover structure, using a sampling frequency of 44.1 kHz. The vocal tract area functions are also set according to [148] to mimic several vowels. For each simulation, the output signals of interest include the subglottal pressure, oral airflow, and radiated sound pressure. The f_o and $H_1 - H_2$ features are computed from the glottal airflow signal; the P_S is derived as the mean of subglottal pressure signal, and the SPL is obtained from the radiated pressure. Notably, the f_o, $H_1 - H_2$ and SPL features are synonymous with those derived from the individual ACC data and the IBIF algorithm.

To make the model output tractable, we focused on the five most commonly used vowel sounds in American English: /I/, /i/, /e/, /AE/, and /A/. We performed 500,000 simulations for each vocal tract configuration. These simulations varied the remaining input model parameters (a_{CT}, a_{TA}, a_{LCA} and P_L) using a random parametric sweep within the ranges specified in Table 1. This approach

ensured a dense and representative exploration of the model's input space. Only simulations that produced sustained oscillatory behavior were retained, resulting in a total synthetic dataset comprising 1.5 million simulations. Each individual simulation was conducted with constant values for muscle activation and lung pressure. After 1 seg of simulation, the last 100 ms were used to compute the f_o, $H_1 - H_2$, P_S and the SPL. To streamline the parameter space, the activation level of the posterior cricoarytenoid muscle (a_{PCA}) was consistently set to 0. We constructed five vowel maps, each representing a distinct vocal tract configuration. These maps consist of comprehensive sets of unique input vectors (a_{CT}, a_{TA}, a_{LCA} and P_L) and their corresponding output vectors (SPL, $H_1 - H_2$, P_S) for each specified vocal tract shape, as illustrated in Figure 3.1.

3.2.3 Inverse mapping procedure

The concept of inverse mapping, in the context of this research, refers to the process of starting from observable acoustic features, such as SPL and $H_1 - H_2$, and using them to estimate the underlying physiological parameters that produced those features. Unlike forward modeling, where model inputs are used to predict outputs, inverse mapping takes the observed outputs (in this case, acoustic features) and works backward to infer the corresponding input parameters of the VF model. This approach bridges the gap between real-world ambulatory data and the physiological conditions responsible for voice production, by mapping

Input parameter	**Acronym**	**Range**
Lung Pressure	P_L	500- 2500 Pa
CT muscle activation	a_{CT}	0-1
TA muscle activation	a_{TA}	0-1
LCA muscle activation	a_{LCA}	0-1
IA muscle activation	a_{IA}	$a_{\mathrm{IA}} = a_{\mathrm{LCA}}$
PCA muscle activation	a_{PCA}	0
Vocal tract	VT	/I/, /i/, /e/, /AE/, /A/

Table 3.1: Input parameters for TBCM simulations.

the acoustic measures back to the model's control variables, such as subglottal pressure and muscle activation.

In this work, inverse mapping is implemented as part of the sample-and-filter method. This method involves aligning the distributions of SPL and $H_1 - H_2$ extracted from ambulatory data with the distributions generated by VF model simulations, thereby allowing us to uncover the physiological conditions that reproduce the observed acoustic patterns.

The integration of experimental and simulated data forms the cornerstone for deriving parameter distributions through this inverse mapping strategy. The process consists of four key steps. First, we pick a point from the SPL and

$H_1 - H_2$ distributions extracted from a participant's ambulatory data. Second, we use a sampling rule that assigns a probability to selecting each vowel, which is determined by the relative occurrence of each vowel in natural spoken English [149, 150]. The probability of selection for each vowel was: 44% for /A/, 21% for /I/, 14% for /i/, 12% for /E/ and 8% for /AE/. Third, from the selected vowel map, we identify the vectors that have SPL and $H_1 - H_2$ values that approximate the points taken from the ambulatory data distributions. The tolerance levels for identifying approximate vectors were 0.3 dB for SPL and 0.2 dB for $H_1 - H_2$ when calculating the distance to the chosen point. These tolerance levels were defined based on the 1% width of the distribution of values of each feature. Fourth, within the subset of filtered results, the final input-output pair that reproduces the sampled point is randomly selected. We iterated this process 20.000 times to acquire a collection of vectors, which not only aligned with the subject-specific feature distributions (SPL, $H_1 - H_2$), but also revealed the model parameter values responsible for producing those features. If there was an individual veridical data point that was not approximated by any SPL - $H_1 - H_2$ combinations, these data points were not included in the final simulated distribution.

3.2.4 Statistical analysis

Similar to the approach used in previous work [16, 18], we compared distributions of SPL, $H_1 - H_2$, muscle activations, and subglottal pressure between

patients and controls (aim one).. The analysis included central tendency and dispersion metrics—mean, median, standard deviation (SD)—as well as measures of distribution shape, namely skewness and kurtosis, for each feature within each subject data set (SPL, $H_1 - H_2$, a_{CT}, a_{TA}, a_{LCA}, and P_S), to total 30 features per subject.

To analyze the matched control-patient paradigm (90 pairs), paired t-tests (parametric data) and Wilcoxon signed-ranks tests (nonparametric data) were used to evaluate differences in 30 resulting features. A Kolmogorov-Smirnov (KS) test assessed the normalcy of each distribution of paired differences (patient minus control). If the KS test was significant (p ¡ .05), a Wilcoxon test was applied; if not, a paired t-test was used. Due to the high number of tests, the alpha significance level was adjusted using the Bonferroni method. Upon finding statistical significance, the Cohen's d effect size was calculated to characterize the magnitude of differences between groups (small when ≤ 0.19, small to medium when 0.20–0.49, medium to large when 0.50–0.79, and large when ≥ 0.80 [147].

For aim two, we examined parameters that were significantly different between the control-patient pairings across five DPI severity ranges. A DPI value was determined for each participant using $H_1 - H_2$ SD and SPL skewness in the calculation first described by [19]. For the analysis, five DPI groups were defined, each representing a probability range of 0.2 within the overall DPI scale ($0 <$ DPI < 1). This ensured that each group had the same probability range and a similar

number of subjects (approximately 36 $\pm$ 2 per group). The average distributions for each parameter were plotted across the incremental levels of DPI using box plots, thus allowing the visualization of any relationships between a parameter and different DPI values.

3.3 Results

3.3.1 Inverse mapping procedure

The inverse mapping procedure was successfully applied for all study subjects, resulting in model-produced SPL and $H_1 - H_2$ distributions that closely matched the distributions based in the ambulatory SPL and $H_1 - H_2$ data. Success in this context was defined as matching a significant proportion of the data points, with a threshold of approximately 60% considered reasonable for successful alignment. Notably, the actual success rate of the inverse mapping, measured by the proportion of matched points between ambulatory and model-derived SPL and $H_1 - H_2$ pairs, was 77.8% ($\pm$6.1%), exceeding our threshold for success. This achievement is particularly noteworthy considering that the numerical model approximates the complex phonatory process, yet it replicated data across multiple individuals over three-quarters of the time. This high success rate highlights a substantial congruence between the model outputs and the real-world voice data regarding these key voice features.

Figure 3.2 shows an example with feature distributions for one selected subject, both ambulatory and model derived. The upper panels show the SPL and H_1-H_2 distributions from both ambulatory data and the inverse mapping, illustrating a robust alignment between the observed and modeled data. The lower panels delineate the distributions of subglottal pressure and muscle activation inferred from the model. The mode of subglottal pressure is observed at approximately 0.7 kPa, with a mean around 0.9 kPa, corroborating with previously reported "comfortable loudness" measurements in controlled environments [10, 11].

Regarding muscle activations, the distribution of the LCA muscle in Figure 3.2 (D) shows a low or nearly zero occurrence of activity below 0.2. This is consistent with the essential role of LCA activation in producing phonation, given its adductor function [36, 151]. Conversely, the CT muscle distribution exhibits a higher prevalence of activation around 0.3, a range typically associated with a "speech" type of voice [152]. The CT muscle also maintains activation across its entire range, reflecting the variability in vocal frequency. Finally, the TA muscle distribution shows activity throughout its range, with a preference for high activation levels. This trend should be interpreted in light of the inverse mapping and filtering scheme applied in this study. Previous experiments and regression analyses [36, 33] have not demonstrated a strong correlation between TA activation and commonly used acoustic or aerodynamic parameters, resulting in a high degree of uncertainty in its estimation. Additionally, high TA activation is known

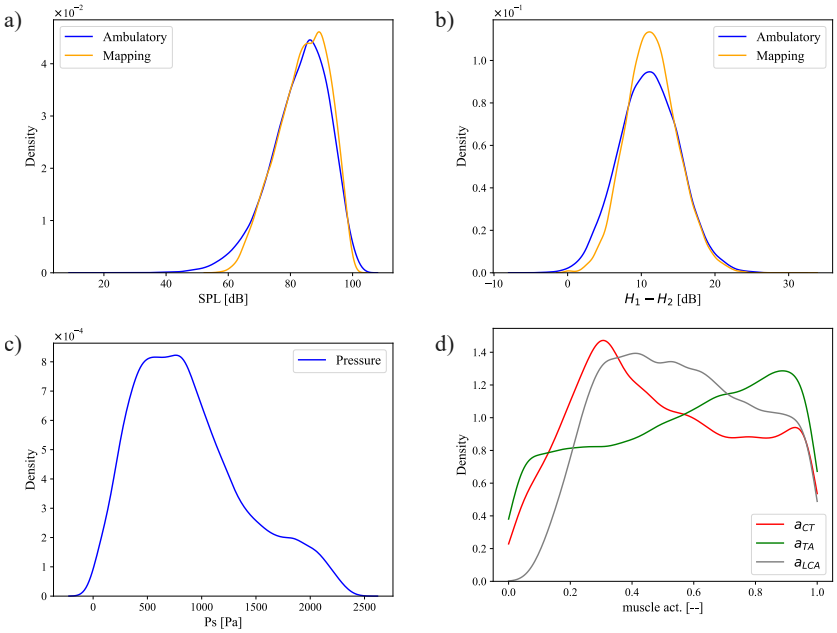

Figure 3.2: One subject distribution of model parameters and features, after applying Inverse Mapping. Acoustic features: (a) SPL: Ambulatory (blue) and inverse mapping (orange) distribution. (b) $H_1 - H_2$: Ambulatory (blue) and inverse mapping (orange) distribution. Model parameters: (c) Subglottal pressure (P_S) distribution. (d) Muscle activation: a_{TA} (red), a_{CT} (green) y a_{LCA} (gray) distribution.

to produce mid-frequency phonation, while low activation can inhibit vocal fold oscillation in certain configurations [36]. These dynamics influence the inverse mapping process, leading to a tendency to select more simulations with high TA activation. These observed trends are consistent across other subjects. Moving forward, we transition to a more population-based analysis to further explore these findings.

3.3.2 Statistical comparison of the PVH and control groups

As outlined in the inverse mapping scheme, measures of central tendency and statistical moments (mean, median, SD, skewness, and kurtosis) were calculated for each subject's distributions (SPL, $H_1 - H_2$, a_{CT}, a_{TA}, a_{LCA}, and P_S), resulting in a total of 30 parameters. Table 3.2 lists the parameters means and SD that statistical testing (KS test, Wilcoxon sum statistic, and Cohen's d) showed were significantly different between the PVH and control groups. Significant differences were found for the two parameters that comprise the DPI ($H_1 - H_2$ SD and SPL skewness), several distributional parameters for LCA muscle activation (mean, median, SD), the mean for CT muscle activation, and the median for P_S Not surprisingly, the two DPI components had the largest values of Cohen's d with $H_1 - H_2$ SD having a large effect size (d=1.12) and SPL skewness having a medium to large effect size (d=0.63). The distributional features of LCA muscle activation showed medium effect sizes (|d|=0.53 to 0.58) and both CT muscle activation (|d|=0.58-0.43) and P_S had small to medium (d=0.4) effective sizes. There were no significant parameters for TA muscle activiation. Conversely, the significant parameters associated with the CT and LCA muscles underscore their correlation with elements of harmonic richness ($H_1 - H_2$). The LCA muscle, as the primary adductor muscle, is crucial for vocal fold closure and posture, while the CT muscle, as the vocal fold tensor, plays a key role in pitch coordination.

Feature	Control (mean ± SD)	Patient (mean ± SD)	KS test (p-value)	Wilcoxon (p-value)	Cohen's d
$H_1 - H_2$ SD	3.77 ± 0.33 dB	3.37 ± 0.38 dB	<0.01	<0.01	1.12
SPL skew	0.01 ± 0.24	-0.15 ± 0.22	<0.01	<0.01	0.63
a_{LCA} mean	0.57 ± 0.01	0.58 ± 0.02	<0.01	<0.01	-0.58
a_{LCA} median	0.56 ± 0.02	0.58 ± 0.02	<0.01	<0.01	-0.54
a_{LCA} SD	0.24 ± 0.01	0.23 ± 0.01	<0.01	<0.01	0.53
a_{CT} mean	0.50 ± 0.02	0.49 ± 0.02	<0.01	<0.01	0.43
P_S median	1.11 ± 0.15 kPa	1.17 ± 0.16 kPa	0.02	<0.01	-0.40

Table 3.2: Features that showed significant differences between controls and patients with PVH, with their p-values and effect sizes.

Additionally, the close relationship between P_S and SPL is further confirmed.

3.3.3 Relationships between significant parameters and DPI

To elucidate the relationship between the distinctive features outlined in Table 3.2 and the DPI, each subject was assigned a DPI score based on their specific $H_1 - H_2$ variability and SPL skewness values [16].

Figures 3.3 - 3.5 display boxplots of the statistically significant parameters (see Table 3.2)) shown across increasing values for DPI. Higher DPI values are

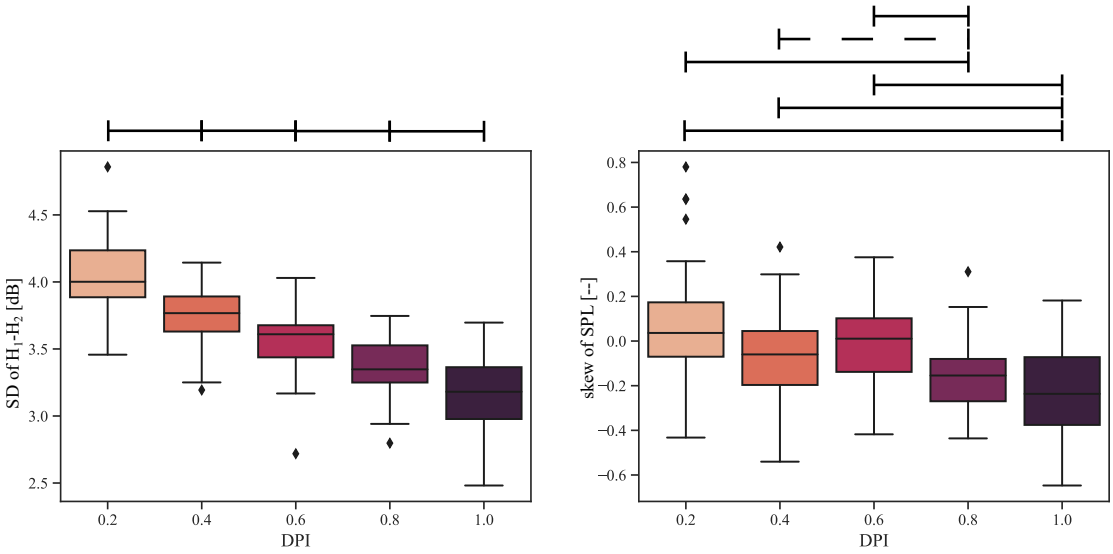

Figure 3.3: (Left) Standard deviation of $H_1 - H_2$. (Right) Skewness of SPL. Colors represent the DPI group (groups of 0.2). Solid (99%) and dash (95%) line represents statistical difference between groups.

generally viewed as reflecting more severe PVH. The solid (99%) and dashed (95%) lines in the head of the boxplots represent statistical difference between groups. All of the displayed parameters showed significant changes when comparing their distributions at a DPI value of 0.2 with those at both 0.8 and 1.0, demonstrating the ability to discriminate between higher control-patient contrasts.

Not surprisingly, the boxplot in Figure 3.3, for $H_1 - H_2$ SD and SPL skewness display significant changes in these two components of the DPI across the entire range of DPI values with the decrease in $H_1 - H_2$ SD being even more consistent than the increase in negative skewness of SPL. These results serve as a simple sanity check.

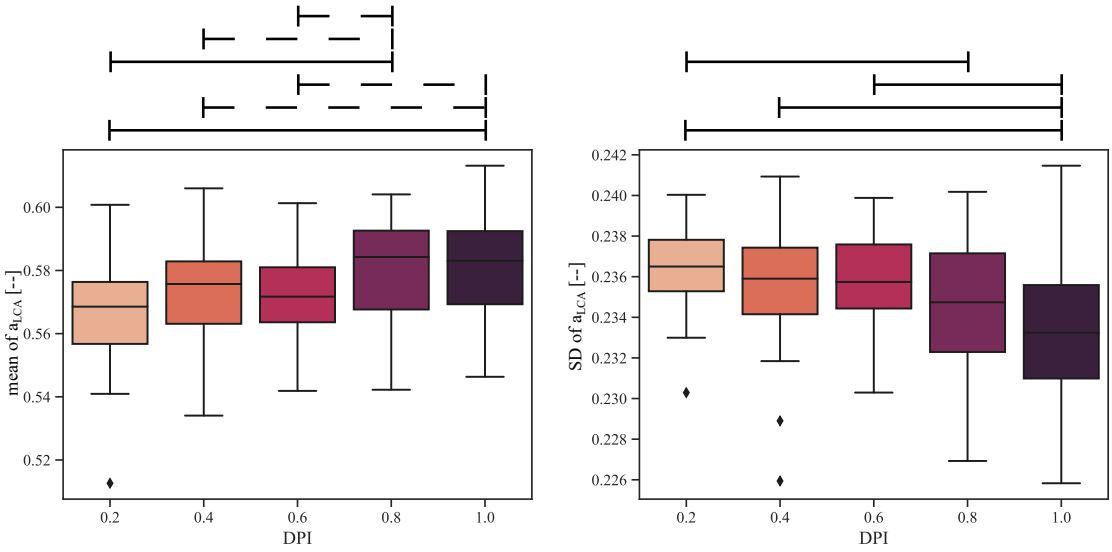

Figure 3.4: (Left) Mean of a_{CT}. (Right) Median of Ps. Colors represent the DPI group (groups of 0.2). Solid (99%) and dash (95%) line represents statistical difference between groups.

Boxplots for the model-derived mean and SD of LCA muscle activation are shown in 3.4. In general, the mean LCA activation level increases significantly and the SD of the LCA activation level decreases significantly with increases in DPI. This can be interpreted as a more vocal fold closure on average and reduced variation towards less vocal fold closure (respectively) as the DPI increases.

Figure 3.5 displays the model derived values for mean CT muscle activation and the median P_S. Overall, there is a significant but inconsistent decrease in the mean CT muscle activation level and a significant but inconsistent increase in median P_S with increases in DPI. Indicating higher aerodynamic forces when the DPI increases.

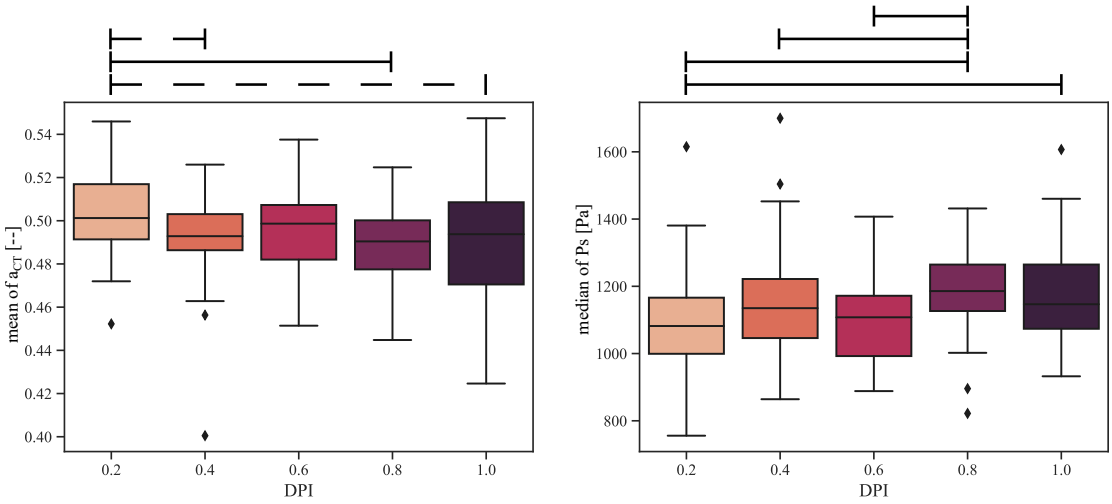

Figure 3.5: (Left) Mean of a_{LCA}. (Right) Standard deviation of a_{LCA}. Colors represent the DPI group (groups of 0.2). Solid (99%) and dash (95%) line represents statistical difference between groups.

Figure 3.6 illustrates the average distribution for both ambulatory and model-based parameters across the two DPI groups: DPI > 0.5 (classified as patients) and DPI < 0.5 (classified as controls), summarizing the results from previous figures. The shaded areas represent the 95% confidence interval, and the gray arrows highlight the direction of increasing DPI scores.

For ambulatory parameters, the $H_1 - H_2$ distribution, Figure 3.6 (A), becomes more concentrated maintaining its tails while the central density heightens, thus diminishing the standard deviation when the DPI value increases. Concurrently, SPL distribution, Figure 3.6 (B), shifts toward elevated SPL values, indicating a skewness alteration. These elements represent what is observed in Figure 3.3

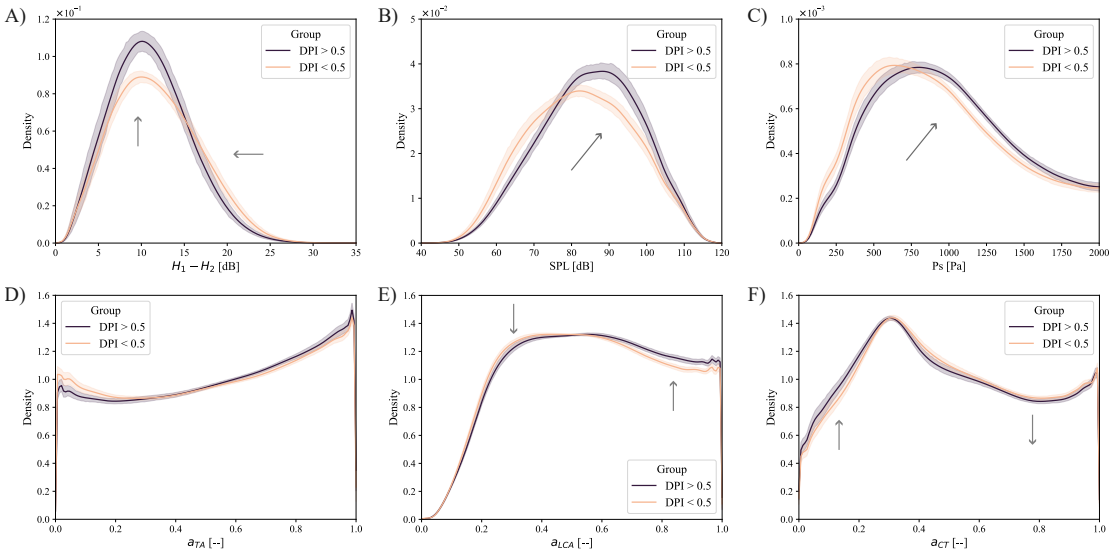

Figure 3.6: Mean distributions of parameters of interest. Acoustic features: (A) $H_1 - H_2$ and (B) SPL distribution. Model parameters: (C) P_S, (D) a_{TA}, (E) a_{LCA} and (F) a_{CT} distribution. The shaded area represents the 95% confidence interval. Gray arrows indicate the direction of DPI increase.

in terms of global parameters and not just their statistical moments. For model-derived parameters, P_S follows a similar trend to SPL, reflecting their correlation, while LCA activation increases with higher DPI scores, as seen in Figure 3.4. CT activation decreases as DPI increases, and TA activation shows no significant variation across DPI categories.

3.4 Discussion

The work presented in this chapter had two hypotheses: (1) Subglottal pressure and muscle activation levels related to adduction/glottal closure will be higher in the patients than controls; (2) The differences in muscle activation and subglottal pressure between patients and controls, from the first hypothesis, will increase or decrease along with the DPI. To evaluate any muscle activation or subglottal pressure values, the TBCM with subglottal and supraglottal tracts needed to replicate distributions of SPL and $H_1 - H_2$. Notably, there was high agreement between the vocal production model predictions and actual ambulatory measurements of SPL and $H_1 - H_2$. The strong consistency between ambulatory data and model predictions lends some validation to using the TBCM and associated lookup tables for estimating physiological parameters from the neck skin acceleration signal. To match real life SPL and $H_1 - H_2$ distributions, the TBCM and lookup tables must accurately reflect the physical processes of vocal production to a nontrivial degree. Regarding muscle activation, the distributions reflect a prevalence of lower activations (below 0.5) for the CT muscle and a predominance of higher activations (above 0.4) for the TA and LCA muscles, replicating the modal phonation pattern described in our previous TBCM paper [44].

The largest differences in estimated physiological measures were patients showing increased average LCA values (d = 0.54-0.58) and decreased LCA variability

(d = 0.53) compared to controls. Excess average LCA activation would adduct the vocal folds more forcefully during voicing. It is thought that normal voicing occurs with the vocal folds barely touching [153]. Patients with PVH increase their risk of phonotrauma by over-adducting their vocal folds and/or voicing in ways that increase tissue-to-tissue contact [154, 146]. Additionally, these patients have lesions in the mid-membranous vocal fold, which would prevent full vocal fold closure during voicing, and further necessitate increased LCA muscle activation/vocal fold adduction to produce clear voicing [6]. Reduced LCA variability corresponds to previous work with the DPI, showing that patients voiced with decreased variation to higher values of $H_1 - H_2$ daily life [19, 16, 17] reflecting a tendency to maintain more abrupt/complete glottal closure. Finally, average LCA increased and LCA variability decreased as the SPL and $H_1 - H_2$ distributions underlying the DPI were classified as more severe, further supporting our interpretation of how LCA muscle activation relates to PVH pathophysiology.

The patients exhibited diminished average CT muscle activation compared to controls, pointing to lower pitches on average. Patients with PVH often voice with reduced pitch variation towards higher pitches and multiple voice therapy approaches incorporate pitch glides and exaggerated prosody to address this observation [155, 156, 157]. Empirically, previous studies identified significantly reduced fo standard deviation in patients' daily life compared to matched controls [12, 128, 16, 19, 17]. This difference normalized after voice therapy, not after sur-

gical removal of the lesions, insinuating that it may be primarily behavioral. The relationship between mean CT activation and the DPI showed that decreased CT activation related to increased potential for phonotrauma [16, 17]. This relationship is counterintuitive because traditional vocal dose measures often rely on the opposite relationship: higher cumulative collision forces as fo (i.e., CT activation levels) increases [158, 133]. However, vocal frequency is known to be motorically redundant as multiple muscles combine to produce different pitches [95]. Perhaps voicing at a specific frequency with less CT muscle activation requires laryngeal configurations with higher potential for phonotrauma, i.e., increased LCA muscle activation.

There was no notable difference in TA muscle activation between patients and controls, as well as minimal changes in TA activation at high DPI levels. While multiple theories underlying PVH hypothesize general increases in intrinsic laryngeal muscle activation levels, the results suggest that PVH is an imbalance among these muscles. Specifically, the imbalance may manifest as a "typical" level of TA activation, "higher than typical" level of LCA activation, and "lower than typical" level of CT activation. Similar imbalances have been found in patients with VH using sEMG intermuscular coherence between the left and right anterior neck muscles [159, 160]. Future research should explore the imbalances between the left and right activation levels of the same muscle. The results show that the proposed methodology, estimate using just SPL and $H_1 - H_2$ distributions, fails

to constrain the behavior of the TA muscle, beyond the uniform distribution of the synthetic data and the prevalence of more simulations with high activation. Therefore, further studies will focus on developing new features that more accurately correlate with its activity. This will lead to more reliable estimates and a deeper understanding of its role in hyperfunction.

Patients exhibited increased average P_S compared to the controls, likely reflecting the increased effort/forces necessary to voice with excessive phonotrauma and/or in the presence of phonotraumatic lesions. Previous work has shown that, compared to controls, patients spend more time at higher amplitudes—i.e., increased negative skewness of the neck skin acceleration amplitude. Perhaps higher overall average subglottal pressures are physiologically necessary to maintain a negatively skewed vocal intensity behavior in daily life. Supporting this interpretation, there was a direct relation between increases in DPI values and increases in P_S.

There are multiple limitations of this study. We acknowledge that the model simulations are bound by a certain level of granularity, as evidenced by a 77% prediction accuracy for SPL and $H_1 - H_2$ values. One potential way to increase the prediction accuracy in future studies includes skipping the transformation of the ambulatory ACC signal into the model glottal airflow signal using the IBIF. While the IBIF transformation allows the ACC to be included in the model, it can introduce noise into the feature calculation [161, 162]. However, the ACC

has been shown to exhibit a strong relationship with airflow features in previous research [11, 138], justifying its inclusion, even with the use of IBIF. Future improvements could explore alternative methods for incorporating the ACC signal into the model, potentially by developing a dedicated model for ACC within the overall phonation model framework. Refining the model resolution could potentially improve its predictive fidelity. Despite this shortcoming, our model shows promising generalizability, simulating vocal features across a varied population with a fair degree of success. The complexity of inverse mapping – deriving muscle activations and pressures from SPL and $H_1 - H_2$ readings – remains a challenging endeavor due to the non-uniqueness of the solution. Our model currently mitigates this challenge through a stochastic process in sample selection. In this study, we aim to delineate the scope of the DPI space estimation and demonstrate how it characterizes PVH. Looking forward, we aim to enhance our methodology by integrating more measured variables to tighten the range of potential solutions and by expanding our simulation library. Future research efforts will be directed towards adopting more sophisticated approaches, including interpolation methods, and leveraging machine learning techniques such as neural networks, to refine parameter estimation and tackle the issue of non-uniqueness with greater precision.

It is important to note that we are currently using a 'healthy' baseline model to estimate parameters in pathological cases. While this might seem limited, it

serves as a valid first approximation because the TBCM has already demonstrated its capacity to capture some of the pathophysiological characteristics of PVH, as evidenced in previous studies [42, 100]. Using a consistent model across both control and PVH patients ensures that we are measuring all subjects with the same set of criteria, creating a fair comparative framework.

However, future work will need to extend beyond this baseline model. More comprehensive simulations could be developed by introducing additional features that better characterize the underlying pathological states. We aim to incorporate a model that more accurately represents PVH by capturing the nuanced effects of vocal fold lesions and nodules. This enhanced model could eventually be coupled with the DPI to establish a continuous scale for assessing the severity of pathology, where parameters associated with nodule characteristics, such as size and stiffness, would change progressively according to DPI values. This would allow for a more nuanced representation of PVH severity, rather than treating it as a binary condition.

3.5 Chapter conclusions

This chapter presented a proof of concept for integrating VF model simulations into an ambulatory monitoring context. The study successfully demonstrated that ambulatory data can be effectively combined with VF models to estimate

underlying physiological parameters, addressing a key objective of this research. It validated the use of the TBCM, achieving a 77.8% success rate in replicating ambulatory SPL - $H_1 - H_2$ data.

The findings emphasize that patients with PVH show higher average activation of the LCA muscle and reduced variability in this activation, suggesting that increased adduction and decreased variability may play a crucial role in the pathophysiology of PVH. Additionally, reduced CT muscle activation was observed, correlating with lower pitch variability in patients, a result consistent with previous clinical observations, offering deeper physiological context to vocal hyperfunction.These results confirm Hypothesis 1 by demonstrating that specific acoustic and aerodynamic signal features are indicative of distinct physiological patterns, particularly reflecting differences in subglottal pressure and muscle activation between PVH patients and controls.

One notable limitation is the challenge in estimating TA muscle behavior. The DPI feature space does not adequately constrain TA muscle activation, even though the sample-filtering scheme performs well for other variables. This raises the point that traditional aerodynamic features, often used in regression approaches, may not correlate strongly with TA muscle activity, underscoring the need for further research and validation of TA behavior with ex vivo and in vivo data. Addressing this limitation will be a priority in future work. Additionally, the complexity of inverse mapping from aerodynamic to physiological parameters

introduces inherent uncertainties, as the solution space is non-unique. Refining the TBCM's simulation capabilities, including increasing simulation input resolution and adding more aerodynamic features in the inverse mapping procedure, will be essential to enhance the accuracy of physiological estimations.

It is worth noting that this study used a 'healthy' vocal fold model (TBCM) to estimate physiological parameters related to pathological conditions. While this serves as a reasonable starting point, particularly given the TBCM's capacity to represent certain aspects of PVH pathophysiology [42, 100], future work will focus on developing models that more accurately reflect the specific characteristics of vocal pathologies. The next chapters will present refined models tailored to represent the compensatory mechanisms and biomechanical features of PVH and NPVH, providing a more detailed framework for understanding these conditions.

This chapter highlights the potential of integrating VF models with ambulatory data to create a comprehensive framework for understanding vocal hyperfunction. This integration not only bridges the gap between ambulatory measurements and physiological insights, but also offers a path for improving clinical assessments through data-driven models.

Chapter 4

Model for vocal fold asymmetry

Considering the aims set out in this thesis, we now focus on developing new models to characterize elements associated with VH. Specifically, this chapter describes the construction of an asymmetric model of the VFs to characterize the biomechanical elements of NPVH, such as increased VF tension and muscular imbalance as key drivers of asymmetry in vocal fold oscillation and the vocal process [131, 112]. The proposed approach builds on a lumped-element model, using the TBCM as its foundation. Previous studies have addressed VF asymmetry by directly modifying the mass and spring components in similar models [50, 51]. Here, the focus is on generating differences between the left and right VFs through variations in laryngeal muscle activation, leveraging the TBCM's built-in muscle control model.

In this model, we also explore the concept of compensation, focusing on how a single parameter, such as muscle activation or subglottal pressure, adjusts to maintain a constant acoustic output (SPL and f_o) despite asymmetry in the VFs. Additionally, two configurations involving joint compensations are examined to

study how muscular imbalance is offset. These configurations are shared between NPVH and PVH models, offering a comparative framework. Finally, the impact of these compensations on other key parameters, such as MFDR and the ratio between collision pressure and subglottal pressure, will be analyzed to characterize the effect of the imbalance on VF collision dynamics.

This chapter provides a brief overview of the transition from the basic TBCM model to an adaptation involving two independent VFs, with the details of the modified equations and formulations provided in Appendix A of this thesis. It examines the effects of muscle imbalance on the viscoelastic properties and the vocal process, illustrated through synthetic and real kymogram graphs showing the edge oscillation of the folds. Within this framework, the chapter introduces a new model for VF asymmetry and explores the concept of muscle imbalance and its role in representing VF oscillation in cases of dysphonia and VF paralysis. The first version of the model and preliminary results were presented at AQL 2023. The current asymmetrical model, along with its equations in the appendix and part of the results detailed in this chapter, has been published in a peer-reviewed journal (JASA).

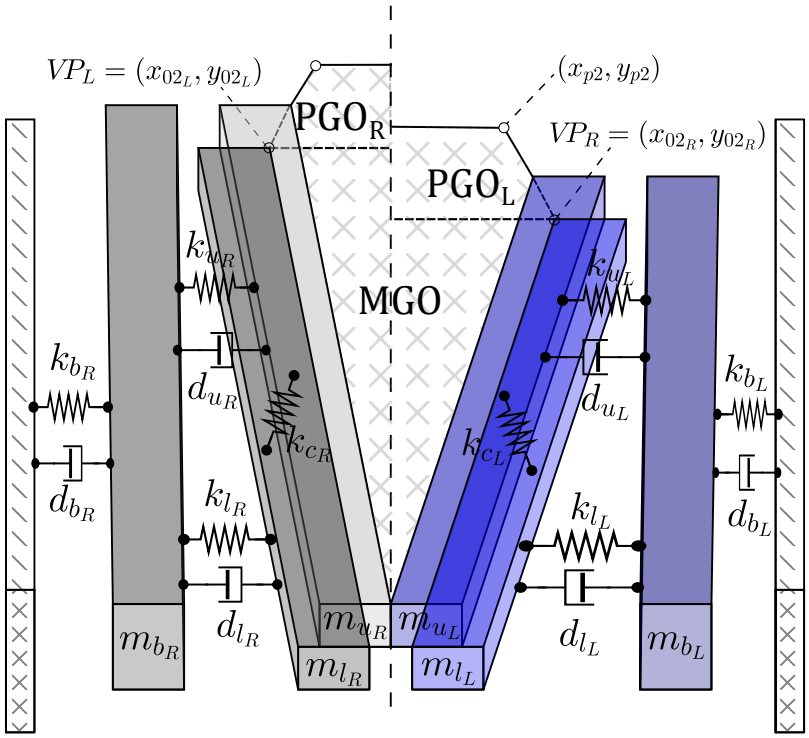

Figure 4.1: Schematic diagram of the VF configuration according to the a-TBCM for an abducted glottal configuration and asymmetrical muscle activation. The scheme shows: The vocal process (VP) defined by its position on the x and y axis (x_{02}, y_{02}). The position of the posterior wall (x_{p2}, y_{p2}). The posterior and membranous component of the glottal area, MGO and PGO. And the components of mass (m), spring (k) and damper (d); for the upper (u), lower (l) and body (b) blocks; for the left (L) and right (R) vocal fold.

4.1 Asymmetrical TBCM (a-TBCM)

The a-TBCM is constructed around the coupling between two independent muscle-controlled TBCMs, as illustrated in Figure 4.1, representing the left and

right VFs. In this scheme, two muscle activation vectors are defined:

$$\vec{a}_i = [a_{\mathrm{IA}}, a_{\mathrm{LCA}}, a_{\mathrm{PCA}}, a_{\mathrm{TA}}, a_{\mathrm{CT}}]_i\,, \quad i \in \{\mathrm{L},\, \mathrm{R}\}, \tag{4.1}$$

where the subscript i denotes the side of the a-TCBM in question, with L and R denoting left and right, respectively. Equal left/right activation vectors correspond to the symmetric control scheme in the TBCM. Then, the posture and the viscoelastic properties of each VF are defined by its own muscle activation. To model the dynamics of the VF, i.e. the movement of the body-cover elements, it is necessary to formulate and solve the coupled equations of motion for the left and right VF (see Appendix A).

4.2 Muscle imbalance for a-TBCM: Bilateral posture and viscoelasticity

Muscle imbalance refers to a slight difference in the activation of the laryngeal muscles between the left and right VF. These varying muscle activations serve as inputs to the control mechanism, influencing the mechanical and geometric properties of the VF. The proposed a-TBCM provides a platform to explore the impact of imbalanced activation in the intrinsic musculature on the geometric and viscoelastic properties, influencing VF oscillations and the aerodynamics of glottal airflow. To establish a reference activation condition for the intrinsic laryngeal

muscles, a simplification was introduced. The activation levels of the primary adductors (LCA and IA muscles) were combined by setting $a_{\mathrm{LCA}} = a_{\mathrm{IA}}$, streamlining the adjustable variables for sustained VF oscillations. Additionally, the PCA activation (a_{PCA}) was consistently set to zero across all simulations, neutralizing its significant abductor effect on glottal posture [44]. For the TA and CT muscles, activations a_{TA} and a_{CT} were calibrated to produce voiced sounds within the fundamental frequency range of 90 to 100 Hz, aligning with the physiological characteristics of a male modal voice. The activation levels were determined based on muscle activation plots reported in [44]. The reference activation for the right side is represented by:

$$\vec{a}_R = [0.6, 0.6, 0.0, 0.6, 0.2]\,. \tag{4.2}$$

To represent the imbalance during voiced phonation, an asymmetry factor $q \in [0.5, 1.5]$ is introduced between the left and right sides in the a-TBCM. A similar idea has been previously applied in [50], henceforth SH95. However, rather than controlling asymmetries in the mass-spring elements through a gain parameter as in SH95, the a-TBCM applies the asymmetry factor directly to the muscle activation levels. These levels jointly influence the posture and configuration of the VF. For comparison with other fold asymmetry approximations, the equivalent of SH95 for the TBCM model is implemented, where:

$$m_L = \frac{m_R}{q} \quad , \quad k_L = q k_R. \tag{4.3}$$

The modified masses and springs correspond to those presented in Figure 4.1, representing the upper, lower, and body blocks. Two primary scenarios of muscle imbalance were explored: an asymmetric activation predominantly affecting glottal adduction (vocal process) and an asymmetry in the biomechanical properties (mass-spring values) of VF. To represent the first case, the asymmetry factor q was applied to a_{LCA} and a_{IA} in the left VF. The muscle activation vector for the left VF is then given by:

$$\vec{a}_L = [q \cdot 0.6, q \cdot 0.6, 0.0, 0.6, 0.2]\,. \tag{4.4}$$

For asymmetries in the biomechanical properties, q was applied to a_{CT} and a_{TA} simultaneously, considering the CT-TA antagonistic relationship in determining the viscoelastic (VE) properties in the VF [95]. The muscle activation vector for the left VF in this scenario is expressed as:

$$\vec{a}_L = [0.6, 0.6, 0.0, q \cdot 0.6, q \cdot 0.2]\,. \tag{4.5}$$

Additionally, individual muscle imbalance in each CT and TA was studied separately, given its significance in characterizing the glottal configuration in the model, based on previous works [95, 98]. This resulted in a total of four configurations of muscular imbalance and the direct imbalance in masses and springs The simulations were performed with a driving pulmonary pressure of 1 kPa. A truncated Taylor-series approximation was implemented to simultaneously solve the differential equations of motion for the six masses, using a sampling frequency

of 44.1 kHz. The glottal tract was defined as the area function corresponding to an /i/ vowel [64], and a subglottal tract as an inverted cone shape [163, 100].

4.3 Measures of muscle imbalance effect

4.3.1 Model derived measures

Several measures were derived from a-TBCM simulations to establish differences between the muscle imbalance scenarios, study the effect of the asymmetric factor on VF vibratory asymmetries, and compare the model with certain clinical cases. These measures include left-right amplitude asymmetry (AA), left-right phase asymmetry (PA), f_o, and open quotient (OQ). Following [51], synthetic kymograms were generated from the right and left VF edge positions over time as illustrated in Figure 4.2. Based on these representations, measures of the VF oscillatory asymmetries can be defined:

$$AA = \frac{A_R - A_L}{W}, \tag{4.6}$$

$$PA = \frac{t_L - t_R}{OP}, \tag{4.7}$$

Note that the normalization factor OP for defining PA follows that of [51], where the focus is on the acoustic effects generated during the open phase. These

measures assess different aspects of oscillatory asymmetries. The left-right amplitude asymmetry, Equation 4.6 [164], is the ratio between the difference and the sum of the maximum excursions for the right (reference) and left (affected) VF amplitude traces within one vibration period. The left-right phase asymmetry, Equation 4.7 [165, 106], is the time delay between the maximum excursions of the left and right VF normalized by the oscillation period.

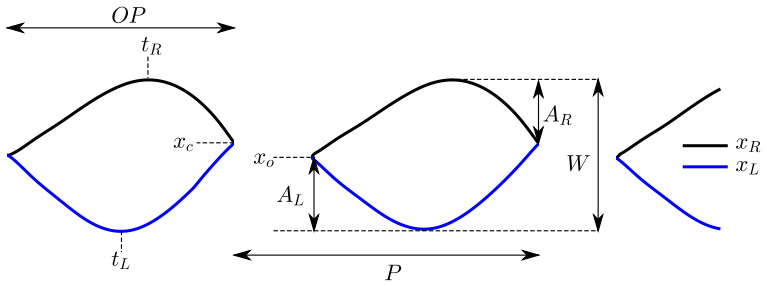

Figure 4.2: Kymogram parametrization: x_L, x_R displacement of left/right VF; t_L, t_R time of maximal left/right VF displacement; x_c, x_o mediolateral position of VFs at the glottal closure/opening; A_L, A_R maximum displacement of left/right VF; W peak-to-peak glottal width; P glottal period; and OP open phase.

The open quotient is also computed from the kymogram as the ratio of the duration of the open phase (OP) to the period (P) for each cycle.

$$OQ = \frac{OP}{P}, \tag{4.8}$$

4.3.2 Clinical cases

In conjunction with the modeled scenarios, this study draws support from in vivo high-speed video (HSV) recordings representing voice disorders. Two female participants aged 25 to 30 years, one with NPVH (muscle tension dysphonia) and another with unilateral VF paralysis, were included. During data acquisition, participants produced a sustained /i/ vowel sound at optimal pitch and loudness levels, adhering to informed consent procedures and Institutional Review Board (IRB) protocols outlined in FONDECYT 1151077 project.

The HSV data collection involved an advanced high-speed camera (SA-X2, Photron, Tokyo, Japan) synchronized with a rigid endoscope (9106, KayPentax, Montvale, NJ) equipped with a 35 mm C-mount adapter and a xenon light source (7152B, KayPentax, Montvale, NJ). The original HSV dataset, comprising approximately 2,670 frames at 8,000 frames per second (fps), underwent pre-processing. This included selecting 600 frames and identifying a 256 $\times$ 256 pixel region of interest containing only the VFs. The refined HSV dataset was then used to segment the glottis area using GlottalImageExplorer [166].

Subsequent to the glottis area segmentation, the digital kymogram (DKG) technique proposed in [51] was applied to visualize VF trajectories. These lateral displacement waveforms, capturing intricate motions of the left and right VFs, were traced along a one-pixel line positioned at the midpoint of the glottal re-

gion's posterior-anterior length. Additionally, the DKG outputs facilitated the automated extraction of amplitude and phase asymmetry metrics for the left and right VFs.

To represent female phonations, geometric quantities of the model were adjusted. Specifically, the resting length (L_g) of the vocal fold was reduced from 16 mm to 10 mm, the resting thickness (T_g) from 3 mm to 2 mm, the ligament depth from 20 mm to 15 mm, the mucosal depth from 20 mm to 15 mm, and the muscle depth from 40 mm to 30 mm. Additionally, the model was coupled with a female /i/ tract, utilizing the area function outlined in [64].

4.4 Compensation mechanism in muscle imbalance

To investigate how muscle imbalance is compensated in the VFs, we designed a simulation framework to study the compensation mechanisms for intensity (SPL) and pitch (f_o). The simulation begins with a reference configuration for SPL and f_o, generated using the symmetric (non-imbalanced) TBCM model. his reference serves as the starting point, providing baseline muscle activation levels and subglottal pressure. The muscular imbalances studied include a viscoelastic imbalance, imbalance isolated to the TA muscle, and a posterior/postural imbalance involving the LCA and IA muscles, as defined in equations (4.5) and (4.4), re-

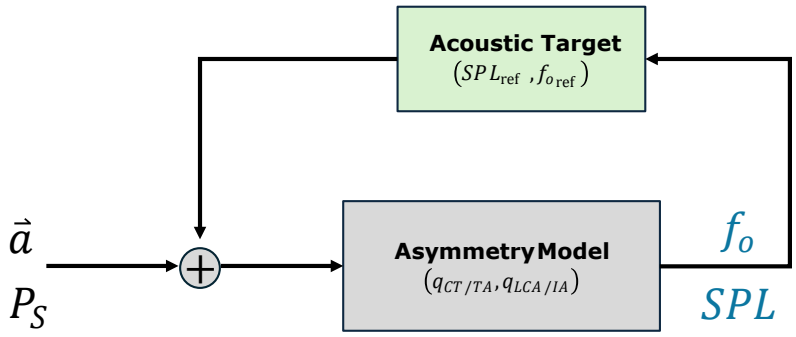

Figure 4.3: Flow diagram of muscle imbalance compensation task.

spectively. For these simulations, a muscular imbalance is adding by 0.1 more in the muscle activity.

The control scheme for adjusting the model is illustrated in Figure 4.3. A Jacobian inverse method is employed to transform the errors in SPL and f_o into changes in muscle activation and subglottal pressure. This strategy is widely used in fields such as robotics and motor control of speech [167, 168, 169], applies a linear approximation to the vocal fold model, allowing for efficient updates to the input parameters.

The linear approximation is represented by the equation:

$$Fx = Fx_0 + \mathbf{J}(x_0)(x - x_0) + O(2), \tag{4.9}$$

where $\mathbf{J}$ is the Jacobian matrix, consisting of partial derivatives $J_{i,j}(x_0) = \frac{\partial F_j}{\partial x_i}\Big|_{x_0}$, and $O(2)$ represents the higher-order terms that are neglected. By focusing on the linear component, we establish a connection between the differences in

the output space and the input space:

$$x - x_0 = +\mathbf{J}^{-1}(x_0)(Fx - Fx_0). \tag{4.10}$$

The inverse of the Jacobian, $\mathbf{J}^{-1}$, is computed using a damped least squares method with a modified Moore-Penrose pseudoinverse [167, 168], which includes a regularization term to ensure the stability of the solution:

$$\mathbf{J}^{-1}(x,t) = \mathbf{J}^T(x,t)\left(\mathbf{J}(x,t)\mathbf{J}^T(x,t) - \gamma^2\mathbf{I}\right)^{-1}, \tag{4.11}$$

Using equation (4.10), the input x is updated iteratively. With $Y_T = Fx_T$ as the target output and $Y = Fx$ as the current output, the update equation becomes:

$$x(t + \Delta t) = x(t) + \alpha\mathbf{J}^{-1}(x,t)(Y_T - Y), \tag{4.12}$$

Several constants—such as α and γ—need to be adjusted in the equations (4.11) and (4.12). In this context, the input vector x consists of the muscle activation parameters a_{CT}, a_{TA}, a_{LCA}, and subglottal pressure P_S, while the output vector F contains f_o and SPL. To maintain consistent magnitudes for the Jacobian components, the input values were normalized. The following parameters were used: $\alpha = 0.1$, $\gamma = 0.2$, and $\epsilon = 0.02$ for the numerical Jacobian calculations. Each simulation step involved 100 iterations.

A major challenge in this process is the multiplicity of laryngeal configurations that can result in the same acoustic output [42, 44]. To address this issue, we supplemented the baseline input by simulating 50 additional points in the input space, applying a random perturbation modulated by 0.1. The input values were normalized, with pressure values ranging from 0.5 to 1.7 kPa. This approach reduces the impact of solution multiplicity, allowing the inverse Jacobian method to explore a broader range of possible compensatory configurations. Additionally, statistical analysis was performed on the resulting compensatory solutions to quantify the variability in muscle activation and pressure adjustments across the 50 sampled configurations. Key metrics such as the mean and standard deviation for each muscle activation parameter (a_{CT}, a_{TA}, a_{LCA}) and P_S were calculated. This analysis provides insights into the robustness of the compensatory mechanisms and how different muscle activations and pressure values can result in equivalent acoustic outputs (SPL and f_o).

To evaluate how the compensated pathological model adapts its control configuration, we compute parameters such as the collision and subglottal pressure quotient and MFDR, which are key indicators of how the vocal folds are affected by muscle imbalance.

Two baseline configurations for muscle activation and lung pressure were considered: one representing an abducted vocal fold posture and the other an adducted posture. The explicit values for each configuration are presented in Table

	a_{CT}	a_{TA}	a_{LCA}	P_L
Abducted	0.1	0.2	0.4	1200
Adducted	0.1	0.2	0.5	1200

Table 4.1: Model input for compensation configuration. muscle activation and lung pressure

4.1.

4.5 Results

4.5.1 Influence of muscle imbalance on VF properties

To illustrate the change in the mechanical properties of the VFs due to the muscle imbalance, Figure 4.4 shows the effect of q on the values of the masses (pointed lines) and springs (solid lines) of the three blocks (upper, lower and body) that compose the VFs, for the four different cases of imbalance described in the method. The effect in mass and spring is represented by the normalized value of these parameters, such normalization is made with respect to the value in the balanced case ($q = 1$). In addition, the asymmetry mechanism introduced in SH95 is included in gray color for comparison.

As shown in Figure 4.4 (a), the asymmetry in the adductor muscles (LCA/IA),

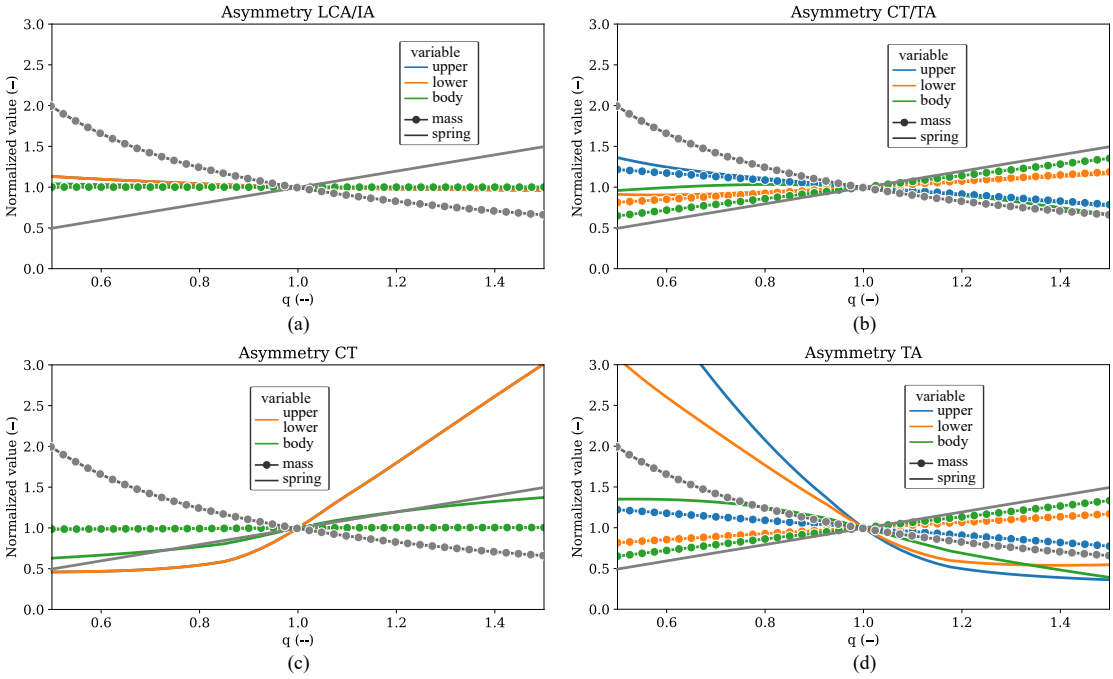

Figure 4.4: Effects on normalized (pointed line) masses and (solid line) springs when varying the joint asymmetry factor q for (a) LCA/IA, (b) CT/TA configurations, and single asymmetry in (c) CT and (d) TA muscle. The legend represents the nomenclature in a-TBCM scheme (see Figure 4.1) form mass and spring: upper, lower, and body blocks. The gray lines describes the effects due to the asymmetry mechanism introduced in SH95.

has a negligible impact on the mass of the blocks, pointed lines, but leads to a minor alteration (<10%) in all the elastic component, solid lines. This is due to the relatively minor role that adductor muscles play in the construction of the folds, although they do contribute slightly to their elongation.

For the stiffness component, the solid lines in Figure 4.4 (c) and (d), the

imbalance in the CT muscle produces a trend in the same direction as the SH95 reference mechanism, k increasing with q, although not in a strictly linear manner in the cover component, orange line. Conversely, the imbalance in the TA muscle produces an effect of decreasing stiffness with increasing q, opposite to the SH95 reference mechanism and the imbalance in the CT muscle.

Regarding the mass component, the pointed lines in Figure 4.4 (c) and (d), imbalance in the CT muscle does not lead to any changes. However, an imbalance in the TA muscle results in a linear change in the mass of the body block, green dots, and redistributes the mass of the cover between the upper and lower blocks, blue and orange dots. This pattern contradicts the direct imbalance in mass and spring values and highlights the intricacy of the laryngeal control mechanism. The CT muscle elongates the fold, impacting its elasticity without adding any mass. In contrast, the TA muscle is responsible for adding mass to the fold and redistributing it between the blocks by altering their dimensions. This generates an effect on both the mass and spring components. Similar trends are described in [95], where the mechanism we used to calculate masses and springs from muscle activation is explained.

The joint imbalance in CT and TA, Figure 4.4 (b), shows the pondered effect of the individual imbalances. The spring component, solid lines, is the sum of the opposing effects, resulting in a decreasing effect with q of small magnitude. The mass component, pointed lines, is the same as in the imbalance in TA, since the

imbalance in CT has no effect on the mass of the blocks.

Figure 4.5 illustrates the impact of q on fundamental frequency (f_o) and posterior glottal opening (PGO) in various imbalance scenarios. Notably, the configuration with an imbalance in LCA and IA muscles exhibits a non-zero posterior gap, resulting in an increased PGO for $q < 1$. This effect arises from the action of the LCA and IA muscles in adducting the VFs. In contrast, other imbalance scenarios do not induce changes in PGO due to a lack of alteration in the adduction degree of the folds. For instance, the SH95 approach, which does not consider vocal posture, exhibits no PGO changes.

Upon examining the influence of the q factor on fundamental frequency, distinct behaviors emerge in comparison to the SH95 model. Similar to the mass and spring figure, Figure 4.5 (b) shows f_o normalized to the balanced case ($q = 1$). In the SH95 model, the q factor proportionally affects f_o, reflecting its design intent. Conversely, in cases of muscle imbalance, different patterns are observed. An imbalance in the TA muscle leads to an inverse relationship with f_o—increasing q results in a decrease in f_o, influenced by the effect of TA on the VF spring. On the other hand, an imbalance in CT exhibits a slight direct correlation with f_o; an increase in q correlates with an increase in f_o, with a defined threshold for the minimum and maximum f_o values. In scenarios involving joint imbalance of two muscles, CT/TA and LCA/IA, the fundamental frequency displays a step-like behavior. At a specific point, there is a 10% variation in its value, indicating a

transition between frequencies.

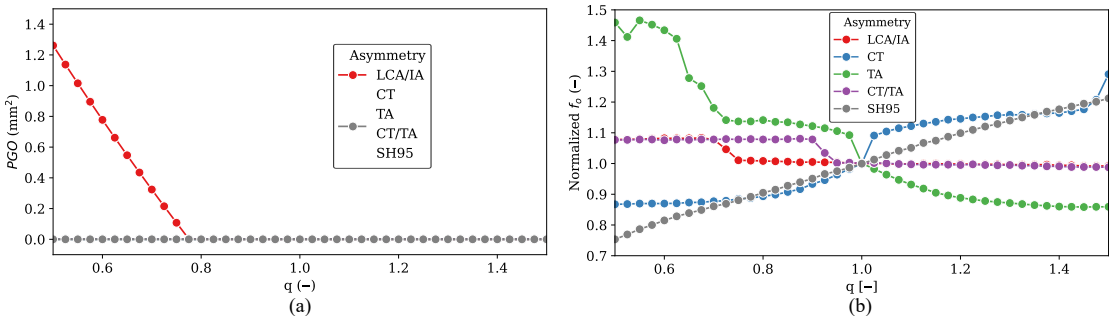

Figure 4.5: Effects on f_o and PGO when varying q for muscle imbalance configurations. SH95 results shown for comparison.

4.5.2 Influence of muscle imbalance on VF oscillation

To examine the impact of muscle imbalance on VF oscillation, Figure 4.6 provides waveforms illustrating synthetic kymograms and glottal airflow. These visualizations encompass the baseline a-TBCM simulation and imbalance scenarios under two configurations: hypofunction ($q = 0.5$) and hyperfunction ($q = 1.5$). The first row displays a synthetic kymograph illustrating the amplitude displacements of the left and right VFs in millimeters. The second row presents glottal airflow expressed in milliliters per second. All signals correspond to a 50 ms simulation. Additionally, Table 4.2 furnishes computed asymmetry measures, including f_o, OQ, AA, and PA for the simulated scenarios.

In the context of direct mass and spring imbalance, as depicted in the second column of Figure 4.6, the hypofunction configuration results in an augmented

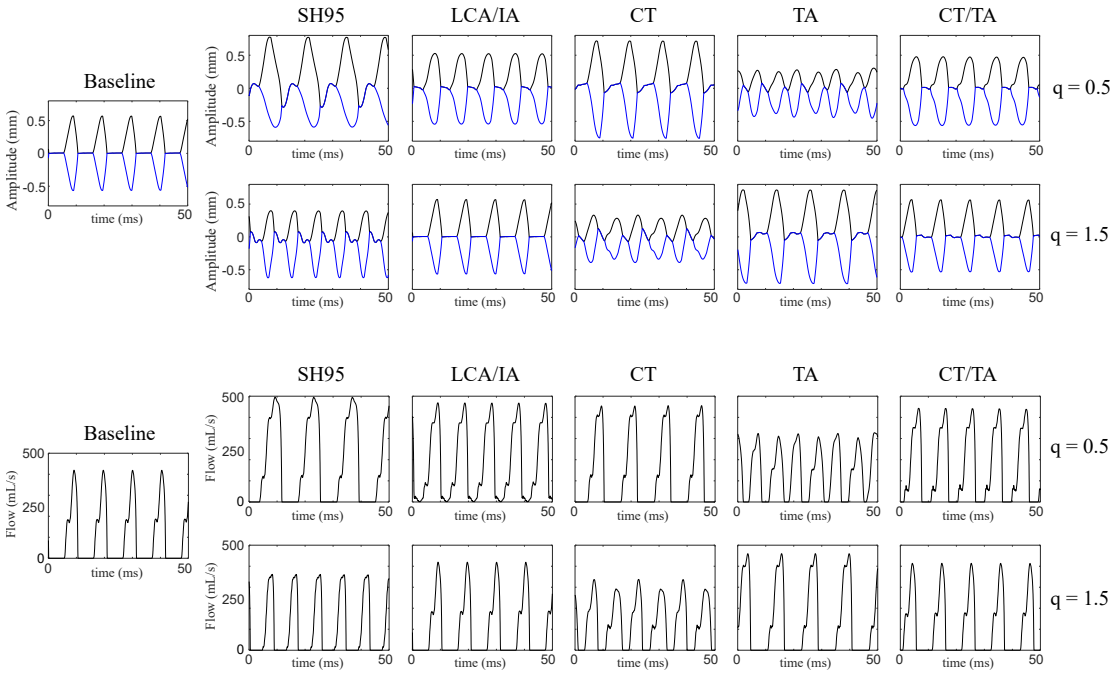

Figure 4.6: Comparison of kymograms and airflows from a-TBCM simulations for asymmetry scenarios and two imbalance conditions: hypofunction ($q = 0.5$) and hyperfunction ($q = 1.5$.)

glottal area, indicated by an increase in airflow. Furthermore, the oscillation amplitude of the affected fold is smaller compared to the unaffected fold, leading to an AA value exceeding 30%. A positive PA value, surpassing 20%, is also evident, accompanied by a decrease in the fundamental frequency during oscillation. In the hyperfunction configuration, this displacement behavior is reversed. Now, the affected fold undergoes greater displacement, reaching its maximum value first, accompanied by an increase in frequency and a decrease in flow amplitude.

In the case of imbalance in the LCA/IA adductor muscles, hypofunction ex-

hibits a slight increase in maximum airflow, along with asymmetry measurements AA and PA below 2%. These measurements indicate that the less adducted (modified) fold exhibits greater displacement. A notable observation in this case is the nonzero airflow value during the closed phase of VF oscillation, aligning with the nonzero PGO value reported in Figure 4.5. In the hyperfunctional configuration, this asymmetry maintains the baseline shape, as, after a certain level of activation in LCA and IA, the folds are already parallel. Therefore, an increase in muscle activation does not induce visible changes in the glottis with asymmetry measures below 1%.

For individually applied CT and TA muscle imbalance, a cross-behavior is observed in terms of hypo- and hyperfunction. In the case of altered CT muscle activity, displacements comparable to the baseline are observed in hypofunction, with greater displacement of the affected fold (AA=-3%) and reduced amplitude oscillation for the hyperfunctional condition. Conversely, in TA imbalance, the reduced amplitude of the folds occurs in hypofunction, while oscillation similar to the baseline is observed in hyperfunction. This underscores the antagonistic nature of this muscle pair in the viscoelastic composition of the VFs, as also observed in the mass and spring curves presented in Figure 4.4. This cross-behavior is clearly evident in the OQ, f_o, and PA values presented in Table 4.2. For individual muscle imbalance, positive values for AA are unattainable, signifying that the affected muscle has a greater amplitude of oscillation.

	Baseline	SH95	LCA/IA	CT	TA	CT/TA
f_o (Hz)	96	72.4	103.6	83.4	140.2	103.5
		116.5	95.3	124	82.6	94.5
OQ (%)	49	60	66.5	51.8	78.7	68.4
		53.9	49.1	72.8	52	49.2
AA (%)	0	31.3	-1.5	-3.1	-21.5	-8.27
		-22.1	0.4	-8.7	-0.5	2.9
PA (%)	0	24.71	-1.9	18.2	-17.5	-5.8
		-21.5	0.44	-16.9	18.2	0.3

Table 4.2: Comparison of asymmetry values and oscillation features between the baseline and the q configurations shown in Figure 4.6. For each feature, the first row is for the hypofunction ($q = 0.5$) and the second row is for the hyperfunction ($q = 1.5$)

In terms of VF oscillation behavior, Figure 4.8 demonstrates the influence of the asymmetry factor q on AA and PA across the examined cases. The conventional imbalance scheme, S&H95, yields positive values for AA and PA when $q < 1$, indicating that the affected fold's oscillation is dampened due to an increase in mass and a decrease in elasticity. Conversely, it intensifies its oscillation and frequency, resulting in negative values of AA and PA for $q > 1$.

When considering scenarios of muscle imbalance, several significant observa-

tions emerge. The imbalance in CT mirrors the conventional scheme's trend for PA, while the imbalance in TA exhibits the opposite trend. This discrepancy can be attributed to the behavior of the elastic component as a function of the q factor, increasing in the CT imbalance and SH95, and decreasing in the TA imbalance. Regarding AA, it proved challenging to find a muscle imbalance scenario that generated the same range as the reference scheme. This stems from the mechanism of muscle activation control and the mass and spring parameters, where not all mass values move in the same direction when varying muscle activation.

In contrast, the imbalance in LCA/IA yields the most minor effect on asymmetry measurements within the studied range, resulting in asymmetries of less than 2%. Meanwhile, the combined imbalance of CT/TA amalgamates the effects of individual imbalance, producing phase asymmetries of approximately 6% and amplitude asymmetries approaching 10%.

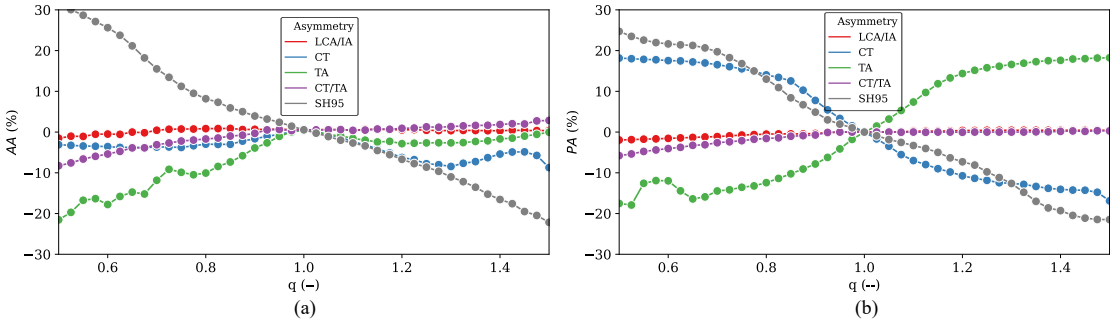

Figure 4.7: Effect of changing q on asymmetry metrics; (Left) AA and (Right) PA, for the different imbalance approaches.

4.5.3 Contrasting a-TBCM against in vivo examples

To demonstrate the efficacy of incorporating muscle imbalance in modeling various clinical scenarios, we present clinical HSV images alongside simulated data generated using the a-TBCM in Figure 4.8. This illustration highlights two distinct cases: unilateral VF paralysis (left) and muscle tension dysphonia (right). Each case features the DKG derived from clinical data, juxtaposed with the synthetic kymogram produced by the a-TBCM. Synthetic DKGs were generated by adjusting amplitude and phase asymmetry measurements heuristically to align with their respective clinical mean and variance values, employing the muscle activation defined in Equation 4.5 as the baseline. The simulation process encompasses the insights detailed in the preceding subsections, accounting for the characteristic oscillatory behavior associated with each pathology. Specifically, for VF paralysis, our aim was to emulate a rigid VF, while for muscle tension dysphonia, we introduced a subtle differentiation between the VFs.

Figure 4.8 highlights the specific imbalance factors employed in generating the synthetic kymogram, providing clarity on the value of q and the muscle to which it is applied. It is worth noting that while this approach may not be entirely singular, a more precise representation could potentially be achieved through model optimization considering the complete kymogram signal. Nevertheless, in both cases, the a-TBCM successfully produces a waveform closely mirroring the actual

kymogram. In the instance of VF paralysis, this results in a fold with minimal amplitude oscillation, out of phase in time. In the case of muscle tension dysphonia, it yields a slightly asymmetric oscillation in both amplitude and phase.

Figure 4.8 also presents corresponding AA and PA measurements from both clinical and synthetic data. Clinical assessments of VF oscillatory asymmetry display noteworthy variability, emphasizing their sensitivity to spatiotemporal resolution and the challenges in obtaining precise measurements of this nature. In the case of unilateral paralysis, both AA and PA values surpass 30%, indicating substantial asymmetry. These asymmetry values are effectively captured by the a-TBCM, which deactivates the TA in the paralyzed fold. Similarly, in the case of muscle tension dysphonia, the a-TBCM enables the representation of such levels of asymmetry by introducing an activation imbalance in the CT and TA simultaneously.

4.5.4 Compensation mechanism for muscle imbalance

As a first step in studying compensation, we present the effect of varying a single control parameter on the key characteristics of interest. Figures 4.9 and 4.10, show the impact of single input compensation on the abducted muscle activation configuration under the presence of imbalance in the TA and LCA muscles, respectively. The summary of the imbalance effect on measured features is presented in Table 4.3, and the summary of one input compensation effects is presented in

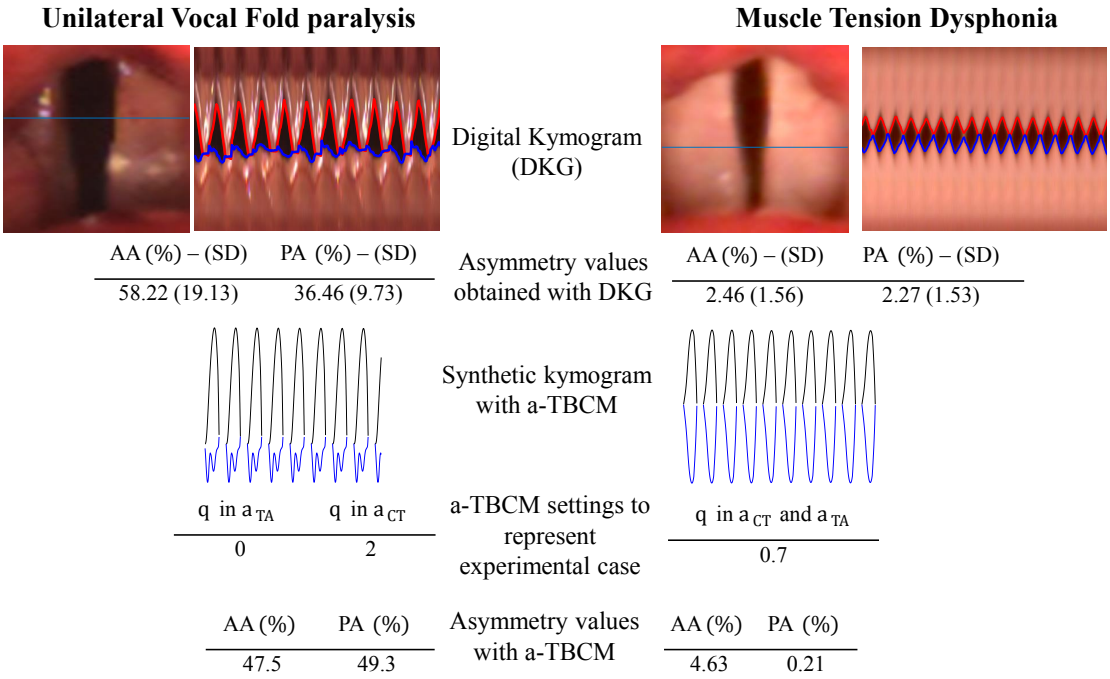

Figure 4.8: Comparison of clínical (HSV and DKG) data and the corresponding simulated responses using a-TBCM: (left) Unilateral VF paralysis and (right) Muscle tension dysphonia.

Table 4.4, which shows common trends across both cases of muscle imbalance.

As extended from the results in Table4.3, for the abducted configuration, muscle imbalance leads to an increase in SPL, with the effect being more pronounced in the LCA imbalance. This is consistent with the adductor role of the LCA and TA muscles [36]. In terms of frequency, the imbalances have opposite effects: the TA imbalance decreases f_o, while the LCA imbalance increases it. This difference is not observed in the previously analyzed adducted configuration, Figure 4.5, but can be explained by the tensor and contractor roles of the LCA and TA muscles

Feature	Abducted		Adducted	
	TA Imb.	LCA Imb.	TA Imb.	LCA Imb.
SPL	+	++	-	o
f_o	-	+	–	o
MFDR	+	+	-	o
$\frac{P_C}{P_S}$	-	+	-	o

Table 4.3: Summary of the imbalance effect on measured features in both muscle activity configuration. ++: high increase. + increase. -: decrease. –: high decrease. o: no effect.

[36]. Both types of imbalances result in increased MFDR, while they have opposing effects on the pressure quotient, due to their impact on the cross-sectional area of the VFs.

For the adducted configuration, LCA imbalance does not seem to significantly affect either frequency or intensity. This can be attributed to the behavior of the LCA muscle, which, above an activation level of 0.5, no longer appears to influence these quantities. However, it does generate slight changes in MFDR and the pressure quotient compared to the balanced configuration. In contrast, TA imbalance leads to a decrease in all measured parameters, with its most notable impact on f_o.

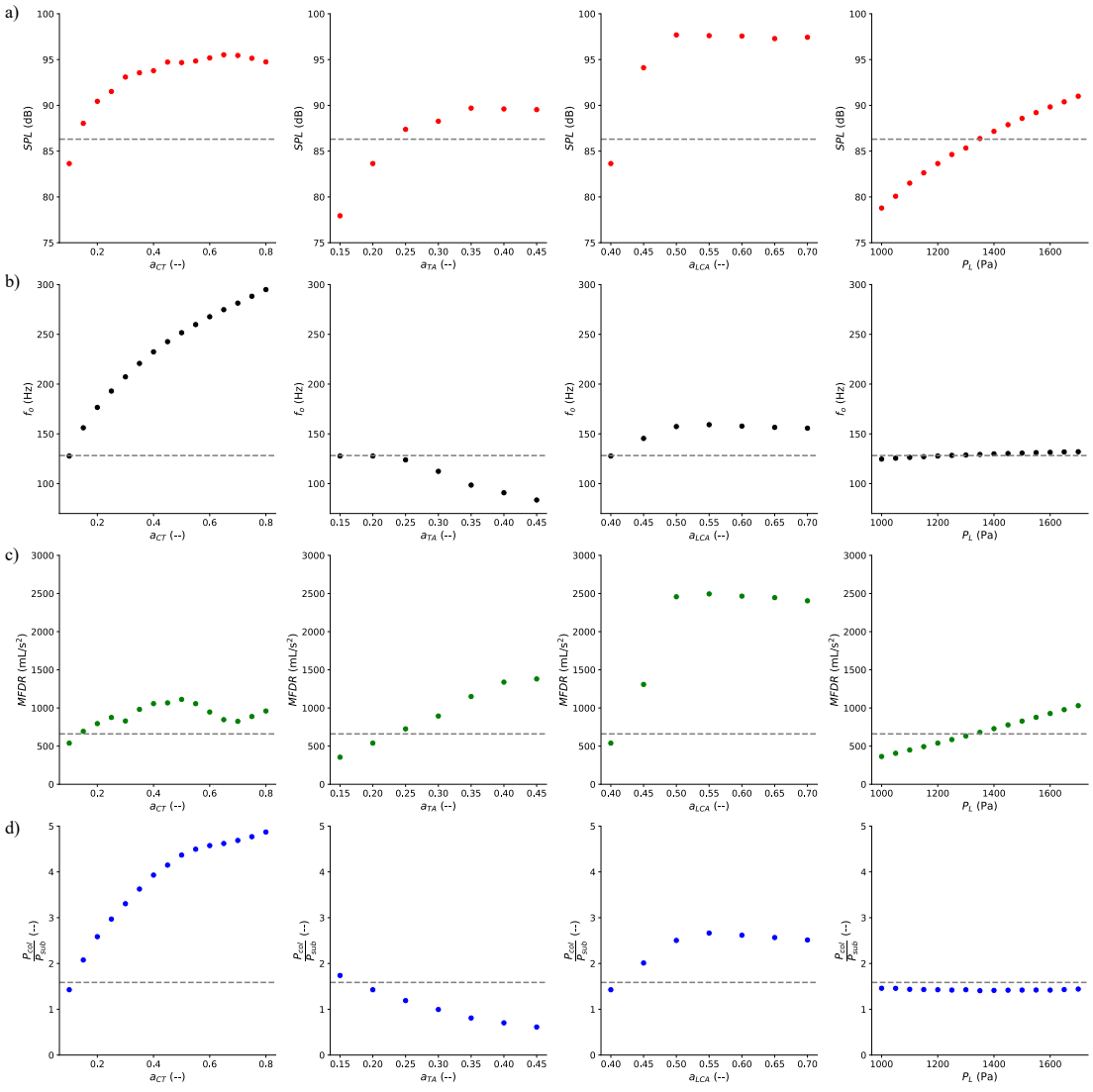

Figure 4.9: Effect of one variable compensation in an abducted VF configuration with imbalance in TA muscle. Effect on: (a) SPL, (b) f_o, (c) MFDR and (d) quotient of collision and subglottal pressure. the columns denote each input variable: muscle activations (a_{CT}, a_{TA}, a_{LCA}) and lung pressure (P_L).

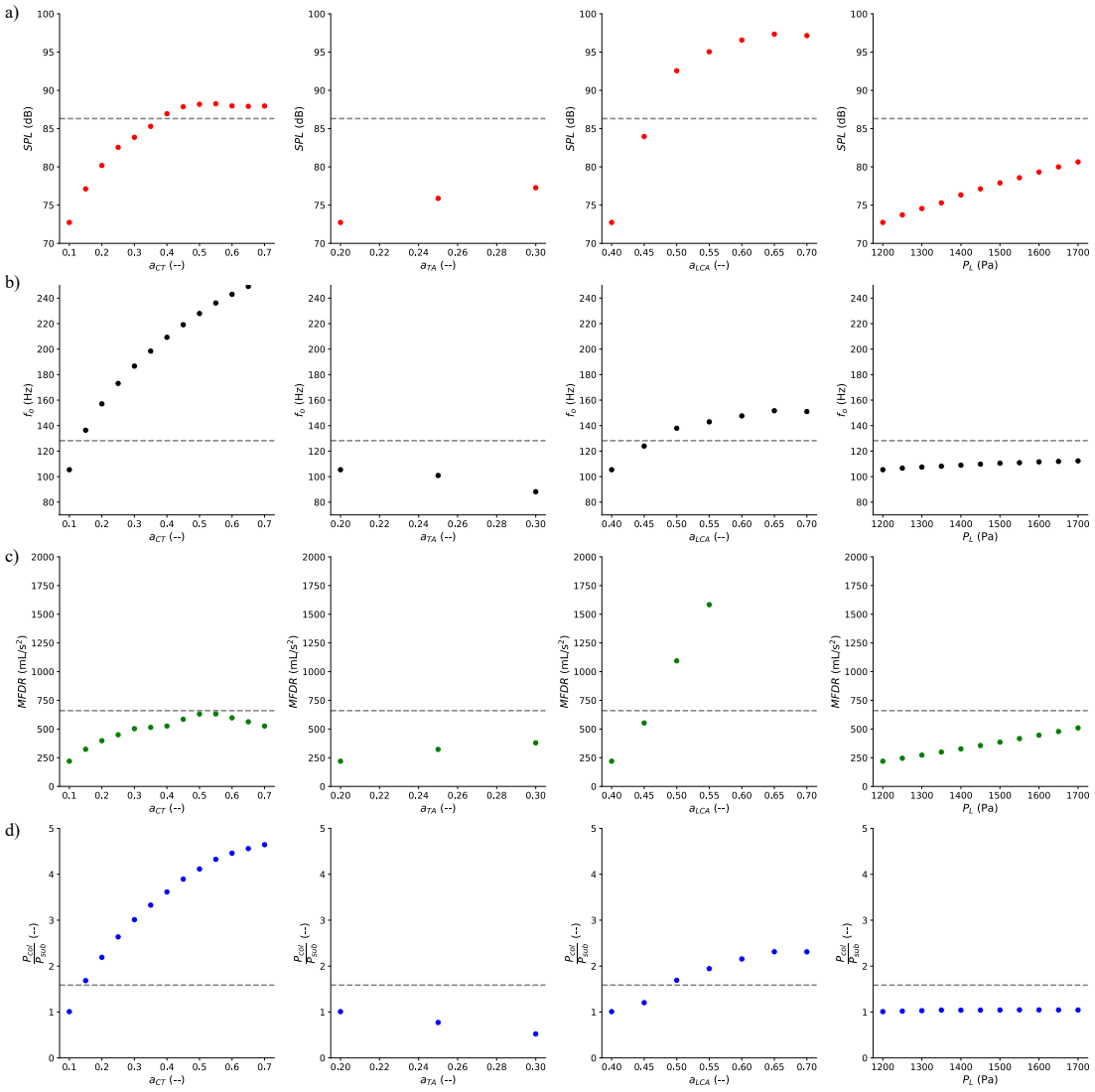

Figure 4.10: Effect of one variable compensation in an abducted VF configuration with imbalance in LCA muscle. Effect on: (a) SPL, (b) f_o, (c) MFDR and (d) quotient of collision and subglottal pressure. the columns denote each input variable: muscle activations (a_{CT}, a_{TA}, a_{LCA}) and lung pressure (P_L).

Feature	CT act.	TA act.	LCA act.	P_L
SPL	+s	+o	+s	++
f_o	++	-	+s	+o
MFDR	+o	+o	+s	++
$\frac{P_C}{P_S}$	++	–	+s	o

Table 4.4: Summary of single compensation effect in muscle imbalance. ++: high increase. + increase. -: decrease. –: high decrease. +s: increase and then saturation. +o: non-consistent increase. o: no effect.

As detailed in Table 4.4, the compensation achieved through CT activation shows strong control over f_o and a minor influence on SPL. This is consistent with the CT's role as a tensor muscle, which also leads to a notable increase in the pressure quotient by reducing the cross-sectional area of the vocal folds. In contrast, compensation via the TA muscle results in slight or inconsistent increases in SPL and MFDR, but a consistent decrease in both f_o and the pressure quotient due to the widening of the vocal folds. The tensor and thickening roles of the CT and TA muscles, as well as their relationship with frequency and intensity, align with previous findings from excised larynx experiments [36, 151].

For the LCA muscle, the compensation follows a uniform pattern across all variables, showing an initial increase followed by saturation at an activation value

near 0.5, where the vocal folds become nearly parallel. Lung pressure (P_L), on the other hand, has a significant effect on MFDR and SPL, underscoring the strong correlation between these variables [10, 33], but it has minimal impact on f_o. Notably, the pressure quotient remains constant, suggesting that changes in collision pressure relative to subglottal pressure are primarily due to alterations in laryngeal configuration—an important factor in the pathophysiology of vocal hyperfunction [119].

To mitigate the effects of muscle imbalances in both abducted and adducted configurations, the following compensation mechanisms could be proposed:

For the abducted configuration, the increase in SPL caused by the imbalances could be compensated by decreasing LCA activation while increasing CT activation or subglottal pressure (P_S) to maintain f_o levels. Such compensation mechanisms have been hypothesized for NPVH, previously referred to as abducted vocal hyperfunction [5, 6]. This compensatory strategy would likely increase MFDR due to the combined effect of CT and P_S, while the LCA would act to decrease it, resulting in a pressure quotient close to the original value as the effects of CT and LCA counterbalance each other.

For the adducted configuration, where TA imbalance dominates, the decrease in measured parameters could be compensated by jointly increasing P_S and CT activation, with the LCA playing a secondary role. In the following section, which presents the results of joint compensation, we will assess whether these compen-

sation mechanisms hold consistent within the model.

Moving on to joint compensation, Tables 4.5 and 4.6 present the results of the compensation process for the abducted and adducted configurations, respectively. These tables compare the values of the control configurations (i.e., the model without imbalance) with the mean and standard deviation of the compensation process using the Jacobian scheme. It is important to note that 50 points close to the initial configuration were simulated to mitigate the effect of solution multiplicity [170].

Key observations include the fact that for both imbalances and VF configurations, P_S is elevated, but MFDR remains similar to the control configuration. This result aligns with clinical data [10].

In terms of adduction degree, for the abducted configuration, the compensation process shows an increase in PGO for LCA imbalance, while maintaining the posterior gap level in the case of TA imbalance. This is consistent with the characterization of NPVH as a form of non-adducted VH [5]. For the adducted configuration, a larger posterior gap was not observed on average, but variability in LCA activation was noted, which could suggest some degree of abduction.

These results indicate that both imbalances can characterize patterns observed in NPVH. Specifically, in subjects with significant abduction, LCA imbalance dominates, while in those with lower abduction but still elevated P_S, TA imbalance takes precedence. It is also worth noting that compensating for these imbalances

Variable	Unit	Control	TA imbalance		LCA imbalance	
			mean	std	mean	std
SPL	dB	86.31	86.31	0.00	86.31	0.01
f_o	Hz	128.20	128.23	0.02	128.23	0.03
CT act.	–	0.1	0.16	0.02	0.17	0.03
TA act.	–	0.2	0.37	0.08	0.35	0.09
LCA act.	–	0.4	0.48	0.17	0.36	0.04
P_S	Pa	897.0	1058.0	193.4	1186.4	150.9
ACFL	mL/s	505.62	464.57	165.31	603.06	94.85
MFDR	L/s^2	660.29	679.42	88.73	1008.86	231.55
P_C/P_S	–	1.58	1.17	0.09	1.40	0.18
PGO	mm^2	1.65	1.18	1.29	2.18	0.40

Table 4.5: Results of compensation mechanism in muscle imbalance for an abducted VF configuration.

Variable	Unit	Control	TA imbalance		LCA imbalance		Nodules	
			mean	std	mean	std	mean	std
SPL	dB	97.83	97.82	0.01	97.81	0.03	86.29	0.36
f_o	Hz	157.50	157.52	0.02	157.51	0.04	128.22	0.54
CT act.	–	0.1	0.18	0.03	0.20	0.03	0.25	0.14
TA act.	–	0.2	0.38	0.07	0.39	0.06	0.48	0.19
LCA act.	–	0.5	0.47	0.12	0.51	0.12	0.52	0.10
P_S	Pa	973.0	1230.9	116.1	1265.9	138.4	1226.8	169.6
ACFL	mL/s	486.49	561.06	52.17	566.89	56.65	593.98	97.08
MFDR	L/s^2	2447.71	2600.45	430.15	2452.98	487.06	2329.47	631.61
P_C/P_S	–	3.1	1.89	0.27	1.90	0.18	3.43	1.07
PGO	mm^2	0	0.49	0.52	0.62	0.58	0.25	0.36

Table 4.6: Results of compensation mechanism in muscle imbalance for an adducted VF configuration.

does not increase the pressure quotient, which may explain why elevated P_S and laryngeal tension do not result in vocal fold trauma in these cases.

4.6 Discussion

The primary aim of this study was to introduce and demonstrate the concept of muscle imbalance as a mechanism for inducing asymmetries within the VFs. We sought to establish its potential in characterizing clinical scenarios by examining the influence of intrinsic laryngeal muscles on the glottis and drawing comparisons with previous approaches for understanding asymmetric VF oscillations.

Our findings lead us to conclude that achieving a direct differentiation scheme akin to SH95 through an imbalance in a single muscle alone is not feasible. The closest approximation is attained through an imbalance in the CT muscle. Additionally, the antagonistic relationship between the CT and TA muscles hinders the replication of the direct imbalance scheme. However, introducing imbalance in both these muscles concurrently, albeit in opposite directions with proportional imbalance in CT and inversely proportional in TA, appears to be a promising avenue towards achieving an analog of the conventional asymmetry mechanism using muscle imbalance. It is worth noting that a comprehensive exploration of this intricacy exceeds the scope of this initial study.

The proposed method for prescribing muscle imbalance provides a systematic

means to introduce asymmetry in the composition and positioning of the VFs. This asymmetry has been studied independently by other researchers. Prior efforts [171, 120, 122, 51] report ranges of amplitude and base asymmetry akin to our findings. Conversely, studies such as [120, 54, 55] report differentiation in the vocal process without altering the mechanical properties of the VFs, resulting in increased minimum glottal flow and decreased sound pressure, much like the effects observed with the imbalance in the LCA and IA adductor muscles.

When considering the results of mechanical parameters and asymmetry measurements collectively, it becomes evident that phase asymmetry aligns with the spring component trends, while mass trends are associated with amplitude asymmetry. This elucidates why the imbalance mechanism does not yield significant amplitude values. Each block follows a distinct mass pattern when subjected to imbalance, resulting in a small AA that deviates from the SH95 scheme, where all blocks experience similar mass effects.

The clinical examples presented in this study serve to illustrate the potential of our approach in emulating actual cases, thereby enhancing our understanding of the origins of differences in VFs, be it in cases of vocal cord paralysis [172, 173] and primary muscle tension dysphonia [174, 175]. While similar insights may exist within other methodologies, our approach integrates muscular activity directly, impacting glottal configuration.

For instance, paralysis is depicted as an inert and minimally mobile fold,

whereas further variations showcase a fold with high mass and rigidity. Notably, this work represents an initial foray into a parameter fitting and optimization problem, without delving into extensive details. Our objective is to demonstrate the utility of this imbalance approach, as exemplified in Figure 4.8.

The compensatory mechanisms identified in this study provide important insights into how VF imbalances, and those that represent NPVH, can be managed. Our findings highlight that muscular imbalances, whether in the LCA or TA, can be compensated effectively through adjustments in muscle activation and P_S while maintaining stable acoustic outputs. Notably, the maintenance of the pressure quotient without significant increases during compensation suggests that VF configuration changes, rather than tension increases, play a crucial role in preserving healthy phonation. This helps explain why patients with NPVH can exhibit high subglottal pressure without developing vocal fold trauma. However, it is essential to recognize the limitations of this study, as it examines a single case. Future research focusing on a broader set of compensatory behaviors across different VF would provide a more comprehensive understanding of these mechanisms, potentially reinforcing these compensatory strategies.

Studying VF oscillation with diverse properties presents several challenges, including the presence of harmonics, as illustrated in Figure 4.6, a topic extensively explored by other researchers [50, 48, 123]. In addition, the a-TBCM has some limitations. This model is a simplified approach to VFs that uses a lumped ele-

ment approach where one continuous block represents the entire anterior-posterior direction. This approach does not allow the assignment of different oscillation patterns to the anterior, medial, and posterior portions of the VF, where other studies have shown different degrees of asymmetry [176, 177]. For an extensive validation of this type of model, a critical consideration is the impact of the chosen reference point, whether it pertains to muscular activation or the specific values assigned to masses and springs.

On the other hand, the muscle imbalance postulate presents a greater challenge for its experimental validation. A first level of validation could be done by experiments with excised larynxes. In this area, there are several works that study the action of the laryngeal muscles on the elongation and posture of the VFs [36, 37, 178]. A possible experiment would be the asymmetric stimulation of the larynx to measure the differences in left and right oscillation, a setup like the one presented in [179], where unfortunately the video analysis is done from the glottal area wave (GAW) and not from the kymogram. The next level of validation would be with intramuscular electromyography. Here there are few works that present the action of the laryngeal muscles, and this has been done symmetrically [39, 180]. Ongoing efforts measure simultaneously the activity of the left and right TA and CT with intramuscular EMG and HSV of the larynx to associate muscle imbalance with asymmetries in vocal fold oscillation, and contrasting with asymmetric VF silicon model.

4.7 Chapter conclusions

This chapter introduced a novel asymmetrical TBCM model aimed at characterizing the effects of muscular imbalance on VF oscillations, relevant to conditions such as NPVH and VF paralysis. The model is built upon the TBCM framework, allowing for direct manipulation of muscle activation to induce asymmetry, contrasting with previous methods that focused solely on altering the mechanical components like mass and spring properties. The introduction of this control mechanism opens new possibilities for representing the biomechanical and physiological effects in VF conditions by muscle imbalance. In terms of findings, the asymmetrical TBCM proved effective in replicating key features of asymmetry in both amplitude and phase across various imbalance conditions. Muscle activation imbalance in the CT and TA muscles showed clear and distinct effects on VF viscoelastic properties, reinforcing the idea that these muscles play critical yet antagonistic roles in determining VF tension. The results confirmed that the imbalance in adductor muscles, such as LCA/IA, had a smaller impact on the mechanical characteristics of the folds but notably affected the PGO, influencing airflow and acoustic properties. A key takeaway from this study is that the proposed approach provides a more physiologically grounded method for modeling VF asymmetry, as it directly incorporates the role of intrinsic laryngeal muscles. However, the study also revealed that a single muscle imbalance was

insufficient to fully replicate the complexity of VF asymmetries observed in clinical cases, particularly in conditions like VF paralysis. The combination of CT and TA imbalances, along with their antagonistic interactions, offers the closest approximation to clinical observations, though further refinement of the model is needed to better capture these dynamics. On the clinical front, comparisons between modeled and real-world data, including high-speed video kymograms, showed promising alignment. This underscores the potential application of this model in simulating clinical pathologies like VF paralysis and muscle tension dysphonia. The ability of the model to mimic the kinematic behavior of the VFs in these conditions demonstrates its utility for understanding and diagnosing voice disorders characterized by VF asymmetry.

The results presented in this chapter begin to reveal key differences in compensatory mechanisms NPVH and PVH, supporting the initial hypothesis of this thesis. The compensation of muscle imbalance in NPVH, characterized by stable pressure quotient values, suggests a protective mechanism that could explain the absence of VF trauma in these patients, despite elevated subglottal pressures. This contrasts with the patterns observed in PVH, where elevated pressure and muscular tension contribute directly to VF damage. These findings mark an important step toward identifying distinct compensatory strategies that differentiate NPVH from PVH, providing a foundation for further investigation into how these mechanisms affect VF health and pathology.

However, it is important to acknowledge the limitations of the current model. The TBCM, as a lumped-element model, simplifies the VF dynamics by not differentiating anterior, medial, and posterior regions of the VFs. This limitation restricts the model's ability to capture finer-grained oscillatory patterns that may arise in real clinical cases. Additionally, while the muscle imbalance approach adds a new layer of realism, experimental validation of these imbalances, particularly through intramuscular EMG or excised larynx studies, remains a significant challenge. Future work will focus on refining the model to better represent VF asymmetry across more varied pathologies, improving its predictive power in clinical settings. Experimental validation, particularly through asymmetric stimulation of excised larynxes or advanced imaging techniques, will be essential to verify the model's assumptions and outcomes. The model's potential for optimization and parameter fitting in clinical cases will also be explored, with the aim of enhancing its relevance for both research and clinical applications in voice disorders.

Chapter 5

Model for vocal fold nodules

This chapter outlines the construction of a VF model that includes nodules, designed to characterize the condition of PVH. As in the previous chapter, the proposed model is built upon the TBCM framework. In this adaptation, nodules are introduced as fixed elements that alter the shape of the VFs and modify their collision behavior, making a balance between physiological description and adding complexity (degrees of freedom) in the model. The detailed equations and modifications to the base model can be found in Appendix B. A comprehensive literature review was conducted to define the operational ranges of the modeled nodule parameters, including size and stiffness. Additionally, the chapter explores the effect of these parameters on the PTP and examines how they influence vocal quality through a compensation scheme based on pitch and intensity. The results demonstrate that in the presence of nodules, the phonatory system compensates for their effects by increasing subglottal pressure and activating laryngeal muscles to maintain desired frequency and intensity levels.

In this context, the chapter presents a TBCM-based model for VF nodules,

analyzes the impact of nodule parameters on the phonation threshold, and demonstrates how laryngeal musculature can compensate for these effects. The trends observed are compared to alternative modeling approaches and experimental data from silicone VF models. The initial version of this model and preliminary findings were presented at VOICE 2023. The extended results described in this chapter are currently being prepared for submission to a peer-reviewed journal.

5.1 Nodule TBCM (n-TBCM)

The Nodule TBCM is constructed based on the TBCM model by adding a fixed mass to the upper block that defines the cover of the VF. Figure 5.1 shows the blocks that define the original TBCM in gray, and the nodules as blue blocks.

This adaptation of the TBCM to include nodules is inspired by previous works [49, 48]. Both studies present lumped-element models of nodules. In the first study [49], nodules are introduced as a change in collision position along with an increase in the mass of the upper block and the coupling spring between the cover masses. In the second study [48], the nodule is an additional block that oscillates and changes the shape of the free edge of the vocal folds.

The proposed adaptation combines both ideas. The nodule is a small mass that changes the shape of the free edge, but this mass does not oscillate; instead, it generates an additional collision force on the vocal folds. The nodule block

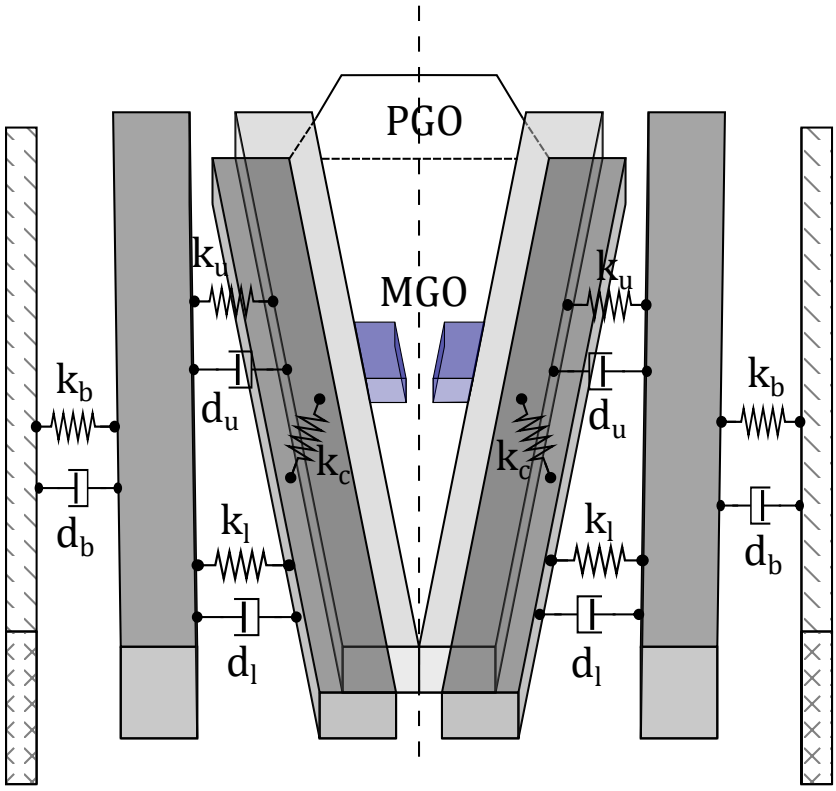

Figure 5.1: Scheme of VF nodule approach. Gray blocks represent the base TBCM scheme. Blue blocks represent the nodules.

increases the mass of the upper block and the coupling spring in the cover.

As an additional element of this model, the position of the nodule is included as a parameter, allowing it to be shifted in the anterior-posterior direction. This parameter is inspired by the work of [52, 53] in finite element models.

Table 5.1 presents the set of parameters characterizing the nodule, along with a description of these parameters and their operating range. This table is constructed based on the literature review presented in Table 2.2. The modifications made to the TBCM model for the construction of this adaptation are detailed in

Table 5.1: Definition of TBCM nodule parameters.

Feature	**Abbr.**	**Range**	**Unit**	**Reference**
Nodule Radious	R_n	0 to 1.5	mm	[7, 52, 114]
Nodule Young Modulus	E_n	10 to 300	kPa	[52, 53, 88, 114]
Nodule Poisson Ratio	ν_n	0.3	–	[92, 115, 52, 53, 97, 88, 118]
Nodule density	ρ_n	0.9 to 1.2	gr/cm^3	[115, 52, 53, 88, 48, 56]
Nodule position	λ_n	0 to 1	–	[52, 53]

Appendix B.

5.2 Effect of nodule parameters

5.2.1 Effects of nodule parameters on phonation threshold

The parameters defining the nodule have an effect on the oscillation of the VFs. To describe this effect, the reference point is the onset of phonation, defined by the PTP and the frequency at the onset ($f_{o,t}$, [114]). This task has been performed for excised larynges [36, 37], and silicone models [114]. The experiment involves gradually increasing the pressure reaching the glottis, with a fixed laryngeal configuration, until the point of phonation onset. PTP is the pressure value at which phonation begins, and $f_{o,t}$ in the frequency of oscillation at that

moment.

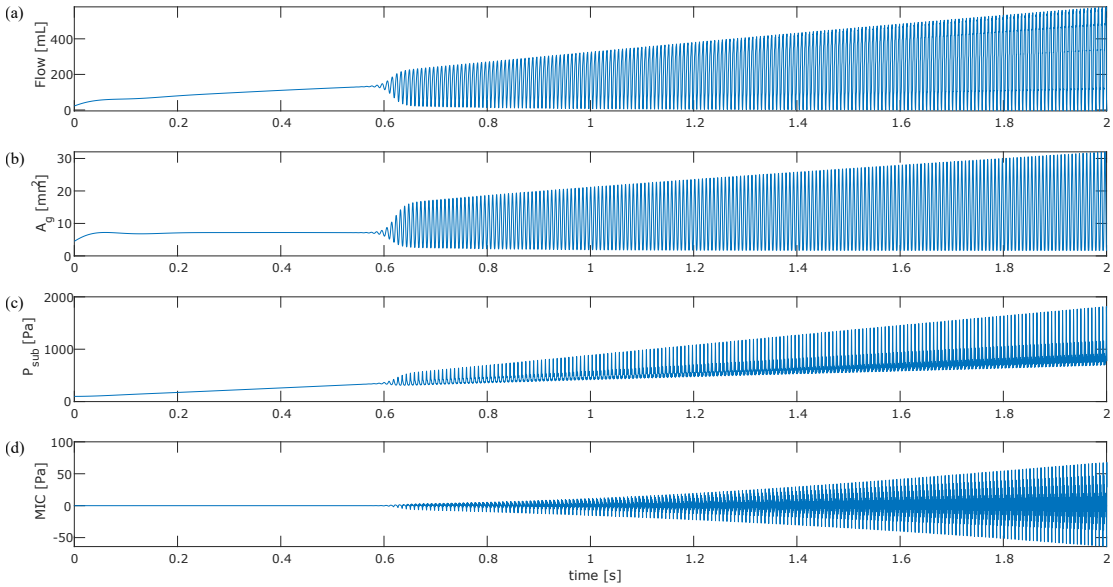

Figure 5.2: Model-based signal in PTP simulation: (a) Gottal iarflow. (b) Glottal area. (c) Subglottal pressure. (d) Microphone signal.

The PTP value indicates the amount of energy required to oscillate the VFs, and its value can be related to vocal efficiency [181]. Figure 5.2 shows a simulation plot of PTP. In plot (c), we can observe how the subglottal pressure gradually increases due to the gradual increase in lung pressure. At the moment when PTP is reached, the oscillation of P_S begins, and similarly, we see that the airflow, glottal area and MIC signal start the oscillation from that point.

To observe the effect of the different nodule parameters on PTP, an initial laryngeal configurations are defined, i.e. a muscle activation vectors:

$$\vec{a} = [0.4, 0.4, 0.0, 0.2, 0.1]\,. \tag{5.1}$$

$$\vec{a} = [0.5, 0.5, 0.0, 0.2, 0.1]\,. \tag{5.2}$$

The selection of the first muscle activation is made to characterize a male modal voice at about 130 Hz with a lung pressure os 1.2 kPa [44]; this configuration also have a posterior gap, to analyze how the nodule affect this type of VF posture; the second muscle activation is chosen to see the effect in no-gap VF posture. To these initial configurations, a base nodule is added whose parameters are given in the first row of Table 5.2. In addition, the nodule parameters are varied as shown in the second row.

	R_n (mm)	ρ_n (gr/cm^3)	E_n (kPa)	λ_n (–)
Baseline	1	1	10	0.5
Range	0.25 - 2	0.5 - 2	2.5 - 20	0 - 1

Table 5.2: Nodule parameters baseline value and range.

For each nodule configuration, PTP, $f_{o,t}$, MFDR and P_C are computed in the threshold instant, consisting of a 2-second simulation at a sampling rate of 44.1 kHz, with lung pressure constantly increased from 100 Pa to 2.5 kPa. Subglottal

and glottal tracts are added to the phonation model for sustained /a/ phonation, resulting in the phonation model simulation shown in the Figure 5.2.

5.3 Compensation mechanism in nodule model

To investigate how a subject compensates for the presence of VF nodules, we designed a simulation scheme to simulate the compensation of intensity (SPL) and pitch (f_o) in the presence of a nodule. The simulation begins with a reference SPL and f_o generated using the nodule-free TBCM model. The muscle activation and subglottal pressure from this reference simulation are used as the starting configuration. The nodules are introduced in the form of fixed masses attached to the VFs, with the following key parameters: ($R_n = 1$mm, $E_n = 10$kPa, $\rho_n = 1$gr/cm^3 and $\lambda_n = 0.5$).

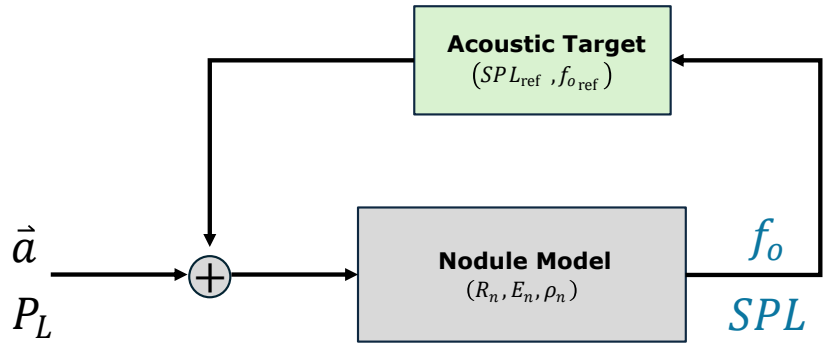

Figure 5.3: Flow diagram of nodule compensation task.

The control scheme for generating model adjustments is represented in Figure 5.3. As in the case of muscle imbalance, the method of inverse Jacobian is

employed to map errors in SPL and f_o into changes in muscle activation and subglottal pressure. For this experiment, we normalized the input variables (muscle activations and subglottal pressure) to maintain consistent Jacobian scaling, with values of $\alpha = 0.1$ and $\gamma = 0.2$, and a numerical Jacobian approximation step size $\epsilon = 0.01$. After compensation is applied, key aerodynamic and acoustic parameters, such as collision pressure (P_C), and MFDR, will be calculated to evaluate the compensation mechanism and characterize the glottal airflow changes due to the presence of nodules.

5.4 Results

5.4.1 Effects of nodule parameters - phonation threshold

Figures 5.4 and 5.5 present the effects of nodule parameters on phonation threshold simulations for the different muscle activation configurations studied. The figures show changes in PTP (row a), f_o (row b), MFDR (row c), and the collision-to-subglottal pressure ratio (row d), for variations in nodule size, Young's modulus, density, and position, across the different columns of the figures.

The first key observation is the negligible effect of nodule density on the studied parameters within an operational range of values around 1 (third column of Figures 5.4 and 5.5). This is consistent with previous modeling studies, such as those in [182], who describe its effect as weak, while other studies, like Deguchi

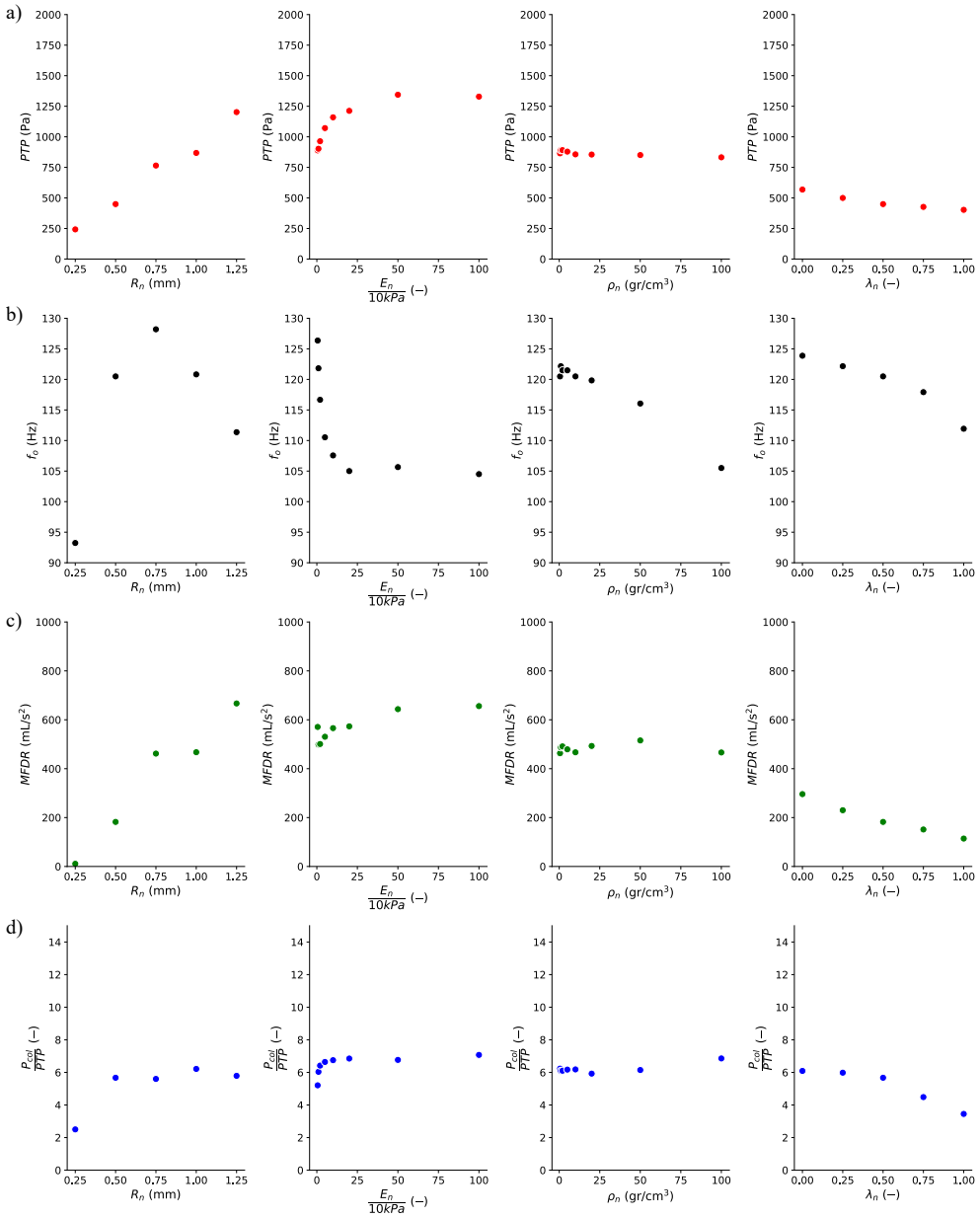

Figure 5.4: Effect of nodule parameter in phonation threshold for an abducted VF configuration: effect on: (a) SPL, (b) f_o, (c) MFDR and (d) quotient of collision and subglottal pressure. the columns denote each of the nodule parameter: size (R_n), elasticity (E_n), density (ρ_n) and position (λ_n).

[52], report only slight effects on airflow amplitude.

Another significant finding from these figures is that the effect of nodule position (fourth column of Figures 5.4 and 5.5) depends on the degree of VF abduction. For example, in Figure 5.5, no effect is observed because the VFs are parallel, whereas in Figure5.4, variations in the measured characteristics appear due to the lower LCA activation, creating a posterior gap. As the nodule moves towards the anterior region ($\lambda < 0.5$), greater contact between the nodules and the VFs occurs, requiring higher pressures (i.e., higher PTP) as the nodule becomes more anterior. This increased pressure also results in an increase in f_o and MFDR for small λ_n. Regarding the collision pressure ratio, as the nodule moves towards a more posterior position, it facilitates VF oscillation and lower pressure values, yielding a better subglottal-to-collision pressure ratio. These results are in line with previous studies that report a strong effect of posture or pre-collision conditions on PTP [182]. However, this finding contrasts with Deguchi's model [52], where higher PTP was observed in the middle zone with a symmetric pattern. The difference arises because that model uses a continuous representation of the vocal folds, allowing for greater deformation of the folds in the middle zone, which is not possible with the lumped-element model used here.

In terms of nodule elasticity, represented by its Young's modulus (second column of the figures), we observe that PTP increases as E_n increases, saturating at extreme stiffness values. This asymptotic behavior is also present in MFDR

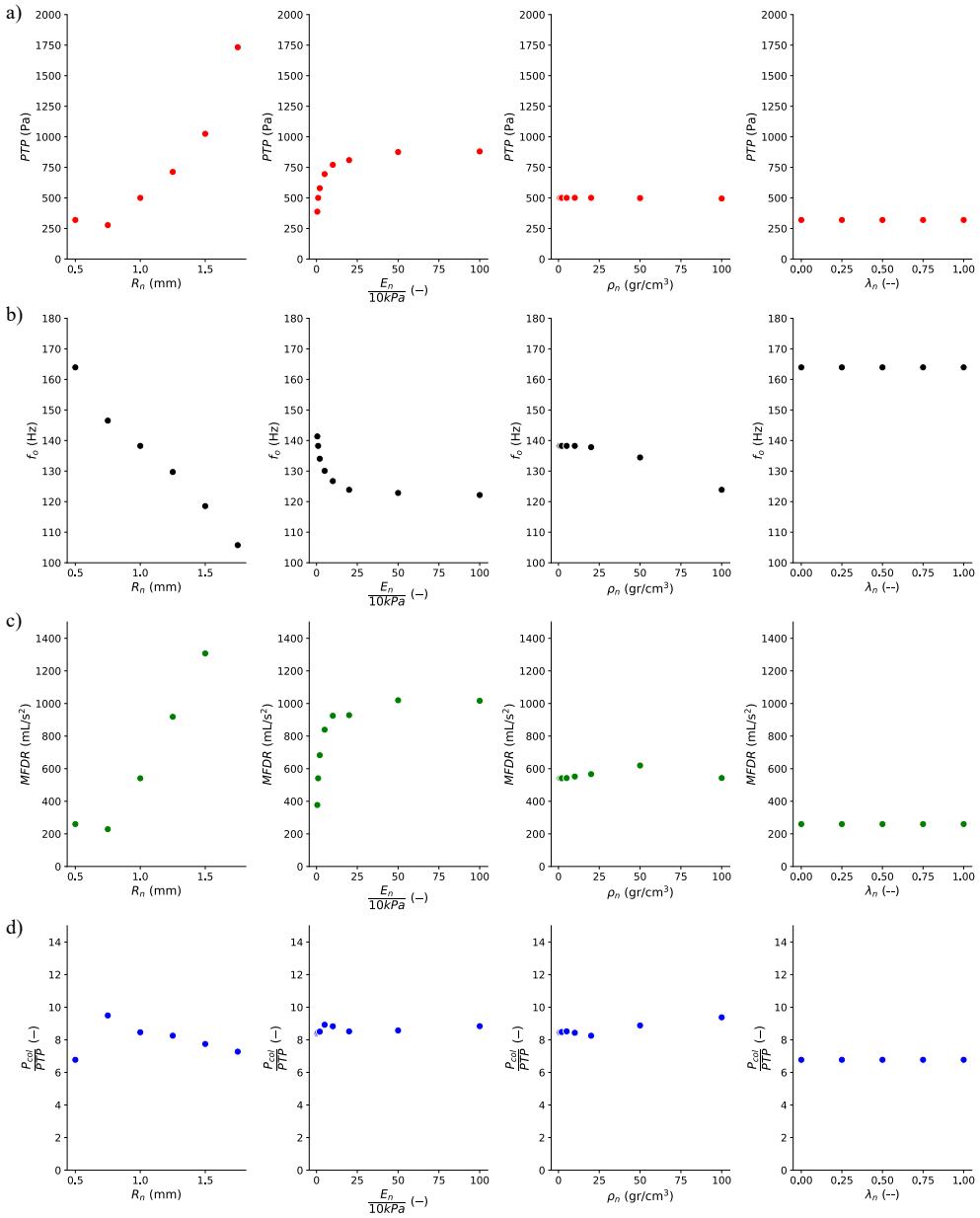

Figure 5.5: Effect of nodule parameter in phonation threshold for an non-abducted VF configuration: effect on: (a) SPL, (b) f_o, (c) MFDR and (d) quotient of collision and subglottal pressure. the columns denote each of the nodule parameter: size (R_n), elasticity (E_n), density (ρ_n) and position (λ_n).

and the pressure ratio. These results align with previous findings from silicone models [114], where no significant effect of stiffness on PTP or f_o was observed. However, in the case of the pressure ratio, those studies reported an increase with stiffness. This difference is due to the way collision pressure is calculated in our model—at the point of maximum collision force when both the upper and lower masses are colliding—rather than when the nodule alone collides. The small area of the nodule can result in a spike in pressure, even though the collision force is minimal.

Finally, when studying the effects of nodule size on phonation threshold, we find the expected results in line with previous studies [114, 53, 52, 182]: an increase in PTP with larger nodule size, along with an increase in MFDR. It is also worth noting that this model yields PTP values within a range consistent with normal phonation, from soft to comfortable levels [29].

These results quantify the effect of the nodule on the initiation of phonation, suggesting that greater pressure is required to compensate for the presence of the nodule.

5.4.2 Effects of nodule parameters - phonation

Another interesting aspect to explore is the effect of the nodule during sustained phonation, i.e., when delivering a consistent lung pressure in the simulations. Figures 5.6 and 5.6 present the effects of the nodule parameters on two

Feature	abducted				adducted			
	R_n	E_n	ρ_n	λ_n	R_n	E_n	ρ_n	λ_n
SPL	–	-	o	+	-	-s	o	o
f_o	-	-	s-	+	-	-s	s-	o
MFDR	-	-	o	+	-	-s	o	o
P_C/P_S	+	+	o	-	+s	+s	o	o

Table 5.3: effects of the nodule parameters on phonation features. ++: high increase. + increase. -: decrease. –: high decrease. s: saturation. +o: non-consistent increase. o: no effect.

muscle activation configurations with a lung pressure of 1.5 kPa. In these cases, PTP is replaced by SPL to quantify vocal intensity. Similar to the asymmetric model, Table 5.3 provides a summary of the effects of the nodule parameters on phonation during sustained vocalization.

The effects of the nodule during sustained phonation complement the observations made under phonation threshold conditions. There is a reduction in SPL, f_o, and MFDR as the size or stiffness of the nodule increases, with this effect being more pronounced in the abducted configuration. Similarly, nodule position has an impact in the abducted configuration but not in the parallel vocal fold configuration. Nodule density, however, does not seem to have any significant

effect. One notable observation is that the pressure ratio is substantially higher compared to the baseline configuration.

These results align with previous simulation studies [52, 53, 182], which demonstrated a decrease in sound intensity and airflow amplitude as nodule size increased. Additionally, the variation in pressure ratio between the two muscle activation configurations may be explained by observations from other authors [119], who noted that different laryngeal configurations can produce similar phonatory conditions but may lead to a higher vocal dose in certain cases.

5.4.3 Compensation mechanism for nodule

The presence of VF nodules introduces significant challenges to maintaining normal vocal function, primarily by altering the biomechanical properties of the VFs and disrupting the typical oscillatory patterns [7, 48]. To better understand how the vocal system compensates for these alterations, we explore how changes in muscle activation and subglottal pressure influence key aerodynamic and acoustic parameters.

Figures 5.8 and 5.8 display the effect of varying the four primary model inputs (muscle activations and P_L) on the four key parameters of interest (f_o, SPL, MFDR, and the pressure quotient) for both the abducted and adducted muscle activation configurations. The summary of each muscle's compensatory action in the presence of a nodule is provided in Table 5.4.

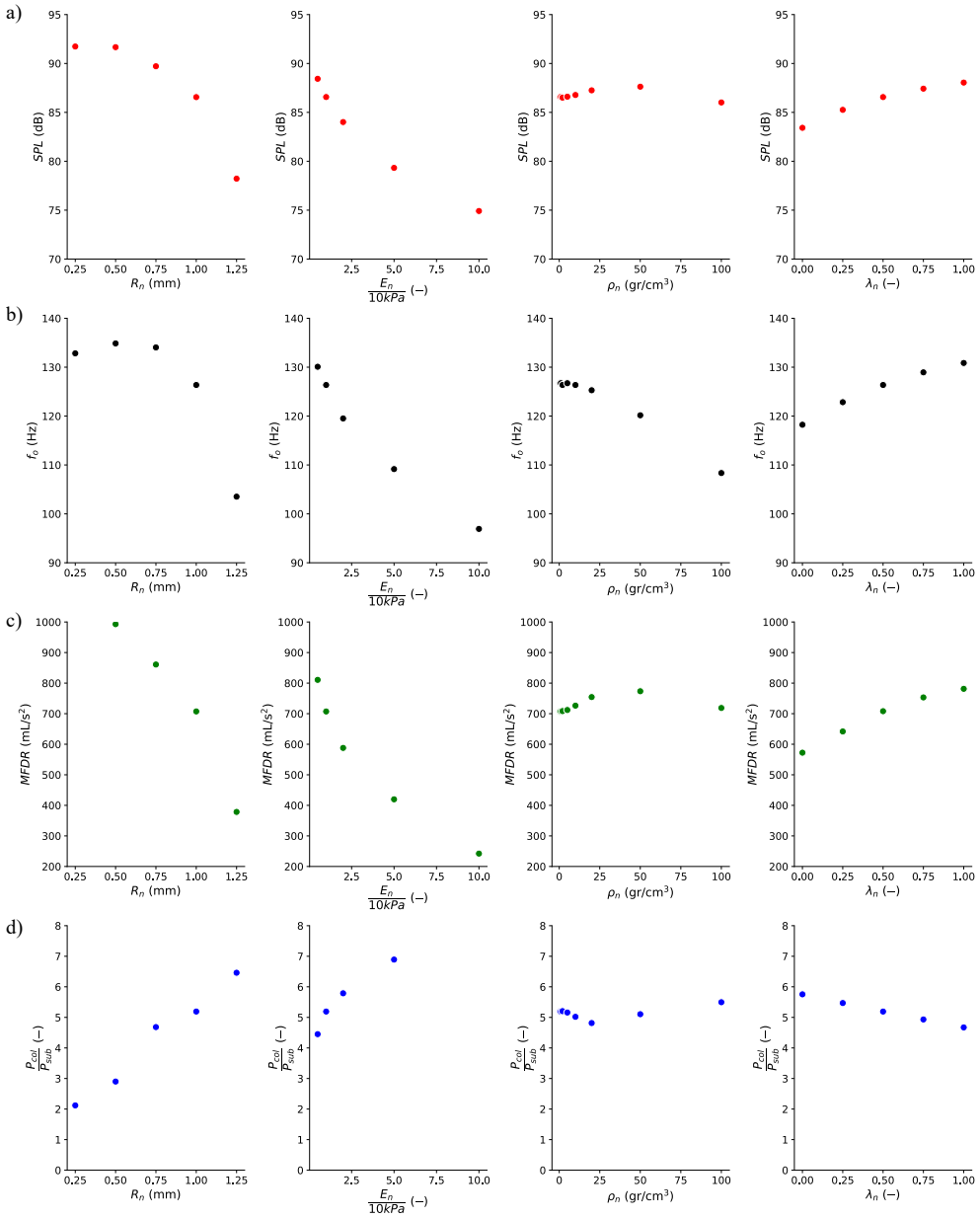

Figure 5.6: Effect of nodule parameter in phonation for an adducted VF configuration: effect on: (a) SPL, (b) f_o, (c) MFDR and (d) quotient of collision and subglottal pressure. the columns denote each of the nodule parameter: size (R_n), elasticity (E_n), density (ρ_n) and position (λ_n).

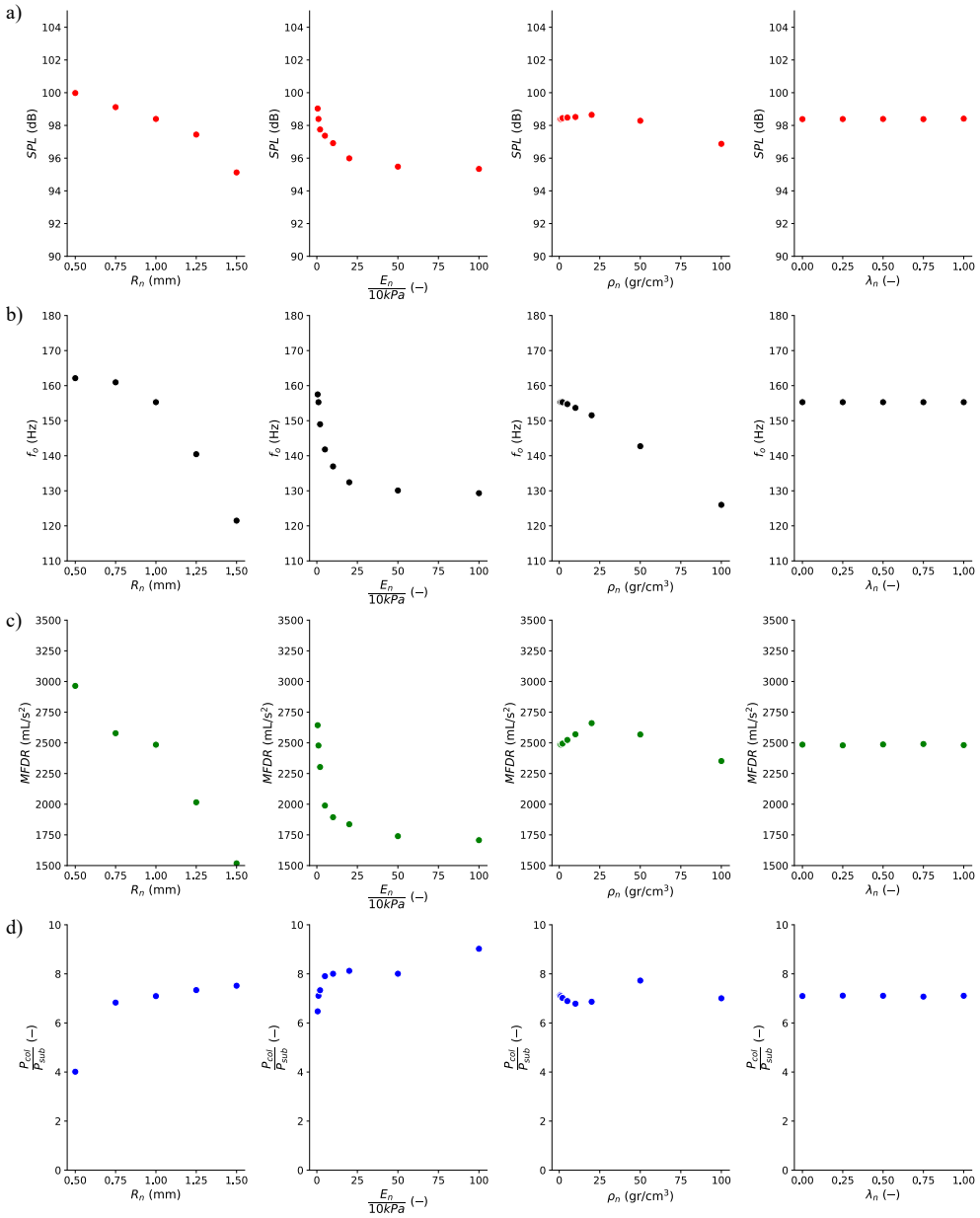

Figure 5.7: Effect of nodule parameter in phonation for an adducted VF configuration: effect on: (a) SPL, (b) f_o, (c) MFDR and (d) quotient of collision and subglottal pressure. the columns denote each of the nodule parameter: size (R_n), elasticity (E_n), density (ρ_n) and position (λ_n).

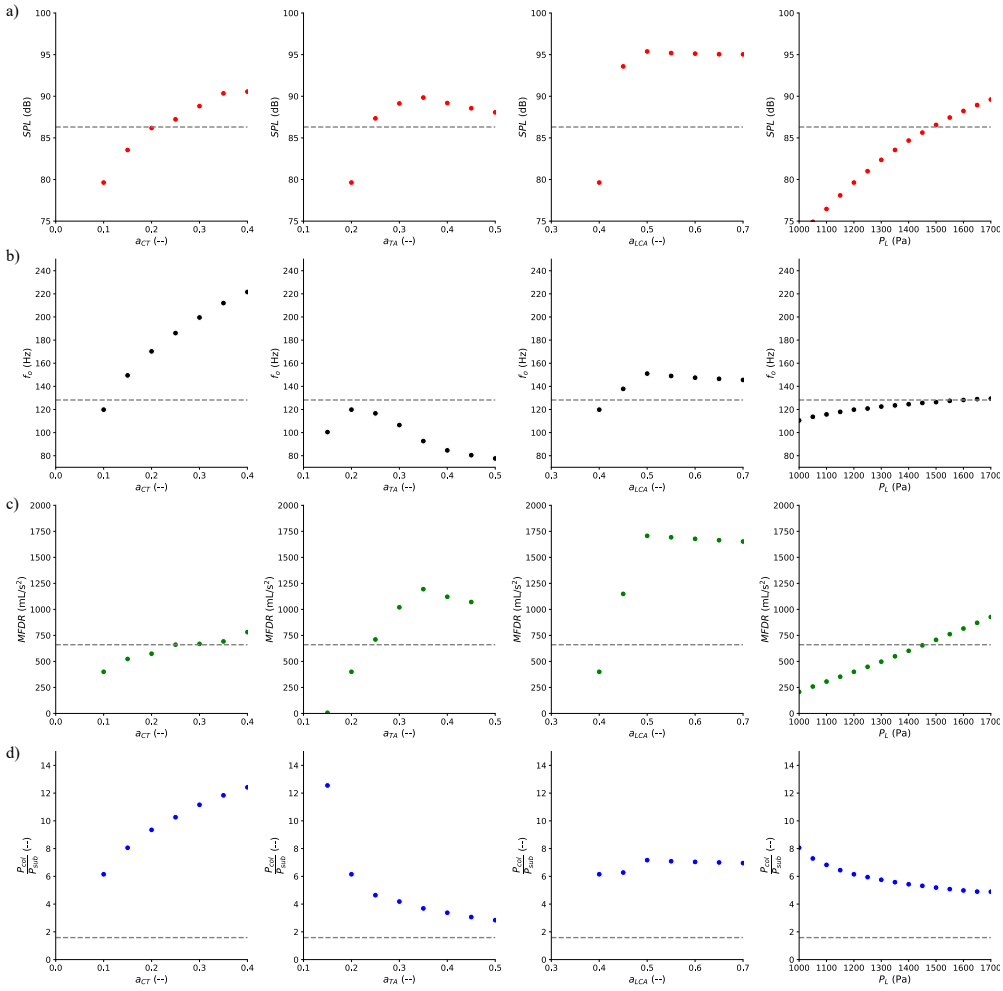

Figure 5.8: Effect of one variable compensation in an abducted VF configuration with nodules. Effect on: (a) SPL, (b) f_o, (c) MFDR and (d) quotient of collision and subglottal pressure. the columns denote each input variable: muscle activations (a_{CT}, a_{TA}, a_{LCA}) and lung pressure (P_L).

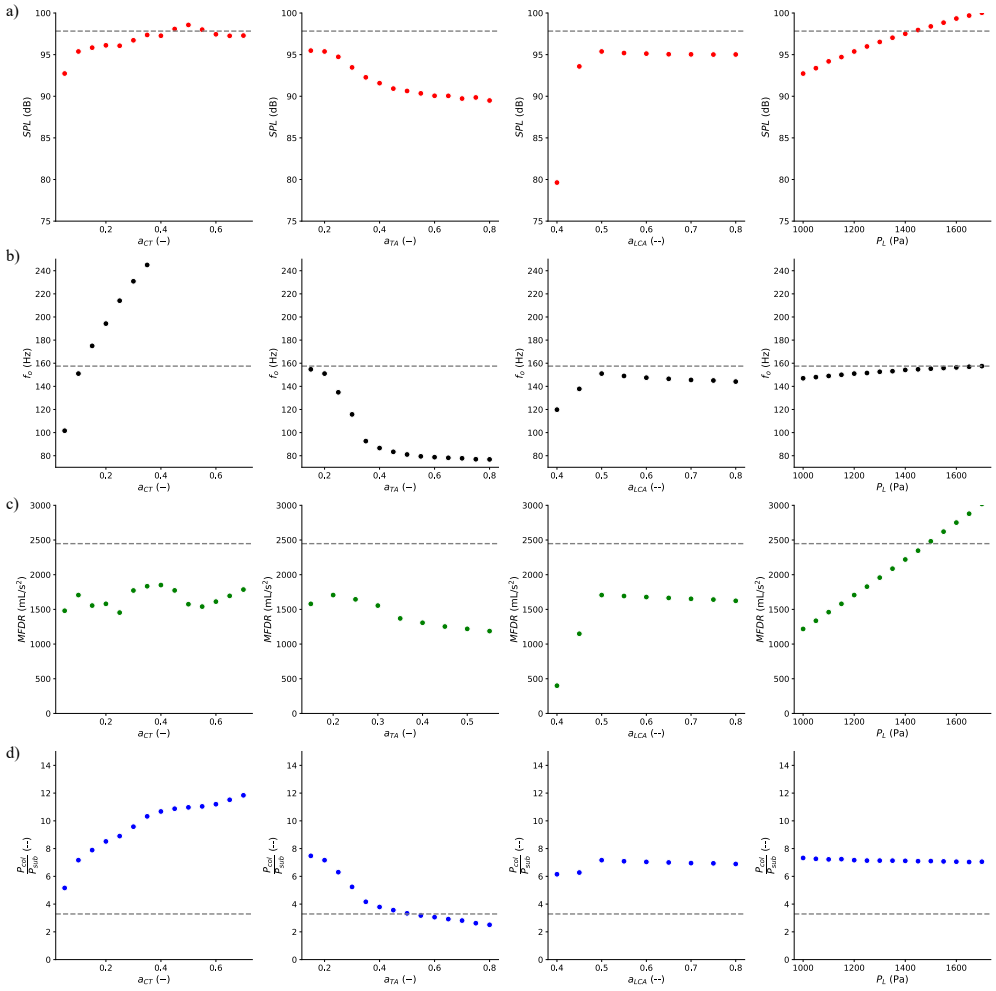

Figure 5.9: Effect of one variable compensation in an non-abducted VF configuration with nodules. Effect on: (a) SPL, (b) f_o, (c) MFDR and (d) quotient of collision and subglottal pressure. the columns denote each input variable: muscle activations (a_{CT}, a_{TA}, a_{LCA}) and lung pressure (P_L)..

For both vocal configurations, increasing P_L and the activation of the CT muscle leads to a noticeable rise in both f_o and SPL. However, the CT muscle proves to be more sensitive to changes in frequency (f_o), while P_L has a stronger impact on SPL, which aligns with previous experimental findings that highlight the role of CT as a tension regulator [36].

In the adducted configuration (Figure 5.9), the LCA muscle activation does not contribute significantly to changes in vocal parameters, as it reaches saturation, meaning its effect is diminished at higher levels of activation. On the other hand, the TA muscle shows only minor contributions across the board. In contrast, for the abducted configuration (Figure 5.8), increasing LCA activation effectively compensates by raising both SPL and f_o, consistent with the adducted hyperfunction profile commonly associated with PVH [5]. These compensatory behaviors mirror the observations made in excised larynx experiments, where the CT muscle increases tension, the TA muscle counteracts it in terms of frequency, and LCA exhibits adducting forces [36].

Shifting from the analysis of individual input effects on acoustic targets and aerodynamic parameters, we now present the results from the compensation experiment using the inverse Jacobian method to adjust for the presence of nodules. Table 5.5 displays the model input values (muscle activations and subglottal pressure), acoustic targets (SPL and f_o), and key aerodynamic variables (ACFL, MFDR, PGO, and the pressure quotient P_C/P_S) for the abducted and adducted

feature	CT act.	TA act.	LCA act.	P_L
SPL	+	o	+s	++
f_o	++	-s	+s	+o
MFDR	o	o	+s	++
P_C/P_S	++	-o	+s	-

Table 5.4: Summary of single compensation effect in nodule. ++: high increase. + increase. -: decrease. –: high decrease. +-s: increase/decrease and saturation. +-o: non-consistent increase/decrease. o: no effect.

configurations.

Several important patterns emerge from this compensatory analysis. The first observation is the marked increase in subglottal pressure, which has been previously reported in clinical data [10]. In both the abducted and adducted configurations, the compensation mechanism leads to a near-closure of the posterior glottal gap (PGO values approaching 0), consistent with what is observed in PVH, previously referred to as adducted VH [5]. There is also heightened activation of the CT and TA muscles, with a high degree of variability, consistent with EMG experiments reporting fluctuating muscle activations in hyperfunctional conditions [159, 160].

From an aerodynamic perspective, the compensated configurations show a

Variable	Unit	abducted			adducted		
		Control	Nodules		Control	Nodules	
			mean	std		mean	std
SPL	dB	86.31	86.29	0.36	97.83	97.84	0.26
f_o	Hz	128.20	128.22	0.54	157.50	157.52	0.37
CT act.	–	0.1	0.25	0.14	0.1	0.28	0.16
TA act.	–	0.2	0.48	0.19	0.2	0.46	0.18
LCA act.	–	0.4	0.52	0.10	0.5	0.50	0.08
P_S	Pa	897.0	1226.8	169.6	973.0	1276.5	200.8
ACFL	mL/s	505.62	593.98	97.08	486.49	590.95	84.14
MFDR	L/s^2	660.29	2329.47	631.61	2447.71	2729.89	403.34
P_C/P_S	–	1.58	3.43	1.07	3.1	3.46	1.02
PGO	mm^2	1.65	0.25	0.36	0	0.25	0.36

Table 5.5: Results of compensation mechanism for nodules.

higher MFDR compared to the control model, aligning with clinical measurements that suggest a greater impact of muscular tension and subglottal pressure in PVH cases [10]. Finally, the pressure quotient is significantly elevated when compensating by eliminating the posterior glottal gap. This, together with the elevated MFDR, suggests that the persistent trauma in PVH can be attributed to the increased collision forces and heightened aerodynamic stress on the Vfs.

These results shed light on the compensatory mechanisms in the presence of VF nodules. The increase in subglottal pressure and the elimination of the posterior gap are key factors in maintaining phonatory function, yet they come at the cost of elevated aerodynamic forces and increased collision pressure.

5.5 Discussion

The results presented indicate that the presence of a nodule leads to incomplete closure in the membranous component of the VF. This effect has been reported in previous modeling efforts of this pathology [49, 48, 57]. Nodule size is identified as the parameter with the most significant influence on phonation. Variations in nodule density do not have a noticeable effect on f_o and PTP. This finding aligns with the results reported in experiments with silicone models [114], where it was observed that while nodule size increases PTP, and y the large scale the elasticy of the nodule does not affect the f_o.

The decrease in f_o for nodule sizes between 0-2 mm was also reported in the finite element models [53, 52], which represented this type of pathology. The air gap is also reflected in the increase of the PTP value as a function of the nodule size, since it is the size of the nodule that defines the size of the gap.

Previous studies have modeled muscle activation under PVH conditions. In [42], the authors presented a compensation scheme in an extended model of the BCM with a triangular shape and posterior gap control, while in [57], the authors demonstrated the impact of varying muscle activations in a polyp model, showing how TA and LCA muscles stabilize VF oscillation. Both studies emphasized the influence of P_S and posterior gap size in generating significant compensatory effects.

Our results align with clinical findings in PVH, where both abducted and adducted configurations exhibit elevated P_S as a compensatory mechanism, consistent with clinical data reporting increased subglottal pressure in PVH [10]. The complete closure of the posterior glottal gap (PGO near 0) in both configurations reflects a key compensatory strategy, where eliminating any gap is crucial to sustain VF oscillation.

A key feature of this compensatory mechanism is the significant increase in MFDR, which mirrors clinical observations of higher MFDR in PVH [5, 10]. The elevated MFDR and increased LCA activation suggest that the vocal system compensates by increasing adduction and tension, which adds stress to the VFs. Ad-

ditionally, the pressure quotient (P_C/P_S) shows a substantial increase, indicating that higher aerodynamic forces are exerted on the VFs during compensation, further explaining the risk of VF trauma in PVH.

Despite its simplified nature, the lumped-element model effectively captures the critical aspects of how nodules impact vocal function and how compensatory strategies exacerbate these effects. However, this study examines only a single case, and a broader exploration of compensatory strategies across different configurations and pathologies would help strengthen these findings.

The current nodule model demonstrates the ability of muscle activation to jointly control VF viscoelasticity and posture in a bio-inspired framework. In individual compensation exercises, P_S and LCA activation emerge as the primary regulators of vocal function, while in joint compensation, LCA activation works together with P_S to maintain vocal quality (f_o and SPL). Increased muscle activation in the membranous portion of the VFs further highlights the compensatory role of TA and CT muscles. These findings are consistent with previous studies [57, 100, 42], showing that muscle activation compensates for both posterior and membranous gaps, with the LCA and TA playing dominant roles.

A notable limitation of this model is the decision to represent the nodule as a static element, focusing on its elastic and geometric contributions rather than adding dynamic degrees of freedom to the VF model. While this approach allowed us to identify trends consistent with previous studies [53, 52, 114, 100], it prevents

the analysis of oscillatory instabilities that more complex models can address [57, 48, 56]. Additionally, it restricts the ability to continue compensating for the membranous gap created by the nodule once the vocal folds are fully adducted. There is also the lack of a mechanism to simulate the occurrence and progression of nodules limits the ability to study their dynamic development over time. Future work should explore this by incorporating nodule formation mechanisms and more flexible VF representation, as has been done in finite element approaches [183, 184].

5.6 Chapter conclusions

The nodule TBCM model developed in this chapter successfully integrates nodules as fixed elements within the VF structure, capturing the effects of such pathologies on phonation dynamics. The primary findings highlight that the size of the nodule significantly influences vocal fold behavior, particularly in terms of PTP, MFDR, f_o and pressure quotient. An increase in nodule size directly leads to an increase in PTP, decreases in SPL, and changes in f_o and MFDR. This is attributed to the incomplete glottal closure caused by the nodule, which not only modifies the collision pattern of the VFs, increasing the pressure quotient, but also introduces a persistent air gap that impacts vocal quality. On the other hand, variations in nodule elasticity and density had minimal impact on phonatory

measures such as PTP or f_o, reinforcing that size is the primary determining factor in the acoustic and aerodynamic changes observed.

The chapter also explored how the phonatory system compensates for the presence of nodules by increasing subglottal pressure and adjusting muscle activation. This results have presented a compensation mechanism in the nodule model that differs significantly from the one observed in the asymmetry model discussed earlier. While both models share an increase in P_S as a compensatory response; the complete closure of the PGO and an adducted VF configuration are critical features in the compensation observed for nodules. Closing the gap is essential for maintaining phonation in cases of PVH, a behavior not seen in the NPVH model, where the presence of a posterior gap does not significantly impact compensation.

Additionally, in the nodule model, a notable increase in MFDR and the pressure quotient was observed, making the VFs more prone to trauma. This contrasts with the findings in the NPVH model, where the pressure quotient did not rise significantly, explaining why vocal fold trauma is not typically observed in NPVH. These differences between the two models underscore the importance of laryngeal posture and aerodynamic dynamics in compensating for these two conditions.

These results validate the hypothesis that compensatory mechanisms differ between PVH and NPVH. Specifically, the closure of the posterior gap, increased MFDR, and elevated pressure quotient in PVH provide a physiological explanation for the development of vocal fold trauma in these patients, aligning with clinical

observations regarding laryngeal posture and increased tension in the VFs.

Looking forward, the development of a dynamic nodule model—capable of representing transient states and fluctuating oscillation—remains an important step. Additionally, understanding the relationship between the mechanical properties of nodules and VF damage, particularly in the context of repeated phonotrauma, could provide a more comprehensive view of the pathological mechanisms underlying PVH. Integrating these elements into future models will enable deeper analysis of the interaction between VF biomechanics and pathological conditions like nodules.

Chapter 6

Future work and conclusions

6.1 Future work

This thesis has focused on the characterization of VH through physiologically inspired VF models. Given the growing interest in the mathematical modeling of voice production, it is natural for future research to extend these foundational elements and explore new avenues for model development and clinical application. One important direction for future research involves the validation of muscle activation levels used in these models. Muscle control, which governs both the posture and viscoelastic properties of the VFs, has been studied extensively, but the current models do not fully capture the precise level of activation needed to generate specific vocal behaviors. Previous research, [95, 98], laid the groundwork for understanding how muscle activation influences VF behavior, while the TBCM [44] has demonstrated how different levels of muscle activation can produce various vocal postures. However, future work will need to refine this aspect by incorporating direct measurements of intrinsic muscle activity, using techniques like intramuscu-

lar electromyography (EMG). This would allow for more accurate normalization values, providing individualized assessments that account for the unique tension patterns present in each subject's larynx.

In addition to validating muscle activation, there is a need to explore the dynamic role of muscle variation over time and its effect on phonation. While this thesis has included some time-varying elements, such as varying lung pressure in PTP calculations, the dynamic nature of phonatory gestures remains underexplored. Modeling vocal onset and offset, for example, could provide significant insights into the biomechanical underpinnings of VH, as parameters like relative fundamental frequency (RFF) have been shown to differ between healthy controls and patients [159, 160]. By simulating these transitions and comparing the results with clinical data, future work could deepen our understanding of how muscle activation patterns influence the onset and progression of vocal disorders.

Another promising avenue for future research is the identification of new model-based features that can serve as indicators of vocal pathology. Several studies have already demonstrated the potential of aerodynamic parameters, such as P_S (P_S), in distinguishing between healthy and hyperfunctional populations [10]. However, related parameters, such as collision pressure and dissipated power during VF oscillation, are more challenging to measure directly. Using the TBCM, we can estimate dissipated power as a function of P_S and explore its potential as a discriminating feature between healthy and pathological voice conditions. Prelim-

inary results suggest a quadratic relationship between dissipated power and P_S, which may provide a powerful indicator of vocal hyperfunction. Additionally, another important direction for future research is finding model-based features that correlate more effectively with TA muscle activation. Currently, estimating TA activation remains inconsistent, and identifying features that are more strongly linked to TA behavior could provide a more reliable and direct way to estimate its role in phonation, offering a clearer picture of its contribution to vocal pathologies.

Building on the work presented in chapter 3, future research could also focus on incorporating pathological VF models into ambulatory analysis. While the current methodology has proven effective for PVH, it is important to explore how the more complex pathological models developed in this thesis could be integrated into ambulatory monitoring frameworks. This would involve developing a systematic approach to determine when to apply a pathological model versus a control model, particularly in cases where the subject's hyperfunctional state remains unclear. By incorporating additional clinical measurements, such as video laryngoscopy, it may be possible to create a continuous scale of vocal health, offering a more nuanced understanding of the transition from healthy to pathological vocal states.

Finally, another area for improvement in the ambulatory analysis framework is the connection between the accelerometer (ACC) signal and the glottal airflow signal. In this thesis, the IBIF inverse filtering scheme was used to connect these signals, but a more integrated approach would involve using the ACC signal as

an output of the VF models themselves. By constructing a multi-layer propagation model that captures the relationship between glottal airflow and neck-skin acceleration, we could directly link model parameters to ambulatory data. This approach would allow for more accurate estimations of VF dynamics in real-world conditions and provide a robust foundation for future clinical diagnostics.

6.2 General conclusions

This thesis has provided a comprehensive exploration of VH through physiologically inspired VF models, establishing a methodological framework that integrates biomechanical and physiological insights into both laboratory and ambulatory settings. The research has made significant contributions across multiple areas, each addressing different aspects of VH and its representation in VF models.

The first major contribution was the adaptation of a stationary VF model, previously validated for sustained phonation, to the analysis of ambulatory data. By employing a statistical sampling scheme, the research demonstrated that VF models can effectively replicate ambulatory data distributions and estimate physiological parameters such as P_S and muscle activation. This supports Hypothesis 1, as the study successfully differentiated between control subjects and PVH patients, showing that PVH is characterized by increased P_S and consistent laryngeal muscle activation. This contribution marks a key advancement in the integration

of biomechanical models into real-world monitoring tools, opening avenues for future work that could include additional features and refine the models for more diverse populations. The success of this methodology opens avenues for future work, particularly in refining the model to better represent diverse populations, including those with specific vocal pathologies. Future research should explore how additional features can be integrated into this ambulatory framework to better capture the complexities of vocal function. An important challenge lies in the estimation of TA muscle activation. The current aerodynamic features struggles to constrain the behavior of the TA muscle accurately, introducing uncertainty into the estimation. This remains a key area for future improvement, as resolving this will allow for a more reliable characterization of TA behavior in VH.

Additionally, the development and testing of new VF models that capture specific elements of VH, particularly asymmetry for NPVH and nodule behavior for PVH, provided key insights into the role of muscle activation in both biomechanical properties and VF posture. The nodule model for PVH used a fixed-mass approach to represent the nodule as a static element, focusing on how it alters VF mechanics. Despite this simplification, the model was able to replicate trends observed in clinical data, as well as in previous silicone models and numerical simulations. Specifically, the model showed how increased adductor muscle activation and P_S are necessary to compensate for the presence of nodules, aligning with clinical observations of elevated P_S in PVH patients. This compensation

strategy leads to greater mechanical stress on the VFs, as evidenced by the elevated MFDR and the increased pressure quotient, both of which are known risk factors for VF trauma.

The asymmetry model for NPVH introduced a new framework for understanding how muscle imbalances affect VF oscillation and lead to compensatory behavior. This study found that NPVH compensation primarily involves increased P_S and laryngeal tension but does not result in a significant increase in adduction, thus validating Hypothesis 2, which posits different compensation mechanisms between NPVH and PVH. The comparison between the two models highlighted key compensatory differences: PVH compensation requires complete closure of the posterior glottal gap and leads to elevated MFDR and a higher pressure quotient, which contributes to VF trauma. In contrast, NPVH compensatory strategies maintain a higher P_S without increasing the pressure quotient, dont need the adduction of the VFs offering an explanation for why VF trauma is absent in NPVH patients despite the presence of elevated P_S and muscle tension.

The compensation study in both models revealed important distinctions between the two types of hyperfunction, particularly regarding the degree of adduction and the pressure quotient. These findings underscore the importance of understanding how different VH conditions influence the underlying compensatory mechanisms and their potential impact on VF health.

Each chapter contributes essential building blocks for the next stages of VF

modeling and clinical application. The ambulatory data integration establishes a foundation for future work, where new acoustic or aerodynamic features could be incorporated into the analysis. The asymmetrical model introduces a framework for exploring muscle imbalance not just in NPVH but also in other vocal pathologies, highlighting the need for dynamic, patient-specific modeling in future research. Lastly, the nodule model provides a simplified but effective way to model VF trauma, which could be expanded with more sophisticated simulations and experimental validation.

By addressing these critical areas, this thesis not only advances the current understanding of VH and its manifestations but also sets the stage for more refined and clinically relevant modeling approaches. Future efforts should focus on validating these models with clinical data, exploring dynamic control schemes, and integrating more detailed physiological measurements, such as intramuscular EMG or high-speed video analysis, to bridge the gap between mathematical models and clinical practice.

Appendix A
Implementation of the a-TBCM

The proposed a-TBCM is, an extension of the TBCM [42, 44], builds upon prior efforts [90, 95, 98, 42, 84]. The proposal moves from a symmetrical to an asymmetrical glottal configuration considering the left and right VFs with two VFs with their independent muscle activation vector.

A.0.1 Posture and biomechanical properties

For each side of the a-TBCM, as shown in Figure 4.1, the dynamic adjustment of vocal posture and VF configuration as a function of activation vector $\vec{a}_i$ in Equation 4.1 follows the methodology in the TBCM. The theoretical development and the implementations were described in [42, 44], and are briefly summarized for completeness.

A system of equations of motion models the laryngeal posturing by describing the relative movements between the arytenoid cartilage and the cricoid cartilage, and between the cricoid cartilage and the thyroid cartilage, in response to the forces in the laryngeal tissues. Vocal process Cartesian coordinates are tracked,

and from this information and the forces in the laryngeal tissues, the adductory displacements, $\Delta x_{u,i}, \Delta x_{l,i}$, and the VF lengths, L_i, for $i \in \{L, R\}$ are obtained (see Figure A.1).

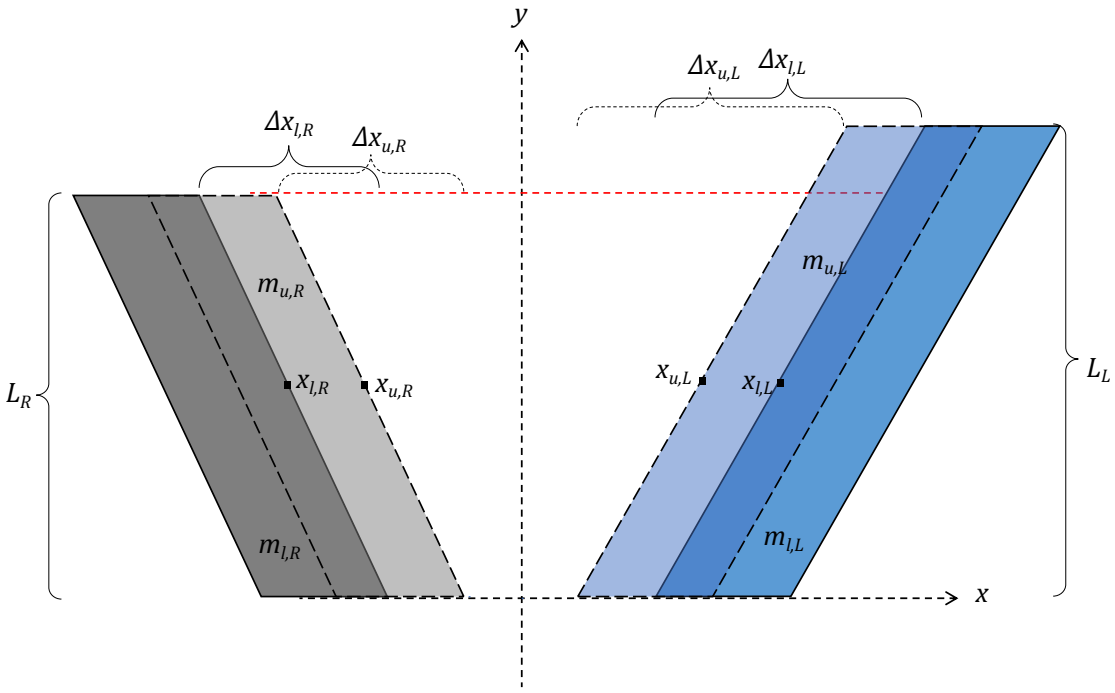

Figure A.1: (Color online) Top view: 2D diagram describing the abducted VFs positioning for the cover blocks in the a-TBCM.

Thereupon, empirical rules are applied for adjusting the geometrical and biomechanical parameters [95, 44] (e.g., thickness, depth, and mass m for each block, the nodal point, the glottal convergence, and the values for the spring k and damping d parameters (see Figure 4.1)) of the left/right TBCM.

A.0.2 Glottal areas calculation

The total glottal area A_g is the contribution of the membrane portion of the VFs (MGO) and the posterior gap (PGO), as shown in the Figure 4.1.

In the a-TBCM, the MGO for upper or lower blocks is computed as follows:

$$A_m = A_{Tr_R} + A_{Tr_L} + A_{Rect}, \tag{A.1}$$

where the subscripts indicate whether the area is triangular, Tr, or rectangular, $Rect$. For the computation of these areas, the fraction of the block that is under collision is introduced as a quantity that simplifies the expressions [84, 42]:

$$\alpha_i = \max\left(0, \min\left(\frac{y_{col}}{L_i}, 1\right)\right) \; i \in \{\text{L, R}\}. \tag{A.2}$$

The collision height y_{col} in Equation A.2 denotes the y-coordinate of the point where the left- and right-edge lines intercept determines the collision height (see the red dot in Figure A.2). There are two possible scenarios for the glottis, as shown in Figure A.2. (top) VF in no collision: ($\alpha = 0$), and (bottom) VFs in collision: ($\alpha > 0$).

The triangular-shaped glottal area is calculated as follows:

$$A_{Tr_i} = (1 - \alpha_i)^2 L_i \frac{\Delta x_i}{2}, \tag{A.3}$$

where factor $(1 - \alpha_i)$ is the non-collision portion of the VF, and the subscript $i \in \{L, R\}$. On the other hand, the rectangular-shaped area is obtained from:

$$A_{Rect} = \max\left(0, x_{B_R} - x_{B_L}\right) \frac{L_R + L_L}{2}. \tag{A.4}$$

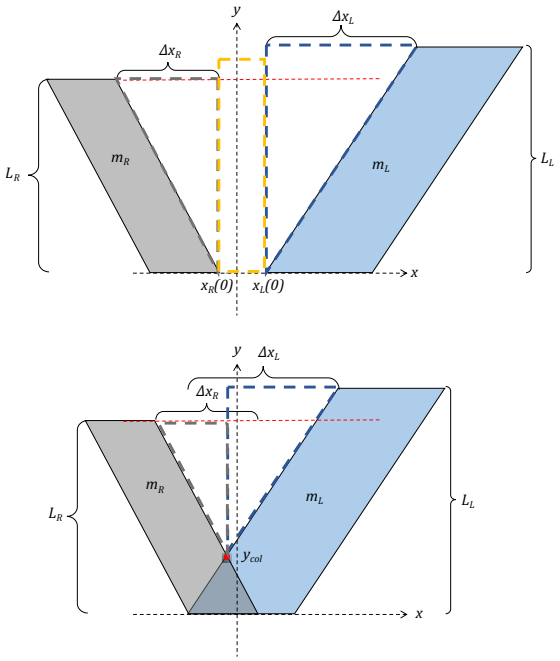

Figure A.2: (Color online) Glottal area scheme for VF blocks: (top) no collision, (bottom) collision.

Note that during VF collision, $A_{Rect} = 0$ in Eq. (A.1).

The cross-sectional area, A_t, and the contact area, A_c, are obtained from the collision fraction, α_i, in a similar way as in [44, 42]:

$$A_{t_i} = (1 - \alpha_i) L_i T_i, \tag{A.5}$$

$$A_{c_i} = \alpha_i L_i T_i, \tag{A.6}$$

where $i \in \{L, R\}$.

The posterior glottal opening, PGO, is defined for each VF following [44], using the cricoarytenoid junction coordinate. In the a-TBCM, the PGO is the

contribution of both VFs:

$$PGO = PGO_R + PGO_L. \tag{A.7}$$

The total glottal area is the sum of the membranous area and the posterior gap:

$$A_g = A_M + PGO, \tag{A.8}$$

where $A_M = \min(A_{m_u}, A_{m_l})$, the minimum between upper and lower MGO, as shown in Figure A.1.

A.0.3 Equations of motion

The asymmetrical VF vibrations are simulated on the basis of coupled left/right systems of equations of motion. Each system simulates the medial-lateral displacements for the upper (u) and lower (l) cover masses and the body (b) mass in the TBCM (see Figure 4.1).

The equations of motion for each VF are:

$$m_{u,i}\ddot{x}_{u,i} = F_{k;u,i} + F_{d;u,i} - F_{k;c,i} + F_{e;u,i} + F_{col;u,i}, \tag{A.9a}$$

$$m_{l,i}\ddot{x}_{l,i} = F_{k;l,i} + F_{d;l,i} - F_{k;c,i} + F_{e;l,i} + F_{col;l,i}, \tag{A.9b}$$

$$m_{b,i}\ddot{x}_{b,i} = F_{k;b,i} + F_{d;b,i} - F_{k;u,i} - F_{d;u,i} - F_{k;l,i} - F_{d;l,i}, \tag{A.9c}$$

where i denotes the considered side (L or R), m is the mass of the block and x is the medial-lateral displacement over time. The right side of the equation presents

the net force acting upon the block due to elastic (k), damping (d), aerodynamic (e), and collision (col) components.

The elastic forces ($F_{k;u,i}$, $F_{k;l,i}$, $F_{k;c,i}$, $F_{k;b,i}$) are modeled through a nonlinear Hooke's law, and the damping forces ($F_{d;u,i}$, $F_{d;l,i}$, $F_{d;b,i}$) are modeled proportional to the velocity. These force components are not affected by the presence of the other VF; the rules for calculation are the same as in [90, 42]. The explicit equation for each force component can be found in the appendix in [42].

The aerodynamic driving forces (F_{eu_i}, F_{el_i}) represents the force that the intraglottal pressure exerts on the VFs. It depends on the glottal flow, glottal and transversal area. Previous works explain how to calculate the intraglottal flow and pressure from the upper/lower glottal area [64]. Using the definition of pressure, the aerodynamic force for the upper/lower block is:

$$F_{aer} = P_{int}(A_g) \times A_t, \tag{A.10}$$

where P_{int} is the intraglottal pressure (constant in the entire upper or lower block) which is a function of glottal area A_g and the sub- and supra-glottal pressures, and A_t is the cross-sectional area that multiplies the pressure to compute the force component, which depends on the geometry and collision fraction α of the VF. The equations from [42, 81] are used to calculate the intraglottal pressure from the glottal area; however, it is necessary to define how to calculate the glottal and cross-sectional areas in the a-TBCM due to the lack of mirror symmetry.

The collision forces ($F_{col;u,i},\ F_{col;l,i}$) depend on the interpenetration distances in the upper/lower cover masses due to the impact between opposing VFs. Therefore, the computation of the collision force requires defining when the collision occurs, and the VF section undergoing the collision.

Similar to [42], the total collision force for the lower or upper blocks is:

$$F_{col;j,i} = -k^*_{col;j} \int_0^{y_{col;j}} \left(d_{j,i}(y) + \eta_{col;j} d^3_{j,i}(y)\right) dy, \tag{A.11}$$

where $i \in \{\text{L, R}\}$, $j \in \{\text{l, u}\}$, $d(y)$ is the interblock penetration distance, y_{col} is the collision height, k^*_{col} is the effective spring collision constant, and η_{col} is a nonlinear coefficient. To calculate the interblock penetration distance $d(y)$, the VF edge is described as a straight line. The x position of the VF is a function of the y coordinate, as follows:

$$x_{j,i}(y) = b_{j,i} y + c_{j,i}, \quad i \in \{\text{L, R}\} \text{ and } j \in \{\text{l, u}\}, \tag{A.12}$$

where b is the slope and c is the intercept, see below. The explicit values for quantities are obtained from Figure A.1 that shows a 2D top view of the cover masses, where L_i is the VF length, $x_{j,i}$ is the mass position and $\Delta x_{i,j}$ is posterior displacement, given the degree of abduction by the muscle activation.

$$b_{j,i} = \begin{cases} -\frac{\Delta x_{j,L}}{L_L} & \text{for } i = L \text{ and } j \in \{\text{l, u}\} \\ \frac{\Delta x_{j,R}}{L_R} & \text{for } i = R \text{ and } j \in \{\text{l, u}\} \end{cases}, \tag{A.13}$$

$$c_{j,i} = \begin{cases} x_{j,L} + \frac{\Delta x_{j,L}}{2} & \text{for } i = L \text{ and } j \in \{\text{l, u}\} \\ x_{j,R} - \frac{\Delta x_{j,R}}{2} & \text{for } i = R \text{ and } j \in \{\text{l, u}\} \end{cases}. \tag{A.14}$$

Without loss of generality, consider one of the cover elements, either upper or lower for both VFs; the upper or lower subscript is removed to have a short notation since the expression and the formulation is equivalent for both blocks. With the mathematical description of the VF edges in Equation A.12, the interblock penetration distance is defined by:

$$d_R(y) = x_R(y) - x_L(y), \tag{A.15}$$

note that $d_L(y) = -d_R(y)$, this denotes the opposite direction in collision forces. Additionally, this distance gives the collision condition:

$$d_R(0) \geq 0. \tag{A.16}$$

For the calculation of the collision height y_{col} in Equation A.11, three possible glottal configurations are considered (see Figure A.3): the posterior portions in both VFs are abducted ($\Delta x_{\mathrm{R}}, \Delta x_{\mathrm{L}} > 0$), the right VF is adducted and the left one is medialized ($\Delta x_{\mathrm{R}} > 0$ and $\Delta x_{\mathrm{L}} = 0$) and *vice versa*, and both VFs are tightly adducted ($\Delta x_{\mathrm{R}}, \Delta x_{\mathrm{L}} = 0$).

For the cases where at least one of the VFs has an adduction degree:($\Delta x_i \neq 0$), the collision height is calculated from the interpenetration distance condition:

$$d_R\left(y_{col}\right) = 0, \tag{A.17}$$

with some algebra:

$$y_{col} = \frac{c_L - c_R}{b_R - b_L} = -\frac{c^*}{b^*}. \tag{A.18}$$

For the case of parallel VFs, i.e., the case for $\Delta x_L = \Delta x_R = 0$. The collision height is calculated simply based on the collision condition and the length of the VF:

$$y_{col} = \begin{cases} \min(L_L, L_R) & \text{if } d_R(0) \geq 0 \\ 0 & \text{if } d_R(0) < 0 \end{cases}. \tag{A.19}$$

The effective spring collision constant in Equation A.11 is computed assuming an in-series spring configuration:

$$\frac{1}{k^*_{col}} = \frac{1}{k_{col_L}} + \frac{1}{k_{col_R}}, \tag{A.20}$$

which comprises the contributions from the left- and right-side collision springs. The values for k_{col_R}, k_{col_L} and η_{col} are computed following [42].

Replacing the Eqs. (A.15) and (A.20) in the integral in Eq. (A.11) the total collision force can be computed by:

$$\begin{aligned} F_{col} = &- \frac{k^*_{col}}{4} y_{col} \left(2c^* + b^* y_{col}\right) \times \\ &\left[2 + \eta_{col} \left(2c^{*2} + 2b^* c^* y_{col} + b^{*2} y^2_{col}\right)\right]. \end{aligned} \tag{A.21}$$

Note that the quantities c^*, b^*, k^*_{col}, and y_{col} have information on the properties of both VFs.

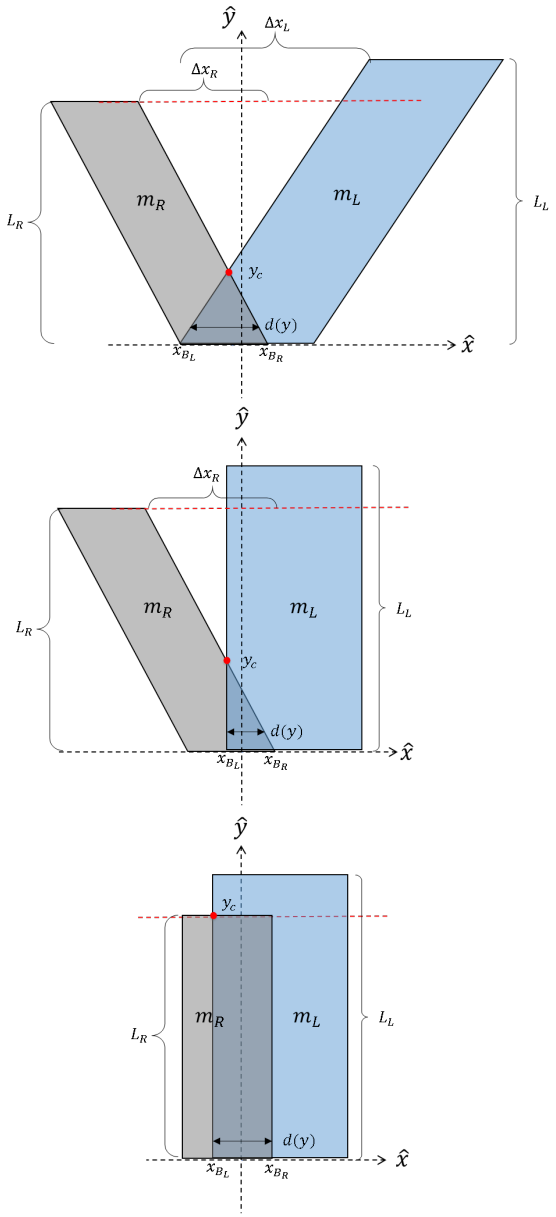

Figure A.3: (Color online) The three collision scenarios in the a-TBCM determined by the left/right posterior displacements: (Top) case $\Delta x_{\mathrm{R}}, \Delta x_{\mathrm{L}} > 0$, (Middle) $\Delta x_{\mathrm{R}} > 0$ and $\Delta x_{\mathrm{L}} = 0$ (and *vice versa*), and (Bottom) $\Delta x_{\mathrm{R}}, \Delta x_{\mathrm{L}} = 0$.

Appendix B
Implementation of the n-TBCM

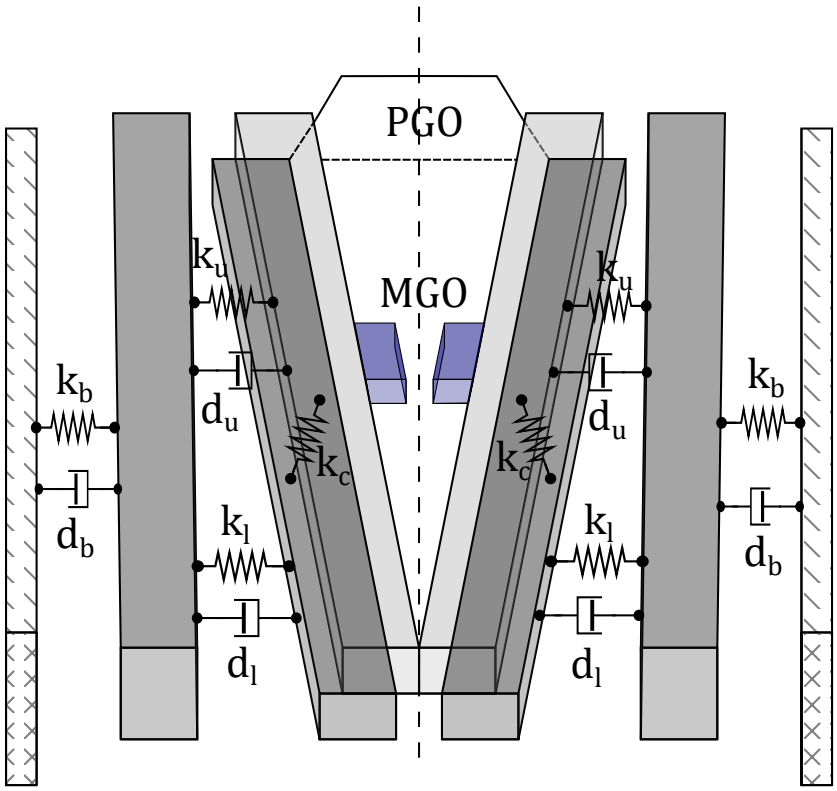

Figure B.1: Vocal fold nodule approach scheme. Gray blocks represent the base TBCM scheme. Blue blocks represent the nodules

Figure B.1 shows the scheme of nodule approach in the TBCM as a fixed shift in length distribution. Figure B.2 shows the sections for nodule implementation

and some important positions in the nodule. In this approach, nodules appear as a new element in the vocal fold. This new element doesn't have movement but changes the collision condition, glottal area calculation and vocal fold mass and stiffness.

Different geometries for the nodule are proposed here: Rectangular, Cylindrical and Spherical. All the definitions and changes that will be shown next can be applied to both vocal fold masses (lower and upper).

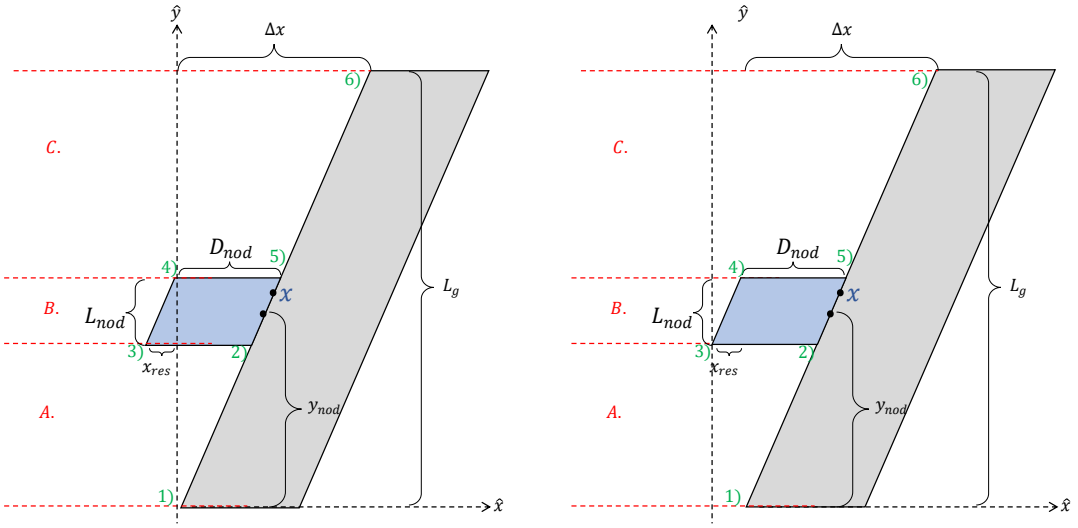

Figure B.2: Vocal fold with nodule as a shift in the length distribution. (Left) Base position . (Right) new rest position

B.1 Model parameters

In our approach for vocal folds nodules. The model add these nodules parameters:

- Nodule position ratio λ, from 0 to 1.
- Nodule nodule length (Radius) $R_n = D_n$, $L_n = 2R_n$.
- Nodule density ρ_n.
- Coupling spring gain β.
- Elastic properties for nodule: Nodule Young Modulus E_n and Poisson Ratio ν_n.

B.2 Definitions

B.2.1 Nodule position

One of the approach parameters is the nodule position y_{nod}:

$$y_{nod} = \frac{L_{nod}}{2} + \lambda \left(L_g - L_{nod} \right) \tag{B.1}$$

where λ is the fraction of position in function of vocal fold length, meaning $\lambda = 0$ bottom and $\lambda = 1$ top.

B.2.2 Points position

For an easier write of the changes in TBCM, we define some positions that will be linked to the equations, the position represented in the Figure B.2.

1. Posterior position.

$$x_1 = x - \frac{\Delta x}{2} \tag{B.2}$$

2. Upper A section position.

   For this position, we start with x_1 position plus a fix distance, this distance is calculated using triangle similarity.

$$x_2 = x_1 + x' \tag{B.3}$$

$$\frac{x'}{y_{nod} - L_{nod}/2} = \frac{\Delta x}{L_g} \rightarrow x' = \frac{\Delta x}{L_g}\left(y_{nod} - L_{nod}/2\right) \tag{B.4}$$

$$x_2 = x - \frac{\Delta x}{2} + \left(y_{nod} - \frac{L_{nod}}{2}\right)\frac{\Delta x}{L_g} \tag{B.5}$$

$$x_2 = x + \Delta x\left[\left(\lambda - \frac{1}{2}\right) - \lambda\frac{L_{nod}}{L_g}\right] \tag{B.6}$$

3. Lower nodule position.

$$x_3 = x_2 - D_{nod} \tag{B.7}$$

$$x_3 = x + \Delta x\left[\left(r - \frac{1}{2}\right) - \lambda\frac{L_{nod}}{L_g}\right] - D_{nod} \tag{B.8}$$

4. Upper nodule position.

   We use another triangle similarity. to move from x_3 to x_4.

$$x_4 = x_3 + x'', \quad \frac{x''}{L_{nod}} = \frac{\Delta x}{L_g} \tag{B.9}$$

$$x_4 = x_3 + \frac{L_{nod}}{L_g}\Delta x \tag{B.10}$$

$$x_4 = x + \Delta x \left[\left(\lambda - \frac{1}{2}\right) + (1-\lambda)\frac{L_{nod}}{L_g}\right] - D_{nod} \tag{B.11}$$

5. Lower C section position.

$$x_5 = x_4 + D_{nod} \tag{B.12}$$

$$x_5 = x + \Delta x \left[\left(\lambda - \frac{1}{2}\right) + (1-\lambda)\frac{L_{nod}}{L_g}\right] \tag{B.13}$$

6. Anterior position.

$$x_6 = x + \frac{\Delta x}{2} \tag{B.14}$$

B.2.3 Section parameters

Figure B.3 shows a section collision and how the area is defined for times in collision $\alpha \neq 0$ and free $\alpha = 0$.

To the area, collision and other equation we use the section parameters: base position x_{base}, length x_{len} and height L. These parameters can be write with the points position.

- Height

$$L = \begin{cases} y_{nod} - \frac{L_{nod}}{2} & \text{for A} \\ L_{nod} & \text{for B} \\ L_g - \left(y_{nod} + \frac{L_{nod}}{2}\right) & \text{for C} \end{cases} \tag{B.15}$$

- Length

$$x_{len} = \begin{cases} x_2 - x_1 & \text{for A} \\ x_4 - x_3 & \text{for B} \\ x_6 - x_5 & \text{for C} \end{cases} \tag{B.16}$$

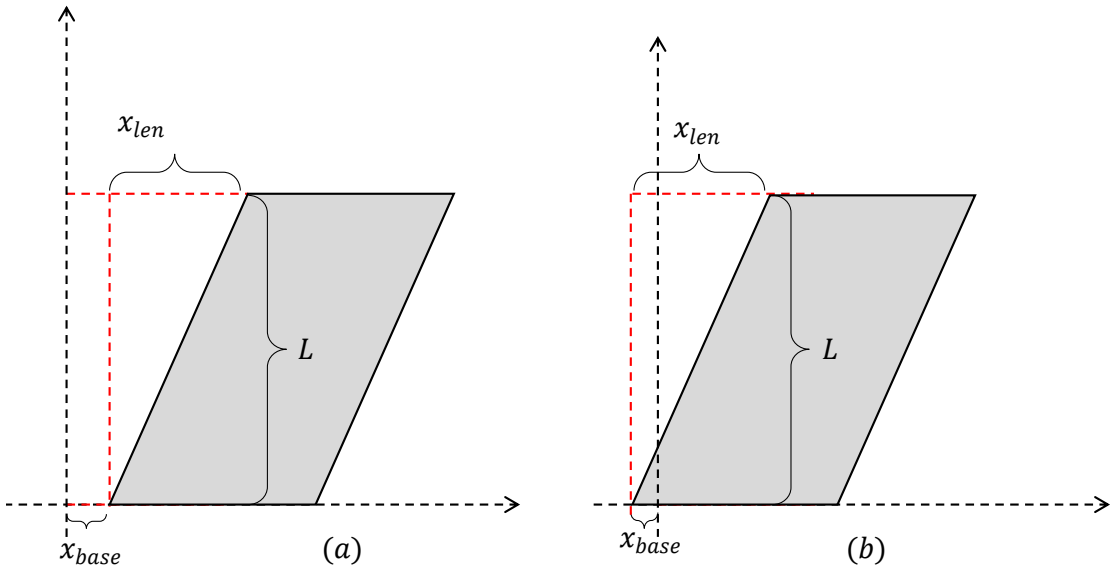

Figure B.3: Vocal fold section area. (a) $\alpha = 0$, (b) $\alpha \neq 0$.

- Base position

$$x_{base} = \begin{cases} x_1 & \text{for A} \\ x_3 & \text{for B} \\ x_5 & \text{for C} \end{cases} \tag{B.17}$$

Replace the value of points position:

- Height

$$L = \begin{cases} \lambda\left(L_g - L_{nod}\right) & \text{for A} \\ L_{nod} & \text{for B} \\ (1-\lambda)\left(L_g - L_{nod}\right) & \text{for C} \end{cases} \tag{B.18}$$

- Length

$$x_{len} = \begin{cases} \lambda\left(1 - \dfrac{L_{nod}}{L_g}\right)\Delta x & \text{for A} \\ \dfrac{L_{nod}}{L_g}\Delta x & \text{for B} \\ (1-\lambda)\left(1 - \dfrac{L_{nod}}{L_g}\right)\Delta x & \text{for C} \end{cases} \tag{B.19}$$

- Base position

$$x_{base} = \begin{cases} x - \frac{\Delta x}{2} & \text{for A} \\ x + \Delta x \left[\left(\lambda - \frac{1}{2} \right) - \lambda \frac{L_{nod}}{L_g} \right] - D_{nod} & \text{for B} \\ x + \Delta x \left[\left(\lambda - \frac{1}{2} \right) + (1 - \lambda) \frac{L_{nod}}{L_g} \right] & \text{for C} \end{cases} \tag{B.20}$$

B.3 Changes in TBCM implementation

B.3.1 Equilibrium position of cover masses

In a tight vocal process like $\Delta x = 0$, the nodule produce a collision in the equilibrium position, for that reason the new relaxing position will be different to posterior position equal to 0.

$$x_{res} = Max \left[0, -x_3 \left(\frac{\Delta x}{2} \right) \right] = Max \left[0, - \left(\Delta x \lambda \left[1 - \frac{L_{nod}}{L_g} \right] - D_{nod} \right) \right] \tag{B.21}$$

this x_{res} is the lower nodule position when the vocal fold is in $x = \Delta x/2$, the Max function is used to keep the displacement equal to zero when the nodule don't cross the media line. So, we define the total equilibrium position and displacement:

$$\hat{x} = x - \left(\frac{\Delta x}{2} + x_{res} \right) \tag{B.22}$$

As a parameter in me implementation we have a switch for this additional rest position. to use or not.

B.3.2 Collision fraction

In figure B.3 appear α as the percentage that indicates how much surface of the mass is actually colliding. To define this proportion, we will do the difference between $\Delta x = 0$ and $\Delta x \neq 0$.

A general way to write it is:

$$\alpha = \begin{cases} (x < 0) & \Delta x = 0 \\ Max\left(0, Min\left(-\frac{x_{base}}{x_{len}}, 1\right)\right) & \Delta x \neq 0 \end{cases} \tag{B.23}$$

The Min function limit the proportion to a maximum value of 1, the Max function regret negative values, this values happen when the section don't cross the y-axis. Replace the value using **section parameters**.

$$\alpha_A = \begin{cases} (x < 0) & \Delta x = 0 \\ Max\left(0, Min\left(-\frac{x - \Delta x/2}{\lambda\left(1 - \frac{L_{nod}}{L_g}\right)\Delta x}, 1\right)\right) & \Delta x \neq 0 \end{cases} \tag{B.24}$$

$$\alpha_B = \begin{cases} (x - D_{nod} < 0) & \Delta x = 0 \\ Max\left(0, Min\left(-\frac{x + \Delta x\left[\left(\lambda - \frac{1}{2}\right) - \lambda\frac{L_{nod}}{L_g}\right] - D_{nod}}{\frac{L_{nod}}{L_g}\Delta x}, 1\right)\right) & \Delta x \neq 0 \end{cases} \tag{B.25}$$

$$\alpha_C = \begin{cases} (x < 0) & \Delta x = 0 \\ Max\left(0, Min\left(-\frac{x + \Delta x\left[\left(\lambda - \frac{1}{2}\right) + (1-\lambda)\frac{L_{nod}}{L_g}\right]}{(1-\lambda)\left(1 - \frac{L_{nod}}{L_g}\right)\Delta x}, 1\right)\right) & \Delta x \neq 0 \end{cases} \tag{B.26}$$

B.3.3 Glottal area

The glottal area will be the sum of the three component, for each one we define the area like a sum of a triangle and a rectangle, as in TBCM implementation. In the figure B.3 we have the two possibilities for a section. $\alpha = 0$ and $\alpha \neq 0$.
The area for the section is:

$$Area = 2(1-\alpha)L\left[Max(0, x_{base}) + (1-\alpha)x_{len}/2\right] \tag{B.27}$$

The factor $(1-\alpha)$ means the portion of h and x_{lon} free. The glottal area for a mass in the vocal fold, is the sum of the three sections area.

The total glottal area in n-TBCM is:

$$A_g = \sum_{i=\text{A, B, C}} 2(1-\alpha_i)L\left[Max(0, x_{base_i}) + (1-\alpha)x_{len_i}/2\right] \tag{B.28}$$

B.3.4 Collision force

To represent the collision force when there is a nodule, we consider it as an additional force.

The nodule collision force is equivalent to the vocal fold collision, resented in equation (2.34):

$$F_{col} = -k_c \int_0^{y_c} \left(x(y) + \eta_c x^3(y)\right) dy, \tag{B.29}$$

using the figure B.3 we describe the position position as:

$$x(y) = x_{base} + \frac{x_{len}}{h} y. \tag{B.30}$$

Solving the integral is:

$$F_{col} = -\frac{1}{4} k_c \alpha \left(2x_{base} + \alpha x_{len}\right) \left(2 + \eta_c \left[2x_{base}^2 + 2\alpha x_{base} x_{len} + \left(\alpha x_{len}\right)^2\right]\right) \tag{B.31}$$

The value for x_{base} and x_{len} for each section are in equation (B.20) and (B.19). k_c is the elastic coefficient not per unit of length, $k_c = \hat{k}_c L_g$. The total collision force is the sum of each section force. To compute this force, we create a function for B.31.

$$F_{col} = f\left(\alpha, x_{base}, x_{len}, k_c, \eta_c\right)$$

For the nodule, the collision force have a maximum value, when the entire nodule is in collision, this is computed with an effective x_{base}: We use the traditional force with a limit in the x_{base} position.

$$x'_{base} = max\left[x_{base}, -D_{nod}\right] \tag{B.32}$$

B.3.5 Collision pressure - contact area

To define the pressure made for the collision force, we need an expression for the contact area, this expression is similar to the glottal area but with thickness dimension.

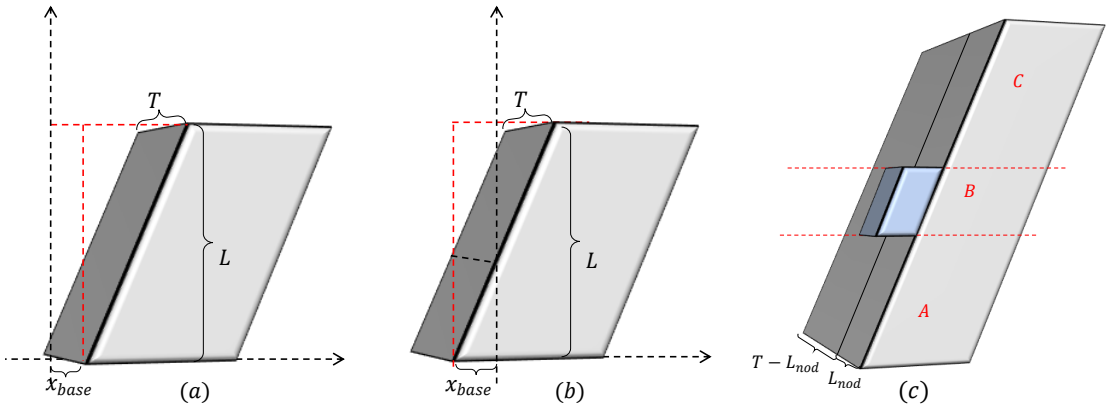

Figure B.4: 3D representation: Section transverse area: (a) $\alpha = 0$, (b) $\alpha \neq 0$. (c) Full block view

Looking the Figure B.4 (c), we understand the contact and transverse area is define into two sections. One with depth $T - L_{nod}$, this section uses TBCM normal equations and the second section uses the nodule distribution in three section.

Contact area for one section is:

$$A_c = \alpha L_g T \tag{B.33}$$

the total contact area is the sum of all sections.

$$\begin{aligned} A_c &= A_{TBCM} + A_{nod} \\ &= \alpha L_g (T - L_{nod}) + L_{nod} \sum_{i=A,B,C} \alpha_i L_i \end{aligned} \tag{B.34}$$

B.3.6 Aerodynamic force - transverse area

The sub-glottal flow do a force in the vocal folds, this force is pressure applied in the transverse area for each fold. For the transverse area we have an expression

similar to contact pressure:

For one section, is equal to (2.39):

$$A_t = (1 - \alpha)LT \tag{B.35}$$

The total transverse area is:

$$A_c = (1 - \alpha)L_g(T - L_{nod}) + L_{nod} \sum_{i=A,B,C} (1 - \alpha_i)L_i \tag{B.36}$$

B.3.7 Vocal fold parameters

The last two changes in the TBCM is the increase of mass and the spring upper block.

- Nodule spring constant, flowwing [95]:

$$k_n = \frac{E_n}{(1 + \nu_n)} D_n \tag{B.37}$$

  Here we use elastic and geometrial properties of the nodule and the relation between shear modulus, Young modulus and Poisson ratio.

- Increase in mass:

$$m' = m + \rho_{nod} D_{nod} L_{nod}^2 \tag{B.38}$$

  Here we use the water density as nodule density reference.

- Increase in coupling spring:

$$k_c' = (1 + \beta)\, k_c \tag{B.39}$$

In this equation define a gain β.

Bibliography

[1] Janet A. Wilson, Ian J. Deary, Audrey Millar, and Kenneth Mackenzie. The quality of life impact of dysphonia. *Clinical Otolaryngology and Allied Sciences*, 27(3):179–182, 2002.

[2] Pinaki Bhattacharya and Thomas Siegmund. The role of glottal surface adhesion on vocal folds biomechanics. *Biomechanics and Modeling in Mechanobiology*, 14(2):283–295, 2014.

[3] 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.

[4] Adrián Castillo, César Casanova, Daniel Valenzuela, and Sebastián Castañón. Prevalencia de disfonía en profesores de colegios de la comuna de Santiago y factores de riesgo asociados. *Ciencia & Trabajo*, 17(52):15–21, 2015.

[5] 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.

[6] 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*, 29(4):2254–2260, 2020.

[7] Lesly Wallis, Cristina Jackson-Menaldi, Wayne Holland, and Alvaro Giraldo. Vocal fold nodule vs. vocal fold polyp: answer from surgical pathologist and voice pathologist point of view. *Journal of Voice*, 18(1):125–129, 2004.

[8] Octavio Garaycochea, Juan Manuel Alcalde Navarrete, Beatriz Del Río, and Secundino Fernández. Muscle tension dysphonia: which laryngoscopic features can we rely on for diagnosis? *Journal of Voice*, 33(5):812–e15, 2019.

[9] Maude Desjardins, Christopher Apfelbach, Marianna Rubino, and Katherine Verdolini Abbott. Integrative review and framework of suggested mechanisms in primary muscle tension dysphonia. *Journal of Speech, Language, and Hearing Research*, 65(5):1867–1893, 2022.

[10] 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.

[11] Víctor M. Espinoza, Daryush D. Mehta, Jarrad H. Van Stan, Robert E. Hillman, and Matías Zañartu. Glottal aerodynamics estimated from neck-surface vibration in women with phonotraumatic and nonphonotraumatic vocal hyperfunction. *Journal of Speech, Language, and Hearing Research*, 63(9):2861–2869, 2020.

[12] 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.

[13] Harold A. Cheyne, Helen M. Hanson, Ronald P. Genereux, Kenneth N. Stevens, and Robert E. Hillman. Development and testing of a portable vocal accumulator. *Journal of Speech, Language, and Hearing Research*, 46(6):1457–1467, 2003.

[14] Peter S. Popolo, Jan G. Svec, and Ingo R. Titze. Adaptation of a pocket pc for use as a wearable voice dosimeter. *Journal of Speech, Language, and Hearing Research*, 48(4):780–791, August 2005.

[15] Juan P. Cortés, Víctor M. Espinoza, Marzyeh Ghassemi, Daryush D. Mehta, Jarrad H. Van Stan, Robert E. Hillman, John V. Guttag, and Matías Zañartu. Ambulatory assessment of phonotraumatic vocal hyperfunction using glottal airflow measures estimated from neck-surface acceleration. *PLoS ONE*, 13(12):e0209017, 2018.

[16] 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.

[17] Jarrad H. Van Stan, Andrew J. Ortiz, Katherine L. Marks, Laura E. Toles, Daryush D. Mehta, James A. Burns, Tiffiny Hron, Tara Stadelman-Cohen, Carol Krusemark, Jason Muise, Annie B. Fox, Charles Nudelman, Steven Zeitels, and Robert E. Hillman. Changes in the daily phonotrauma index following the use of voice therapy as the sole treatment for phonotraumatic vocal hyperfunction in females. *Journal of Speech, Language, and Hearing Research*, 64(9):3446–3455, 2021.

[18] 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, Jason Muise, Annie B. Fox-Galalis, Charles Nudelman, Steven Zeitels, and Robert E. Hillman. 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.

[19] 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.

[20] John M. Heinz and Kenneth N. Stevens. On the derivation of area functions and acoustic spectra from cinéradiographic films of speech. *The Journal of the Acoustical Society of America*, $36(5_{S}upplement) : 1037–1038, May1964$.

[21] Johan Sundberg, Ronald Scherer, Markus Hess, Frank Müller, and Svante Granqvist. Subglottal pressure oscillations accompanying phonation. *Journal of Voice*, 27(4):411–421, 2013.

[22] Randall L Plant and Allen D Hillel. Direct measurement of subglottic pressure

and laryngealresistance in normal subjects and in spasmodic dysphonia. *Journal of Voice*, 12(3):300–314, 1998.

[23] Janwillem Van Den Berg. Direct and indirect determination of the mean subglottic pressure; sound level, mean subglottic pressure, mean air flow, subglottic power and efficiency of a male voice for the vowel (a). *Folia Phoniatrica et Logopaedica*, 8(1):1–24, 1956.

[24] Peter Ladefoged and Norris P McKinney. Loudness, sound pressure, and subglottal pressure in speech. *The Journal of the Acoustical Society of America*, 35(4):454–460, 1963.

[25] 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.

[26] Stellan Hertegård, Jan Gauffin, and Per-Åke Lindestad. A comparison of subglottal and intraoral pressure measurements during phonation. *Journal of Voice*, 9(2):149–155, 1995.

[27] Kazutomo Kitajima and Fumika Fujita. Estimation of subglottal pressure with intraoral pressure. *Acta Oto-Laryngologica*, 109(5-6):473–478, 1990.

[28] Ingo R. Titze and Johan Sundberg. Vocal intensity in speakers and singers. *The Journal of the Acoustical Society of America*, 91(5):2936–2946, 1992.

[29] Jon Z. Lin, Victor M. Espinoza, Katherine L. Marks, Matias Zanartu, and Daryush D. Mehta. Improved subglottal pressure estimation from neck-surface vibration in healthy speakers producing non-modal phonation. *IEEE Journal of Selected Topics in Signal Processing*, 14(2):449–460, 2020.

[30] Katherine L. Marks, Jonathan Z. Lin, James A. Burns, Tiffiny A. Hron, Robert E. Hillman, and Daryush D. Mehta. Estimation of subglottal pressure from neck surface vibration in patients with voice disorders. *Journal of Speech, Language, and Hearing Research*, 63(7):2202–2218, 2020.

[31] Pablo Gomez, Anne Schutzenberger, Marion Semmler, and Michael Dollinger. Laryngeal pressure estimation with a recurrent neural network. *IEEE Journal of Translational Engineering in Health and Medicine*, 7:1–11, 2019.

[32] Zhaoyan Zhang. Estimating subglottal pressure and vocal fold adduction from the produced voice in a single-subject study (l). *The Journal of the Acoustical Society of America*, 151(2):1337–1340, 2022.

[33] Emiro J. Ibarra, Jesús A. Parra, Gabriel A. Alzamendi, Juan P. Cortés, Víctor M. Espinoza, Daryush D. Mehta, Robert E. Hillman, and Matías Zañartu. Estimation of subglottal pressure, vocal fold collision pressure, and intrinsic laryngeal muscle activation from neck-surface vibration using a neural network framework and a voice production model. *Frontiers in Physiology*, 12, 2021.

[34] Juan P. Cortés, Jon Z. Lin, Katherine L. Marks, Víctor M. Espinoza, Emiro J. Ibarra, Matías Zañartu, Robert E. Hillman, and Daryush D. Mehta. Ambulatory monitoring of subglottal pressure estimated from neck-surface vibration in individuals with and without voice disorders. *Applied Sciences*, 12(21):10692, 2022.

[35] Varun Varadarajan, Joel H. Blumin, and Jonathan M. Bock. State of the Art of Laryngeal Electromyography. *Current Otorhinolaryngology Reports*, 1(3):171–177, 2013.

[36] 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.

[37] Dinesh K Chhetri and Soo Jin Park. Interactions of subglottal pressure and neuromuscular activation on fundamental frequency and intensity. *The Laryngoscope*, 126(5):1123–1130, 2016.

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

[39] 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.

[40] 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.

[41] 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.

[42] 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.

[43] Brad H. Story. *Physiologically-Based Speech Simulation Using an Enhanced Wave-Reflection Model of the Vocal Tract*. PhD thesis, University of Iowa, Iowa City, 1995.

[44] 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.

[45] Michael Döllinger, Zhaoyan Zhang, Stefan Schoder, Petr Šidlof, Bogac Tur, and Stefan Kniesburges. Overview on state-of-the-art numerical modeling of the phonation process. *Acta Acustica*, 7:25, 2023.

[46] Carlos Calvache, Leonardo Solaque, Alexandra Velasco, and Lina Peñuela. Biomechanical models to represent vocal physiology: a systematic review. *Journal of Voice*, 37(3):465–e1, 2023.

[47] Byron D. Erath, Matías Zañartu, 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.

[48] 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.

[49] Jeff Kuo, Robert E. Hillman, and Eva B. Holmberg. Modeling and analysis of speakers with vocal-fold nodules. *The Journal of the Acoustical Society of America*, $102(5_{S}upplement) : 3207–3207, 1997$.

[50] 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.

[51] Daryush D. Mehta, Matías Zañartu, Thomas F. Quatieri, Dimitar D. Deliyski, and Robert E. Hillman. Investigating acoustic correlates of human vocal fold vibratory phase asymmetry through modeling and laryngeal high-speed videoendoscopy. *The Journal of the Acoustical Society of America*, 130(6):3999–4009, 2011.

[52] Shinji Deguchi, Yuki Kawahara, et al. Simulation of human phonation with vocal nodules. *American Journal of Computational Mathematics*, 1(03):189, 2011.

[53] Raymond Greiss, Joana Rocha, and Edgar Matida. Validation of a finite element code for a continuum model of vocal fold vibration under the influence of a sessile polyp. *Canadian Acoustics*, 43(1):13–23, 2015.

[54] Robin A Samlan, Brad H Story, Andrew J Lotto, and Kate Bunton. Acoustic and perceptual effects of left–right laryngeal asymmetries based on computational modeling. *Journal of Speech, Language, and Hearing Research*, 57(5):1619–1637, 2014.

[55] Robin A Samlan and Brad H Story. Influence of left–right asymmetries on voice quality in simulated paramedian vocal fold paralysis. *Journal of Speech, Language, and Hearing Research*, 60(2):306–321, 2017.

[56] Jônatas Santos, Jugurta Montalvao, and Israel Santos. Improved model for vocal

folds with a polyp with potential application. In *Interspeech*, pages 1386–1390, 2020.

[57] Mingjun Ji, Boquan Liu, Jack Jiang, Matthew R. Hoffman, Jinwei Lan, and Jin Fang. Effect of controlled muscle activation in a unilateral vocal fold polyp setting on vocal fold vibration. *Applied Sciences*, 12(23):12486, 2022.

[58] Charles J. Nudelman, Andrew J. Ortiz, Annie B. Fox, Daryush D. Mehta, Robert E. Hillman, and Jarrad H. Van Stan. Daily phonotrauma index: An objective indicator of large differences in self-reported vocal status in the daily life of females with phonotraumatic vocal hyperfunction. *American Journal of Speech-Language Pathology*, 31(3):1412–1423, 2022.

[59] Minoru Hirano, Yuki Kakita, Koichi Ohmaru, and Shigejiro Kurita. Structure and mechanical properties of the vocal fold. In *Speech and Language*, volume 7, pages 271–297. Elsevier, 1982.

[60] Gunnar Fant. *Acoustic theory of speech production: with calculations based on X-ray studies of Russian articulations*. Number 2. Walter de Gruyter, 1971.

[61] Ingo R. Titze. Nonlinear source–filter coupling in phonation: Theory. *The Journal of the Acoustical Society of America*, 123(5):2733–2749, 2008.

[62] Jorge C. Lucero, Kélem G. Lourenço, Nicolas Hermant, Annemie Van Hirtum, and Xavier Pelorson. Effect of source–tract acoustical coupling on the oscilla-

tion onset of the vocal folds. *The Journal of the Acoustical Society of America*, 132(1):403–411, 2012.

[63] Rajat Mittal, Byron D. Erath, and Michael W. Plesniak. Fluid dynamics of human phonation and speech. *Annual Review of Fluid Mechanics*, 45(1):437–467, 2013.

[64] Brad H Story and Ingo R Titze. Parameterization of vocal tract area functions by empirical orthogonal modes. *Journal of Phonetics*, 26(3):223–260, 1998.

[65] Ingo R Titze and Daniel W Martin. *Principles of voice production*. Acoustical Society of America, 1998.

[66] Mohammad Eslami, Christiane Neuschaefer-Rube, and Antoine Serrurier. Automatic vocal tract landmark localization from midsagittal mri data. *Scientific Reports*, 10(1):1468, 2020.

[67] Shinji Maeda. A digital simulation method of the vocal-tract system. *Speech Communication*, 1(3-4):199–229, 1982.

[68] Peter Birkholz. Modeling consonant-vowel coarticulation for articulatory speech synthesis. *PloS ONE*, 8(4):e60603, 2013.

[69] Bernd J Kröger. Computer-implemented articulatory models for speech production: A review. *Frontiers in Robotics and AI*, 9:796739, 2022.

[70] 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.

[71] Kenneth N Stevens, Stanley Kasowski, and C Gunnar M Fant. An electrical analog of the vocal tract. *The Journal of the Acoustical Society of America*, 25(4):734–742, 1953.

[72] Subhasmita Sahoo and Aurobinda Routray. A novel method of glottal inverse filtering. *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, 24(7):1230–1241, 2016.

[73] Matías Zañartu. Influence of acoustic loading on the flow-induced oscillations of single mass models of the human larynx. Master's thesis, School of Electrical and Computer Engineering, Purdue University, West Lafayette, IN, 2006.

[74] James L Flanagan. *Speech analysis synthesis and perception*, volume 3. Springer Science & Business Media, 2013.

[75] Mazin G. Rahim. *Artificial Neural Networks for Speech Analysis/synthesis*. Chapman & Hall, 1994.

[76] Brad H Story. Mechanisms of voice production. *The handbook of speech production*, pages 34–58, 2015.

[77] 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.

[78] 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.

[79] Byron D. Erath and Michael W. Plesniak. The occurrence of the Coanda effect in pulsatile flow through static models of the human vocal folds. *The Journal of the Acoustical Society of America*, 120(2):1000–1011, 2006.

[80] Xiaojian Wang, Xudong Zheng, Ingo R Titze, Anil Palaparthi, and Qian Xue. Examining the quasi-steady airflow assumption in irregular vocal fold vibration. *Applied Sciences*, 13(23):12691, 2023.

[81] 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.

[82] Steven D Gray, Fariborz Alipour, Ingo R Titze, and Thomas Hale Hammond. Biomechanical and histologic observations of vocal fold fibrous proteins. *Annals of Otology, Rhinology & Laryngology*, 109(1):77–85, 2000.

[83] Marie D. Dubois, Martin Ch, and Jean M. Prades. Cellular, extracellular and

histological analysis of the vocal folds: correlation between structure and function. *Revue de Laryngologie-Otologie-Rhinologie*, 128(5):267–271, 2007.

[84] Peter Birkholz, Bernd J Kröger, and Christiane Neuschaefer-Rube. Synthesis of breathy, normal, and pressed phonation using a two-mass model with a triangular glottis. In *Interspeech*, pages 2681–2684, 2011.

[85] Livija Cveticanin. Review on mathematical and mechanical models of the vocal cord. *Journal of Applied Mathematics*, 2012(1):928591, 2012.

[86] Zhaoyan Zhang, Jody Kreiman, Bruce R. Gerratt, and Marc Garellek. Acoustic and perceptual effects of changes in body layer stiffness in symmetric and asymmetric vocal fold modelsa). *The Journal of the Acoustical Society of America*, 133(1):453–462, 2013.

[87] 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.

[88] Jack J. Jiang, Carlos E. Diaz, and David G. Hanson. Finite element modeling of vocal fold vibration in normal phonation and hyperfunctional dysphonia: Implications for the pathogenesis of vocal nodules. *Annals of Otology, Rhinology & Laryngology*, 107(7):603–610, 1998.

[89] 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.

[90] 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.

[91] Fariborz Alipour-Haghighi and Ingo R. Titze. Elastic models of vocal fold tissues. *The Journal of the Acoustical Society of America*, 90(3):1326–1331, 1991.

[92] Fariborz Alipour and Sarah Vigmostad. Measurement of vocal folds elastic properties for continuum modeling. *Journal of Voice*, 26(6):816–e21, 2012.

[93] Roger W. Chan and Ingo R. Titze. Viscoelastic shear properties of human vocal fold mucosa: Measurement methodology and empirical results. *The Journal of the Acoustical Society of America*, 106(4):2008–2021, 1999.

[94] Dinesh K. Chhetri, Zhaoyan Zhang, and Juergen Neubauer. Measurement of Young's modulus of vocal folds by indentation. *Journal of Voice*, 25(1):1–7, 2011.

[95] 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.

[96] Mikhail Wadie, Stewart I. Adam, and Clarence T. Sasaki. Development, anatomy,

and physiology of the larynx. In *Principles of Deglutition: A Multidisciplinary Text for Swallowing and its Disorders*, pages 175–197. Springer, 2013.

[97] Ingo R. Titze. *The Myoelastic Aerodynamic Theory of Phonation*. National Center for Voice and Speech, 2006.

[98] 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.

[99] 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.

[100] 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.

[101] Jennifer Oates and Alison Winkworth. Current knowledge, controversies and future directions in hyperfunctional voice disorders. *International Journal of Speech-Language Pathology*, 10(4):267–277, 2008.

[102] Janina K. Casper and Rebecca Leonard. *Understanding voice problems: A phys-*

iological perspective for diagnosis and treatment. Lippincott Williams & Wilkins, 2006.

[103] Petros D. Karkos and Maxwell McCormick. The etiology of vocal fold nodules in adults. *Current Opinion in Otolaryngology & Head and Neck Surgery*, 17(6):420–423, 2009.

[104] Melda Kunduk and Andrew J. McWhorter. True vocal fold nodules: the role of differential diagnosis. *Current Opinion in Otolaryngology & Head and Neck Surgery*, 17(6):449–452, 2009.

[105] Rebecca Leonard. Voice therapy and vocal nodules in adults. *Current Opinion in Otolaryngology & Head and Neck Surgery*, 17(6):453–457, 2009.

[106] Jörg Lohscheller, Ulrich Eysholdt, Hikmet Toy, and Michael Dollinger. Phonovibrography: mapping high-speed movies of vocal fold vibrations into 2-d diagrams for visualizing and analyzing the underlying laryngeal dynamics. *IEEE Transactions on Medical Imaging*, 27(3):300–309, 2008.

[107] Katherine Verdolini, Clark A Rosen, and Ryan C Branski. *Classification manual for voice disorders-I*. Psychology Press, 2014.

[108] Barbara L. Herrington-Hall, Linda Lee, Joseph C. Stemple, Kim R. Niemi, and Molly Miller McHone. Description of laryngeal pathologies by age, sex, and occu-

pation in a treatment-seeking sample. *Journal of Speech and Hearing Disorders*, 53(1):57–64, 1988.

[109] Samantha Kridgen, Robert E Hillman, Tara Stadelman-Cohen, Steven Zeitels, James A Burns, Tiffiny Hron, Carol Krusemark, Jason Muise, and Jarrad H Van Stan. Patient-reported factors associated with the onset of hyperfunctional voice disorders. *Annals of Otology, Rhinology & Laryngology*, 130(4):389–394, 2021.

[110] 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.

[111] 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.

[112] Nancy Pearl Solomon. Vocal fatigue and its relation to vocal hyperfunction. *International Journal of Speech-Language Pathology*, 10(4):254–266, 2008.

[113] 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.

[114] Mohsen Motie-Shirazi, Matías Zañartu, Sean D. Peterson, Daryush D. Mehta, Robert E. Hillman, and Byron D. Erath. Effect of nodule size and stiffness on phonation threshold and collision pressures in a synthetic hemilaryngeal vocal fold model. *The Journal of the Acoustical Society of America*, 153(1):654, 2023.

[115] 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.

[116] Kerstyn Comley and Norman Fleck. The compressive response of porcine adipose tissue from low to high strain rate. *International Journal of Impact Engineering*, 46:1–10, 2012.

[117] Young B Min, Ingo R Titze, and Fariborz Alipour-Haghighi. Stress-strain response of the human vocal ligament. *Annals of Otology, Rhinology & Laryngology*, 104(7):563–569, 1995.

[118] Haoxiang Luo, Rajat Mittal, Xudong Zheng, Steven A. Bielamowicz, Raymond J. Walsh, and James K. Hahn. An immersed-boundary method for flow–structure interaction in biological systems with application to phonation. *Journal of Computational Physics*, 227(22):9303–9332, 2008.

[119] 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.

[120] Christian Dresel, Patrick Mergell, Ulrich Hoppe, and Ulrich Eysholdt. An asymmetric smooth contour two-mass model for recurrent laryngeal nerve paralysis. *Logopedics, Phoniatrics, Vocology*, 31(2):61–75, 2006.

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

[122] Qi Xue, Rajat Mittal, Xudong Zheng, and Steven Bielamowicz. A computational study of the effect of vocal-fold asymmetry on phonation. *The Journal of the Acoustical Society of America*, 128(2):818–827, 2010.

[123] David E. Sommer, Byron D. Erath, Matías Zañartu, and Sean D. Peterson. The impact of glottal area discontinuities on block-type vocal fold models with asymmetric tissue properties. *The Journal of the Acoustical Society of America*, 133(3):EL214–220, 2013.

[124] Zhaoyan Zhang. Vibration in a self-oscillating vocal fold model with left-right asymmetry in body-layer stiffness. *The Journal of the Acoustical Society of America*, 128(5):EL279–285, 2010.

[125] Hye Rhyn Chung, Yoonjeong Lee, Neha K Reddy, Zhaoyan Zhang, and Dinesh K Chhetri. Effects of thyroarytenoid activation induced vibratory asymmetry on voice acoustics and perception. *The Laryngoscope*, 134(3):1327–1332, 2024.

[126] Alessio Carullo, Alberto Vallan, and Arianna Astolfi. Design issues for a portable vocal analyzer. *IEEE Transactions on Instrumentation and Measurement*, 62(5):1084–1093, 2013.

[127] Jeff Searl and Angela Dietsch. Testing of the vocalog vocal monitor. *Journal of Voice*, 28(4):523.e27–523.e37, 2014.

[128] Jarrad H. Van Stan, Daryush D. Mehta, Steven M. Zeitels, James A. Burns, Anca M. Barbu, and Robert E. Hillman. Average ambulatory measures of sound pressure level, fundamental frequency, and vocal dose do not differ between adult females with phonotraumatic lesions and matched control subjects. *Annals of Otology, Rhinology & Laryngology*, 124(11):864–874, 2015.

[129] Robert E. Hillman, James T. Heaton, Asa Masaki, Steven M. Zeitels, and Harold A. Cheyne. Ambulatory monitoring of disordered voices. *Annals of Otology, Rhinology & Laryngology*, 115(11):795–801, 2006.

[130] Qiang Chen, Yunxuan Zhang, Chenyu Sun, Li Liu, Xiaoyan Luo, Hua Wang, and Sili Ni. Case report: Deterioration of infantile hemangioma related to oral

or nebulized administration of β2-ar agonist: Three cases reports. *Frontiers in Oncology*, 12:1000099, 2022.

[131] Eric J. Hunter and Ingo R. Titze. Variations in intensity, fundamental frequency, and voicing for teachers in occupational versus nonoccupational settings. *Journal of Speech, Language, and Hearing Research*, 53(4):862–875, 2010.

[132] Katlyn Elizabeth McGrattan, Alicia Hofelich Mohr, Ellen Weikle, Kayla Hernandez, Katie Walsh, Jinhee Park, Sara E. Ramel, Michael K. Georgieff, Kelly Dietz, Kyle Dahlstrom, John Lindsay, and Suzanne Thoyre. Establishing normative values for healthy term infant feeding performance: Neonatal eating assessment tool-mixed, oral feeding scale, and early feeding skills assessment. *American Journal of Speech-Language Pathology*, 32(6):2792–2801, 2023.

[133] Ingo R. Titze and Eric J. Hunter. Comparison of vocal vibration-dose measures for potential-damage risk criteria. *Journal of Speech, Language, and Hearing Research*, 58(5):1425–1439, 2015.

[134] 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.

[135] Matias Zanartu, Julio C. Ho, Daryush D. Mehta, Robert E. Hillman, and

George R. Wodicka. Subglottal impedance-based inverse filtering of voiced sounds using neck surface acceleration. *IEEE Transactions on Audio, Speech, and Language Processing*, 21(9):1929–1939, 2013.

[136] Jarrad H. Van Stan, Andrew J. Ortiz, Dagmar Sternad, Daryush D. Mehta, Chuanbing Huo, and Robert E. Hillman. Ambulatory voice biofeedback: Acquisition and retention of modified daily voice use in patients with phonotraumatic vocal hyperfunction. *American Journal of Speech-Language Pathology*, 31(1):409–418, 2022.

[137] Jarrad H. Van Stan, James Burns, Tiffiny Hron, Steven Zeitels, Bharat A. Panuganti, Phillip R. Purnell, Daryush D. Mehta, Robert E. Hillman, and Hamzeh Ghasemzadeh. Detecting mild phonotrauma in daily life. *The Laryngoscope*, 133(11):3094–3099, 2023.

[138] Daryush D. Mehta, Víctor M. Espinoza, Jarrad H. Van Stan, Matías Zañartu, and Robert E. Hillman. The difference between first and second harmonic amplitudes correlates between glottal airflow and neck-surface accelerometer signals during phonation. *The Journal of the Acoustical Society of America*, 145(5):EL386–EL392, 2019.

[139] Kenneth N Stevens. *Acoustic phonetics*, volume 30. MIT press, 2000.

[140] Joaquín Sepúlveda, Jesús Parra, Emiro Ibarra, Mauricio Araya, Patricio

De La Cuadra, and Matías Zañartu. Estimation of physiological vocal features from neck surface acceleration signals using probabilistic bayesian neural networks. 2024.

[141] Jorge Andrés Gómez-García, Laureano Moro-Velázquez, and Juan Ignacio Godino-Llorente. On the design of automatic voice condition analysis systems. part i: Review of concepts and an insight to the state of the art. *Biomedical Signal Processing and Control*, 51:181–199, 2019.

[142] Zhaoyan Zhang. Estimation of vocal fold physiology from voice acoustics using machine learning. *The Journal of the Acoustical Society of America*, 147(3):EL264–EL270, 2020.

[143] Anders Löfqvist, Björn Carlborg, and Peter Kitzing. Initial validation of an indirect measure of subglottal pressure during vowels. *The Journal of the Acoustical Society of America*, 72(2):633–635, 1982.

[144] Amanda S. Fryd, Jarrad H. Van Stan, Robert E. Hillman, and Daryush D. Mehta. Estimating subglottal pressure from neck-surface acceleration during normal voice production. *Journal of Speech, Language, and Hearing Research*, 59(6):1335–1345, 2016.

[145] Katherine L. Marks, Jonathan Z. Lin, Annie B. Fox, Laura E. Toles, and Daryush D. Mehta. Impact of nonmodal phonation on estimates of subglottal

pressure from neck-surface acceleration in healthy speakers. *Journal of Speech, Language, and Hearing Research*, 62(9):3339–3358, 2019.

[146] Daryush D. Mehta, James B. Kobler, Steven M. Zeitels, Matías Zañartu, Emiro J. Ibarra, Gabriel A. Alzamendi, Rodrigo Manriquez, Byron D. Erath, Sean D. Peterson, Robert H. Petrillo, and Robert E. Hillman. Direct measurement and modeling of intraglottal, subglottal, and vocal fold collision pressures during phonation in an individual with a hemilaryngectomy. *Applied Sciences*, 11(16):7256, 2021.

[147] Jacob Cohen. *Statistical Power Analysis for the Behavioral Sciences*. Academic Press, 1988.

[148] 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.

[149] Adam C. Lammert, Jennifer Melot, Douglas E. Sturim, Daniel J. Hannon, Richard DeLaura, James R. Williamson, Gregory Ciccarelli, and Thomas F. Quatieri. Analysis of phonetic balance in standard english passages. *Journal of Speech, Language, and Hearing Research*, 63(4):917–930, 2020.

[150] Jesús A Parra, Carlos Calvache, and Matías Zañartu. Assessing phoneme distribution for speech modeling. In *18th International Symposium on Medical Information Processing and Analysis*, volume 12567, pages 366–372. SPIE, 2023.

[151] Dinesh K. Chhetri, Juergen Neubauer, and Elazar Sofer. Influence of asymmetric recurrent laryngeal nerve stimulation on vibration, acoustics, and aerodynamics. *The Laryngoscope*, 124(11):2544–2550, 2014.

[152] Andrew M. Vahabzadeh-Hagh, Pranati Pillutla, Zhaoyan Zhang, and Dinesh K. Chhetri. Dynamics of intrinsic laryngeal muscle contraction. *The Laryngoscope*, 129(1), 2018.

[153] Katherine Verdolini, David G. Druker, Phyllis M. Palmer, and Hani Samawi. Laryngeal adduction in resonant voice. *Journal of Voice*, 12(3):315–327, 1998.

[154] Byron D. Erath, Matías Zañartu, and Sean D. Peterson. Modeling viscous dissipation during vocal fold contact: the influence of tissue viscosity and thickness with implications for hydration. *Biomechanics and Modeling in Mechanobiology*, 16(3):947–960, 2017.

[155] Jackie Gartner-Schmidt and Amanda I. Gillespie. Conversation training therapy: let's talk it through. *Seminars in Speech and Language*, 42(1):32–40, 2021.

[156] Nelson Roy, Diane M. Bless, Dennis Heisey, and Charles N. Ford. Manual circumlaryngeal therapy for functionaldysphonia: An evaluation of short- and long-term treatment outcomes. *Journal of Voice*, 11(3):321–331, 1997.

[157] Katherine Verdolini. *Lessac-Madsen Resonant Voice Therapy Clinician Manual*. Oxfordshire, UK: Plural Publishing Inc, 2008.

[158] Leena Rantala and Erkki Vilkman. Relationship between subjective voice complaints and acoustic parameters in female teachers' voices. *Journal of Voice*, 13(4):484–495, 1999.

[159] Cara E Stepp, Robert E Hillman, and James T Heaton. Use of neck strap muscle intermuscular coherence as an indicator of vocal hyperfunction. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 18(3):329–335, 2010.

[160] 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.

[161] Juan P. Cortés, Gabriel A. Alzamendi, Alejandro J. Weinstein, Juan I. Yuz, Víctor M. Espinoza, Daryush D. Mehta, Robert E. Hillman, and Matías Zañartu. Kalman filter implementation of subglottal impedance-based inverse filtering to estimate glottal airflow during phonation. *Applied Sciences*, 12(1):401, 2021.

[162] Arturo Morales, Juan I. Yuz, Juan P. Cortés, Javier G. Fontanet, and Matías Zañartu. Glottal airflow estimation using neck surface acceleration and low-order kalman smoothing. *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, 31:2055–2066, 2023.

[163] Matías Zañartu. *Acoustic coupling in phonation and its effect on inverse filtering*

of oral airflow and neck surface acceleration. PhD Thesis, School of Electrical and Computer Engineering, Purdue University, West Lafayette, IN, 2010.

[164] Qingjun Qiu, HK Schutte, Lide Gu, and Qilian Yu. An automatic method to quantify the vibration properties of human vocal folds via videokymography. *Folia Phoniatrica et Logopaedica*, 55(3):128–136, 2003.

[165] Heather Shaw Bonilha, Dimitar D. Deliyski, and Terri Treman Gerlach. Phase asymmetries in normophonic speakers: visual judgments and objective findings. *American Journal of Speech-Language Pathology*, 17(4):367–376, 2008.

[166] Peter Birkholz. Glottalimageexplorer–an open source tool for glottis segmentation in endoscopic high-speed videos of the vocal folds. *Studientexte zur Sprachkommunikation*, pages 39–44, 2016.

[167] Yoshihiko Nakamura, Hideo Hanafusa, and Tsuneo Yoshikawa. Task-priority based redundancy control of robot manipulators. *The International Journal of Robotics Research*, 6(2):3–15, 1987.

[168] Frank H. Guenther and Daniele Micci Barreca. *Neural Models for Flexible Control of Redundant Systems*, page 383–421. Elsevier, 1997.

[169] 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.

[170] Mark L. Latash. The bliss (not the problem) of motor abundance (not redundancy). *Experimental Brain Research*, 217(1):1–5, 2012.

[171] Weili Jiang, Xudong Zheng, and Qian Xue. Computational modeling of fluid–structure–acoustics interaction during voice production. *Frontiers in Bioengineering and Biotechnology*, 5:7, 2017.

[172] Chandra M Ivey. Vocal fold paresis. *Otolaryngologic Clinics of North America*, 52(4):637–648, 2019.

[173] Philip W Tipton, Dale C Ekbom, Amy L Rutt, and Jay A van Gerpen. Vocal fold "paralysis": an early sign in multiple system atrophy. *Journal of Voice*, 34(6):940–944, 2020.

[174] Ming-Wang Hsiung and Yu-Che Hsiao. The characteristic features of muscle tension dysphonia before and after surgery in benign lesions of the vocal fold. *ORL; journal for oto-rhino-laryngology and its related specialties*, 66(5):246–254, 2004.

[175] Martin L Spencer. Muscle tension dysphonia: a rationale for symptomatic subtypes, expedited treatment, and increased therapy compliance. *Perspectives On Voice And Voice Disorders*, 25(1):5–15, 2015.

[176] Jakub Malinowski, Wioletta Pietruszewska, Magdalena Kowalczyk, and Ewa Niebudek-Bogusz. Value of high-speed videoendoscopy as an auxiliary tool in differentiation of benign and malignant unilateral vocal lesions. *Journal of Cancer Research and Clinical Oncology*, 150(1):10, 2024.

[177] Heather Shaw Bonilha, Dimitar D Deliyski, Joanna Piasecki Whiteside, and Terri Treman Gerlach. Vocal fold phase asymmetries in patients with voice disorders: A study across visualization techniques. *American Journal of Speech-Language Pathology*, 21(1), 2012.

[178] Michael Döllinger, David A Berry, and Stefan Kniesburges. Dynamic vocal fold parameters with changing adduction in ex-vivo hemilarynx experiments. *The Journal of the Acoustical Society of America*, 139(5):2372–2385, 2016.

[179] Patrick Schlegel, David A. Berry, and Dinesh K. Chhetri. Analysis of vibratory mode changes in symmetric and asymmetric activation of the canine larynx. *PLoS ONE*, 17(4):e0266910, 2022.

[180] Allen D. Hillel. The study of laryngeal muscle activity in normal human subjects and in patients with laryngeal dystonia using multiple fine-wire electromyography. *The Laryngoscope*, 111(S97):1–47, 2001.

[181] Shaghauyegh S. Azar and Dinesh K. Chhetri. Phonation threshold pressure

revisited: Effects of intrinsic laryngeal muscle activation. *The Laryngoscope*, 132(7):1427–1432, 2021.

[182] 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.

[183] Jonathan J Deng, Byron D Erath, Matías Zañartu, and Sean D Peterson. The effect of swelling on vocal fold kinematics and dynamics. *Biomechanics and Modeling in Mechanobiology*, 22(6):1873–1889, 2023.

[184] Mohamed A Serry, Gabriel A Alzamendi, Matías Zañartu, and Sean D Peterson. An euler–bernoulli-type beam model of the vocal folds for describing curved and incomplete glottal closure patterns. *Journal of the Mechanical Behavior of Biomedical Materials*, 147:106130, 2023.