The Voice...can be heard only in silence, and as long as the heart is filled with the clamor of desire the silver tones of the Voice cannot be heard...and not all earth's tumult will be able to deafen us to the majestic rhythm of that Voice, that Voice that reverberates throughout the Eternities as the tides of Being thunder upon the beaches of the worlds.

Sri Krishna Prem, The Yoga of the Bhagavad Gita

Accept the view that nothing in nature is useless...In short, consider problems on their own merits when attacking them. Avoid deviating to secondary concerns that distract attention and weaken analytical powers. In struggling with nature, the biologist, like the astronomer, must look beyond the earth he lives on and concentrate on the serene universe of ideas, where the light of truth will eventually shine.

Santiago Ramon y Cajal, Advice for a Young Investigator

University of Alberta

Theoretical Framework for Modeling Ingressive Phonation

by

Michael Vincent Brougham

A thesis submitted to the Faculty of Graduate Studies and Research in partial fulfillment of the requirements for the degree of

Master of Science

Department of Mechanical Engineering

©Michael Vincent Brougham
Fall 2012
Edmonton, Alberta

Permission is hereby granted to the University of Alberta Libraries to reproduce single copies of this thesis and to lend or sell such copies for private, scholary or scientific research purposes only. Where the thesis is converted to, or otherwise made available in digital form, the University of Alberta will advise potential users of the thesis of these terms.

The author researves all other publication and other rights in association with the copyright in the thesis and, except as herein before provided, neither the thesis nor any substantial portion thereof may be printed or otherwise reproduced in any material form whatsoever without the author's prior written permission.

To my parents for inspiring me to study the cycles of nature and people

Abstract

Researchers in vocal acoustics have used computer simulations of single and multi-mass models of the human vocal folds to study human phonation for over 40 years. They have successfully given insight into different voice qualities and registers as well as the irregular phonation associated with pathologies. This study hypothesizes that current techniques for simulating egressive phonation will also lead to self-oscillation of the vocal folds upon changing the pressure and flow orientation. Through symmetrical reasoning, well-established aeroelastic relations for multi-mass vocal fold models are generalized to cases where the lung pressure is negative and the air flows into the trachea. The pitch, flow rates, and phonation ranges are compared between the ingressive and egressive conditions. A modal analysis is used to understand the differences between the two approaches. This study shows that current techniques can be utilized to simulate ingressive phonation and provides a unified framework for both kinds of phonation.

Acknowledgements

This body of work would not have been possible had not Dr. Brian Fleck provided me with a way to study both engineering and musical principles together. Vocal acoustics is such an interdisciplinary field of study incorporating fluid dynamics, control theory, mechanics, vibration and signal processing and he was able to provide feedback, suggestions, support and perspective on every front along the entire way.

My biomedical engineering course with Dr. Jason Carey and his subsequent supervision were invaluable in beginning and completing this study. The work began as a course project with him and I am grateful for the encouragement and knowledge of human physiology he shared with me which informed my understanding of the vocal folds. His attention to detail and assistance with managing the project helped keep me focused and moving steadily forward toward completion.

I am especially grateful to Dr. Laurier Fagnan for his mentorship over these past few years. It is inspiring to learn vocal principles from someone so dedicated to singing and its power to heal and sustain individual people, families and entire communities.

I am grateful to the Chorale Saint-Jean and the students and staff of the Faculté Saint-Jean for hosting me during my research studies.

My musical collaborators and close friends here in Edmonton are a blessing in my life and always inspiring and encouraging me to follow my heart at all times.

I am especially grateful to Rhéa and her family for their love, humour, support and patience. Special thanks also to D.B. and S.J-B.

Finally, I want to thank my family members and old friends for their love, patience and support despite the geographic distance from Edmonton. I am looking forward to spending more time with them now and making up for being away for so long.

Table of Contents

1 Introduction **1**

1.1 Motivation 1

1.2 General tissue behavior 1

1.3 Vocal tract and subglottal resonators 7

1.3.1 Uniform tubes 7

1.4 Thesis objectives and scope 9

1.5 Thesis overview 10

2 Ingressive Phonation using a Multi-Mass Model **11**

2.1 Introduction 11

2.2 Methods 13

2.2.1 Overview 13

2.2.2 Tissue Mechanics 14

2.2.3 Aerodynamics 17

2.2.4 Waveguides and acoustic Pressures 22

2.2.5 Lip radiation 25

2.2.6 Aeroacoustic Coupling 26

2.2.7 Modal analysis 28

2.2.8 Aerodynamic state feedback 30

2.2.9 Computational algorithm 35

2.2.10 Tissue parameters 38

2.2.11 Parameter space: range of lung pressures and pre-phonatory half-widths 39

2.2.12 Post-simulation analysis 41
2.3 Results and Discussion 42
2.4 Conclusion 61

3 Summary and outlook **63**
3.1 Summary 63
3.2 Future work 65
3.3 Closing remarks 65

Bibliography **66**

Appendices **72**

A Equivalency of state-space representation of the tissue dynamics to Newton's second law of motion **73**

List of Tables

2.1 Tokuda et al. (2010) parameters. $L_g = 1.4$, $k_{12} = 1.0$ [N/m],k_{23} = 0.5 [N/m][1] 39

List of Figures

1.1 3D view of glottis - 2 Mass Model after Lous 1998 [2] 4
1.2 3D view of glottis - 3 Mass Model after Story and Titze 1995 [3] 5
1.3 3D view of glottis - 4 Mass Model after Tokuda et al. 2010 [1] 6

2.1 Full System showing subglottal and supraglottal resonators, larynx and possible flow directions. The lung and mouth pressures are shown along with the pressures upstream and downstream of the larynx. The system is based on the Tokuda et al. model [1]. The larynx is comprised of two symmetrical vocal folds each represented by a body mass and three cover masses connected by springs and dampers. The half-heights along the profile are shown along with the vocal fold length. 14
2.2 Free body diagram of a cover mass 17
2.3 Aerodynamic forces and reaction forces [after Lous et al.[2]]. The forces are actually applied to both sides of the glottis but only the bottom fold is shown with force for comprehension purposes. 18
2.4 Algorithm flowchart. See section 2.2.9 for details. 37
2.5 In vacuo eigenanalysis of a 4 mass system with $P_L = 0Pa$, $P_M =$ 0 Pa, $x_0 = 0.036cm$. (a) Natural frequency eigenmodes of a four mass model ($\mathbf{v}_{eigen}$) in comparison to the simulation extrema ($\mathbf{v}_{sim}$); (b) Eigenvalues of each eigenmode on imaginary versus real component axes, colors correspond to eigenmodes of (a); (c) Power spectrum analysis graph of the 4 mass model. . . . 43

2.6 Aero feedback eigenanalysis. Eigenmodes still shown *in vacuo*. $P_L = -100$ top, -120 middle and $-140Pa$ bottom, $P_M = 0$ Pa, $x_0 = 0.036cm$. (a) Natural frequency eigenmodes of a four mass model ($\mathbf{v}_{eigen}$) in comparison to the simulation extrema ($\mathbf{v}_{sim}$); (b) Eigenvalues of each eigenmode on imaginary versus real component axes; (c) Power spectrum analysis graph of the 4 mass model. 44

2.7 Time history of key variables for egression for four-mass model, source not coupled with waveguides. $0Pa < P_L < 2000Pa$, $P_M = 0Pa$, $x_0 = 0.036cm$. a) Spectrogram ; b) volumetric flow rate $[cm^3/s]$; c) Vocal fold aperture areas $[cm^2]$ 47

2.8 Key measures for egression for four-mass model, source not coupled with waveguides. $0Pa < P_L < 2000Pa$, $P_M = 0Pa$, $x_0 = 0.036cm$. a) Sound pressure level $[dB]$; b) Pitch $[hz]$; c) Mucosal wave phase delay $[ms]$; d) Largest real eigenvalue component ; e) Correlation of simulation extrema ($\mathbf{v}_{sim}$) to *in vacuo* eigenmodes ($\mathbf{v}_{eigen}$) 48

2.9 Time history of key variables for ingression for four-mass model, source not coupled with waveguides. $0Pa > P_L > -2000Pa$, $P_M = 0Pa$, $x_0 = 0.036cm$. a) Spectrogram ; b) volumetric flow rate $[cm^3/s]$; c) Vocal fold aperture areas $[cm^2]$ 50

2.10 Key measures for ingression for four-mass model, source not coupled with waveguides. $0Pa > P_L > -2000Pa$, $P_M = 0Pa$, $x_0 = 0.036cm$. a) Sound pressure level $[dB]$; b) Pitch $[hz]$; c) Mucosal wave phase delay $[ms]$; d) Largest real eigenvalue component ; e) Correlation of simulation extrema ($\mathbf{v}_{sim}$) to *in vacuo* eigenmodes ($\mathbf{v}_{eigen}$) 51

2.11 Time history of key variables for ingression for four-mass model, source not coupled with waveguides. $0Pa > P_L > -800Pa$, $P_M = 0Pa$, $x_0 = 0.036cm$. a) Spectrogram ; b) volumetric flow rate $[cm^3/s]$; c) Vocal fold aperture areas $[cm^2]$ 53

2.12 Key measures for ingression for four-mass model, source not coupled with waveguides. $0Pa > P_L > -800Pa$, $P_M = 0Pa$, $x_0 = 0.036cm$. a) Sound pressure level $[dB]$; b) Pitch $[hz]$; c) Mucosal wave phase delay $[ms]$; d) Largest real eigenvalue component ; e) Correlation of simulation extrema ($\mathbf{v}_{sim}$) to *in vacuo* eigenmodes ($\mathbf{v}_{eigen}$) 54
2.13 Time history of key variables for ingression for four-mass model, source coupled with waveguides. $0Pa > P_L > -800Pa$, $P_M = 0Pa$, $x_0 = 0.036cm$. a) Spectrogram ; b) volumetric flow rate $[cm^3/s]$; c) Vocal fold aperture areas $[cm^2]$ 55
2.14 Key measures for ingression for four-mass model, source coupled with waveguides. $0Pa > P_L > -800Pa$, $P_M = 0Pa$, $x_0 = 0.036cm$. a) Sound pressure level $[dB]$; b) Pitch $[hz]$; c) Mucosal wave phase delay $[ms]$; d) Largest real eigenvalue component ; e) Correlation of simulation extrema ($\mathbf{v}_{sim}$) to *in vacuo* eigenmodes ($\mathbf{v}_{eigen}$) 56
2.15 Time history of key variables for ingression for four-mass model, source not coupled with waveguides, tissue critically damped. $0Pa > P_L > -800Pa$, $P_M = 0Pa$, $x_0 = 0.036cm$. a) Spectrogram ; b) volumetric flow rate $[cm^3/s]$; c) Vocal fold aperture areas $[cm^2]$ 57
2.16 Key measures for ingression for four-mass model, source not coupled with waveguides, tissue critically damped. $0Pa > P_L > -800Pa$, $P_M = 0Pa$, $x_0 = 0.036cm$. a) Sound pressure level $[dB]$; b) Pitch $[hz]$; c) Mucosal wave phase delay $[ms]$; d) Largest real eigenvalue component ; e) Correlation of simulation extrema ($\mathbf{v}_{sim}$) to *in vacuo* eigenmodes ($\mathbf{v}_{eigen}$) 58
2.17 Observed eigenvalue trajectories for egressive (top) and ingressive (bottom) phonation with equilibrium matrices indicated . 60

Nomenclature

Matrices are represented by bold uppercase letters. Vectors are indicated by lowercase bold letters. The corresponding elements of matrices and vectors are in lower case.

Main Notation

A	Area
$\mathbf{A}$	System matrix
$\hat{\mathbf{A}}$	Augmented system matrix incorporating aerodynamic feedback
F	Force or linear frequency or formant
K	Stiffness matrix
L_g	vocal fold dorsal length along y axis
M	Mass matrix
P_M	Mouth Pressure
P_L	Lung Pressure
P_A	Atmospheric Pressure
P_s	Pressure below glottis

R	Damping matrix
T	thickness along z axis
Z	Acoustic impedance
a_g	glottal area [cm^2]
b	Backward delay line and also body
b_e	Backward delay line pressure in the supraglottal waveguide
b_s	Backward delay line pressure in the subglottal waveguide
c	Speed of sound in air
col	Collisional
f	Forward delay line
f_e	Forward delay line pressure in the supraglottal waveguide
f_s	Forward delay line pressure in the subglottal waveguide
e	Epilaryngeal
h	Height (distance between the folds)
k	Stiffness
$k_{c,ij}$	Stiffness of coupling spring between mass i and mass j
m	Mass
q	Vocal fold tension modifier

p_e^+	Acoustic perturbation above glottis leaving glottis (towards pharynx)
p_e^-	Acoustic perturbation above glottis incident on glottis (from pharynx)
p_s^+	Acoustic perturbation below glottis incident on glottis (from trachea)
p_s^-	Acoustic perturbation below glottis leaving glottis (towards trachea)
r	Reflection coefficient
sup	Supraglottal
sub	Subglottal
t	Time
u	Flow
v	Flow velocity
x	Distance from glottal midline along the dorsalventral axis
$\mathbf{x}$	Displacements from prephonatory distances
y	Lateral distance along the left-right axis
z	Rostral distance along the anteroposterior axis
$_0$	Resting or initial
$_c$	Coupling spring
$(\cdot)^T$	Transpose

$(\cdot)_{ij}$	ij^{th} element of a matrix
$\dot{(\cdot)}$	First time derivative
$\ddot{(\cdot)}$	Second time derivative

Greek Symbols

Ω	Aggregate pressure difference across glottis
ω	Angular frequency
γ	Propagation loss factors in the waveguides
δ	Boundary layer thickness
λ	Eigenvalue
ρ	Density
ς	Damping parameter
$\Im$	Imaginary component
$\Re$	Real component

Abbreviations

$FVFs$	False vocal folds
$TVFs$	True vocal folds
$vent$	Ventricular
sep	Referring to jet separation

Chapter 1

Introduction

1.1 Motivation

Researchers in vocal acoustics have used computer simulations of single and multi-mass models of the human vocal folds to reproduce and understand human phonation for over 40 years [4]. They have successfully given insight into different voice qualities and singing registers as well as the irregular phonation associated with pathologies. Despite their widespread use, they have not yet been applied to sounds generated by a negative lung gauge pressure and reversal of airflow which lead to ingressive phonation. Given the frequency of occurrence of ingressive sounds in human communication [5], the question arises as to whether current egressive models are still applicable with slight modifications or if an entirely new framework is required for simulating this approach to phonation.

1.2 General tissue behavior

The human vocal cords are a pair of soft tissues found within the larynx [6] and situated between two acoustical resonators: the supra- and subglottal tracts [7]. The application of air pressure of sufficient magnitude below the folds can lead to their deformation and flow-induced oscillations [8]. This contributes to a time-varying pressure field within the vocal tract which is perceived as

sound. The vibration of the tissue can be examined through its equations of motion.

In terms of its displacement x, the behavior of a single oscillator with mass m may be predicted by a balance of forces [9, 10]:

$$\underbrace{F}_{\text{driving/external}} = \underbrace{m\ddot{x}}_{\text{inertial}} + \underbrace{r\dot{x}}_{\text{dampening}} + \underbrace{kx}_{\text{elastic}} \tag{1.1}$$

This equation expresses that the forces external to a body are counteracted by the internal forces of inertia, dampening and elastic deformation which depend on the acceleration, velocity and displacement of the oscillator respectively.

To model the oscillatory phenomena within the human phonatory system, independent coordinates can be assigned both to points or regions between, on or within the vocal fold tissue to describe its changing configuration over time, as well as to the aeroacoustical conditions of the air upstream and downstream of the folds. In the case of the folds, these coordinates may describe the position and velocity of several tissue regions as well as the flow of air between the folds. Meanwhile, coordinates can be ascribed to the pressures or flows within regions of the resonators. The number of such independent coordinates is known as the *degrees of freedom* of the system [11] and the set of these coordinates characterizes the state of the system. A system will often execute motions such as vibrations in characteristic patterns known to engineers as *eigenmodes* with growth rates and frequencies determined by their respective complex-valued *eigenvalues*. An important result of classical mechanics is that a body with N degrees of freedom may have at most N such resonant modes [12]. For a given phenomenon of interest, a major question in vocal acoustic research is precisely how many degrees are required to capture the behavior of interest?

Early researchers explored two extremes for modeling individual vocal folds: finite-element methods [13] or large arrays of masses [14] vs. one- to two-lumped mass models [4, 15]. The latter provided computational advantages and was justified in that the first few eigenmodes of the multi-element simulations accounted for the bulk of the material dynamics [16] making addi-

tional elements unnecessary [17]. As such, a relatively low number of elements were attributed material properties to reproduce the low-order modes [17], a standard analysis technique in vibrational analysis of structures [11]. The question as to how many elements are sufficient to model a behavior of interest persists to this day and techniques continue to range from a discretized near-continuum to a single mass.

A behavior of interest that constitutes a major area of ongoing research is that of phonation onset which is studied in terms of the conditions under which oscillations prevail over energy losses [18].

In order for the vocal tissue oscillations to be sustained, the system must allow for the net transfer of energy from the flow to the folds to overcome losses [19]. For this to occur, there must be at least two eigenmodes [18] and therefore two degrees of freedom in the system to allow for a temporal asymetry in the energy transfer over the course of a cycle [20].The earliest vocal models had one degree of freedom representing the lateral tissue displacement and another degree for the acoustic loading of the vocal tract [21]. These are known as single-mass models which did not oscillate without the vocal tract loading [7].

Titze [22] provided an explanation of possible energy transfer mechanisms in terms of the phase relationships between the velocity of the folds, the density of air downstream of the glottis, and the changing shape of the folds' surfaces. He described two complementary mechanisms for phonation. In the first mechanism, the mass of air downstream of the folds lags behind the tissue motion because of the former's inertia. During the glottal opening phase, the air in this region gains mass and density and therefore exerts a pressure force in the same direction as the folds motion. By the time the folds begin closing because of a restoring spring force, the air column has shifted, leaving a low pressure region adjacent to the folds thereby facilitating their closure. Once the folds are closed, the pressure below the folds would be restored contributing to their eventual re-opening. The other mechanism was based on the observation of a surface wave travelling along the vocal fold surface, from bottom to top, known as the *mucosal wave*. During opening, the folds first separated from their fully-contacted state at the bottom of the folds and the result-

ing convergent channel was maintained over the opening phase causing higher outward pressures. During closing, the lower portion still led the downstream portion movement, forming a divergent channel and lower or even negative gauge pressures. The convergent-divergent channel shape alternated in phase with the fold motion allowing for energy transfer. Although the folds had an alternating shape in his model, it was still a single-mass model described by the position and velocity of each fold's midpoint; the mucosal wave delay was a prescribed control parameter [23]. This model did however elucidate and contrast two velocity-dependent energy transfer mechanisms.

Two-mass vocal fold models had been explored earlier than Titze's paper, most notably by Ishizaka and Flanagan [15]. In such models, the mucosal wave resulted from the independent motion of two masses coupled to each other by a spring [3]. Later refinements gave the vocal fold surface a smooth profile [2]. In this model, shown in Figure 1.1, two masses given identical material properties and geometry. The model is symmetric about the midpoint of the masses. It is of interest to the present study as any difference in behavior between ingressive and egressive phonation would be due to the lack of symmetry in terms of the differences in aerodynamic and acoustic conditions.

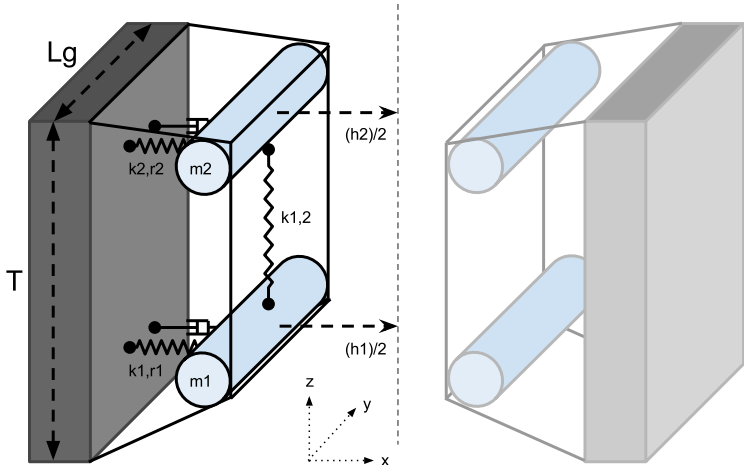


Figure 1.1: 3D view of glottis - 2 Mass Model after Lous 1998 [2]

It has been argued that the early one or two lumped mass models were over-simplifications [3]. Although promising, the single mass models could not independently model the convergent-divergent surface phenomenon within

the glottis, if at all, whereas the two mass models aggregated several distinct tissue layers into a single mass with uniform stiffness, density and damping [24]. Given these inherent limitations, models with more than 2 masses were eventually explored [21].

Story and Titze [3] proposed a 3-mass model (Figure 1.2) that modeled both the pliable-mobile cover via two masses of roughly equal properties and the body with a larger and stiffer mass behind the cover [3]. The left-right vocal folds were mirror images, so no lateral asymmetries were investigated in the paper. They were able to successfullly model different vocal registers such as chest voice and falsetto yet needed to rely on alternative parameter settings to produce distinct registers.

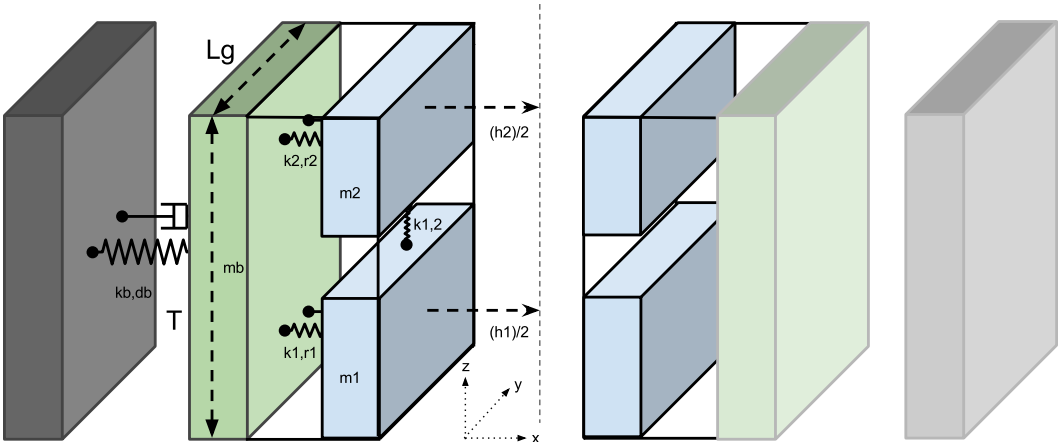


Figure 1.2: 3D view of glottis - 3 Mass Model after Story and Titze 1995 [3]

A more recent model of the vocal folds, proposed by Tokuda et al. [1] and depicted in Figure 1.3, divides the vibrating part of the vocal folds into a body mass and three cover masses. The dampers r_1, r_2 and r_3 dissipate the energy in the system. The body-cover springs k_1, k_2 and k_3 act to maintain a distance between the body mass and the cover masses while the body spring k_b acts between the body and the immobile part of the vocal folds attached to the laryngeal wall. The cover springs $k_{1,2}$ and $k_{2,3}$ provide stiffness solely along the x-axis despite appearing vertical. A unique element of their model is the presence of 3 cover masses, the third of which is significantly lighter than the other two. The researchers successfully modeled register breaks and

other phenomena with a single parameter set using this configuration [1], [25]. The reader is referred to [21] for a comprehensive survey of lumped-parameter models.

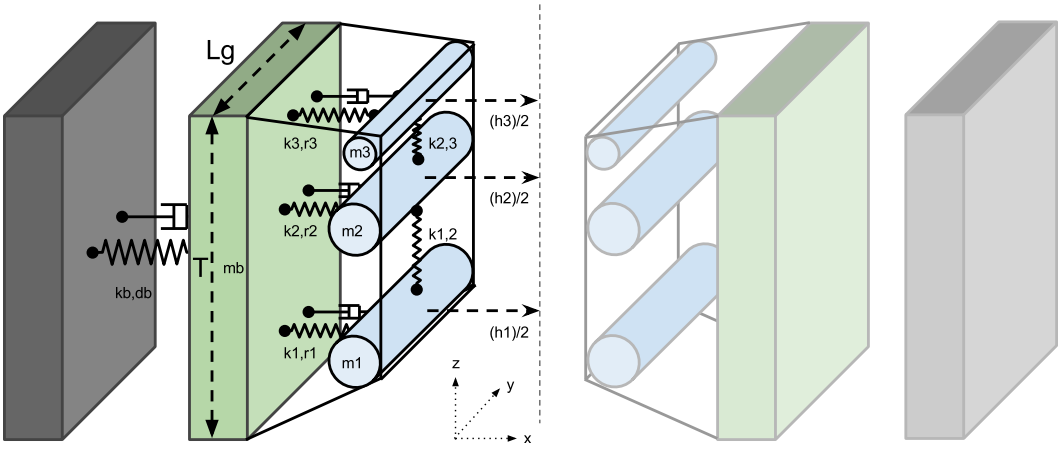


Figure 1.3: 3D view of glottis - 4 Mass Model after Tokuda et al. 2010 [1]

While the ability of these models to model vocal phenomena has improved considerably, the understanding of the conditions required for phonation have also developed. Returning to the balance of forces equation (1.1), since the energy transfer in Titze's (1988) model [22] results from the effect of velocity it is known as a *negative damping* model [23] in analogy with the positive damping term r which results in the dissipation of energy [10]. Newer models have been explored which consider the potential of displacement-dependent phenomena to cause energy gain by the system [23]. In particular, Zhang developed a continous model and later contrasted a fuller set of phenomena by which the folds can be set and kept in motion [23]. He explored the possibility for two of the resonant eigenmodes of the tissue to synchronize their frequency prior to the onset of vibration. In the absence of air pressure, these eigenvalues have zero or negative growth rates, however the effect of air pressure and various flow effects may lead to two eigenvalues meeting (in the complex plane) at the same frequency and one of the eigenvalues gaining a positive growth rate [23]. When at least one of the eigenvalues in a system has a positive growth rate, the system becomes exponentially unstable meaning that a disturbance will lead to either a growing oscillation about an equilibrium configuration

[23] or a static divergence from that configuration [8, 18]. In the first case, the vibration of the folds will result in phonation. While Zhang mentioned the possibility of and even observed the latter case of a static divergence, it was not central to his observations and analysis [18]. In the case of a dynamic instability, there is therefore a clear connection between an eigenvalue gaining a positive growth rate and phonation being possible. As such, to answer the question as to whether ingressive phonation can be modeled may amount to considering under what conditions at least one eigenvalue of the system assumes a positive growth rate. In the present study, this heuristic is developed for a system with an arbitrary number of surface masses and applied specifically to a very recent model, namely Tokuda's four-mass model [1] in terms of a tissue geometry parameter set and acoustical resonators.

1.3 Vocal tract and subglottal resonators

In contrast to the range of glottal modeling approaches, the treatment of vocal tract acoustics is more standard though equally relevant to our considerations.

1.3.1 Uniform tubes

The simplest approximation of the supra- and sub-glottal resonators is a cylindrical tube of length L_{tube} with constant cross-sectional area. This study assumes longitudinal plane waves propagating in one-dimension. Such waves propagate as pulses of air rarefactions and compressions along the tube that reflect off of the ends of the tube. The new wave crests superimpose with reflected ones and cause regions of amplitude cancellation or amplification. At certain source frequencies, known as resonances, a standing wave will be established as certain locations along the tube will be associated with a pressure node or antinode. The resonant frequencies of such tubes depend on whether the ends are open or closed.

If the tubes are open or closed at both ends, the resonant frequencies for its n modes ($n = 1, 2, 3, ...$) are given by [26]:

$$f_n = n\frac{c_{air}}{2L_{tube}} \tag{1.2}$$

If the tubes are open at one end and closed at the other, the resonant frequencies are given by [26]:

$$f_n = (2n-1)\frac{c_{air}}{4L_{tube}} \tag{1.3}$$

The latter represents an ideal case of the supraglottal resonator. The lips are approximated as an open end while the glottal aperture is much smaller than the tube's cross-sectional area and therefore treated as closed. Taking the length of the vocal tract to be $17.5\,cm$ and the speed of air to be $35,000\,cm/s$ suggests resonant frequencies at $500\,Hz$, $1500\,Hz$, $2500\,Hz$ and so on. These resonant frequencies are known as the formants of the vocal tract. Actual vocal tracts deviate significantly from a uniform tube and the resonant frequencies can be further shifted by a speaker's articulators to produce the formants associated with different vowels such as /a/, /e/, /i/, /o/, /u/.

As a first approximation, the subglottal resonator can be treated as a $20\,cm$ tube open at the inferior end and closed at the glottal end [27]. Assuming the same air speed, the previous calculations with equation (1.3) now provide resonances at $437.5\,Hz$, $1312.5\,Hz$, $2187.5\,Hz$, $3062.5\,Hz$, $3937.5\,Hz$, $4812.5\,Hz$. This approach is suitable for the 2nd and 3rd resonances but underestimates the first [27]. Lulich provides a more accurate equation as follows [27]:

$$Sg_n = (2n-1)\frac{c_{air,body}}{4h/k_a} \tag{1.4}$$

Here $c_{air,body} = 35,900\,cm/s$ is "the free-field speed of sound in saturated air at body temperature" is given as appropriate for humid air at body temperature, h is the person's height in cm, $k_a = 8.508$ is a scaling factor for curve fitting and Sg_n are the subglottal resonances. They define an acoustic length for the subglottal resonator via the ratio $L_a = h/k_a$ which for their 50 participants was roughly $20\,cm$ on average. For the 2nd and 3rd subglottal resonances, they conclude:"that a rigid-walled, uniform tube with a length equal to about $1/k_a$ times the height of any given speaker is an anatomically appropriate model for estimating the frequencies of Sg_2 and Sg_3. The rapid expansion of the area function beyond 20 cm below the glottis essentially acts as an infinite baffle, so that the tube can be considered open at this distal end."

The scenario for Sg1 is complicated by the resonant frequencies of the subglottal walls which cause a shift in the wave propagation velocities [27]. Given the latter deviation, Chapter 2 of this study adopts the subglottal model proposed by Zanartu [7] as it successfully reproduces the first three subglottal resonances through an expansion section below $20\,cm$ rather than a fully uniform tube.

Perhaps the most common method is to segment the tract into a series of cylinders with equal thickness and different radii [28]. The acoustic wave generated by flow through the glottis is then propagated from each cylindrical section to the next in a sequential fashion [28]. Computational savings are realized by simultaneously back-propagating the reflected wave in each segment for every time step [28]. Oftentimes either the subglottal section or the supraglottal section is ignored to concentrate on the interaction between one half of the tract with the glottis. Additional algorithmic complexity arises in modeling the lengthening of the tract, lip rounding and filtering of certain frequencies and pressure attenuation from heat and momentum losses to the compliant walls [7]. Another approach of note that includes these effects is to treat the vocal tract column as a pipe and consider the glottis to be valve. This produces a water hammer effect and greatly reduces the computational cost in accounting for the vocal tract as was performed by Sciamarella et al. [29].

1.4 Thesis objectives and scope

The central objective of this study was to determine whether ingressive phonation could be simulated via a lumped-parameter approach and if so under what conditions. The choice of physiological conditions was limited to the prephonatory gap between the vocal folds and the lung pressure. Several parameter scans were performed to investigate the roles of these two variables. The spring stiffness multiplier q was set to a value of 1 such that the tension of the folds was kept constant.

This study hypothesizes that ingressive phonation can be interpreted through traditional lumped- parameter methods with the following modifications: a) a negative gauge pressure in the lungs and b) reversal of airflow.

The scope of the work will be limited as follows, although the importance of these aspects to a more general understanding of ingressive phonation is recognized:

- the two folds are assumed to be identical in their geometry, material properties and motions such that calculations are required for only one of them
- the false vocal folds are not incorporated in the analysis
- the moving masses will only oscillate along a single axis toward or away from their counterparts in the other vocal fold
- the tissue stiffnesses are approximated as being linear
- the subglottal and supraglottal tracts will be treated as straight waveguides
- the nasal cavity is not included
- the flow will be treated as non-viscous and time-invariant
- the prephonatory profile of the folds is assumed to be rectangular (neither divergent nor convergent)

This study will show that current techniques can in fact be utilized to simulate ingressive phonation and thereby provides a unified theoretical framework for both kinds of phonation.

1.5 Thesis overview

Chapter 2 presents a generalization of multi-mass models with tissue oscillation along the lateral axis (adduction/abduction) of the folds. The aerodynamics of the folds are coupled to sub- and supraglottal resonators, the latter with lip radiation. Chapter 3 concludes the study, reviews the key findings and suggests future work based on the new framework. An appendix is also provided to show the equivalency between the state-space representation used in Chapter 2 and Newton's laws of motion.

Chapter 2

Ingressive Phonation using a Multi-Mass Model

2.1 Introduction

Researchers in vocal acoustics have used computer simulations of single and multi-mass models of the human vocal folds to reproduce and understand human phonation for over 40 years [30]. They have successfully given insight into different voice qualities and singing registers as well as the irregular phonation associated with pathologies. Despite their widespread use, they have not yet been applied to sounds generated by a negative lung gauge pressure and reversal of airflow which lead to ingressive phonation [31]. According to Kelly and Fisher in 1999, "phonatory modeling and theory in general have not included the mechanisms by which vocal fold oscillation can be initiated or sustained by ingressive airflow [31]." They point out that one of the only attempts to study a model of phonation with airflow in both directions was carried out by van den Berg, Zantema, and Doornenbal in 1957 [31, 32].

The occurrence of ingressive sounds in human communication [5] justifies further theoretical studies to complement experimental ones. The utility of ingressive phonation studies goes beyond linguistics and pathologies to understanding laryngeal tissue mechanics and aerodynamics. While the top portion of the vocal folds is visible with high-speed laryngoscopy, the bottom portion is not. By reversing the direction of airflow and consequently the tissue motion, the motion of the folds in regular phonation from the perspective below the folds can be inferred [33]. As Orlikoff states: "renewed examination of phona-

tion produced using ingressive airflow may yield important insight regarding mucosal and laryngeal airway dynamics [34]."

Experimentalists have observed several key differences between egressive and ingressive phonation. Childers and Larar [33] used ultrahigh-speed laryngeal films combined with EGG to compare a subject phonating with an /i/ vowel with both flow directions. In ingression phonation, they observed the mucosal wave traveling in the opposite direction from regular phonation and noted mild vocal fry in ingressions with a double beat pattern. Kelly and Fisher [31], whose participants also phonated with an /i/ vowel, likewise noted the reversal of the mucosal wave direction. In contrast to the anterior glottal chink present in egressive phonation, they observed the glottis closing anteriorly and a large posterior glottic opening for all of their participants. In common with regular phonation, they observed that ingressive phonation occurs with the true vocal folds. Orlikoff et al. recorded microphone, EGG and airflow signals and observed decreased vocal fold contact, significantly higher fundamental frequency (F0) and greater absolute airflow by 48.5% relative to egressive phonation [34].

Given these significant differences and the shared occurence of true vocal fold vibration, the question arises as to whether current egressive models are still applicable with slight modifications or if an entirely new framework is required for simulating this inverse approach to phonation.

This study hypothesizes that current techniques for simulating egressive phonation will also lead to self-oscillation of the vocal folds upon changing the pressure and flow orientation which lead to ingression. Through symmetrical reasoning, well-established aeroelastic relations for both multi-mass and single-mass vocal fold models are generalized to cases where the lung pressure is less than atmospheric and the air flows into the trachea. The resulting pressure forces are then used for simulating the response of the vocal fold masses in terms of their motion.

The pitch, flow rates, and range of phonation are compared between the ingressive and egressive conditions. A modal analysis is also used to understand the differences in motion between the two approaches.

2.2 Methods

2.2.1 Overview

This paper considers the dynamic behavior of the full system shown in Figure 2.1. The system is based directly on Tokuda et al.'s model [1]. It is composed of three components: a valve (the glottis) at $z = 0$ and two waveguide resonators on either end of it [29]. The valve opens and closes based on the vibratory motion of the vocal fold tissue [29]. The glottal air flow depends on the valve's aperture size and the pressures upstream and downstream from it [35]. An increasing flow causes a rise in the downstream pressure and drop in the upstream pressure [22]. Such pressure variations propagate as waves towards the ends of the upstream and downstream tubes and undergo reflections and transmissions at the interfaces of pairs of sections with different cross-sectional area [36]. The subglottal resonator on the left represents the airways below the larynx including the lungs and trachea [37]. The supraglottal resonator corresponds to the pharynx and oral cavity [38]. Lip radiation is included and acts to raise the higher frequencies exiting the mouth and to reflect mainly low frequencies back towards the glottis [39]. In their model, Tokuda et al. [1] did not include the nasal cavity, piriform sinus or the epilaryngeal constriction and the supraglottal waveguide was assigned a constant cross-section of 3 cm^2. The model is rich enough to consider the coupling between the source and its two filters (resonators) as the pressures upstream and downstream to it are composed of both the static lung and oral pressures along with the dynamic acoustic pressures which affect the flow through the folds and the pressure distribution along them [36].

What is unique about the present analysis is the consideration of cases where the static oral pressure is non-atmospheric and instances where the static lung pressure is lower than the static oral pressure or even sub-atmospheric, in other words a weak partial vacuum. The only other variations to the system considered in [1] are the inclusion of reflection coefficients in the glottal flow calculation as suggested by Titze [35], since the glottis has a finite area and should therefore not be approximated as a fully closed end, and an increment of the damping ratios by 1.0 upon collision as suggested by Ishizaka and Flanagan

[15].

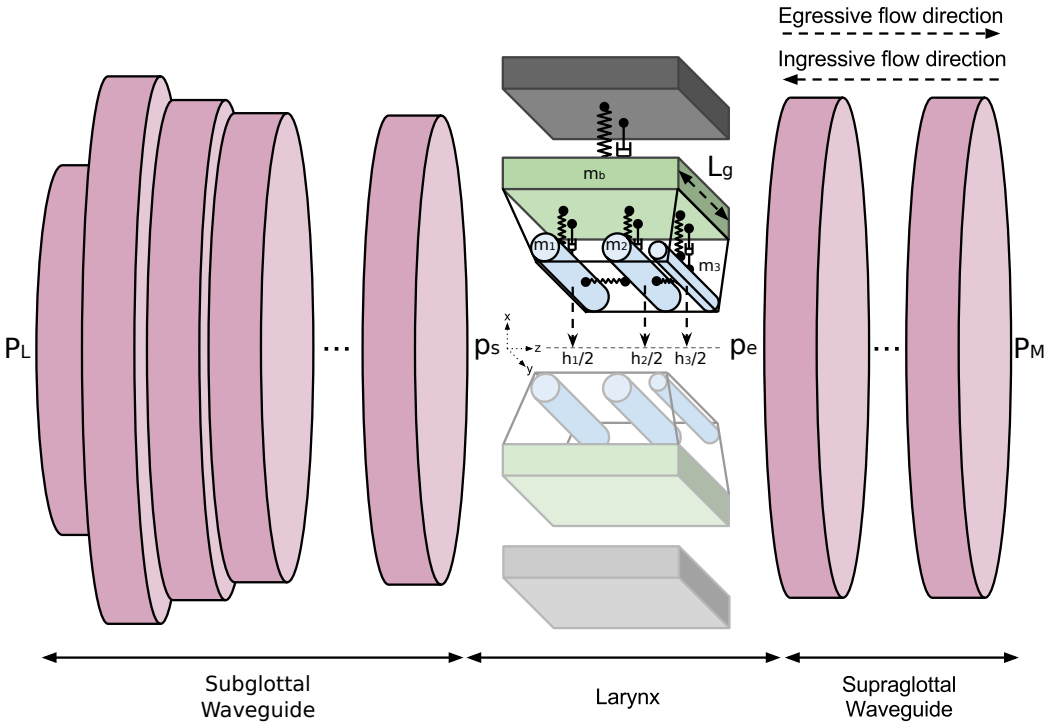


Figure 2.1: Full System showing subglottal and supraglottal resonators, larynx and possible flow directions. The lung and mouth pressures are shown along with the pressures upstream and downstream of the larynx. The system is based on the Tokuda et al. model [1]. The larynx is comprised of two symmetrical vocal folds each represented by a body mass and three cover masses connected by springs and dampers. The half-heights along the profile are shown along with the vocal fold length.

2.2.2 Tissue Mechanics

While the earliest vocal fold simulations used single mass oscillators to reproduce phonation, two or more mass models were eventually explored [21]. A recent model of the glottis, proposed by Tokuda et al. [1] and depicted in Figure 1.3, divides the vibrating part of the vocal folds into a body mass and three cover masses. The dampers r_1, r_2 and r_3 dissipate the energy in the system. The body-cover springs k_1, k_2 and k_3 act to maintain a certain distance between the body mass and the cover masses while the body spring k_b acts between the body and the immobile part of the vocal folds attached

to the laryngeal wall. The cover springs $k_{1,2}$ and $k_{2,3}$ provide stiffness solely along the x-axis despite appearing vertical. The air pressures upstream and downstream of the folds generally exert a force away from the midline and serve to push the folds apart during the opening phase of the glottal cycle. A unique element of their model is the presence of 3 cover masses, the third of which is significantly lighter than the other two. The researchers have successfully modeled register breaks and other phenomena with a single parameter set using this configuration [1], [25].

In considering ingressive phonation, it is difficult to anticipate in advance how many masses are required to reflect a given phenomena of interest. The general case with an arbitrary number of cover masses with a single body mass is considered.

For a general oscillator with n-degrees of freedom:

$$\mathbf{M}\ddot{\mathbf{x}} + \mathbf{R}\dot{\mathbf{x}} + \mathbf{K}\mathbf{x} = \mathbf{F}(t) \tag{2.1}$$

The vectors $\mathbf{x}$, $\dot{\mathbf{x}}$, and $\ddot{\mathbf{x}}$ represent the oscillatory displacements, velocities and accelerations of $n+1$ control point masses respectively noting that the body mass motion is included in these vectors: $\mathbf{x} = [x_1\, x_2 \ldots x_n\, x_b]^T$. The positive x direction is associated with a movement away from glottal midline along the dorsalventral axis.

Taken together, these represent the system's state at a given point in time:

$$\mathbf{w} = \begin{bmatrix} \mathbf{x} \\ \dot{\mathbf{x}} \end{bmatrix} \tag{2.2}$$

With initial conditions:

$$\mathbf{x}(t_0) = \mathbf{x_0}\ ;\ \dot{\mathbf{x}}(t_0) = \dot{\mathbf{x}}_\mathbf{0}\ ;\ \mathbf{w_0} = \begin{bmatrix} \mathbf{x_0} \\ \dot{\mathbf{x}}_\mathbf{0} \end{bmatrix} \tag{2.3}$$

The actual position of the masses relative to the midline $h_i/2$ is the sum of the displacements x_i and their corresponding prephonatory half-heights x_{0i}:

$$\mathbf{h} = 2\begin{bmatrix} x_{01} \\ x_{02} \\ x_{03} \\ \vdots \\ x_{0n} \end{bmatrix} + 2\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_n \end{bmatrix} \tag{2.4}$$

During a collision, the force pushing apart the folds is modeled as being proportional to their linear interpenetration past the midline following Lous et al. [2]. The proportionality constant is given by the collisional springs k_c which are taken to be three times the magnitude of the cover-to-body springs k_i [1, 2]. As such, the force exerted is expressed by:

$$F_{ci} = -k_c(h_i/2) = -k_c(x_{0i} + x_i) \quad h_i < 0 \tag{2.5}$$

The dampening depends on the mass, stiffness and dampening ratio ζ_i of each component "i":

$$r_i = 2\zeta_i\sqrt{m_i k_i} \tag{2.6}$$

As shown in Figure 2.2, the total forces acting on each mass include the aerodynamic pressure force, the dampening force, the spring force to the body and the spring forces to its adjacent masses which gives:

$$\begin{aligned} \underbrace{\sum F_i}_{\text{total}} = \underbrace{F_{aero,i}}_{\text{air}} - \underbrace{r_i(\dot{x}_i - \dot{x}_b)}_{\text{dampening}} - \underbrace{k_i(x_i - x_b)}_{\text{body-cover}} - \underbrace{k_c(h_i/2)}_{\text{collisional}} \\ - \underbrace{k_{i-1,i}(x_i - x_{i-1})}_{\text{lower mass}} - \underbrace{k_{i,i+1}(x_i - x_{i+1})}_{\text{upper mass}} \end{aligned} \tag{2.7}$$

For the body mass:

$$\underbrace{\sum F_b}_{\text{total}} = -\underbrace{r_b(\dot{x}_b)}_{\text{dampening}} - \underbrace{k_b(x_b)}_{\text{body}} - \underbrace{\sum_{i=1}^{n} r_i(\dot{x}_b - \dot{x}_i)}_{\text{dampening}} - \underbrace{\sum_{i=1}^{n} k_i(x_b - x_i)}_{\text{stiffness}} \tag{2.8}$$

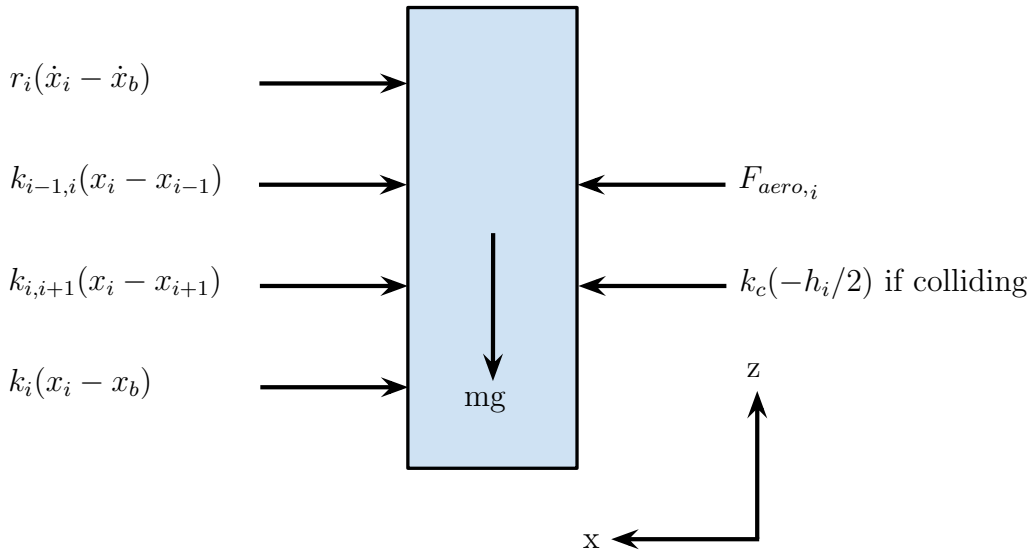


Figure 2.2: Free body diagram of a cover mass

Using Newton's 2nd law of motion we can derive the acceleration of the masses to determine their trajectory in state space:

$$\ddot{x}_i = \frac{\sum F_i}{m_i} \tag{2.9}$$

$$\ddot{x}_b = \frac{\sum F_b}{m_b} \tag{2.10}$$

2.2.3 Aerodynamics

We return our attention to $\mathbf{F}(t)$, the matrix of applied forces to each mass, in equation (2.1). In a lumped-parameter glottal model, the aerodynamic pressure profile along the vocal fold surface supplies the energy behind these forces and is of central importance. What is unique about our approach is that no prior assumption is made as to the direction of the flow u and the variable can therefore take on both positive and negative values to represent flow from and into the trachea respectively.

To derive the aperture areas at any two points, we first express the flow

channel height as a piecewise linear function [1]:

$$h_{i,i-1}(z,t) = \frac{h_i(t) - h_{i-1}(t)}{z_i - z_{i-1}}(z - z_{i-1}) + h_{i-1}(t) \tag{2.11}$$

here $h_{i,i-1}$ provides the height function within $(z_i \leq z \leq z_{i-1})$

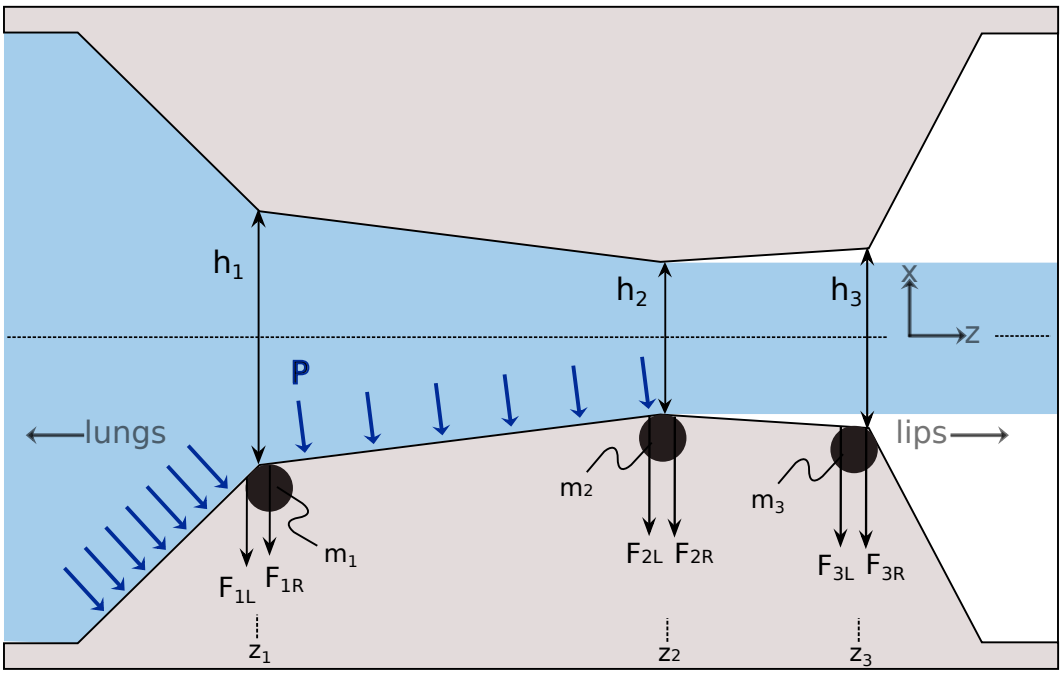


Figure 2.3: Aerodynamic forces and reaction forces [after Lous et al.[2]]. The forces are actually applied to both sides of the glottis but only the bottom fold is shown with force for comprehension purposes.

Each mass supports two reaction forces as a result of the aerodynamic forces acting on the massless plates to its left and right. The right reaction force of one mass and the left reaction force of the adjacent mass support the total aerodynamic force bearing down on the plate between them:

$$F_{i-1R} + F_{iL} = L_g \int_{z_{i-1}}^{z_i} P(z)dz \tag{2.12}$$

Treating mass $i-1$ as the moment origin gives the following:

$$F_{iL}(z_i - z_{i-1}) = L_g \int_{z_{i-1}}^{z_i} (z - z_{i-1}) P(z) dz$$
$$F_{iL} = \int_{z_{i-1}}^{z_i} L_g \frac{z - z_{i-1}}{z_i - z_{i-1}} P(z) dz \tag{2.13}$$

Treating mass $i+1$ as the moment origin gives the following:

$$F_{iR}(z_{i+1} - z_i) = L_g \int_{z_i}^{z_{i+1}} (z_{i+1} - z) P(z) dz$$
$$F_{iR} = \int_{z_i}^{z_{i+1}} L_g \frac{z_{i+1} - z}{z_{i+1} - z_i} P(z) dz \tag{2.14}$$

As such, the total reaction force at each mass can be evaluated as follows: [1, 2]:

$$F_{aeroi}(t) = \int_{z_{i-1}}^{z_i} L_g \frac{z - z_{i-1}}{z_i - z_{i-1}} P(z,t) dz + \int_{z_i}^{z_{i+1}} L_g \frac{z_{i+1} - z}{z_{i+1} - z_i} P(z) dz \tag{2.15}$$

To derive the pressure relation between any two points within the larynx, the simplest approximation is a steady, laminar, incompressible, inviscid and lossless airflow given by Bernoulli's equation [40, 41] where v is the average velocity:

$$P_1 + \frac{1}{2}\rho v_1^2 = P_2 + \frac{1}{2}\rho v_2^2 \tag{2.16}$$

For any two points with cross-sectional areas A_1 and A_2, with volumetric flow rates u_1 and u_2, mass balance gives [41]

$$u = u_1 = A_1 v_1 = u_2 = A_2 v_2 \tag{2.17}$$

such that Bernoulli's equation can be written as:

$$P_1 + \frac{1}{2}\rho \frac{u_1^2}{A_1^2} = P_2 + \frac{1}{2}\rho \frac{u_2^2}{A_2^2} \tag{2.18}$$

By continuity of flow:

$$P_1 + \frac{1}{2}\rho\frac{u^2}{A_1^2} = P_2 + \frac{1}{2}\rho\frac{u^2}{A_2^2} \tag{2.19}$$

Bernoulli's equation is invoked up to the point of jet separation at which point the pressure is set to the jet pressure [42]. In their model, Tokuda et al. [1] set the jet pressure equal to the pressure downstream of the glottis, which in the case of egressive phonation is the oral pressure. In the present analysis, the jet pressure is equated to the pressure downstream of the glottis regardless of flow direction.

$$P(z) + \frac{1}{2}\rho\frac{u^2}{a_i^2} = P_{jet} + \frac{1}{2}\rho\frac{u^2}{a_g^2} \tag{2.20}$$

$$P(z) = P_{jet} + \frac{1}{2}\rho u^2\left[\frac{1}{a_g^2} - \frac{1}{a_i^2}\right] \tag{2.21}$$

Since the glottal areas can be expressed as the heights by the length of the vocal folds [36]:

$$a_i = 2(x_i + x_{i0})L_g = h_i L_g \tag{2.22}$$

this gives an expression for the glottal area mentioned above based on the minimum height [1]:

$$a_g = \begin{cases} (h_{min}L_g), & h_{min} > 0 \\ 0, & h_{min} < 0 \end{cases} \tag{2.23}$$

It is important to note that the glottal aperture area a_g provided here differs from the a_{min} defined by [1] which equates to $h_{min}L_g$ regardless of whether the folds are in collision or not. The present definition of a_g is considered preferable as it lessens the requirements for specifying conditionals when h_{min} is negative by making it implicit.

In the standard case where the pressure is lower in the epilarynx than in the trachea, the pressure $P(z)$ at any point along the vocal fold profile can therefore be determined via [36]:

$$P_s > P_e : P(z) = \begin{cases} P_e + \frac{\rho}{2}\left(\frac{u}{L_g}\right)^2 \left(\frac{1}{(h_{min})^2} - \frac{1}{(h(z))^2}\right), & \text{if } z < z_{sep} \\ P_e, & \text{if } z \geq z_{sep} \end{cases} \tag{2.24}$$

Here z_{sep} refers to the position along z at which the flow is predicted to separate from the vocal fold surface as a jet. For simplicity, this is taken as the vertical point of minimum height between the folds [1, 43].

Based on symmetrical reasoning, the following approach is proposed to determine the pressure profile $P(z)$ during ingressive phonation:

$$P_s < P_e : P(z) = \begin{cases} P_s, & \text{if } z \leq z_{sep} \\ P_s + \frac{\rho}{2}\left(\frac{u}{L_g}\right)^2 \left(\frac{1}{(h_{min})^2} - \frac{1}{(h(z))^2}\right), & \text{if } z > z_{sep} \end{cases} \tag{2.25}$$

Within a collision's timeframe, it has been suggested that the dynamics of the flow boundary layer could affect the rate of flow decrease [43] but since this model uses a tophat (uniform) velocity profile, this effect is not built into the simulation. In fact, any consideration whatsoever of the dynamics of the flow field from variable geometry could be misleading. Clearly, the true flow field would be complex and three-dimensional, but to avoid the extreme complexity of full three-dimensional direct numerical simulation of the Navier-Stokes equations resolving the smallest flow structures to the Kolmogorov length scale, one must draw the line, balancing the mutually exclusive advantages of simplicity and rigour. In this simplified analysis, the flow is arrested when the vocal folds come into contact. Regardless of the prior flow direction, the pressure at a given point will then be determined by:

$$h_{min} < 0 : P(z) = \begin{cases} P_s, & \text{if } z \leq z_{h_{min}} \\ P_e, & \text{if } z > z_{h_{min}} \end{cases} \tag{2.26}$$

2.2.4 Waveguides and acoustic Pressures

The pressures on the epilaryngeal and subglottal sides of the glottis, p_e and p_s, are taken to be the sum of the static pressures, P_M and P_L and the superposition of the backward and forward traveling pressure waves [36]. While many researchers ignore P_M [1], as the mouth pressure is near-atmospheric and therefore set to zero, this study includes it for the sake of generality:

$$p_e = p_e^+ + p_e^- + P_M \tag{2.27}$$

$$p_s = p_s^+ + p_s^- + P_L \tag{2.28}$$

The forward and backward traveling waves are modeled using waveguides based on the Kelly-Lochbaum model [28].

The sub- and supraglottal resonators are each modeled by their own waveguide. Each waveguide consists of a series of coaxial cylinders of equal thickness and different cross-sectional areas [44]. The latter implies changes in the acoustic impedance leading to reflection and transmission of the acoustic waves at the interface between each pair of cylinders [45].

The $forward$ direction is here defined as moving away from the glottis and either towards the lungs or the lips. The *backward* direction is defined as moving towards the glottis. On the supraglottal side, + and − refer to the $forward$ and *backward* directions respectively. The notation is reversed on the subglottal side; + referring to a wave traveling towards the glottis. A uniformity in notation is thus maintained in that + always refers to waves traveling in the direction from the lungs towards the lips and − in the opposite direction in keeping with [35]. For a "half-sample delay" waveguide model that provides calculation stability by ensuring updated information is transported exactly one node distance per time step (neglecting Doppler effect), the segment length is given by [28]:

$$L_{seg} = c/(2F_s) \tag{2.29}$$

This dictates a segment length of 0.397 cm if the speed of sound is 35000 cm/s and the sampling frequency is 44.1 kHz [28]. The typical adult male vocal tract

is approximately 17.5 *cm* [46] thus requiring 44 segments. The areas for the 44-segment supraglottal waveguide follow the Story *et al.* data based on MRI studies of a single adult male [47]. Losses and lip radiation are included as was done in previous studies [7]. The subglottal waveguide is based on the 64-segment model proposed by Zanartu et al. [7]. The segments are numbered according to the forward direction. As such, the index "1" denotes the first segment adjacent to the glottis for both waveguides whereas "44" and "64" refer to the last segment (lips/lungs) for the supra- and subglottal waveguides respectively.

The reflection coefficient for a wave traveling from segment i to segment $i+1$ is [28]:

$$r_i = \frac{(A_i - A_{i+1})}{(A_i + A_{i+1})} \tag{2.30}$$

There are three limiting cases of interest. The first is representative of cases where the subsequent tube is much larger than the first tube and the wave essentially moves into an open space as is the case at the lips [28]. The second case is when the subsequent tube has no area and the propagating wave sees only a wall as for example a closed glottis [28]. The last case is when the two tubes have the same area and the wave does not scatter.

$$r_{open} = \lim_{A_{i+1}\to\infty} \frac{(A_i - A_{i+1})}{(A_i + A_{i+1})} = -1 \tag{2.31}$$

$$r_{closed} = \lim_{A_{i+1}\to 0} \frac{(A_i - A_{i+1})}{(A_i + A_{i+1})} = 1 \tag{2.32}$$

$$r_{neutral} = \lim_{A_{i+1}\to A_i} \frac{(A_i - A_{i+1})}{(A_i + A_{i+1})} = 0 \tag{2.33}$$

The partial pressures impinging on the glottis are taken from the backward waveguide lines at the previous time-step:

$$p_s^+ = b_s(1) \quad p_e^- = b_e(1) \tag{2.34}$$

The tracheal and epilaryngeal fluid properties are assumed to be nearly

equivalent for this study as others have done [1]:

$$\rho_e = \rho_s = \rho_{air} \;\; ; \;\; c_e = c_s = c_{air} \tag{2.35}$$

The forward pressures are derived via the assumption of flow continuity above and below the glottis [48, 35]:

$$p_e^+ = r_e p_e^- + u_n(\frac{\rho_e c_e}{A_{e,1}}) \quad p_s^- = r_s p_s^+ - u_n(\frac{\rho_s c_s}{A_{s,1}}) \tag{2.36}$$

These are used as inputs into the waveguides via the initial segments:

$$f_e(1) = p_e^+ \quad f_s(1) = p_s^- \tag{2.37}$$

The following expressions for the scattering equations are based on [28] but with the inclusion of propagation loss factors according to [49]. The equations apply to the scattering at junctions within subglottal tract by replacing the e subscript with s.

$$f_{e,(i+1)} = (1 + r_{e,(i)}) f_{e,(i)} \gamma_{e,(i)} - r_i b_{e,(i+1)} \gamma_{e,(i+1)} \tag{2.38}$$

$$b_{e,(i)} = r_{e,(i)} f_{e,(i)} \gamma_{e,(i)} + (1 - r_{e,(i)}) b_{e,(i+1)} \gamma_{e,(i+1)} \tag{2.39}$$

The propagation loss factors for each section are calculated according to [7]:

$$\alpha_{e,(i)} = \frac{0.0112}{\sqrt{A_{e,(i)}}} \tag{2.40}$$

$$\gamma_{e,(i)} = e^{-L_{seg}\alpha_{e,(i)}} \tag{2.41}$$

The reflection coefficients result from the relative areas of the glottis and the adjacent sections of the waveguides [35]:

$$r_s = \frac{(A_{s1} - a_g)}{(A_{s1} + a_g)} \quad r_e = \frac{(A_{e1} - a_g)}{(A_{e1} + a_g)} \tag{2.42}$$

If the glottis is closed ($a_g = 0$) then the reflection coefficients are close to 1

corresponding to a closed tube. This is in contrast to the last segment in the subglottal waveguide, where the lungs are modeled as not reflecting ($r_{lung} = 0$) and the lips which are either treated as an open section ($r_{lips} = -1$) or with radiation. As a_g increases as the folds separate, the glottal section will behave more like a neutral end.

2.2.5 Lip radiation

As an alternative to assuming an open end at the lip segment of the supraglottal waveguide, the following equations are used [28]. These convert Flanagan's infinite plain baffle model to the discrete-sampling domain via the bilinear transform [28]. These equations provide the output lip pressure P_{out} and the backward traveling wave pressure b_M based on their previous values P_{prev}, b_{prev}, along with the forward traveling component f_M and its former value f_{prev}.

$$b_M = \frac{1}{b_2}(f_M a_2 + f_{prev} a_1 + b_{prev} b_1) \tag{2.43}$$

$$P_{out} = \frac{1}{b_2}(P_{prev} b_1 + f_M(b_2 + a_2) + f_{prev}(a_1 - b_1)) \tag{2.44}$$

with filter coefficients according to [28]:

$$\begin{aligned} a_2 &= -R' - L' + R'L' \\ a_1 &= -R' + L' - R'L' \\ b_2 &= R' + L' + R'L' \\ b_1 &= -R' + L' + R'L' \\ R' &= \frac{128}{9\pi^2} \\ L' &= 2F_s \frac{8a}{3\pi c} \end{aligned} \tag{2.45}$$

2.2.6 Aeroacoustic Coupling

Titze [48] provides a relation between the transglottal pressure and the flow that assumes negligible air inertia within the glottis and that viscous losses can be included in a pressure coefficient, k_t:

$$P_{trans} = k_t \frac{1}{2} \rho \left(|u| \, u / a_g^2 \right) \tag{2.46}$$

Recognizing that the transglottal pressure P_{trans} is defined by $p_s - p_e$ [48]

$$\left[p_s^+ + p_s^- + P_L \right] - \left[p_e^+ + p_e^- + P_M \right] = k_t \frac{1}{2} \rho \left(|u| \, u / a_g^2 \right)$$

$$\left[p_s^+ + \left\{ r_s p_s^+ - \frac{u \rho c}{A_s} \right\} + P_L \right] - \left[p_e^- + \left\{ r_e p_e^- + \frac{u \rho c}{A_e} \right\} + P_M \right] = k_t \frac{1}{2} \rho \left(|u| \, u / a_g^2 \right)$$

$$\left[(1 + r_s) p_s^+ - (1 + r_e) p_e^- \right] + \left[P_L - P_M \right] - \rho c u \left[\frac{1}{A_s} + \frac{1}{A_e} \right] = k_t \frac{1}{2} \rho \left(|u| \, u / a_g^2 \right) \tag{2.47}$$

Defining A_{eff} through:

$$A_{\text{eff}} = \left[\frac{1}{A_s} + \frac{1}{A_e} \right]^{-1} \tag{2.48}$$

We arrive at the following 2nd-order polynomial in u

$$k_t \frac{1}{2} \frac{\rho}{a_g^2} \left(|u| \, u \right) + \rho c \Big(\frac{1}{A^{\text{eff}}} \Big) u - \left\{ P_L - P_M + (1 + r_s) p_s^+ - (1 + r_e) p_e^- \right\} = 0 \tag{2.49}$$

It is appropriate at this point to propose a new term Ω defined as:

$$\Omega \equiv \left\{ P_L - P_M + (1 + r_s) p_s^+ - (1 + r_e) p_e^- \right\} \tag{2.50}$$

The polynomial can be re-written in quadratic form:

$$\left[k_t \frac{1}{2} \frac{\rho}{a_g^2} \right] |u| \, u + \left[\frac{\rho c}{A_{\text{eff}}} \right] u - \Omega = 0 \tag{2.51}$$

Bernoulli's equation cannot suggest the direction of the flow. The direction of the flow is imposed based on the relative magnitudes of P_L and P_M. If P_L is the greater of the two, then the flow is taken to be forward and equation (2.51) is solved assuming $u > 0$ giving the standard result [36] with the reflection coefficients at the glottis approximated by their closed-end values $r_e \approx r_s \approx 1$:

$$u_g = \frac{a_g c}{k_t}\left\{-\left(\frac{a_g}{A_{\text{eff}}}\right) \pm \left[\left(\frac{a_g}{A_{\text{eff}}}\right)^2 + \frac{2k_t}{\rho c^2}(P_L - P_M + 2p_s^+ - 2p_e^-)\right]^{1/2}\right\} \tag{2.52}$$

By symmetry, if P_L is less than P_M, then the flow would be in the reverse direction and equation (2.51) can be solved assuming $u < 0$. The two cases can be summarized as follows

$$u_g = \begin{cases} \dfrac{a_g c}{k_t}\left\{-\left(\dfrac{a_g}{A_{\text{eff}}}\right) + \left[\left(\dfrac{a_g}{A_{\text{eff}}}\right)^2 + \dfrac{2k_t}{\rho c^2}\Omega\right]^{1/2}\right\}, & \text{if } P_L > P_M \\ -\dfrac{a_g c}{k_t}\left\{-\left(\dfrac{a_g}{A_{\text{eff}}}\right) + \left[\left(\dfrac{a_g}{A_{\text{eff}}}\right)^2 - \dfrac{2k_t}{\rho c^2}\Omega\right]^{1/2}\right\}, & \text{if } P_L < P_M \end{cases} \tag{2.53}$$

When the resonators are specified as not acting, the expressions for the total pressure difference and the flow simplify to the following equation. This approximation is also valid when the coupling between the tracts and the glottal source are low as when their resonant frequencies are far apart [50, 48, 2].

$$\Omega = \{P_L - P_M\} \tag{2.54}$$

$$u_g = \begin{cases} a_g\left\{\left[\dfrac{2\Omega}{\rho}\right]^{1/2}\right\}, & \Omega > 0 \\ -a_g\left\{\left[\dfrac{-2\Omega}{\rho}\right]^{1/2}\right\}, & \Omega < 0 \end{cases} \tag{2.55}$$

2.2.7 Modal analysis

It is possible to anticipate the vibratory patterns (eigenmodes) and frequencies of the system under consideration. The masses (**M**), dampening coefficients (**R**) and stiffnesses (**K**) can be put in matrix form with the assumption of proportional (Rayleigh) dampening:

$$\mathbf{M} = \begin{bmatrix} m_1 & & & & \\ & m_2 & & & \\ & & \ddots & & \\ & & & m_n & \\ & & & & m_b \end{bmatrix} \tag{2.56}$$

$$\mathbf{R} = \begin{bmatrix} r_1 & & & & \\ & r_2 & & & \\ & & \ddots & & \\ & & & r_n & \\ & & & & r_b \end{bmatrix} \tag{2.57}$$

For the 4-mass model:

$$\mathbf{K} = \begin{bmatrix} (k_1 + k_{12}) & -k_{12} & 0 & -k_1 \\ -k_{12} & (k_2 + k_{12} + k_{23}) & -k_{23} & -k_2 \\ 0 & -k_{23} & (k_3 + k_{23}) & -k_3 \\ -k_1 & -k_2 & -k_3 & (k_b + k_1 + k_2 + k_3) \end{bmatrix} \tag{2.58}$$

The forces acting from the exterior of the folds on the masses are:

$$F_{exi} = F_{aero,i} + k_c(-h_i/2) \tag{2.59}$$

Noting that there is no aerodynamic or collisional force applied to the body mass ($F_{exb} = 0$), the external force vector can be defined as $\mathbf{F} = [F_{ex1}\, F_{ex2} \ldots F_{exn}\, 0]^T$.

To simplify the analysis, the second-order differential equation 2.1 can be put in first-order form [51]. With $\mathbf{I}$ representing the $(n+1)by(n+1)$ identity matrix, Equation (2.1) can be rewritten into state space as [52]:

$$\begin{bmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{0} & \mathbf{M} \end{bmatrix} \begin{bmatrix} \dot{\mathbf{x}} \\ \ddot{\mathbf{x}} \end{bmatrix} = \begin{bmatrix} \mathbf{0} & \mathbf{I} \\ -\mathbf{K} & -\mathbf{R} \end{bmatrix} \begin{bmatrix} \mathbf{x} \\ \dot{\mathbf{x}} \end{bmatrix} + \begin{bmatrix} \mathbf{0} \\ \mathbf{F} \end{bmatrix} \tag{2.60}$$

Left multiplication by the inverse of the mass matrix gives [52]:

$$\frac{\mathrm{d}}{\mathrm{d}t} \begin{bmatrix} \mathbf{x} \\ \dot{\mathbf{x}} \end{bmatrix} = \begin{bmatrix} \mathbf{0} & \mathbf{I} \\ -\mathbf{M}^{-1}\mathbf{K} & -\mathbf{M}^{-1}\mathbf{R} \end{bmatrix} \begin{bmatrix} \mathbf{x} \\ \dot{\mathbf{x}} \end{bmatrix} + \begin{bmatrix} \mathbf{0} \\ \mathbf{M}^{-1}\mathbf{F} \end{bmatrix} \tag{2.61}$$

If a matrix $\mathbf{A}$ is defined according to [51]:

$$\mathbf{A} = \begin{bmatrix} \mathbf{0} & \mathbf{I} \\ -\mathbf{M}^{-1}\mathbf{K} & -\mathbf{M}^{-1}\mathbf{R} \end{bmatrix} \tag{2.62}$$

and $\mathbf{B}$ as:

$$\mathbf{B} = \begin{bmatrix} \mathbf{0} \\ \mathbf{M}^{-1}\mathbf{F} \end{bmatrix} \tag{2.63}$$

the dynamics of the system can then be derived from:

$$\dot{\mathbf{w}} = \mathbf{A}\mathbf{w} + \mathbf{B} \tag{2.64}$$

with $\mathbf{w}$ referring to the state vector as in equation (2.2).

According to [52] the matrix $\mathbf{A}$ is of central importance in characterizing the system's behavior in the absence of external forces ($\mathbf{F} = \mathbf{0}$). For the non-collided *in vacuo* case

$$\dot{\mathbf{w}} = \mathbf{A}\mathbf{w} \tag{2.65}$$

Its eigenvalues λ_i and eigenvectors $\mathbf{z}$ have the following relationship:

$$\lambda_i \mathbf{z} = \mathbf{A}\mathbf{z} \tag{2.66}$$

These eigenvectors correspond to the normal modes of vibration of the tissue in the absence of external forces. The eigenvalues correspond to an-

gular frequencies. Their linear frequencies f_i, from which vocal pitch can be predicted, can be derived via [25]:

$$f_i = |\Im(\lambda_i)| / 2\pi \tag{2.67}$$

Here $\Im$ refers to the imaginary component of the eigenvalue.

2.2.8 Aerodynamic state feedback

Equation (2.64) is a state-space description of a nonlinear system [53] and may be written in the following form:

$$\dot{\mathbf{w}} = \mathbf{g}(\mathbf{w}, \mathbf{B}) \tag{2.68}$$

The stability of the system can be evaluated by considering perturbations $\delta\mathbf{w}$ and $\delta\mathbf{B}$ around operating points defined by $\mathbf{g}(\mathbf{w_{eq}}, \mathbf{B_{eq}}) = \mathbf{0}$ [53].

$$\mathbf{w} = \mathbf{w_{eq}} + \delta\mathbf{w} \quad \mathbf{B} = \mathbf{B_{eq}} + \delta\mathbf{B} \tag{2.69}$$

$$\frac{\mathrm{d}}{\mathrm{d}t}(\mathbf{w_{eq}} + \delta\mathbf{w}) = \frac{\mathrm{d}}{\mathrm{d}t}(\delta\mathbf{w}) = \mathbf{g}(\mathbf{w_{eq}} + \delta\mathbf{w}, \mathbf{B_{eq}} + \delta\mathbf{B}) \tag{2.70}$$

$$(\dot{\delta\mathbf{w}}) = \mathbf{g}(\mathbf{w_{eq}}, \mathbf{B_{eq}}) + \left.\frac{\partial\mathbf{g}}{\partial\mathbf{w}}\right|_{(\mathbf{w_{eq}}, \mathbf{B_{eq}})} \delta\mathbf{w} + \left.\frac{\partial\mathbf{g}}{\partial\mathbf{B}}\right|_{(\mathbf{w_{eq}}, \mathbf{B_{eq}})} \delta\mathbf{B} + \dots \tag{2.71}$$

The first term is zero by definition while the two Jacobians can be evaluated from equation (2.64) [53]:

$$(\dot{\delta\mathbf{w}}) = \mathbf{A}\delta\mathbf{w} + \mathbf{I}\delta\mathbf{B} \tag{2.72}$$

According to [8] the forces arising from the Bernoulli effect depend on the cross-sectional areas along the channel and thereby the positions of the masses. Indeed, considering equations (2.59), (2.15), (2.25), and (2.24), the external forces acting on the masses, whether collisional or aerodynamic, result from their heights and therefore their displacements. They also depend on the

flow, however the flow in turn depends on the mass displacements through the glottal area term along with the Ω term. In essence then, the external forces acting on the tissue are a function of the state $\mathbf{w}$ and the pressure Ω term and the following can be written:

$$\delta\mathbf{B} \approx \frac{\partial\mathbf{B}}{\partial\mathbf{w}}\bigg|_{(\mathbf{w_{eq}},\Omega_{eq})}\delta\mathbf{w} + \frac{\partial\mathbf{B}}{\partial\Omega}\bigg|_{(\mathbf{w_{eq}},\Omega_{eq})}\delta\Omega \tag{2.73}$$

The lung pressure is the reference input to the system while the oral pressure is set to atmospheric. The only perturbation to it is from the acoustics of the waveguides. An appropriate choice for Ω_{eq} is therefore the reference pressure input in the no-waveguide case without acoustic feedback: $\Omega_{eq} = \{P_L - P_M\}$. As such, the perturbations can be attributed to the acoustic feedback effect. Hence, the case without waveguides reduces to:

$$\delta\mathbf{B} \approx \frac{\partial\mathbf{B}}{\partial\mathbf{w}}\bigg|_{(\mathbf{w_{eq}},\Omega_{eq})}\delta\mathbf{w} \tag{2.74}$$

Substituting the latter expression into the differential equation 2.72 gives:

$$(\dot{\delta\mathbf{w}}) = \mathbf{A}\delta\mathbf{w} + \mathbf{I}\frac{\partial\mathbf{B}}{\partial\mathbf{w}}\bigg|_{(\mathbf{w_{eq}},\Omega_{eq})}\delta\mathbf{w} \tag{2.75}$$

$$(\dot{\delta\mathbf{w}}) = \left[\mathbf{A} + \mathbf{I}\frac{\partial\mathbf{B}}{\partial\mathbf{w}}\bigg|_{(\mathbf{w_{eq}},\Omega_{eq})}\right]\delta\mathbf{w} = \hat{\mathbf{A}}\delta\mathbf{w} \tag{2.76}$$

Recognizing that:

$$\frac{\partial\mathbf{B}}{\partial\mathbf{w}} = \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{M}^{-1}\dfrac{\partial\mathbf{F}}{\partial\mathbf{x}} & \mathbf{M}^{-1}\dfrac{\partial\mathbf{F}}{\partial\dot{\mathbf{x}}} \end{bmatrix} \tag{2.77}$$

with the components of the above two Jacobians being given by $\dfrac{\partial F_i}{\partial x_j}$ and $\dfrac{\partial F_i}{\partial \dot{x}_j}$, an augmented system matrix can then be written as:

$$\hat{\mathbf{A}} = \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{M}^{-1}\left(\dfrac{\partial\mathbf{F}}{\partial\mathbf{x}}\bigg|_{(\mathbf{w_{eq}},\Omega_{eq})} - \mathbf{K}\right) & \mathbf{M}^{-1}\left(\dfrac{\partial\mathbf{F}}{\partial\dot{\mathbf{x}}}\bigg|_{(\mathbf{w_{eq}},\Omega_{eq})} - \mathbf{R}\right) \end{bmatrix} \tag{2.78}$$

$$(\dot{\delta\mathbf{w}}) = \hat{\mathbf{A}}\delta\mathbf{w} \tag{2.79}$$

The last two equations express the feedback of the aerodynamic forces as a result of the tissue's state [54]. Since the coefficients of $\hat{\mathbf{A}}$ are real, its eigenvalues, if complex, will each have their own conjugate [54]. Whereas consideration of the system matrix $\mathbf{A}$ provides the natural frequencies and eigenmodes of the system, an eigenvalue analysis of $\hat{\mathbf{A}}$ reveals the effect of the feedback from the aerodynamic forces on the masses as a result of their state. In particular, if the real part of any of the eigenvalues are greater than zero ($\Re(\lambda) > 0$), then the operating point is unstable and tissue oscillations will grow leading to phonation [23]. The eigenvalues may not all be complex and an eigenvalue and its conjugate may meet on and diverge along the real axis with one of them potentially taking on a positive value [8]. Such static divergence instabilities have been observed in vocal fold simulations [18].

What remains is to evaluate the change in the aerodynamic forces on each mass with respect to change in state variables. Rather than using a smooth geometry, Tokuda et al. [25] treated the masses as flat plates such that the air pressure across a mass was constant. To simplify the present analysis, the force on each mass will be approximated by the pressure at the location of the mass times its medial surface area:

$$F_i \approx \int_{0.5(z_i+z_{i-1})}^{0.5(z_i+z_{i+1})} L_g P(z) dz \approx L_g d_i P(z_i) \approx L_g d_i P_i \tag{2.80}$$

The approximation is taken to be an equality in the remainder of this study.

As before, Bernoulli's equation is invoked up to the point of jet separation at which point the pressure is set to the jet pressure [42]. The jet pressure is set equal to the downstream static pressure [1, 55]. In the present analysis, the jet pressure is equated to the pressure downstream of the glottis regardless of flow direction. Following [55] but with the jet pressure generalized as potentially non-zero:

$$P_i + \frac{1}{2}\rho\frac{u^2}{a_i^2} = P_{jet} + \frac{1}{2}\rho\frac{u^2}{a_g^2} \tag{2.81}$$

$$P_i = P_{jet} + \frac{1}{2}\rho u^2 \left[\frac{1}{a_g^2} - \frac{1}{a_i^2}\right] \tag{2.82}$$

$$\frac{\partial P_i}{\partial a_j} = \frac{\partial P_{jet}}{\partial a_j} + \frac{1}{2}\rho \left\{ 2u\frac{\partial u}{\partial a_j}\left[\frac{1}{a_g^2} - \frac{1}{a_i^2}\right] + u^2\left[\frac{-2}{a_g^3}\frac{\partial a_g}{\partial a_j} + \frac{2}{a_g^3}\frac{\partial a_i}{\partial a_j}\right]\right\} \tag{2.83}$$

Rewriting the expression for the flow (2.55) with the minus sign accounting for the condition of reversed flow and deriving the terms with flow:

$$u = \pm a_g\sqrt{\frac{2\,|\Omega|}{\rho}} \tag{2.84}$$

$$\frac{\partial u}{\partial a_j} = \pm\frac{\partial a_g}{\partial a_j}\sqrt{\frac{2\,|\Omega|}{\rho}} \tag{2.85}$$

$$u\frac{\partial u}{\partial a_j} = \frac{2\,|\Omega|}{\rho}\frac{\partial a_g}{\partial a_j}a_g \tag{2.86}$$

$$u^2 = \frac{2\,|\Omega|}{\rho}(a_g)^2 \tag{2.87}$$

Substituting the flow terms leads to the following expression for the changes in pressure with respect to changes in the areas:

$$\frac{\partial P_i}{\partial a_j} = \frac{\partial P_{jet}}{\partial a_j} + 2\,|\Omega|\left\{\frac{a_g}{(a_i)^2}\left[\frac{a_g}{a_i}\frac{\partial a_i}{\partial a_j} - \frac{\partial a_g}{\partial a_j}\right]\right\} \tag{2.88}$$

In terms of the forces:

$$\frac{\partial F_i}{\partial x_j} = L_g d_i \frac{\partial P_i}{\partial a_j}\frac{\partial a_j}{\partial x_j} \tag{2.89}$$

$$\frac{\partial a_j}{\partial x_j} = 2L_g \tag{2.90}$$

$$\frac{\partial F_i}{\partial x_j} = 2d_i(L_g)^2\left(\frac{\partial P_{jet}}{\partial a_j} + 2\,|\Omega|\left\{\frac{a_g}{(a_i)^2}\left[\frac{a_g}{a_i}\frac{\partial a_i}{\partial a_j} - \frac{\partial a_g}{\partial a_j}\right]\right\}\right) \tag{2.91}$$

Since the state variables are independent of each other, their partials will be unity when referrring to the same index and zero otherwise.

$$\frac{\partial a_i}{\partial a_j} = \frac{\partial a_i}{\partial x_i}\frac{\partial x_i}{\partial x_j}\frac{\partial x_j}{\partial a_j} = 2L_g\frac{\partial x_i}{\partial x_j}\frac{1}{2L_g} = \delta_{ij} \tag{2.92}$$

$$\left.\frac{\partial F_i}{\partial x_j}\right|_{(\mathbf{w_{eq}},\Omega_{eq})} = 4d_i(L_g)^2\left(\left\{\frac{a_g}{(a_i)^2}\left[\frac{a_g}{a_i}\delta_{ij} - \delta_{min,j}\right]\right\}\right)|\Omega_{eq}| \tag{2.93}$$

Here δ_{ij} refers to the Kronecker delta. Substituting the partials of F with respect to the state variables into equation (2.76) enables the calculation of the augmented system matrix and its eigenvalues providing an understanding of the stability of the feedback system.

Two extreme cases need to be considered: a rectangular channel with $a_g = a_1 = a_2 = a_3$ and a convergent channel with $a_g = a_1, a_3 > a_2 > a_1$ in the case of ingression and $a_g = a_3, a_1 > a_2 > a_3$ in the case of egression. Considering the subset of the Jacobian referring to the cover masses, the ingressive convergent channel leads to

$$\left.\frac{\partial \mathbf{F}}{\partial \mathbf{x}}\right|_{(\mathbf{w_{eq}},\Omega_{eq})} = 4d_i(L_g)^2|\Omega_{eq}|\begin{bmatrix} 0 & 0 & 0 \\ -\frac{a_1}{a_2^2} & \frac{a_1^2}{a_2^3} & 0 \\ -\frac{a_1}{a_3^2} & 0 & \frac{a_1^2}{a_3^3} \end{bmatrix} \text{ Matrix “C”} \tag{2.94}$$

while the rectangular channel leads to

$$\left.\frac{\partial \mathbf{F}}{\partial \mathbf{x}}\right|_{(\mathbf{w_{eq}},\Omega_{eq})} = 4d_i(L_g)^2\frac{|\Omega_{eq}|}{a_1}\begin{bmatrix} 0 & 0 & 0 \\ -1 & 1 & 0 \\ -1 & 0 & 1 \end{bmatrix} \; a_1 < a_3 \text{ Matrix “R13”} \tag{2.95}$$

$$\left.\frac{\partial \mathbf{F}}{\partial \mathbf{x}}\right|_{(\mathbf{w_{eq}},\Omega_{eq})} = 4d_i(L_g)^2\frac{|\Omega_{eq}|}{a_3}\begin{bmatrix}1 & 0 & -1\\ 0 & 1 & -1\\ 0 & 0 & 0\end{bmatrix} \text{ if } a_3 < a_1 \text{ Matrix "R31"} \quad (2.96)$$

Given the row of zeros, these matrices are degenerate and would, on their own, lead to eigenmodes corresponding to rigid-body motions [11]. The spring constants in the **K** matrix offset this degeneracy resulting in eigenvalues with non-zero frequencies. However at high Ω_{eq} values, the augmented matrix would be increasingly dominated by the aerodynamic feedback terms and therefore rigid-body motions. This effect should be less pronounced for the convergent height profiles since the denominators $a_2, etc.$ are greater than the numerators suggesting that rigid-body motions would only result at much greater Ω_{eq} values.

It should be noted that only the aerodynamic state feedback dependence on the displacement was included in this analysis as Zhang highlighted its importance over negative damping (velocity-dependent) terms [18]. Future work should consider the influence of purely negative damping in determining the conditions leading to phonation onset.

2.2.9 Computational algorithm

The equations for the tissue mechanics, aerodynamics and aeroacoustic coupling were coded in Matlab version 7.11.0 (R2010b) with 64-bit precision.

The algorithm simulates both the tissue movement and the changing acoustic wave pressure distribution in the airways in synchrony. As stated previously, the sampling frequency is set to 44.1 kHz [7] such that the time step between events is 0.0227 ms.

For each time step, the vocal fold medial height profile is derived from the prephonatory half-heights and the displacements via equations (2.4) and (2.11). The glottal area is then calculated from the product of the vocal fold length by the minimum height (2.23). The z position corresponding to the minimum height is taken to be the location of jet separation. The area is used

in calculating the reflection coefficients at the glottis (2.42). The acoustic wave pressures incident on the glottis are derived from the backward components of the waveguides (2.34). The latter are taken together with the lung and mouth pressures to form the aggregate pressure difference (2.50). If this value is greater than zero, the positive flow equation is employed, otherwise the negative flow equation determines the flow value (2.53). The flow value scaled by the acoustic impedance of the initial waveguide sections is combined with the wave component reflected from the glottis to give the pressure components entering the waveguides (2.36) and (2.37). At this point, the waveguide pressure distribution is updated for both the forward and backward delay lines with losses (2.39). The lip radiation is also calculated with the pressure exiting the last section of the oral waveguide being stored in analogy to a microphone signal. The part not exiting the waveguide is fed backwards towards the glottis. The calculated pressures upstream and downstream from the glottis involve the static lung and mouth pressures respectively along with the incident and reflecting acoustic wave components (2.27) and (2.28). The pressure distribution along the medial section of the folds is calculated based on the flow, the total upstream and downstream pressures and the height profile (2.24) and (2.25). The corresponding aerodynamic forces on each mass (2.15) are combined with collisional forces at locations of negative height to determine the acceleration of the masses (2.7). The corresponding accelerations are used by a fourth-order Runge-Kutta solver to derive the positions and velocities for the masses in the next time step. The algorithm is presented in symbolic flowchart form in Figure 2.4.

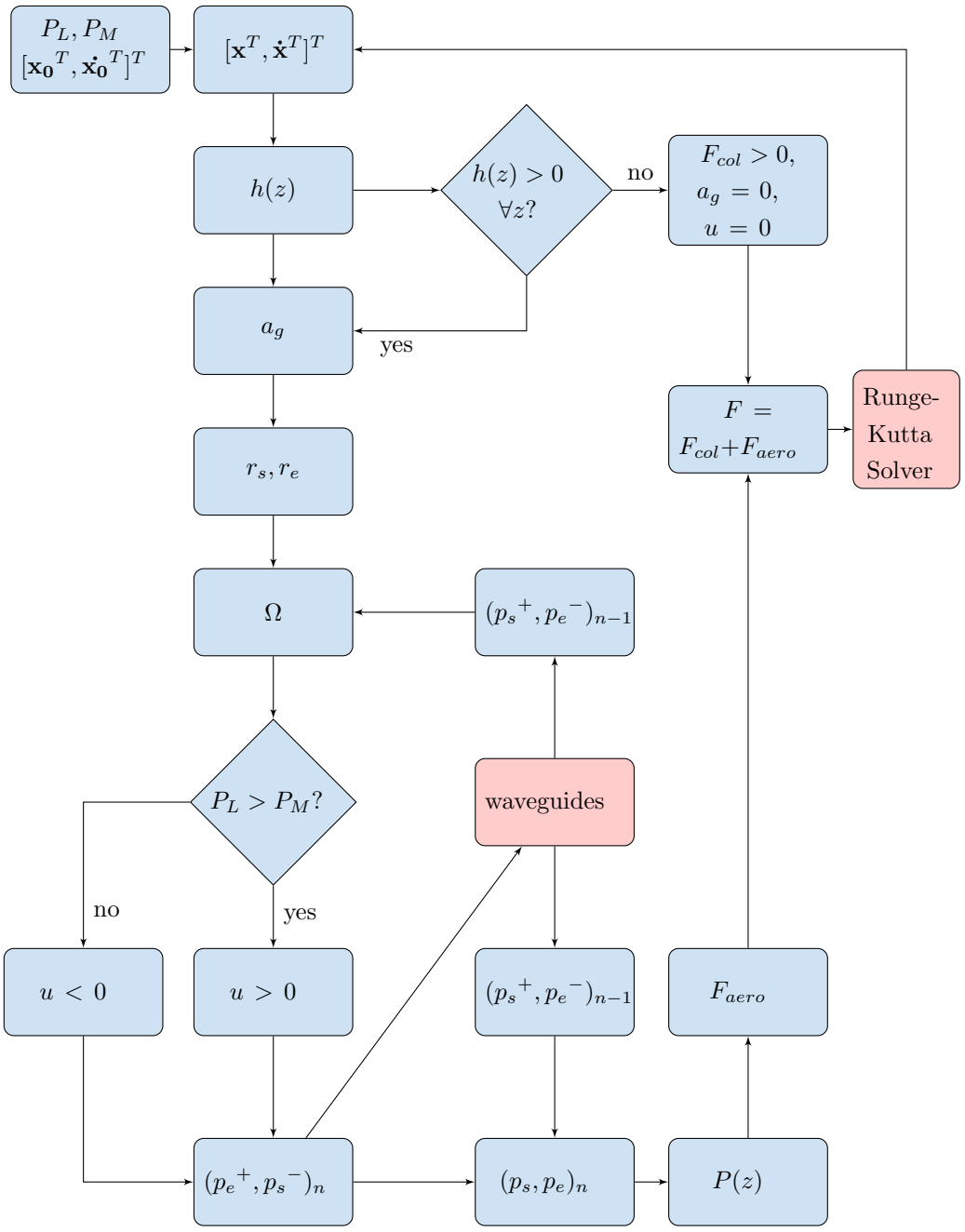


Figure 2.4: Algorithm flowchart. See section 2.2.9 for details.

The differential equation (2.65) together with the initial conditions (2.3) form an initial value problem. This problem is here solved using a 4th-order classical Runge-Kutta solution.

Equations (2.65) and (2.3) may be rewritten as follows:

$$\dot{\mathbf{w}} = f(\mathbf{w}), \quad \mathbf{w}(\mathbf{t_0}) = \mathbf{w_0} \tag{2.97}$$

An estimate for $\mathbf{w}$ at the next timestep $n+1$ can be calculated from the following system of equations where h is the timestep size [56]:

$$\begin{aligned}
\mathbf{w_{n+1}} &= \mathbf{w_n} + (1/6)(\mathbf{k_1} + 2\mathbf{k_2} + 2\mathbf{k_3} + \mathbf{k_4}) + O(h^5) \\
t_{n+1} &= t_n + h \\
&\text{where } \mathbf{k_1}\ \mathbf{k_2}\ \mathbf{k_3}\ \mathbf{k_4} \text{ are derived from:} \\
\mathbf{k_1} &= hf(t_n, \mathbf{w_n}) \\
\mathbf{k_2} &= hf(t_n + h/2, \mathbf{w_n} + \mathbf{k_1}/2) \\
\mathbf{k_3} &= hf(t_n + h/2, \mathbf{w_n} + \mathbf{k_2}/2) \\
\mathbf{k_4} &= hf(t_n + h, \mathbf{w_n} + \mathbf{k_3})
\end{aligned} \tag{2.98}$$

For the next timestep, the results of the calculation, $\mathbf{w_{n+1}}$ and t_{n+1}, serve as the new initial conditions to derive the subsequent state estimations $\mathbf{w_{n+2}}$ and t_{n+2}. This process is repeated iteratively until $t = t_{final}$ to estimate the time-course of the state.

A 4th-order classical Runge-Kutta solver provided by Mathworks [57] was used to solve the differential equation (2.65) subject to the initial conditions (2.3). This is a non-adaptive solver with a constant step-size, which is here equated to the sampling period of the waveguides, 0.023 ms based on a 44.1 kHz sampling rate. The function determines the derivative of the state vector and integrates the differential equation to arrive at the state for the next time step 0.023 ms later.

2.2.10 Tissue parameters

The parameters for the lumped-mass tissue were set as detailed by Tokuda et al. for their four-mass model [1] and summarized in Table 2.1. The prephonatory displacement of the cover masses was doubled as ingression is associated with a greater abduction of the folds [31]. The same displacement was also applied to egression for consistency. In contrast to [1], the initial values for

the displacements were all set to zero.

The uniform tube area for the supraglottal tract was set to 3.0 cm^2 [58, 1] while the subglottal tract areas were set according to [7].

Table 2.1: Tokuda et al. (2010) parameters. $L_g = 1.4$, $k_{12} = 1.0$ [N/m],$k_{23} = 0.5$ [N/m][1]

z [cm]	**h** [cm]	**m** [g]	**k** [N/m]	ζ
0	1.8	0	0	0
0.05	0.072	0.009	6.0	0.1
0.2	0.072	0.009	6.0	0.4
0.275	0.072	0.003	2.0	0.4
0.3	1.8	0	0	0
body	N.A.	0.05	30.0	0.4

2.2.11 Parameter space: range of lung pressures and pre-phonatory half-widths

Lausted et al. found that the maximum static pressures generated by the respiratory muscles range from -66 to 61 cmH_2O in women and -97 to 97 cmH_2O in men [59]. Comfortable expiratory phonation is usually performed at approximately 6 cmH_2O [60]. For conditions of loud voicing, the average values are 9.0 cmH_2O for men and 8.2 cmH_2O for women [61]. The following equation can be used to estimate a corresponding value for inspiratory phonation [61]:

$$\Delta P = RU \tag{2.99}$$

The resistance term R incorporates both viscous losses and dynamic resistance while ignoring the air's inertia [61]. By reversing the airflow through a cast model of the human larynx, van den Berg et al. found that the air resistance for ingression was 15 percent higher than in egression [32]. Meanwhile, Orlikoff found that airflow was 48.5 percent greater in ingression [34]. Taken together, the pressure drop for ingressive phonation for a corresponding egressive pressure drop can be estimated via

$$\Delta P_{IN} = R_{IN}U_{IN} = (1.15)R_{EX}(1.485)U_{EX} = (1.71)P_{EX} \tag{2.100}$$

Thus, the average to loud pressure drop values for exhalation of 6 to 9.0 cmH_2O would correspond roughly to -10.3 to -15.4 cmH_2O. Extending slightly beyond the latter value the minimum lung pressure considered in this study is -20 cmH_2O. To facilitate comparison, 20 cmH_2O is the maximum egressive lung pressure in this study. The -20 to 20 cmH_2O range should accomodate the range of subglottal pressures that occured in previous studies on inspiratory phonation where participants were instructed to produce comfortable phonation rather than loud voicing [31, 34].

The pre-phonatory half-height is the half-distance between the vocal folds in the absence of a pressure drop across them [22]. Above a certain threshold distance, the folds no longer oscillate. By fitting a two-mass model to experimental frequency and flow data for both men and women, Lucero et al. [62] found a maximum prephonatory distance of $x_0 = 0.14$ cm above which oscillation ceased. A value of 0 was used at the lower end as in their study. The case of pressed phonation is not considered in this study. Tokuda et al. used a nominal value of 0.018 cm for the prephonatory displacement [1].

Taken together, the lung pressures in this study range from -20 to 20 cmH_2O while the prephonatory half-heights range from 0 to 0.15 cm. These ranges include the default values typically used in lumped-parameter simulations of 8 cmH_2O and $x_0 = 0.02$ cm [63].

2.2.12 Post-simulation analysis

According to Tokuda et al. (2007) [25], the simulated extreme positions and velocities for each mass $\mathbf{v}_{sim}$ can be contrasted against the set of eigenmodes of the system matrix $\mathbf{v}_{eigen}$ when each are normalized by their maximum values. In their study, they invite the reader to do a visual comparison between the two. In this study, correlations values were calculated between the following two vectors based on the simulation results and the eigenmode analysis over all eigenmodes i:

$$\begin{aligned}\mathbf{v}_{sim} = \{&(x_1)_{max} - (x_1)_{min}, (\dot{x}_1)_{max} - (\dot{x}_1)_{min}, \cdots, \\ &(x_b)_{max} - (x_b)_{min}, (\dot{x}_b)_{max} - (\dot{x}_b)_{min}\}_{sim}\end{aligned} \tag{2.101}$$

$$\mathbf{v}_{eigen,i} = \{x_1, \dot{x}_1, x_2, \dot{x}_2, x_3, \dot{x}_3, x_b, \dot{x}_b\}_i \tag{2.102}$$

The eigenvalues based on aerodynamic state feedback were calculated for each lung pressure and the largest real value $Re(\lambda)$ over the sets of both oscillatory and zero-frequency (rigid-body) modes were plotted. The pitch in hertz was calculated as the inverse of the period between maximum values in the glottal aperture area. The phase delay was estimated as the difference in milliseconds between the peak values of the displacements of mass three to mass one. In other words, if the third mass trailed behind the first mass as generally observed for egressive phonation [2], the value was positive and represented a lag time. A negative value would indicate a lead time of mass three over that of mass one. The areas of the three cover masses and the volumetric flow rates were also plotted.

The pressure output at the lips $Pout$ was retained and used to generate a spectrogram along with a sound pressure measure as follows:

$$Pout_{rms} = \sqrt{(1/N) \sum_{n=n_{start}}^{n_{start}+N} (Pout_n)^2} \tag{2.103}$$

$$SPL_{[dB]} = 10log10(Pout_{rms}/20\mu Pa)^2 - 30\,dB - 20log_{10}(15cm/10cm) \tag{2.104}$$

The reduction by 30 dB was found empirically by Zanartu to convert between the sound pressure at the lips, as generated by simulations such as these, and a distance 10 cm from the lips where a microphone would be placed [64]. The further reduction via the last term accounts for the level $15cm$ from the lips for ease of comparison with [61]. In the foregoing, $N = 4096$ is the total number of data points retained and $n_{start} = 1$.

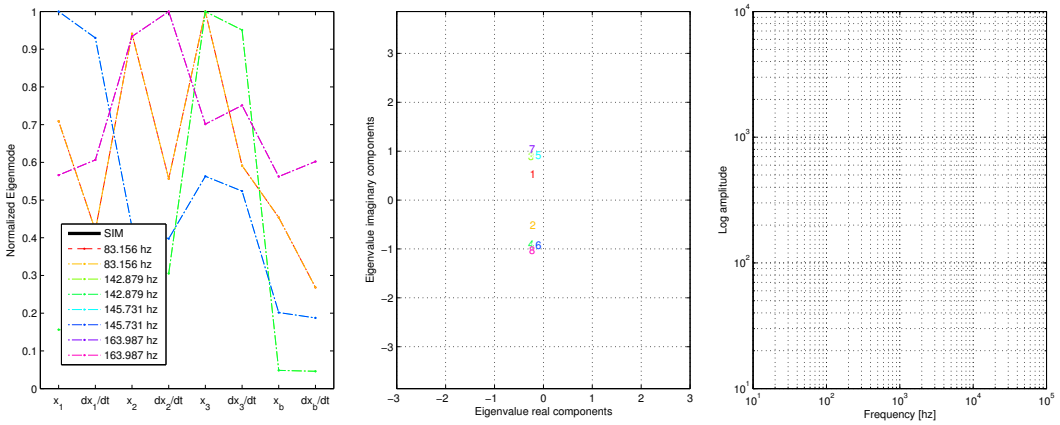
2.3 Results and Discussion

The central objective of this study was to determine whether ingressive phonation could be simulated via a lumped-parameter approach and if so, under what conditions. The choice of physiological conditions was limited to the prephonatory gap $2x_0$ and the lung pressure P_L. Several parameter scans were performed to investigate the roles of these two variables. The spring stiffness multiplier q was set to a value of 1 such that the tension of the folds was kept constant.

The system was first simulated without the waveguides affecting the pressures near the glottis. As such, the upstream pressure was equated to the lung pressure while the downstream pressure was set to atmospheric. These approximations are valid for fundamental frequencies well below the first formant [65]. The waveguides still affected the pressure signal leaving the glottis to generate the expected pressure at the lips, P_{out}, for which sound pressure levels were calculated.

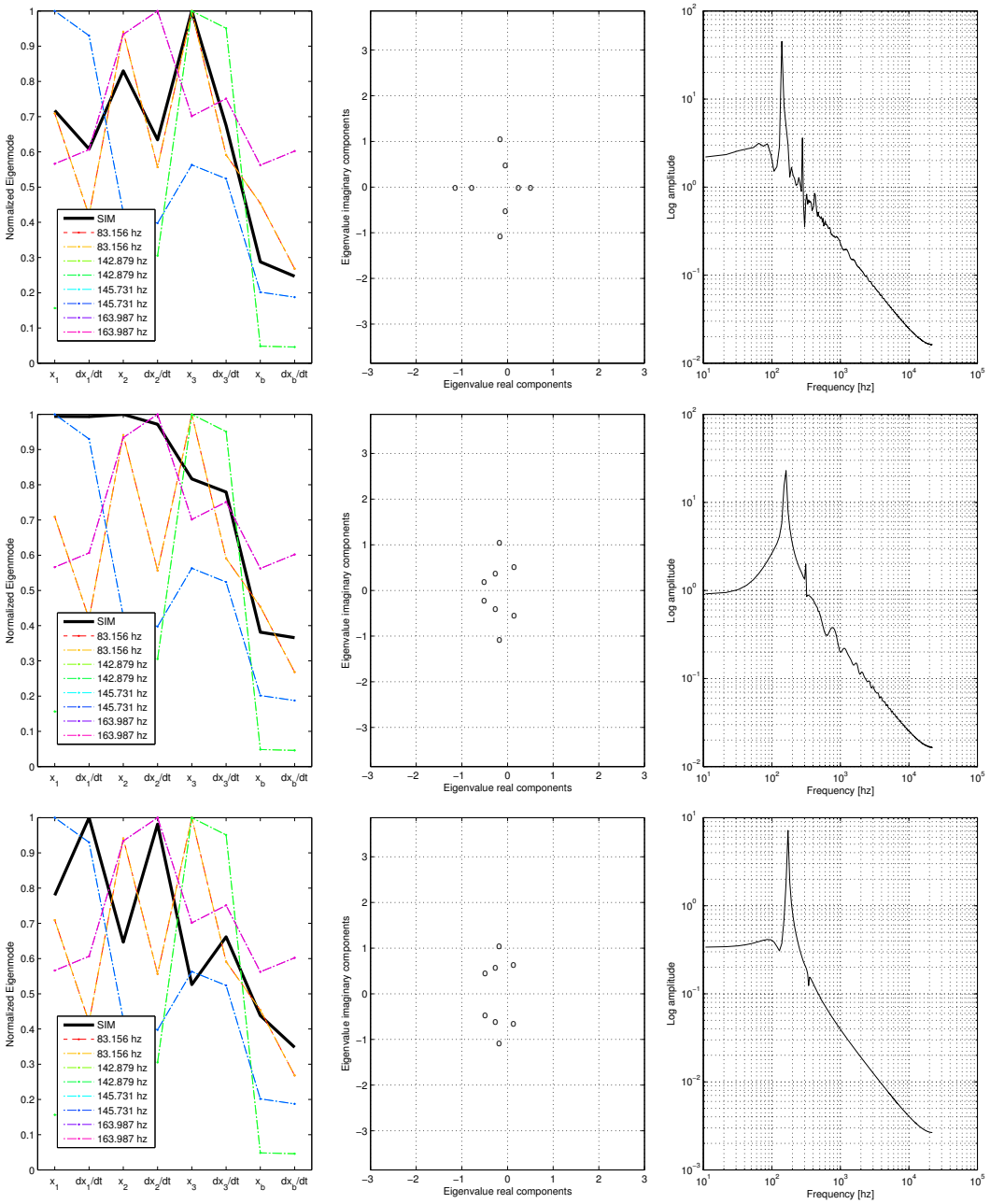
In vacuo analysis is depicted in Figure 2.5. From Figure 2.5(a) we can see the four natural frequencies of the associated eigenmodes, $\mathbf{v}_{eigen}$, of the 4-mass model as defined by equation (2.102) and $\mathbf{v}_{sim}$ as per equation (2.101). From this figure it can be seen that the lowest frequency corresponds to the most in-phase potential motion. The in-phase motion will dictate the overall lateral motion and thereby strongly influences the aperture of the glottis and the flow through it [66]. The highest frequencies involves more out-of-phase motions [12]. The out-of-phase motions allow for the mucosal wave and the possibility of an alternating converging and diverging channel profile [67, 50]. Of note, the software presents eight eigenmodes, but four distinct pairs of eigenmodes.

The sim line is the difference of the extrema of each degree of freedom; this result line will be compared with those of Figure 2.6. Figure 2.5(b) shows that the system is stable since none of the eigenvalues have crossed to the positive side of the real components axis. The distance of these eigenvalues from the real line determines the frequency according to equation (2.67). Figure 2.5(c) is the power spectrum of the analysis. Given the absence of lung pressure, no vibration occurs and there is no signal energy.



(a) (b) (c)

Figure 2.5: In vacuo eigenanalysis of a 4 mass system with $P_L = 0Pa$, $P_M = 0$ Pa, $x_0 = 0.036cm$. (a) Natural frequency eigenmodes of a four mass model ($\mathbf{v}_{eigen}$) in comparison to the simulation extrema ($\mathbf{v}_{sim}$); (b) Eigenvalues of each eigenmode on imaginary versus real component axes, colors correspond to eigenmodes of (a); (c) Power spectrum analysis graph of the 4 mass model.



(a) (b) (c)

Figure 2.6: Aero feedback eigenanalysis. Eigenmodes still shown *in vacuo*. $P_L = -100$ top, -120 middle and $-140Pa$ bottom, $P_M = 0$ Pa, $x_0 = 0.036cm$. (a) Natural frequency eigenmodes of a four mass model ($\mathbf{v}_{eigen}$) in comparison to the simulation extrema ($\mathbf{v}_{sim}$); (b) Eigenvalues of each eigenmode on imaginary versus real component axes; (c) Power spectrum analysis graph of the 4 mass model.

In Figure 2.6 are presented the eigenanalysis for three different test cases, where $P_L = -100$ top, -120 middle and $-140Pa$ bottom. In decreasing the pressure, the general trend from the $\mathbf{v}_{sim}$ curve of the (a) subfigures, from an in-phase to an out-of-phase motion. It is known that out-of-phase motions generally correspond to higher eigenfrequencies [12] and a corresponding rightward shift is seen in the power spectrum of subfigures (c). In subfigures (b), in comparison with *in vacuo* case, the pair of eigenvalues are moving towards the real half plane. In the case of the top figure, subfigure (b) the degeneracy is lost for point 5 and 6 when the two meet on the real axis. The subsequent movement of one of these towards the positive real half plane is called a static divergence [8] and is considered a rigid-body mode because it has a zero-frequency component [11]. Although one eigenvalue has adopted a very large positive growth rate as seen in later Figures 2.10, 2.12 and 2.14, it can be seen that several other eigenvalues have also shifted towards the real half plane. For the middle and bottom subfigures (b) it can be seen that there are no eigenvalues on the real axis and that there are no rigid body modes. However there are oscillatory eigenvalues in the positive real plane, suggesting that the system is unstable under aerodynamic feedback.

In comparison with Figure 2.5, the *in vacuo* case study, we see that dropping the pressure leads to an increasingly out of phase motion and therefore an increasingly higher frequency. Also, the aerodynamic feedback destabilizes the *in vacuo* eigenvalue frequencies. Also important to note was the observation of static divergence not seen in the *in vacuo* cases. It was observed that upon continued dropping of the pressure, the eigenvalues return to the negative real plane.

The following series of figure pairs beginning with Figure 2.7 show the system time history as a result of raising or lowering the lung pressure in equal increments or decrements. For each pressure level, 4096 data points were simulated at a sampling rate of $44.1\,kHz$. The second figure in each pair, such as Figure 2.8, shows certain measures corresponding to each lung pressure level. In the first figure of each pair, subgraph a) is a spectrogram providing the short-time Fourier transform of the glottal area signal a_g. The signal was generated at $44.1\,kHz$ and subsequently downsampled by a factor of

4. The spectrogram used Hamming windows of length 1024 overlapping by 800 data points. Subgraph b) shows the volumetric flow rate $[cm^3/s]$ of air passing through the glottis while subgraph c) shows the vocal fold aperture areas $[cm^2]$ at each of the masses along with the glottal area a_g in black. In the second figure of each pair, subgraph a) presents the sound pressure level in decibels measured $15\,cm$ from the lips based on the P_{out} signal and equation (2.104). The next subgraph b) provides the pitch in hertz based on a_g signal. The period was determined by the average spacing in seconds between the maxima of a_g. The pitch was then calculated from the inverse of the period. Subgraph (c) presents the mucosal wave phase delay in milliseconds. For a given lung pressure increment, the nearest peak in the x_3 signal to each peak in the x_1 signal was first found. For each such pairing, the time instant of the x_1 peak was subtracted from that of the x_3 peak to generate the time delay. The time delays were then averaged across the pairings to arrive at the mucosal wave phase delay. This approach allowed for the possibility of both positive and negative delays depending on whether the x_1 peaks preceded or lagged behind the x_3 peaks in each pairing. Subgraph d) presents the largest real eigenvalue components in two sets of modes, namely the oscillatory modes and the rigid-body modes with zero frequency. The latter modes are not readily apparent in this figure yet show up prominently in subsequent figures referring to ingressive phonation. Positive values of the real eigenvalue components imply positive growth rates of the corresponding eigenmodes and system instability. Finally, subgraph e) shows the correlation of simulation extrema ($\mathbf{v}_{sim}$) to *in vacuo* eigenmodes ($\mathbf{v}_{eigen}$). Values may range between -1 and 1 with higher values indicating a greater resemblance of the simulation extrema to a particular eigenmode. The correlation was based on the off-diagonal values of Matlab's 'corrcoef' function.

Figure 2.7 shows the system time history as a result of raising the lung pressure in increments of 100 Pa. At the outset, the lung pressure is nil and there is no vocal fold motion. As seen in subgraph d) of Figure 2.8, the real parts of all of the *in vacuo* eigenvalues are less than zero and the system is exponentially stable. With a rise in the lung pressure, at least one eigenvalue adopts a positive growth rate and phonation onset occurs. This remains the case

throughout the rest of the simulation. Consequently, the phonation region in the spectrogram in subgraph a) of Figure 2.7 extends to the highest pressure at 2000 Pa. It can be seen in c) of Figure 2.8 that the phase delay is positive throughout the range of phonation.

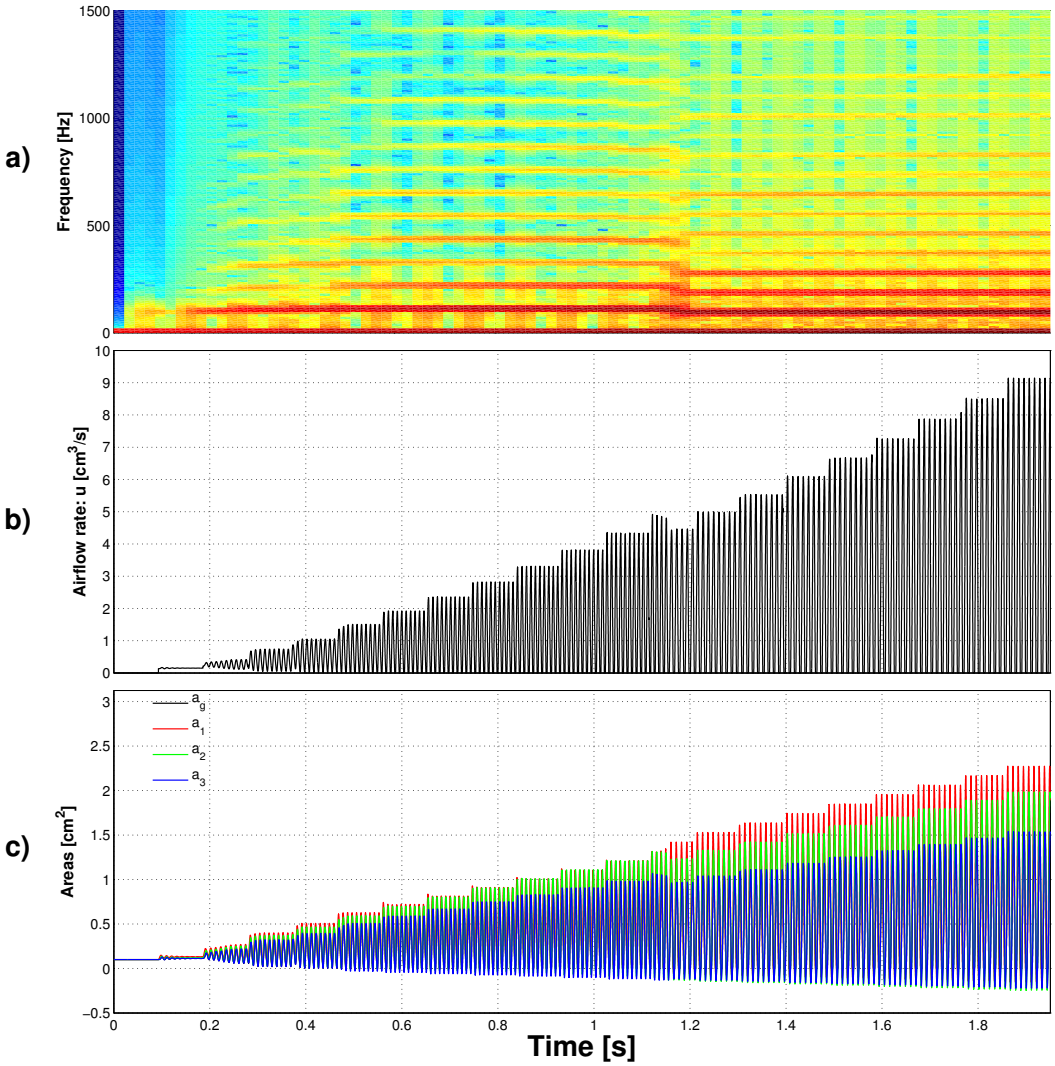


Figure 2.7: Time history of key variables for egression for four-mass model, source not coupled with waveguides. $0Pa < P_L < 2000Pa$, $P_M = 0Pa$, $x_0 = 0.036cm$. a) Spectrogram ; b) volumetric flow rate $[cm^3/s]$; c) Vocal fold aperture areas $[cm^2]$

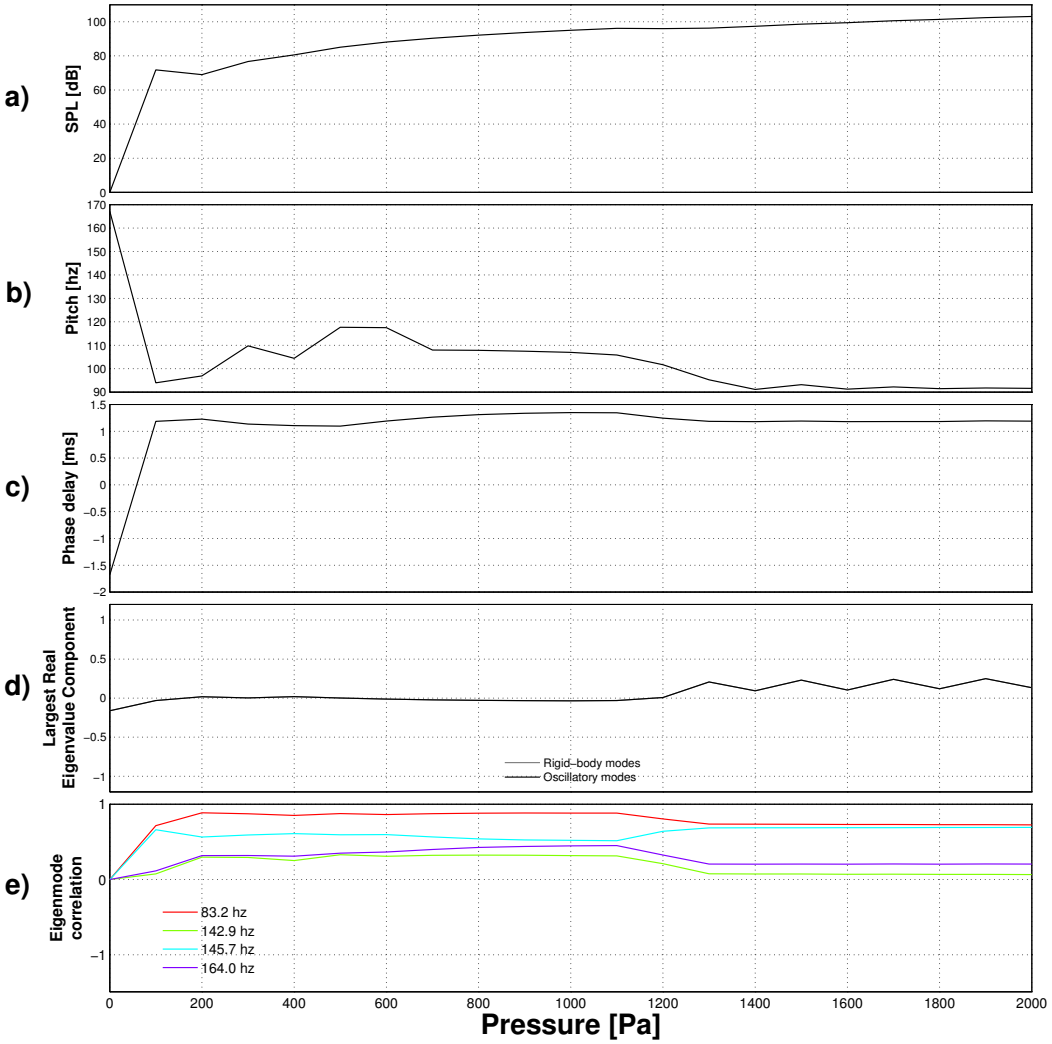


Figure 2.8: Key measures for egression for four-mass model, source not coupled with waveguides. $0Pa < P_L < 2000Pa$, $P_M = 0Pa$, $x_0 = 0.036cm$. a) Sound pressure level $[dB]$; b) Pitch $[hz]$; c) Mucosal wave phase delay $[ms]$; d) Largest real eigenvalue component ; e) Correlation of simulation extrema ($\mathbf{v}_{sim}$) to *in vacuo* eigenmodes ($\mathbf{v}_{eigen}$)

Figure 2.9 shows the system time history as a result of lowering the lung pressure in decrements of 100 Pa leading to ingression. As with egression, the *in vacuo* case is exponentially stable according to subgraph d) of Figure 2.10. According to subgraph c) in Figure 2.9, the folds move from a rectangular shape, with all of the areas overlapping, to a convergent one with $a_1 < a_3$. The masses undergo a rotation such that the first mass is in a nearly closed

state while the other two masses move apart from their counterparts in the other fold. A convergent state with respect to the flow direction results. This is reflected in the eigenanalysis which predicts a rigid-body (zero-frequency) mode of rotation in d) of Figure 2.10. At the same time, the real part of at least one oscillatory mode increases, eventually becoming positive. The vibration of the folds is more sustained according to b) of Figure 2.9 and the sound pressure level at the lips becomes pronounced in a) of Figure 2.10. As the pressure is further dropped, the largest real eigenvalue part steadily decreases in d) of Figure 2.10 and the system is increasingly more stable. Consequently, the oscillations are damped out in b) of Figure 2.9.

The system returns to stability as the pressure is lowered since the folds are increasingly drawn together at the first mass, as seen in c) of Figure 2.9, arresting the flow and therefore the aerodynamic feedback. With the reduction in the aerodynamic feedback, the eigenvalues of the overall mechanical-aerodynamic system seen in d) of Figure 2.10 return to their *in vacuo* values and corresponding exponential stability.

It can also be seen in b) of Figure 2.10 that the frequency of the glottal area generally rises. Part of this rise may be accounted for by the dominance of a higher frequency out-of-phase eigenmode over the lower frequency in-phase one in subgraph e). The pitch exceeds that of all of the *in vacuo* modes in Figure 2.5 however suggesting that the aerodynamic feedback effectively raises the stiffness of the system.

In the region of phonation, the mucosal wave in c) of Figure 2.10 is observed to travel in the reverse direction with the lower mass trailing the upper mass. As such, the phase delay is negative.

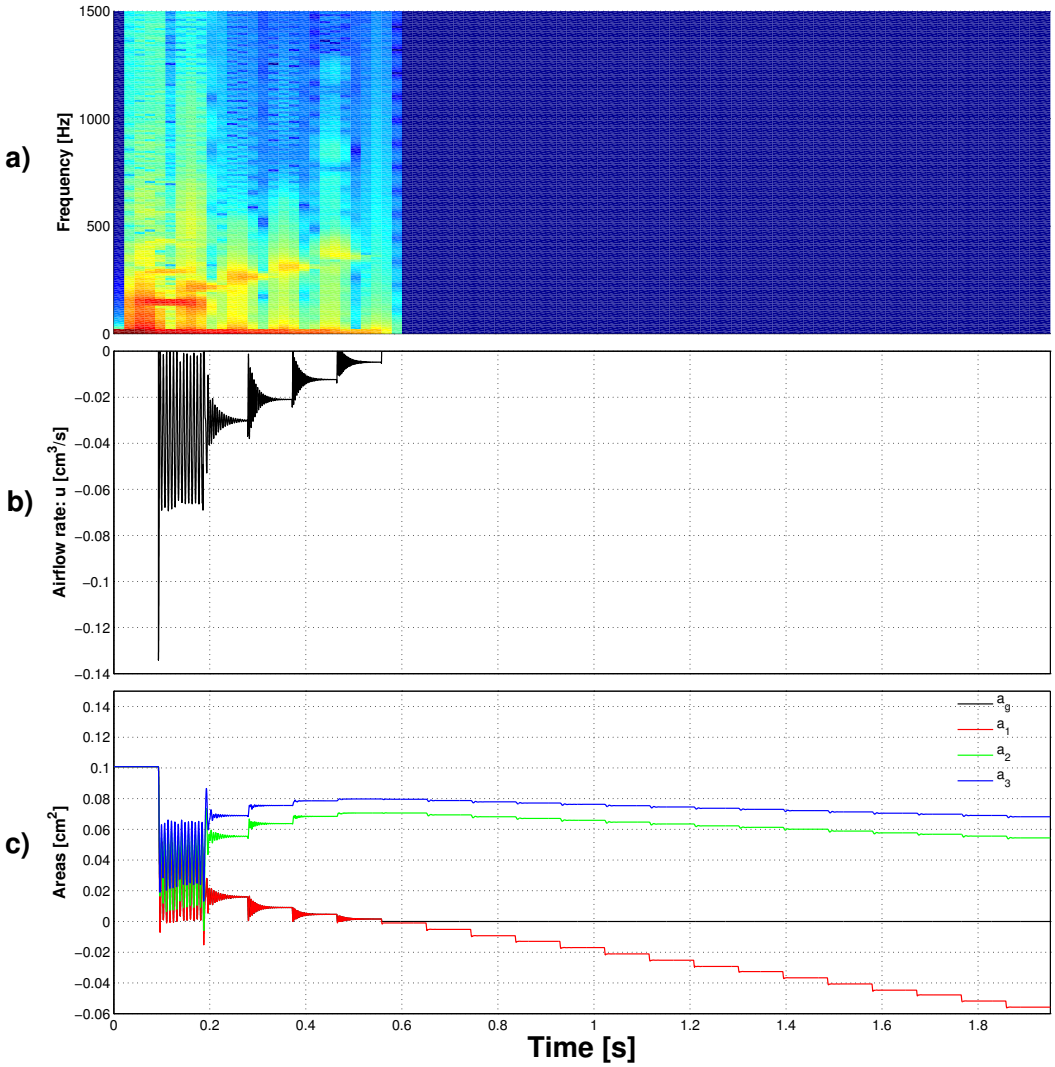


Figure 2.9: Time history of key variables for ingression for four-mass model, source not coupled with waveguides. $0Pa > P_L > -2000Pa$, $P_M = 0Pa$, $x_0 = 0.036cm$. a) Spectrogram ; b) volumetric flow rate $[cm^3/s]$; c) Vocal fold aperture areas $[cm^2]$

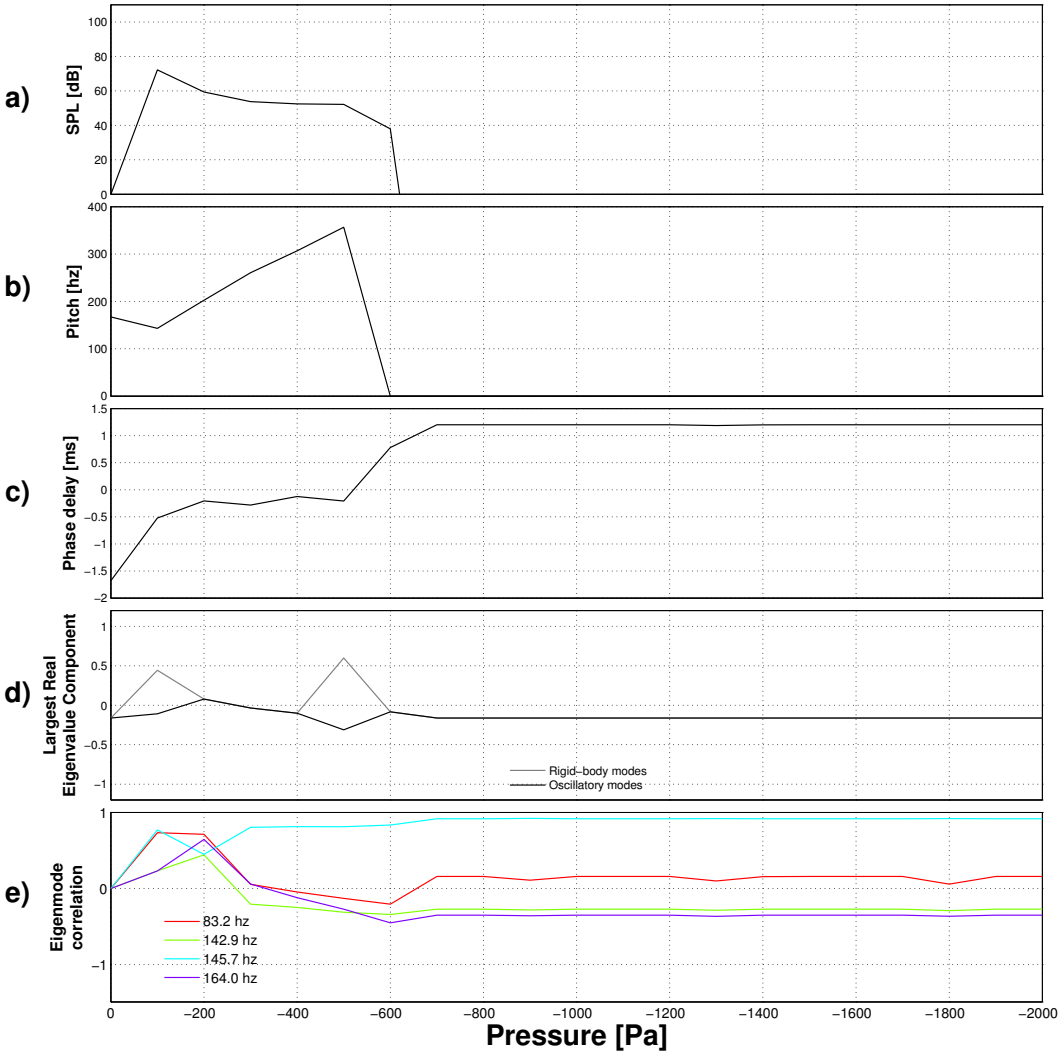


Figure 2.10: Key measures for ingression for four-mass model, source not coupled with waveguides. $0Pa > P_L > -2000Pa$, $P_M = 0Pa$, $x_0 = 0.036cm$. a) Sound pressure level $[dB]$; b) Pitch $[hz]$; c) Mucosal wave phase delay $[ms]$; d) Largest real eigenvalue component ; e) Correlation of simulation extrema ($\mathbf{v}_{sim}$) to *in vacuo* eigenmodes ($\mathbf{v}_{eigen}$)

In the next six figures, 2.11 through 2.16, the above analysis is repeated but with decrements of 40 Pa to focus on the smaller range of phonation found in ingression.

With the first decrement, the folds are drawn inwards in c) of Figure 2.11 and oscillate about a new equilibrium point but the vibration is damped out b); the eigenvalues are still within the left-half of the complex plane in subgraph

d) of Figure 2.12 and the system remains stable. Subgraph c) of Figure 2.11 shows that towards the second decrement the cover masses undergo a rotation as before in Figure 2.9. A convergent state with respect to the flow direction results. This rotation is reflected in the eigenanalysis which predicts a rigid-body (zero-frequency) mode of rotation in subgraph d) of Figure 2.12. At the same time, the real part of at least one oscillatory mode increases, eventually becoming positive. The vibration of the folds is more sustained in b) of Figure 2.11 and the sound pressure level at the lips becomes pronounced in a) of Figure 2.12. As the pressure is further dropped, the largest real eigenvalue part steadily decreases d) of Figure 2.12 and the system is increasingly more stable. The oscillations are therefore damped out as seen in b) of Figure 2.11. Within the range of phonation seen in the spectrogram a) of Figure 2.11, a general rise in pitch and reversed mucosal wave are again observed in subgraphs b) and c) of Figure 2.12 respectively.

Figure 2.13 shows the effect of the waveguide coupling and the strengthening of the secondary phonatory burst towards full glottal closure by the waveguides. The oral tract resonator is known to have a first resonance around 500 Hz [68] and the vocal fold oscillation appears to entrain to and be boosted by the tract's resonance. For the real-component eigenvalue plot in subgraph c) of figuref 2.12 and 2.14, the ordinate axis is limited such that subtle changes about zero can be clearly seen. This leads to certain parts of the rigid-body curves being cut-off.

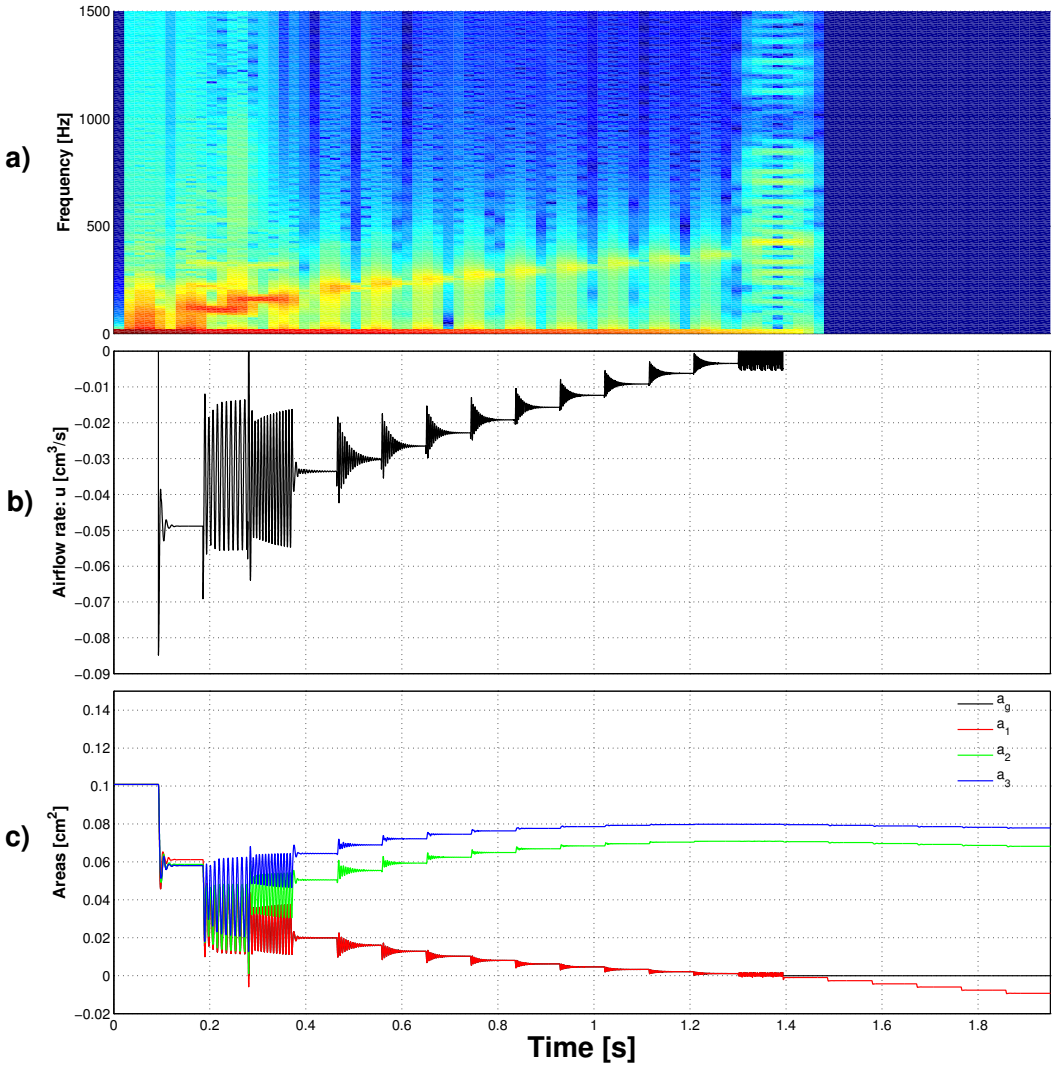


Figure 2.11: Time history of key variables for ingression for four-mass model, source not coupled with waveguides. $0Pa > P_L > -800Pa$, $P_M = 0Pa$, $x_0 = 0.036cm$. a) Spectrogram ; b) volumetric flow rate $[cm^3/s]$; c) Vocal fold aperture areas $[cm^2]$

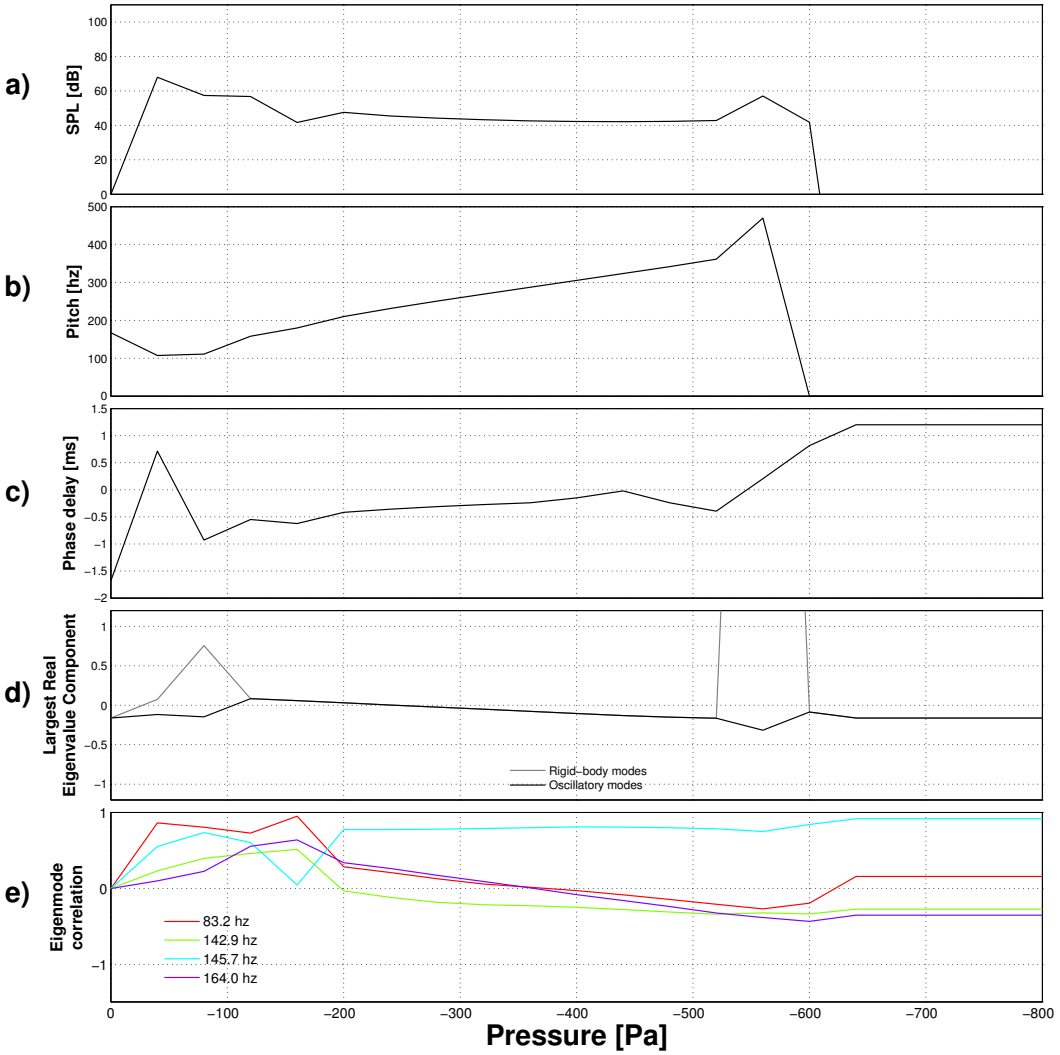


Figure 2.12: Key measures for ingression for four-mass model, source not coupled with waveguides. $0Pa > P_L > -800Pa$, $P_M = 0Pa$, $x_0 = 0.036cm$. a) Sound pressure level $[dB]$; b) Pitch $[hz]$; c) Mucosal wave phase delay $[ms]$; d) Largest real eigenvalue component ; e) Correlation of simulation extrema ($\mathbf{v}_{sim}$) to *in vacuo* eigenmodes ($\mathbf{v}_{eigen}$)

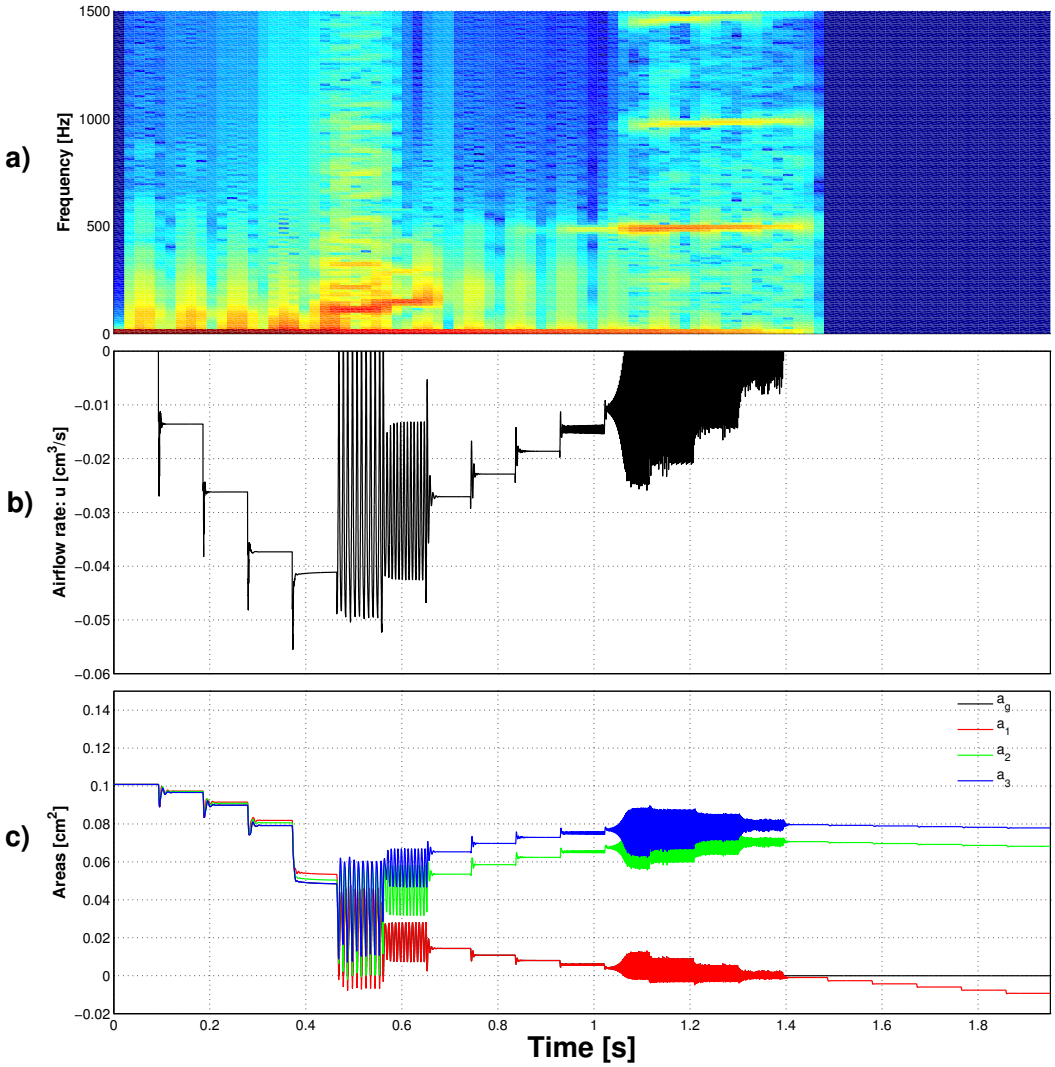


Figure 2.13: Time history of key variables for ingression for four-mass model, source coupled with waveguides. $0Pa > P_L > -800Pa$, $P_M = 0Pa$, $x_0 = 0.036cm$. a) Spectrogram ; b) volumetric flow rate $[cm^3/s]$; c) Vocal fold aperture areas $[cm^2]$

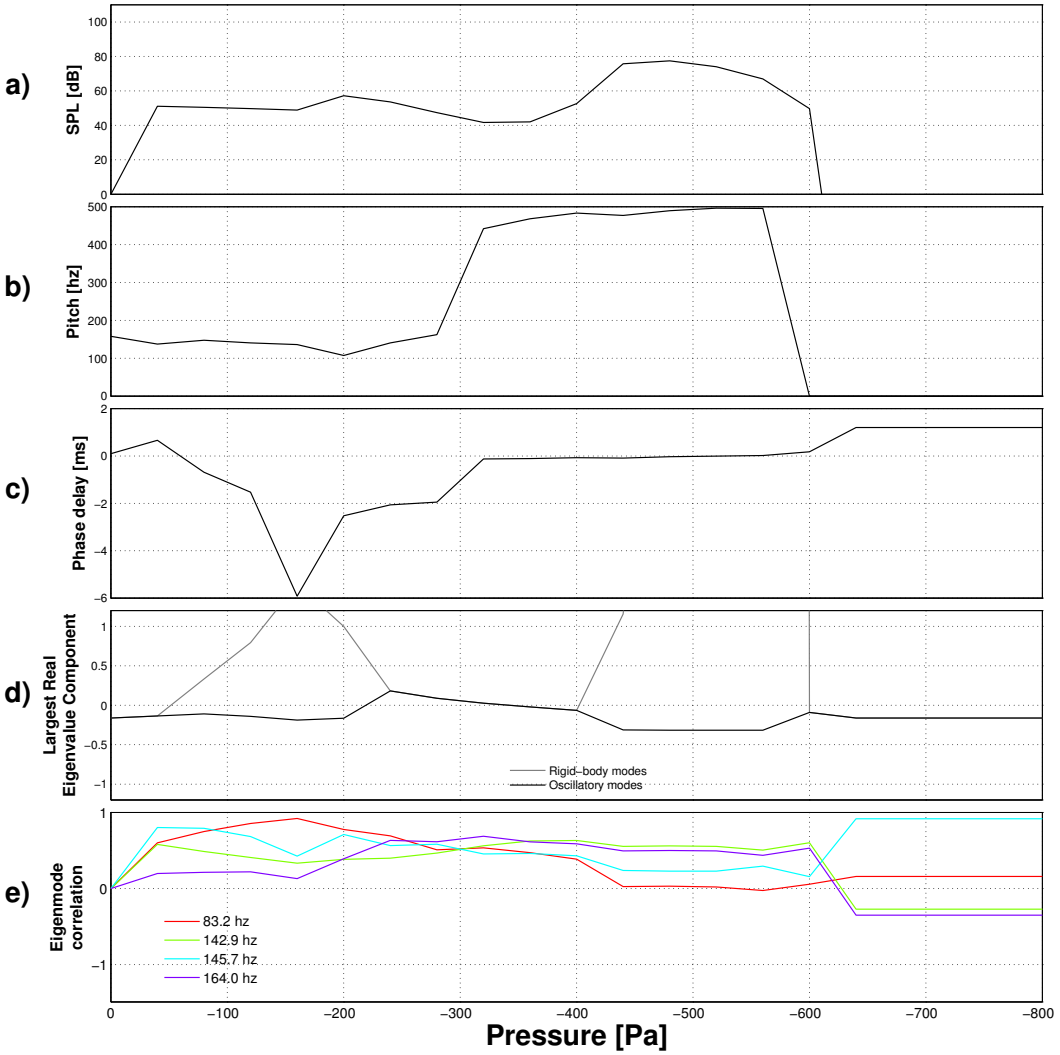


Figure 2.14: Key measures for ingression for four-mass model, source coupled with waveguides. $0Pa > P_L > -800Pa$, $P_M = 0Pa$, $x_0 = 0.036cm$. a) Sound pressure level $[dB]$; b) Pitch $[hz]$; c) Mucosal wave phase delay $[ms]$; d) Largest real eigenvalue component ; e) Correlation of simulation extrema ($\mathbf{v}_{sim}$) to *in vacuo* eigenmodes ($\mathbf{v}_{eigen}$)

Figure 2.15 shows the system time history as a result of lowering the lung pressure in decrements of 40 Pa but with the damping parameter set to 1.0 for all of the masses, the critical damping condition. This removes the oscillations seen in Figure 2.11 and provides a clearer picture as to the effect of lowering the pressure in a step-like manner. The lowest pressure is -800 Pa since the phonation region at this level of vocal fold adduction is limited as found in

figure 2.9. It can be seen that the folds start out in a rectangular configuration with the upstream and downstream cross-sectional areas roughly equal in subgraph c) of Figure 2.15. A shift occurs however, and the folds adopt a convergent profile with respect to flow direction. None of the oscillatory eigenvalues cross the zero-axis in subgraph d) of Figure 2.16, as expected for critical damping. Rigid-body modes persist however.

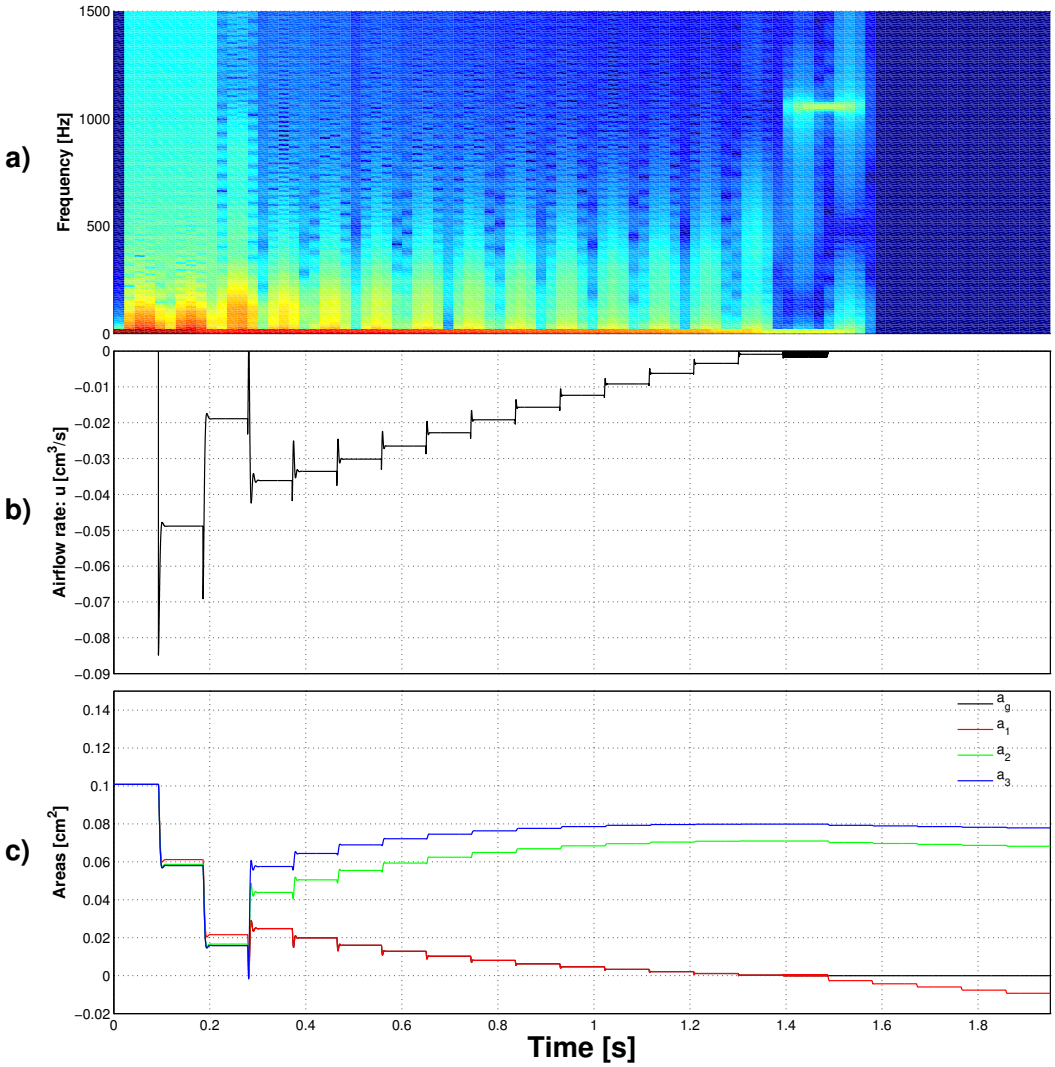


Figure 2.15: Time history of key variables for ingression for four-mass model, source not coupled with waveguides, tissue critically damped. $0Pa > P_L > -800Pa$, $P_M = 0Pa$, $x_0 = 0.036cm$. a) Spectrogram ; b) volumetric flow rate $[cm^3/s]$; c) Vocal fold aperture areas $[cm^2]$

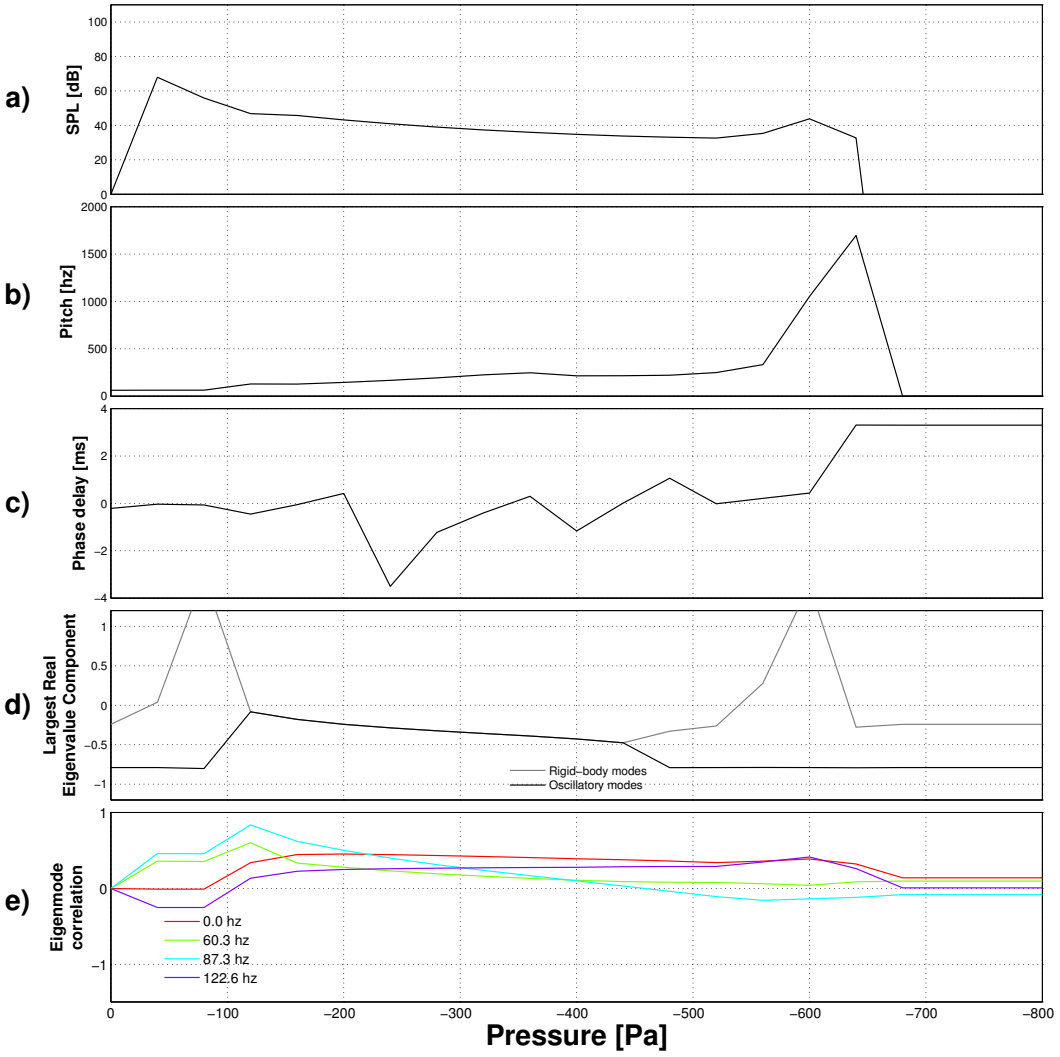


Figure 2.16: Key measures for ingression for four-mass model, source not coupled with waveguides, tissue critically damped. $0Pa > P_L > -800Pa$, $P_M = 0Pa$, $x_0 = 0.036cm$. a) Sound pressure level $[dB]$; b) Pitch $[hz]$; c) Mucosal wave phase delay $[ms]$; d) Largest real eigenvalue component ; e) Correlation of simulation extrema ($\mathbf{v}_{sim}$) to *in vacuo* eigenmodes ($\mathbf{v}_{eigen}$)

The aim of this study was to determine whether ingressive phonation could be simulated via a lumped-parameter approach and if so, under what conditions. As there is found to be a range of pressures for which there is sustained oscillation in the foregoing ingression simulations, it can be said that phonation can indeed be modeled in this way. The stability analysis predicts the range of parameters where the oscillations of the folds are not completely damped

out. The correspondence between the ranges of phonation and the parts of the $Re(\lambda)$ curve where at least one eigenvalue adopts a positive growth rate suggests that the aerodynamic state feedback model is applicable.

The model was able to simulate both egression and ingression with only a single parameter change, namely the lung pressure. While the aerodynamic equations were reflected upon flow reversal, the dynamics of ingressive phonation appear to be quite distinct from their egressive counterpart. The pitch is consistently higher as was found by [34]. In trying to explain the higher pitch, experimentalists have argued that the vocal fold tension must be greater [31, 34], however the tension has not been changed at all in the present simulations. There was also found to be a reduced range of phonation in ingression. Since the energy transfer to the folds is from the fluid flow, it is reasonable to expect that the system would not be able to gain energy as efficiently as the flow is arrested by a narrowing glottis. The rapid sustained vocal fold oscillations just prior to closure have not yet been fully explained. It is premature to assert that they are an artifact of the simulation. Figure 2.17 provides a potential explanation for a decreased range of phonation for ingression relative to that of egression. In egression (top of figure), two eigenvalues synchronize with one adopting an increasingly positive growth rate and thus de-stabilizing the system towards phonation [18, 23]. By contrast, in ingression the folds move towards a lack of stability via a static divergence along the real axis. The static divergence causes a rotation that shifts the folds from a rectangular equilibrium position to a convergent one. In this new convergent state, at least one eigenvalue then adopts a non-zero frequency and oscillation results. However, upon further lowering the pressure and drawing the downstream masses closer together, the flow is arrested and the aerodynamic feedback is suppressed causing a return to stability and loss of phonation.

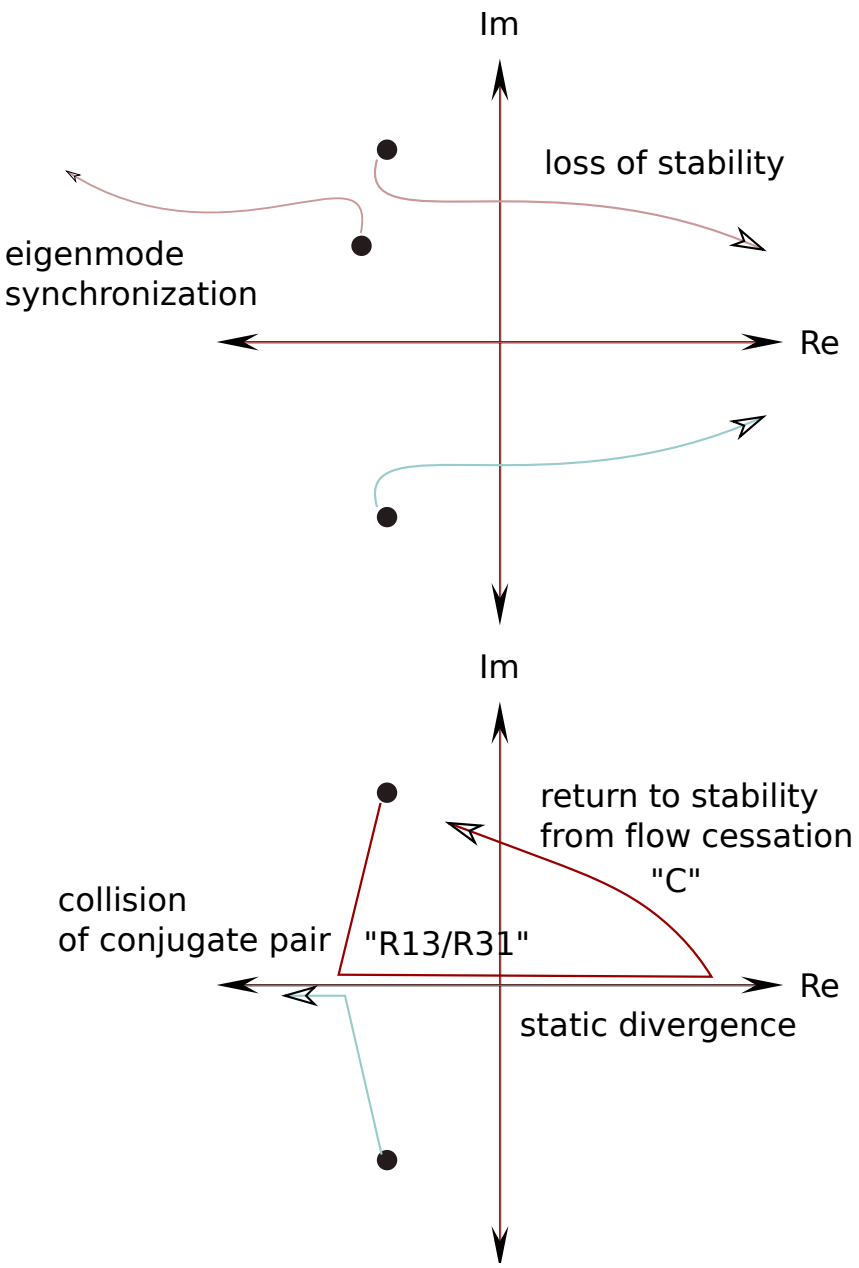


Figure 2.17: Observed eigenvalue trajectories for egressive (top) and ingressive (bottom) phonation with equilibrium matrices indicated

It is important to note that the mucosal wave direction was not imposed on the simulation, but emerged from it as suggested by [23]. It was consistently found that for ingressive phonation, the phase delay was negative in the first (leftmost) oscillation region. This is in keeping with observations

by experimentalists [33, 31]. This requires an update of Titze's assertion that "Surface-wave propagation in the mucosa (cover) must be from bottom to top. Reversal of this direction would not produce oscillation." [22]. It can perhaps be said that, where the mucosal wave is the source of energy transfer, the upstream mass must lead the downstream mass whether the flow is from or towards the lungs. The secondary region of oscillation in ingression overturns this assertion however since the phase delay is again positive in this region, suggesting an energy transfer mechanisms that requires further elucidation.

2.4 Conclusion

It was found that the equations, geometry and tissue parameters used by previous authors for simulating egressive phonation could be extended to simulate ingressive phonation without major theoretical or parameter adjustments beyond symmetrical reasoning. The range of P_L values for which phonation was observed was significantly reduced for ingression relative to that for egression and were found to be $-100\,Pa < P_L < -200\,Pa$ and $P_L > 200\,Pa$ respectively for a prephonatory half-width of $0.036\,cm$. This could be accounted for by the suction force causing the folds to remain closed increasingly as P_L was reduced and arresting flow-induced oscillations. The frequency in ingression was typically higher than in egression as was found by Orlikoff [34]. However, while the latter author speculated that this had to do with greater tension in the folds, the present results and eigenmode vibrational analyses suggest that increased frequency may be a general result of ingressive phonation, as tension was not varied in this study. Orlikoff also observed increased flow in ingression [34] in contrast to present findings. The difference can be explained by the possibility of inspiratory muscle posturing leading to a greater prephonatory half-width and therefore a larger glottal area and increased flow. There may also be a contribution to the area and thus the flow rate from the large posterior glottic opening seen in ingression [31]. The latter's contribution to the flow was not modeled here.

This study highlights the need for more experimental data comparing ingression and egression in terms of sound pressure levels, flow rates, pitch, lung

pressure and resting position of the folds. As well, other variables may be considered in the study of ingression including the glottal angle and the tension parameter.

This study shows that current techniques can in fact be utilized to simulate ingressive phonation and thereby provides a unified theoretical framework for both kinds of phonation.

Chapter 3

Summary and outlook

3.1 Summary

The purpose of this study was to establish a general theoretical framework encompassing both ingressive and egressive phonation. The specific objective was to determine whether ingressive phonation could be simulated via lumped-parameter approach and if so under what conditions. This was accomplished by considering a recently developed four-mass lumped-parameter model [1] that has been used to simulate egression and by further exploring its behavior upon reversal of lung pressure and flow direction. The equations for the flow were established through symmetrical reasoning. The system was then simulated with the lung pressure either raised or dropped in a step-like manner and the results in terms of the mucosal phase delay, pitch, sound-pressure level, spectrogram, areas, and volumetric flow rates were recorded. It was found that phonation did result in ingression but for a narrower region of the parameter space then was the norm for egression.

A new aerodynamic feedback analysis was developed to explain the potential destabilizing effect of the flow upon the folds. The aerodynamic forces were cast as a function of the state of the system and the input lung pressure. It was found that the parameter regions of phonation for both egression and ingression corresponded to regions where at least one eigenvalue adopted a positive growth rate and non-zero frequency. For egression, the system moved towards instability without any of the eigenvalues touching the real axis. For ingression, eigenvalues met on the real axis with at least one moving into

the positive real half-plane. This static divergence caused a rotation of the folds' surfaces such that the equilibrium position changed from rectangular to convergent. With a further decrease in pressure, the folds were increasingly drawn together throttling the flow and suppressing aerodynamic feedback and thereby phonation.

The study was limited in several ways. A main limitation was that the folds were assumed to have the same shape and properties regardless of flow direction and pressure polarity. It is likely that the folds would not only be drawn inward by negative pressures but also downwards towards the lungs. Only the first effect was modeled in this study. As well, the distribution of the mass density and stiffness properties of the folds may change along with the nodal points of vibration [69]. This could be modeled by changing the values of the mass and vertical position assigned to each oscillator along the cover. A detailed finite-element model would give a clearer picture as to the potential changes in oscillatory behavior from a negative pressure. A more detailed aerodynamic simulation of the fluid flow via the Navier-Stokes equations would also be appropriate with consideration of the flow's time dependency and viscosity at small aperture sizes.

The main contributions of this work are threefold: a unified framework for both ingressive and egressive phonation highlighting distinctions between the two, a generalized treatment of vocal folds to allow for an arbitrary number of cover masses, and an aerodynamic feedback analysis capable of predicting phonation parameter ranges. Since ingressive phonation essentially inverts the vocal folds' motion, reversing the directions of horizontal and vertical wave propagation [31], future experimental studies including EGG analysis and high-speed cameras coupled with the theoretical analysis and results of this study, could provide greater insight into vocal fold dynamics not only in ingression but in egression as well [33, 34]. As Orlikoff states: "renewed examination of phonation produced using ingressive airflow may yield important insight regarding mucosal and laryngeal airway dynamics [34]." Such insight could inform both vocal pedagogy and clinical treatments for vocal fold pathologies.

3.2 Future work

Future work based on this study could consider the following:

- the effect of changing the prephonatory half-widths as a result of muscle activation [69]
- the incorporation of the false vocal folds
- the ability of the masses to oscillate along more than one axis including lateral, vertical and horizontal motions
- the nonlinear tissue stiffness
- the role of the cover coupling springs in determining the mucosal wave and consequent energy transfer
- a more detailed subglottal tract and a rounded non-uniform supraglottal tract with inclusion of the nasal tract
- a prephonatory profile of the folds with either a divergent or convergent shape
- a more detailed prediction of the flow separation vertical location along the folds

In particular, future work towards understanding the rapid sustained oscillations at very small aperture sizes would be justified.

3.3 Closing remarks

The modeling and analysis presented in this work drew together theory and techniques from several disciplines including vocal science, fluid dynamics, control theory, mechanics, vibration and signal processing. It is remarkable that such a commonplace human activity as phonation, typically explored from an arts or medical perspective, should involve so many branches of engineering. This highlights the ongoing need for greater dialogue between the arts, medicine and science and the availability of generalist engineering education.

Bibliography

[1] I. Tokuda, M. Zemke, M. Kob, and H. Herzel, "Biomechanical modeling of register transitions and the role of vocal tract resonators", J. Acoust. Soc. Am. **127**, 1528–1536 (2010).

[2] N. Lous, G. Hofmans, R. Veldhuis, and A. Hirschberg, "A symmetrical two mass vocal fold model coupled to vocal tract and trachea, with application to prothesis design", Acta Acust. **84**, 1135–1150 (1998).

[3] B. Story and I. Titze, "Voice simulation with a body-cover model of the vocal folds", J. Acoust. Soc. Am. **97**, 1249–1260 (1995).

[4] J. L. Flanagan and L. L. Landgraf, "Self-oscillating source for vocal-tract synthesizers", IEEE Transactions on audio and electroacoustics **AU-16**, 57–64 (1968).

[5] R. Eklund, "Pulmonic ingressive phonation: Diachronic and synchronic characteristics, distribution and function in animal and human sound production and in human speech", Journal of the international phonetic association **38**, 235–324 (2008).

[6] P. Šidlof, "Fluid-structure interaction in human vocal folds", Ph.D. thesis, Faculty of Mathematics and Physics, Charles University in Prague (2007).

[7] M. Zanartu, L. Mongeau, and G. Wodicka, "Influence of acoustic loading on an effective single mass model of the vocal folds", J. Acoust. Soc. Am. **121** (2007).

[8] M. Heil and A. Hazel, "Fluid-structure interaction in internal physiological flows", Annu. Rev. Fluid. Mech. **43**, 141–162 (2011).

[9] W. Boyce and R. DiPrima, *Elementary differential equations and boundary value problems*, 7th edition edition (John Wiley Sons, Inc.) (2001).

[10] J. Den Hartog, *Mechanical vibrations*, 4th edition edition (Dover) (1985).

[11] W. Thomson and M. Dahleh, *Theory of vibration with applications*, 5th edition edition (Prentice Hall) (1998).

[12] A. H. Benade, *Fundamentals of Musical Acoustics*, 2nd, reprint, illustrated edition (Dover) (1990).

[13] F. Alipour, D. Berry, and I. Titze, "A finite-element model of vocal-fold vibration", J. Acoust. Soc. Am. **108**, 3003–3012 (2000).

[14] I. Titze and W. Strong, "Normal modes in vocal cord tissues", J. Acoust. Soc. Am. **57**, 736–744 (1975).

[15] K. Ishizaka and J. L. Flanagan, "Synthesis of voiced sounds from a two-mass model of the vocal cords", The Bell System Technical Journal **51**, 1233–1268 (1972).

[16] D. Berry, H. Herzel, I. Titze, and K. Krischer, "Interpretation of biomechanical simulations of normal and chaotic vocal fold oscillations with empirical eigenfunctions", J. Acoust. Soc. Am. **95**, 3595–3604 (1994).

[17] I. Steinecke and H. Herzel, "Bifurcations in an asymmetric vocal-fold model", J. Acoust. Soc. Am. **97** (1995).

[18] Z. Zhang, J. Neubauer, and D. Berry, "Physical mechanisms of phonation onset: a linear stability analysis of an aeroelastic continuum model of phonation", J. Acoust. Soc. Am. **122**, 2279–2295 (2007).

[19] S. Smith and S. Thomson, "Effect of inferior surface angle on the self-oscillation of a computational vocal fold model", J. Acoust. Soc. Am. **131**, 4062–4075 (2012).

[20] S. Thomson, L. Mongeau, and S. Frankel, "Aerodynamic transfer of energy to the vocal folds", J. Acoust. Soc. Am. **118**, 1689–1700 (2005).

[21] B. Kröger and P. Birkholz, "A survey of self-oscillating lumped-element models of the vocal folds", in *In B.J. Kröger, P. Birkholz (eds.) Studientexte zur Sprachkommunikation: Elektronische Sprachsignalverarbeitung (TUDPress, Dresden), pp. 47-58* (2011).

[22] I. R. Titze, "The physics of small-amplitude oscillation of the vocal folds", J. Acoust. Soc. Am. **83**, 1536–1552 (1988).

[23] Z. Zhang, "On the difference between negative damping and eigenmode synchronization as two phonation onset mechanisms", J. Acoust. Soc. Am. **129**, 2163–2167 (2011).

[24] B. Story, "An overview of the physiology, physics and modeling of the sound source for vowels.", Acoust. Sci. Tech. **23** (2002).

[25] I. Tokuda, J. Horacek, J. Svec, and H. Herzel, "Comparison of biomechanical modeling of register transitions and voice instabilities with excised larynx experiments", J. Acoust. Soc. Am. **122**, 519–531 (2007).

[26] C. Forster, *Musical Mathematics; On the art and science of acoustic instruments* (Chronicle Books, San Francisco) (2010).

[27] S. Lulich, A. Alwan, H. Arsikere, J. Morton, and M. Sommers, "Resonances and wave propagation velocity in the subglottal airways", J. Acoust. Soc. Am. **130**, 2108–2115 (2011).

[28] S. Mathur, B. Story, and J. Rodriguez, "Vocal-tract modeling: Fractional elongation of segment lengths in a waveguide model with half-sample delays.", IEEE Transactions on audio, speech, and language processing. (2006).

[29] D. Sciamarella and G. Artana, "A water hammer analysis of pressure and flow in the voice production system", Speech Communication **51**, 344–351 (2009).

[30] J. L. Flanagan and L. L. Landgraf, "Self-oscillating source for vocal tract synthesizers", IEEE Trans. Audio Electroacoust. AU-16, 5764 (1968).

[31] C. Kelly and K. Fisher, "Stroboscopic and acoustic measures of inspiratory phonation", Journal of Voice **13**, 389–402 (1999).

[32] J. van den Berg, J. Zantema, and P. Doornenbal, Jr., "On the air resistance and the bernoulli effect of the human larynx", J. Acoust. Soc. Am. **29**, 626–631 (1957).

[33] D. Childers and J. Larar, "Electroglottography for laryngeal function assessment and speech analysis", IEEE Transactions on Biomedical Engineering **BME-31** (1984).

[34] R. Orlikoff, R. Baken, and D. Kraus, "Acoustic and physiologic characteristics of inspiratory phonation", J. Acoust. Soc. Am. **102**, 1838–1845 (1997).

[35] I. Titze and A. Worley, "Modeling source-filter interaction in belting and high-pitched operatic male singing", J. Acoust. Soc. Am. **126**, 1530–1540 (2009).

[36] M. Kob, "Physical modeling of the singing voice; doctoral dissertation", Ph.D. thesis, RWTH Aachen University, Aachen (2002).

[37] I. R. Titze, *Principles of Voice Production*, 1st edition edition (Prentice-Hall Inc) (1994).

[38] P. Lieberman and S. Blumstein, *Speech physiology, speech perception, and acoustic phonetics*, 1st edition edition (Cambridge University Press) (1998).

[39] H. Rasilo, "Estimation of vocal tract shape trajectory using lossy kelly-lochbaum model", Ph.D. thesis, Aalto University School of Science and Technology Faculty of Electronics, Communications and Automation (2010).

[40] H. Yamaguchi, *Engineering Fluid Mechanics*, 1st edition edition (Springer) (2008).

[41] S. Li, R. Scherer, M. Wan, S. Wang, and H. Wu, "Numerical study of the effects of inferior and superior vocal fold surface angles on vocal fold pressure distributions", J. Acoust. Soc. Am. **119**, 3003–3010 (2006).

[42] I. Steinecke and H. Herzel, "Bifurcations in an asymmetric vocal-fold model", J. Acoust. Soc. Am. **97**, 1874–1884 (1995).

[43] M. Zemke, "Biomechanical-aerodynamical modeling of register transitions of the human voice", Ph.D. thesis, Humboldt University of Berlin, Institute for Theoretical Biology (2008).

[44] H. Strube, "The meaning of the kelly-lochbaum acoustic-tube model", J. Acoust. Soc. Am. **108**, 1850–1855 (2000).

[45] D. Murphy, A. Kelloniemi, J. Mullen, and S. Shelley, "Acoustic modeling using the digital waveguide mesh", Signal Processing Magazine, IEEE **24**, 55–66 (2007).

[46] J. Benesty, M. Sondhi, and Y. Huang, *Springer Handbook of Speech Processing*, 1st edition edition (Springer) (2007).

[47] B. Story and I. Titze, "Parameterization of vocal tract area functions by empirical orthogonal modes", Journal of Phonetics **26**, 223–260 (1998).

[48] I. R. Titze, "Parameterization of the glottal area, glottal flow, and vocal fold contact area", J. Acoust. Soc. Am. **75**, 570–580 (1984).

[49] R. Fraile, M. Kob, J. Godino-Llorente, N. Senz-Lechn, V. Osma-Ruiz, and J. Gutirrez-Arriola, "Physical simulation of laryngeal disorders using a multiple-mass vocal fold model", Biomedical Signal Processing and Control **7**, 65–78 (2012).

[50] I. R. Titze, "Nonlinear source-filter coupling in phonation: theory", J. Acoust. Soc. Am. **123**, 2733–2749 (2008).

[51] D. Wagg and S. Schafer, *Nonlinear Vibration with Control; For Flexible and Adaptive Structures*, Solid Mechanics And Its Applications, Volume 170, 1st edition (Springer) (2010).

[52] J. Horacek and J. Svec, "Aeroelastic model of vocal-fold-shaped vibrating element for studying the phonation threshold", Journal of Fluids and Structures **16**, 931–955 (2002).

[53] R. Vaccaro, *Digital control: A state-space approach*, 1st edition edition (Mcgraw-Hill College) (1995).

[54] G. Dullerud and F. Paganini, *A course in robust control theory: A convex approach*, 1st edition edition (Springer Science+Business Media.Inc.) (2010).

[55] P. Mergell, H. Herzel, and I. Titze, "Irregular vocal-fold vibrationhigh-speed observation and modeling", J. Acoust. Soc. Am. **108**, 2996–3002 (2000).

[56] W. Press, *Numerical recipes in C: the art of scientific computing*, 2nd edition (Cambridge University Press) (1992).

[57] http://www.mathworks.com/support/tech–notes/1500/1510.html Accessed: 28/03/2012.

[58] S. Lowell and B. Story, "Simulated effects of cricothyroid and thyroarytenoid muscle activation on adult-male vocal fold vibration", J. Acoust. Soc. Am. **120**, 386–397 (2006).

[59] C. Lausted, A. Johnson, W. Scott, M. Johnson, and D. Coyne, K.M. Coursey, "Maximum static inspiratory and expiratory pressures with different lung volumes", BioMedical Engineering OnLine **5** (2006).

[60] M. de Vries, H. Schutte, and G. Verkerke, "Determination of parameters for lumped parameter models of the vocal folds using a finite-element method approach", J. Acoust. Soc. Am. **106** (1999).

[61] K. Stevens, *Acoustic Phonetics*, 1st edition edition (The MIT Press) (1998).

[62] J. Lucero and L. Koenig, "Simulations of temporal patterns of oral airflow in men and women using a two-mass model of the vocal folds under dynamic control", J. Acoust. Soc. Am. **117**, 1362–1372 (2005).

[63] J. Lucero and L. Koenig, "On the relation between the phonation threshold lung pressure and the oscillation frequency of the vocal folds (l)", J. Acoust. Soc. Am. **121**, 3280–3283 (2007).

[64] M. Zanartu, "Acoustic coupling in phonation and its effect on inverse filtering of oral airflow and neck surface acceleration", Ph.D. thesis, School of Electrical and Computer Engineering, Purdue University (2010).

[65] H. Hatzikirou, W. Fitch, and H. Herzel, "Voice instabilities due to source-tract interactions", Acta Acustica united with Acustica **92**, 468–475 (2006).

[66] D. Berry, "Mechanisms of modal and nonmodal phonation.", Journal of Phonetics **29**, 431–450 (2001).

[67] F. Avanzini, "Simulation of vocal fold oscillation with a pseudo-one-mass physical model", Speech Communication **50**, 95–108 (2008).

[68] L. Raphael, G. Borden, and K. Harris, *Speech science primer: physiology, acoustics, and perception of speech*, 5th edition edition (Lippincott Williams Wilkins) (2007).

[69] I.R.Titze and B. Story, "Rules for controlling low-dimensional vocal fold models with muscle activation", J. Acoust. Soc. Am. **112**, 1064–1076 (2002).

Appendix A

Equivalency of state-space representation of the tissue dynamics to Newton's second law of motion

Newton's second law form for the behavior of the oscillators:

$$m_1\ddot{x}_1 + r_1\dot{x}_1 + k_1x_1 + (x_1 - x_2)k_{12} = F_1 \tag{A.1}$$

$$m_2\ddot{x}_2 + r_2\dot{x}_2 + k_2x_2 + (x_2 - x_1)k_{12} = F_2 \tag{A.2}$$

Proof that the following is equivalent to Newton's second law:

$$\frac{\mathrm{d}}{\mathrm{d}t}\begin{bmatrix}\mathbf{x}\\ \dot{\mathbf{x}}\end{bmatrix} = \mathbf{A}\begin{bmatrix}\mathbf{x}\\ \dot{\mathbf{x}}\end{bmatrix} + \begin{bmatrix}\mathbf{0}\\ \mathbf{M}^{-1}\mathbf{F}\end{bmatrix} \tag{A.3}$$

where $\mathbf{A}$ is defined by:

$$\mathbf{A} = \begin{bmatrix}\mathbf{0} & \mathbf{I}\\ -\mathbf{M}^{-1}\mathbf{K} & -\mathbf{M}^{-1}\mathbf{R}\end{bmatrix} \tag{A.4}$$

Positional vector:

$$\mathbf{x} = \begin{bmatrix}x_1\\ x_2\end{bmatrix} \tag{A.5}$$

The state is the combination of both positions and velocities:

$$\begin{bmatrix} \mathbf{x} \\ \dot{\mathbf{x}} \end{bmatrix} = \begin{bmatrix} x_1 \\ x_2 \\ \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} \tag{A.6}$$

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \ddot{x}_1 \\ \ddot{x}_2 \end{bmatrix} = \frac{\mathrm{d}}{\mathrm{d}t} \begin{bmatrix} x_1 \\ x_2 \\ \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} \tag{A.7}$$

$$\mathbf{M} = \begin{bmatrix} m_1 & 0 \\ 0 & m_2 \end{bmatrix} \tag{A.8}$$

$$\mathbf{R} = \begin{bmatrix} r_1 & 0 \\ 0 & r_2 \end{bmatrix} \tag{A.9}$$

$$\mathbf{K} = \begin{bmatrix} k_1 + k_{12} & -k_{12} \\ -k_{12} & k_2 + k_{12} \end{bmatrix} \tag{A.10}$$

$$\mathbf{M}^{-1}\mathbf{R} = \begin{bmatrix} 1/m_1 & 0 \\ 0 & 1/m_2 \end{bmatrix} \begin{bmatrix} r_1 & 0 \\ 0 & r_2 \end{bmatrix} = \begin{bmatrix} r_1/m_1 & 0 \\ 0 & r_2/m_2 \end{bmatrix} \tag{A.11}$$

$$\mathbf{M}^{-1}\mathbf{K} = \begin{bmatrix} (k_1 + k_{12})/m_1 & -k_{12}/m_1 \\ -k_{12}/m_2 & (k_2 + k_{12})/m_2 \end{bmatrix} \tag{A.12}$$

$$\mathbf{F} = \begin{bmatrix} F_1 \\ F_2 \end{bmatrix} \tag{A.13}$$

$$\mathbf{M}^{-1}\mathbf{F} = \begin{bmatrix} 1/m_1 & 0 \\ 0 & 1/m_2 \end{bmatrix} \begin{bmatrix} F_1 \\ F_2 \end{bmatrix} = \begin{bmatrix} F_1/m_1 \\ F_2/m_2 \end{bmatrix} \tag{A.14}$$

Recalling the definition for the **A** matrix:

$$\mathbf{A} = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ -(k_1 + k_{12})/m_1 & k_{12}/m_1 & -r_1/m_1 & 0 \\ k_{12}/m_2 & -(k_2 + k_{12})/m_2 & 0 & -r_2/m_2 \end{bmatrix} \tag{A.15}$$

$$\frac{\mathrm{d}}{\mathrm{d}t}\begin{bmatrix} x_1 \\ x_2 \\ \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ \dfrac{-(k_1 + k_{12})}{m_1} & \dfrac{k_{12}}{m_1} & -\dfrac{r_1}{m_1} & 0 \\ \dfrac{k_{12}}{m_2} & \dfrac{-(k_2 + k_{12})}{m_2} & 0 & -\dfrac{r_2}{m_2} \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ \dfrac{F_1}{m_1} \\ \dfrac{F_2}{m_2} \end{bmatrix} \tag{A.16}$$

Applying the time derivative and multiplying out the matrices for each row of **A**:

$$\frac{-(k_1 + k_{12})}{m_1}x_1 + \frac{k_{12}}{m_1}x_2 - \frac{r_1}{m_1}\dot{x}_1 + \frac{F_1}{m_1} = \ddot{x}_1 \tag{A.17}$$

$$\frac{k_{12}}{m_2}x_1 + \frac{-(k_2 + k_{12})}{m_2}x_2 + -\frac{r_2}{m_2}\dot{x}_2 + \frac{F_2}{m_2} = \ddot{x}_2 \tag{A.18}$$

Which are identical to the Newton's second law expressions as required.