INFLUENCE OF ACOUSTIC LOADING ON THE FLOW-INDUCED OSCILLATIONS OF SINGLE MASS MODELS OF THE HUMAN LARYNX

A Thesis

Submitted to the Faculty

of

Purdue University

by

Matías Zañartu Salas

In Partial Fulfillment of the

Requirements for the Degree

of

Master of Science in Engineering

May 2006

Purdue University

West Lafayette, Indiana

ACKNOWLEDGMENTS

I would like to acknowledge my advisors, Prof. Luc Mongeau, for inviting me to participate in the voice production project and encouraging me to achieve this goal, and Prof. George Wodicka, for his interest and support along the way. I would also like to thank my advisory committee members, Prof. Alex Francis and Prof. Thomas Talavage for their guidance during the time I have spent in Purdue. Special thanks to Prof. Mary Harper for her support, particularly at the beginning of this project.

I gratefully acknowledge the support provided by the Fulbright Program and the Institute of International Education to pursue this Master degree. Being a Fulbright fellow has been a great honor and a wonderful experience. Part of this project was funded by the National Institute on Deafness and Other Communication Disorders NIDCD, National Institutes of Health (NIH), grant number R01 DC05788.

I would like to thank my fellow students in our research group, especially Doug Cook, Jong Beom Park and Byron Erath for the inspiring discussions and interactions. I would also like to thank Shohei Shibata for his help during the implementation of the source model.

Most of all, I would like to acknowledge the endless support and understanding of my beloved wife, Sandra. Her support and love have made this thesis and journey possible.

TABLE OF CONTENTS

Page

Page

Page

LIST OF TABLES

LIST OF FIGURES

Figure Page

Figure Page

Figure Page

Figure Page

LIST OF SYMBOLS

F	Forward acoustic pressure [Pa]
B	Backward acoustic pressure [Pa]
Z	Acoustic impedance [Ns/m^3]
A	Area of a tract section [m^2]
F_s	Sampling frequency [Hz]
r	Reflection coefficient
ω	Angular frequency [rad/s]
z^{-1}	Unit delay in the z-domain
α	Attenuation factor [cm^{-1}]
ζ	Shape factor
γ	Propagation loss factor
R_v	viscous friction losses factor [cgs Ω^{-1} cm^{-1}]
G_t	heat conduction losses factor [cgs Ω^{-1} cm^{-1}]
G_w	yielding walls vibrations losses factor [cgs Ω^{-1} cm^{-1}]
ρ	Density of air [kg/m^3]
c	Speed of sound [m/s]
m	Mass of one vocal fold [kg]
k	Spring constant [N/m]
b	Damping constant [kg/s]
ς	Damping ratio
E	Young modulus of the vocal folds [kPa]
μ	Poisson ratio of the vocal folds
y	Position of the mass at the glottal exit [m]
y_o	Equilibrium position of the mass [m]
x	Axial flow direction
v	Velocity of the mass [m/s]
d	Depth of the folds [m]
L	Length of the folds [m]
p_{up}	Upstream steady pressure (subglottal) [Pa]

p_{dn}	Downstream steady pressure (supraglottal) [Pa]
Δp	Steady transglottal pressure drop [Pa]
p'_{up}	Upstream acoustic pressure (subglottal) [Pa]
p'_{dn}	Downstream acoustic pressure (supraglottal) [Pa]
$p_{cl}(x)$	Centerline pressure as function of position [0, d] [Pa]
$p_{wall}(x)$	Wall pressure over the folds as function of position [0, d] [Pa]
u_{up}	Upstream flow velocity (subglottal) [m/s]
$u_{cl}(x)$	Flow velocity at the wall as function of position [0, d] [m/s]
u_{cl}	Centerline flow velocity at the glottal opening [m/s]
$u_{wall}(x)$	Flow velocity at the wall as function of position [0, d] [m/s]
Q	Volumetric flow rate or volumetric flow velocity [m^3/s]
A_s	Subglottal area before the glottis [m^2]
A_e	Supraglottal area right after the glottal opening (epilarynx) [m^2]
$A_g(x)$	Glottal area as a function of position [0, d] [m^2]
A_g	Glottal area at the glottal opening [m^2]
$f(x)$	Geometry of the folds as function of distance [0, d] [m]
$c_d(t)$	Time-varying orifice discharge coefficient

ABSTRACT

Zañartu Salas, Matías. M.S.E., Purdue University, May 2006. Influence of Acoustic Loading on the Flow-Induced Oscillations of Single Mass Models of the Human Larynx. Major Professors: Luc Mongeau and George R. Wodicka, School of Electrical and Computer Engineering.

Linear models of voice production assume that the vibrations of the vocal folds are independent of the sound waves radiated in the subglottal tract and the vocal tract. In normal phonation, a significant level of interaction between these systems is expected. In this project, three-way interactions between sound waves in the subglottal and supraglottal tracts, the vibrations of the vocal folds, and laryngeal flow were investigated. Nonlinear pressure-flow and impact forcing relations were used, with a focus on the effects of realistic acoustic loading on the self-sustained oscillation of the vocal folds. MRI-based shapes for the supraglottal tract were used, along with an equivalent acoustical model for the subglottal tract. The influence of planar sound waves in the tracts was modeled based on a wave reflection analog technique. An effective single degree-of-freedom model of the vocal folds described the source. The effects of a mucosal wave on the vocal folds were incorporated using a time-varying discharge coefficient, and a flow model based on Bernoulli's equation. The effects of collision were included using non-linear impact forces in the system. Although a single degree of freedom model was used, results comparable with higher order models were obtained. The model showed that collision forces significantly affect the oscillation of the system. The relative importance of acoustic loading and the time-varying flow resistance for fluid-structure energy transfer was established for various configurations. The results suggest that both acoustic coupling and aerodynamic adjustments may be important factors that contribute to self-sustained oscillations during normal phonation. Acoustic loading might be more significant than the effects introduced by the orifice discharge coefficient. The influence of the subglottal tract appears to be significant, but not as determinant as that of the vocal tract.

1. INTRODUCTION

This chapter introduces the motivation for the research, indicating the problem and its significance. The objectives of the research are also stated. Finally, a description of the organization of the thesis is provided.

1.1 Introduction and Motivation for the Research

Phonation is a key component of speech, the primary tool for human communication. Engineering has played an important role in understanding, modeling, and reproducing the physics of phonation. Since the early work of Van den Berg [53] and Fant [10], the study of voice production has continued to evolve. Early models of voice production have been useful in the development of speech synthesis, speech recognition and speech coding. Current research aims at a more accurate description of laryngeal behavior. A better understanding of phonation may improve upon previous applications, and in addition lead to the early detection and better treatment of voice pathologies, the development of bioimplants, the optimization of voice prosthesis, voice enhancements for singers, better tools for voice simulation and speech perception, or phonosurgical modeling.

One of the first models of voice production, the linear source-filter theory has been a leading concept in speech production [10]. This theory assumes that the source and the tracts are acoustically independent, i.e. the transglottal pressure is independent of the sound field in the airways. More recent studies have demonstrated that strong fluid-sound interactions take place during normal phonation [49]. However, no detailed study of the factors that control this interaction has been performed yet. Hence, important questions in voice production still remain unanswered. The consequences of including fluid-sound interactions in voice production include the potential impact on voice efficiency, the spectral content of the voice signature, and modified stability characteristics resulting in a change in oscillation frequency and onset pressure.

To develop models that may allow for example the influence of surgery or pathological conditions to be modeled, accurate and detailed models of the tissue and the airflow are needed. Ongoing research [45] uses the finite element method to model the complex fluid-structure interactions within the larynx. Such models often use an incompressible flow formulation, and ignore acoustic loading effects. Simplified, unrealistic boundary conditions such as constant pressure or a constant velocity are imposed at the inflow and outflow boundaries of the computational domain. One additional motivation for the present study was to investigate the use of the time domain methods to be included as boundary conditions in future finite element simulations and other continuum models of the fluid flow through the glottis.

1.2 Research Objectives

The purpose of this research was to investigate the factors that control voice production, in particular the coupling between the vibration of the vocal folds, sound field in the subglottal and supraglottal tracts, and the airflow through the larynx. Specific goals have been established. The first goal was to implement a scheme to compute the sound pressure in the airways, including the subglottal and the supraglottal tracts. The approach was a static representation of the systems during the production of non-nasal vowels. The scheme was expected to produce an audible output, and to compute the instantaneous acoustic pressures upstream and downstream of the glottis. Since most useful vocal folds models require a time marching solution, a time domain representation of the acoustic loading is required. Given that acoustic models are normally in the frequency domain, this requirement can be met by either an inverse filtering method or a time domain based scheme.

The second goal was the implementation of a simplified model of the vocal folds vibration to allow investigations of the time marching scheme for the acoustics of the tracts. Model requirements included the possibility of fluid-sound interaction and physically based fluid-structure interactions. The source model was required to perform self-sustained oscillation without the presence of acoustic loading.

The third and more general goal was to investigate, using both previously developed models, the role and importance of acoustic loading in phonation. Thus, the time domain representation of the sound field in the tracts was used to impose acoustic loading in the numerical simulation of the source model. This approach allows the effects

of both supraglottal and subglottal tracts on self-sustained oscillations to be compared with the effects of aerodynamic pressure forces and impact forces present in phonation.

1.3 Organization of the Thesis

This thesis is organized as follows. In Chapter 2, the literature in the field of voice modeling is summarized. Non-linear models of voice production are discussed. The most significant source models along with models for the acoustics of the subglottal and supraglottal tracts are described. The criteria for the selection of the specific design parameters used are defined. In Chapter 3, the wave reflection analog technique used to model the acoustic loading is described. Key details that could be improved in future work are described. In Chapter 4, the development of the vocal fold model is discussed, along with the analytical expressions that define the source model. In Chapters 5, a series of preliminary studies of the acoustics of the tract is presented. The accuracy of the wave reflection analog is evaluated. Details of the subglottal tract design are discussed. Linear and non-linear models for the acoustic coupling between subglottal and supraglottal tracts are compared. In Chapter 6, preliminary results for the source model with no acoustic loading are presented. The effects of the time-varying orifice discharge coefficient are discussed. An analysis of the forces in presence is presented. The main assumptions and limitations of the model are reviewed and discussed. In chapter 7, the role of acoustic loading in self-sustained oscillations is illustrated for both subglottal and supraglottal tracts. The effects of acoustic coupling are discussed along with other aerodynamics effects observed in human phonation. In Chapter 8, the major conclusions and implications of this study are summarized. Recommendations for future research are made.

2. LITERATURE REVIEW

This chapter reviews the literature in the field of voice production. The basic principles of voice production are presented to create a foundation for the topics discussed in the document. The most significant source models for phonation, along with models for the acoustics of the subglottal and supraglottal tracts are described. The factors that led to the specific model design used in this study are discussed.

2.1 Basics of Voice Production

The primary organs of phonation, the vocal folds, are two tissue formations located within the larynx. Each vocal fold is composed of various types of tissue, forming a rather complex layered structure. Tissue layers can be categorized in terms of a cover, a transition, and a body. The glottis is a constriction between the two vocal folds, being a few millimeters wide during phonation. Figure 2.1 illustrates the location of the vocal folds within the larynx. The glottis is nearly symmetric.

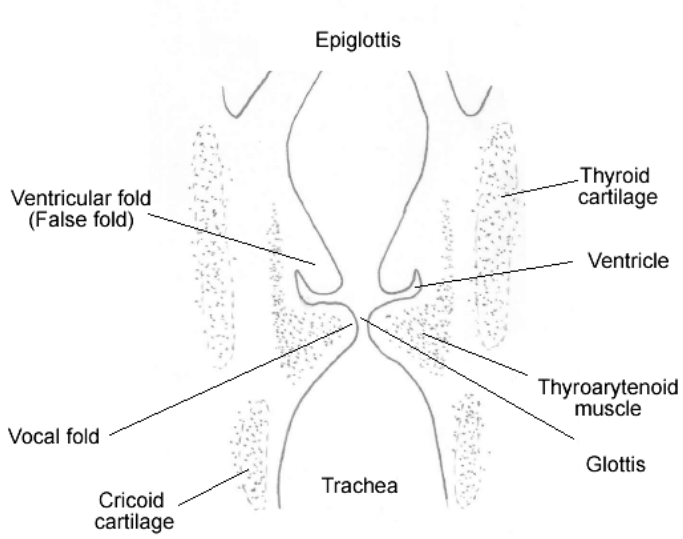


Fig. 2.1 Coronal section of the larynx (adapted from Stevens [39]).

It is normally convenient to divide the voice production system using the glottis as a midline. Hence, the complete system is divided into subglottal, laryngeal and supraglottal structures. The total subglottal pressure right below the glottis is also referred to as the upstream pressure, and the total supraglottal pressure right above the glottis is referred as the downstream pressure.

In normal voice production, the subglottal system supplies energy to the airflow. The laryngeal and the supraglottal structures modulate the airflow. Sound is produced by the net streamwise force applied on the fluid by the vocal folds walls. Flow modulation may occur either at the glottal constriction, in areas surrounding the constriction, or along the supraglottal pathway. Periodic modulations at the glottis constriction were considered. In this case, the energy provided by the subglottal system (the lungs in particular) produces airflow resulting in a net pressure difference across the glottis. The aerodynamic pressure on the vocal folds wall exerts a lateral force on the surface of the vocal folds, causing the surface to move outward and allowing the air to pass through the glottis. As soon as the airflow begins, the pressure differential is reduced (reaching sometimes negative values), diminishing the forces in the lower portion of the folds, which initiates an inward movement of the surface. During inward and outward motions, the upper and lower portions of the folds are delayed with respect to each other. This is referred as the "mucosal wave" effect. This produces convergent and divergent glottal shapes during phonation. During inward motion (where collision between the folds can happen), the upstream pressure exerts a force on the lower surfaces of the folds, and the cycle repeats. A cross-section of the glottis in the coronal plane is shown in Figure 2.2 at different instants during one cycle. Note that the glottis profile alternates between convergent, Figure 2.2c, and divergent, Figure 2.2e, shapes during the open phase.

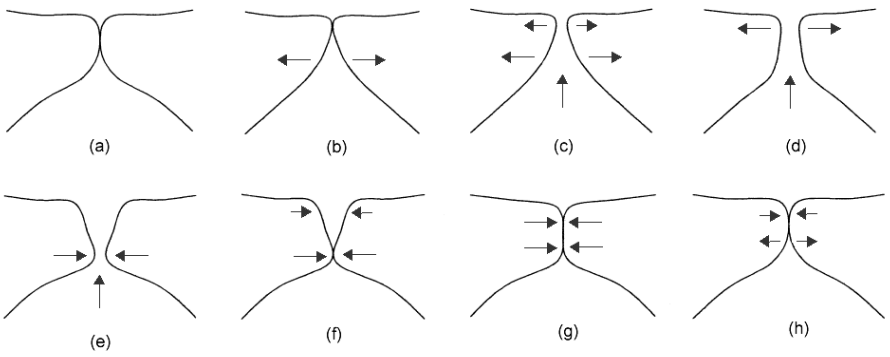


Fig. 2.2 Vocal folds vibration cycle (adapted from Story [44]).

Although the phonation process appears to be relatively simple, it involves complex anatomy, muscular control, fluid-structure interactions, fluid flow oscillations, flow-sound interactions, dynamics, impact, and adhesion. The voice apparatus may be viewed as a non-linear oscillator governed by physical laws involving complex interplay between several variables. Since the forcing involved is steady, the system oscillations are "self-sustained", i.e. they involve a net energy transfer between flow and tissue vibrations. These highly complex oscillations produce an acoustic signal rich in harmonics, normally referred as a voiced sound.

2.2 Acoustics of the Subglottal and Supraglottal Tracts

The linear source-filter theory developed by Fant [10] assumes that the vocal folds vibrate independently from the vocal tract. This implies that the filter properties of the vocal tract do not affect the source. The imposed source signal is filtered by the vocal tract. Resonance modes, referred to as "formants", color the source creating a variety of voiced sounds. Note that the acoustics of the subglottal tract is neglected in this approach.

Many acoustic models of the vocal tract have been developed. Early on, the primary interest was speech synthesis. Initially, uniform tubes and coupled resonators were used to describe the behavior of the vocal tract. Fant [10] showed that a system with three tube sections and a radiation impedance allowed differentiation between vowels. The formant frequencies were computed by combining models for sound in uniform tubes and Helmholtz resonators. To produce a synthesized output, analog electronic circuits were required. The resulting synthesized signals were far from realistic. Aiming to improve this scheme, Stevens and House [38] used larger numbers of tube sections using again electronic circuit analogies. Variations of this design were introduced by Flanagan *et al.* [13] [19] [20] to account for source-load interactions and better models for the loss mechanisms. These authors proposed a model consisting of cascaded T-sections that accounted for losses in each tube section, and for the properties of the subglottal tract. These models, referred to as electrical transmission line models, required iterative numerical solutions. This traditional approach to model the acoustics of both supraglottal and subglottal tracts allows the production of realistic speech sounds, but is ineffective at handling the fast transitions occurring during normal speech.

Searching for time-domain based methods to describe the acoustics of the tracts, Kelly and Lochbaum [22] designed a wave reflection analog approach to speech synthesis. The technique follows the multi-tube approach used by Stevens and House in [38], but the sound field inside the tubes is computed using reflection coefficients. The technique performs computations synchronously with the acoustic wave propagation. Several improvements have been introduced by Liljencrants [26], Rahim [34], Story [41] and others since its original design. The most recent development of this intrinsically time-domain technique includes many novel features such as nose and skin radiation, yielding wall, viscous effects, heat conduction losses, coupling with the nasal tracts, and others. Perhaps one of the most attractive benefits of the wave reflection analog technique is its capacity to account for geometric details. Highly detailed tract designs are modeled with the same level of efficiency as simple ones. For this reason, particular attention has been paid to realistic tract shapes. Story *et al.* [43] created a database of magnetic resonance based images of the vocal tract for different vowels and speech sounds. Story stated [41] that no other modeling technique has been able to positively correlate geometrical improvements. Overall, this technique has been successful for the simulation of natural sounding vowels. The same approach, however, is problematic for consonants, which has limited its use for articulatory speech synthesis. The technique introduces natural delays that are cause and effect of physical laws during the propagation of the wave inside the tracts. It has found important applications in modeling the acoustic coupling with self-sustained models of the vocal folds. Direct comparisons between simulated and real speech sounds produced by a subject "donor" of MRI vocal tract shapes have shown the accuracy of the wave reflection analog technique [43].

Since neither the electrical transmission line nor the wave reflection analog methods produce acceptable speech signals, other techniques have been used in speech synthesis. Klatt [24] designed a highly detailed "formant synthesizer" that produces quality speech utterances and a text-to-speech software. Although considered an improvement over previous methods, the formant synthesizer does not involve articulatory models of the vocal tract. Efforts to link Klatt's formant synthesizer to articulatory models have been made, resulting in so-called "hybrid speech synthesizers". Nevertheless, articulatory synthesis never took off since formant synthesis was better understood at that time. Following Klatt's synthesizer, the need to produce better text-to-speech devices has led to the advent of "concatenative speech synthesis". This technique concatenates recorded speech segments using prosodic elements to make the resulting

sound more natural. This has been the leading technique in industrial applications. Articulatory synthesis and analytical models of the acoustics of the vocal tract have found applications in other research areas such as phonetic acoustics, speech recognition, speech coding, and speech perception.

In contrast with the vocal tract, the acoustics of the subglottal tract have received less attention. Recently, the early work of Ishizaka [21] and Flanagan [12] has been revisited [10] [16] [31] [2] [55]. One goal has been the search for clinical-oriented applications for the diagnosis of respiratory diseases. With a different perspective, efforts to include the effects of the acoustics of the subglottal tract in speech production are still under development. Stevens [10] has shown that the resonances observed for the subglottal tract can be present in radiated speech sounds. Such coupling between tracts has also been observed by Harper [16]. Experimental work by Austin and Titze [2] has suggested that the subglottal tract could affect the vibration of the vocal folds. Little has been done in that direction since, very recent studies notwithstanding [31].

2.3 Mathematical Models of Voice Production

Early models of the vocal folds vibrations were based on the myoelastic-aerodynamic theory from Van den Berg [53]. According to this theory, fluid pressure forces and tissue elasticity were the main cause of vocal fold oscillations. Based on this principle, the first model of the vocal folds was the one-mass model from Flanagan and Landgraf [13]. It was constructed based on empirical expressions from Van den Berg [52] for the fluid pressure forces. The resulting force was based on a linear distribution of the acoustic pressures upstream and downstream, in the glottis. This concept is attractive because it produces representative oscillations of the vocal folds and allows interaction with the sound in the supraglottal tract. The subglottal tract was modeled as a constant pressure source. Since self-oscillation was governed by acoustic feedback, the absence of the vocal tract made it impossible to obtain self-sustained oscillations in the model. Excised larynx experiments have shown that self-sustained oscillation can be obtained in absence of a vocal tract. That led to a revision of the oscillatory theory of vocal fold vibration. Ishizaka and Flanagan [19] included the effects the delay between the upper and lower portion of the vocal folds during oscillation, creating the well known two-mass model of the vocal folds. This model allowed for acoustic feedback and self-sustained oscillations with or without the presence of acoustic loading. The model used a fluid

pressure driving force that changed depending on the shape of the model. Bernoulli's equation is used to compute the wall pressure when the glottis has a convergent shape. When the glottis has a divergent shape, the flow separates, resulting in the establishment of a jet regime, and the force is determined from the supraglottal pressure only. Although the flow model is oversimplified, the model has been widely accepted as a valid representation of the vocal fold vibrations. Several multi-mass models have been developed since, to better account for the asymmetry between the lower and upper parts of the folds and the presence of a mucosal wave. Titze and Story [42] developed a "cover model", which adopted the main features of Ishizaka and Flanagan's model but added a third degree of freedom that helped the modeling of the mucosal wave. The model corrected discontinuity problems observed in the original two-mass model. In addition, it allowed the effects of acoustic loading on the volumetric flow rate to be considered based on a wave reflection analog approach. The flow equation (taken from [47]) was shown to be stable and able to handle time-varying flow and irregular geometries.

High order models with thousands of degree of freedom have also been used to study the oscillatory characteristics of the vocal folds. These models have been used to provide additional insight on the phonation mechanisms that are not captured by low order models. Modal analysis and fluid-structure interactions have been studied using the finite elements method. Berry and Titze [3] have emphasized the importance of the lower-order eigenmodes in the vocal folds to produce self-oscillation. Cook [6] suggested that three-dimensional models of the vocal folds are required to reproduce the vibration characteristics of phonation. Using a three-dimensional model, Hunter et al. [17] have studied the role of the thyroarytenoid and other muscles in posturing prior to phonation. Little attention has been paid to the acoustic coupling in high order models. Only recently, Alipour [1] presented a model that combines vocal fold tissue mechanics, laryngeal aerodynamics, and acoustic interactions with the subglottal and supraglottal tracts. The results of this study confirm the trends observed using low order models. Potential applications of higher order interactive models include the diagnostic and treatment of pathological conditions, and pre-operative predictions of the effects of surgery on voice quality.

2.4 Experimental Studies on Voice Production

Invaluable information is gained from the direct observation of human vocal folds, excised canine vocal folds and physical models of the vocal folds. Techniques to examine human vocal folds *in vivo* normally use a laryngoscope and a stroboscope to observe the vibrations of the vocal folds. Kymography provides a continuous image of the motion of a portion of the vocal folds surface, rather than single frame pictures such as those obtained via stroboscopy. Certain irregularities that may not be captured using stroboscope may be detected with a kymograph [46].

Due to similarities in size and gross structure with human physiology, canine excised larynges have been widely studied [46]. The aim in this case has been to measure the oscillation pattern and flow behavior. To produce self-sustained oscillations, a subglottal section is added to a flow supply. To allow optical access, no supraglottal loading is used in this type of experiment. Several factors affect the accuracy and repeatability of excised larynx experiments. For this reason, physical models of the vocal folds have been developed. Rigid models of the vocal folds have led to a better understanding of the flow in the larynx. Measurements using these models have been performed to quantify intraglottal pressure distributions [52] [36], to identify different sound sources [30], to observe the difference between different intraglottal angles [37] [25], to validate the quasi-steady assumption [58], to study the Coanda effect [8], and to measure orifice discharge coefficients [32] [58]. More recently, flexible synthetic models that resemble more closely the human vocal folds have been used [45] to study oscillatory patterns. Most of the current physical models use an idealized vocal fold geometry, referred to as M5, developed by Scherer [37].

Measures of the mechanical properties of the vocal folds play a significant role in the construction of both mathematical and synthetic models. The constants used in low order models are based on material properties. The values are normally validated by observing oscillatory behavior. For high order models, synthetic physical models and the development of bio-implants, more accurate expressions of the material properties are required. Variability among individuals, changes in tissue properties over time, and variability between tissue samples causes many problems, which are being investigated in an effort to obtain accurate mechanical properties of vocal folds tissue [3] [57] [4] [5].

2.5 Non-Linear Models of Voice Production

Extensive research has been done on modeling the interaction between the source and the vocal tract. The fluid-sound interactions observed during voice production are illustrated as acoustic feedbacks in Figure 2.3.

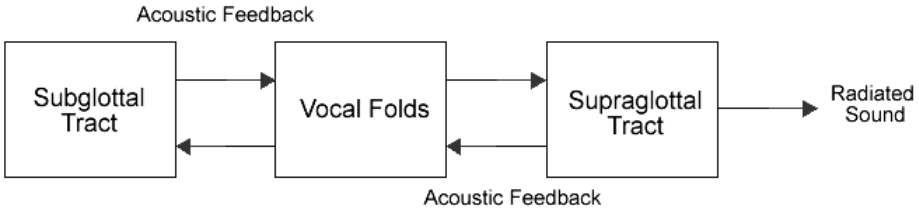


Fig. 2.3 Diagram of the voice production mechanism

Analytical models of these interactions have been reported by Rothenberg [35], Fant [10] [11], Titze [48], and Lucero [27] among others. The acoustic impedance of the vocal tract has been characterized. The impedance of the vocal tract is defined as the ratio between the acoustic pressure and the particle velocity at the glottis. This quantity is a complex number that depends on frequency, tract geometry and its configuration. Acoustic theory for plane waves in tubes states that for frequencies matching a resonance of the tract, the impedance is purely resistive (dissipating energy). For frequencies slightly below the resonance frequencies, the impedance is mainly inductive (positively reactive); and for frequencies above a formant the impedance is primarily compliant (negatively reactive). A detailed analysis of the impedance of the vocal tract was performed by Story and Titze [49]. The impedance of the supraglottal tract was argued to be primarily governed by the impedance of the lower tract, the epilarynx. From Rothenberg [35], this impedance is mostly inertive during normal speech. The length of the epilarynx and its cross-sectional area appear to be the main factors in controlling the impedance of the vocal tract, the latter being more sensitive. Narrow entry sections in the vocal tract result in a high impedance, comparable with the impedance of the glottis, which favors flow-sound interactions. Fant [11] and Rothenberg [35] have shown that under these conditions, flow-sound interactions affect the volumetric flow rate waveform. Skewing, ripple or depressions in the open phase, and short-term variations of formant bandwidth and frequency are observed in cases with strong interactions.

Titze [48] suggested that an inertive supraglottal loading could be the result of the air column within the glottis acting like a mass that is accelerated and decelerated, altering the pressures. Thus, phonation is facilitated by the combination of differential Bernoulli pressure and the delayed vocal tract response. In general, for both aerodynamic forces in the glottis and acoustic loading forces, forces that are in phase with the velocity of the mass are favorable to phonation. Initially, this theory was used to explain particular singing techniques [51]. More recently, Mergell *et al*. [29] showed that the vocal tract facilitates instabilities in the oscillation, occurring at lower lung pressures and for frequencies close to the formants. In other applications of this theory, Chan and Titze [4] have showed that the phonation threshold pressure decreased with the inertive impedance of the vocal tract, and increased with the viscous shear modulus or dynamic viscosity of the vocal fold cover. This conclusion supports the potential biomechanical benefits of certain bio-implants for repairing voice disorders.

In contrast with the wealth of information on the flow-sound interaction for the vocal tract, the topic of subglottal tract impedance lies in a relative vacuum. The role of the subglottal acoustic impedance in the vibrations of the vocal folds has not been explored. Based on previous experimental work [2], Neubauer *et al.* are currently working in this direction [31]. Preliminary results have shown that a minimum subglottal tract length is required to obtain self-sustained oscillations. Different subglottal tract configurations were found to affect the phonation threshold pressure.

3. MODEL FOR THE ACOUSTICS OF THE SUBGLOTTAL AND SUPRAGLOTTAL TRACTS

This chapter introduces the model developed to describe the acoustic properties of both the supra and subglottal tracts. As discussed previously, a time domain based scheme that allows for the evaluation of realistic tract geometries is needed. The scheme must allow flexibility when coupled with different types of self-oscillating source models - lumped element or finite element based - using a time marching algorithm. A wave reflection analog approach was chosen. The scheme is based on the guidelines and notation presented by Story [41], which constitutes the most recent version of this technique for voice production and speech synthesis.

3.1 General Assumptions and Approximations

The wave reflection analog technique models the propagation of sound in the subglottal and supraglottal tracts under a one-dimensional planar wave assumption. The three-dimensional tract shapes are discretized into a finite number of cylindrical sections having equal length but different cross-sectional areas. Sound pressure waves across these cylinders are tracked over time. Under the assumption that the waves are planar, the acoustic pressure remains the same radially across each section. The surface area of each section determines reflected wave amplitudes. Eriksson [9] showed that the planar wave assumption is valid below a cutoff frequency, given by

$$f_{cutoff} = \frac{0.5861c}{2a} \quad , \tag{3.1}$$

where c is the speed of sound and a is the radius of the cylinder. Above this cutoff frequency, higher order modes create radial variations in the acoustic pressure across the cylinder, hence violating the plane waves assumption. As can be noticed from equation

(3.1), the cutoff frequency is lower for large areas. Cross sections of 10-15 cm^2 may be found for some vowels in a large male vocal tract. For such extreme cases, the cutoff is between 4700-5700 Hz. Thus, the wave reflection analog approach may be considered accurate below 5000 Hz. For typical configurations of sub- and supra-glottal tracts, the cross-sectional area is normally less than one third of the values aforementioned. Thereby these modeling approaches are accurate over a broader range of frequencies.

Cross-sectional areas are defined using an area function for both sub and supra glottal tracts. The effects of curvature, mainly observed in the vocal tract, are not considered. Tube bends can either increase or decrease the phase velocity of acoustic waves around the bending region, which can alter the resonances of the tube. However, as discussed by Story [41], these effects may be neglected.

3.2 Scattering Equations

It is assumed that the fluid is homogeneous and isotropic, for which the speed of sound is a constant that depends only on temperature. Under these assumptions, the well-known plane wave equation for each tube section is given by

$$\frac{\partial^2 P_k}{\partial x^2} = \frac{1}{c^2}\frac{\partial^2 P_k}{\partial t^2} \quad , \tag{3.2}$$

where x is the direction of propagation, P_k is the acoustic pressure in the k-th section of the tube, and c is the speed of sound. The solution of this equation can be expressed as the superposition of forward and backward traveling acoustic pressures in the form

$$P_k(x,t) = p_k^+(x-ct) + p_k^-(x+ct) \quad , \tag{3.3}$$

where superscript "+" indicates a wave traveling in the $+x$ or forward direction (from left to right or towards the mouth) and the superscript "-" indicates a wave traveling in the $-x$ direction or backward (from right to left or towards the lungs). Note that the values of the pressures considered in this solution are real numbers. This type of solution was selected because the wave reflection analog is in fact a wave digital filter, where all complex effects are described through delays of real-valued expressions. Therefore, all acoustic variables propagated across the tracts are real quantities.

The forward and backward traveling pressure waves are denoted [41]

$$F_k = p_k^+(x-ct) \quad \text{, and} \tag{3.4}$$

$$B_k = p_k^+(x+ct) \quad , \tag{3.5}$$

respectively. Thus, at any time and for any k-th section, the total acoustic pressure is given by

$$P_k = F_k + B_k \quad . \tag{3.6}$$

Moreover, taking the difference between partial waves and considering the local characteristic acoustic impedance Z_k, the volumetric flow rate through each section in [m^3/s] can be obtained as

$$Q_k = \frac{1}{Z_k}\left(F_k - B_k\right) \quad , \tag{3.7}$$

where the characteristic acoustic impedance in [Ns/m^3] for plane waves inside a tube varies only with the area of each section, and is expressed as

$$Z_k = \frac{\rho c}{A_k} \tag{3.8}$$

where A_k is the area of the k-th section in [m^2], ρ is the density of air in [kg/m^2] and c is the speed of sound in [m/s].

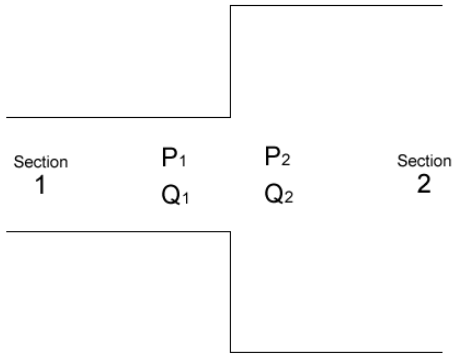


Fig. 3.1 Two-tube connection

The scattering equations yield transmitted and reflected acoustic pressures from the incident acoustic pressures for each junction between two tube sections. The acoustic pressures are computed simultaneously for all junctions. Thus, it is sufficient to present the scattering equations for one junction. The same procedure is used for all sections at every time step. The total pressure and the volumetric flow rate must be continuous at each tube junction. For the junction presented in Figure 3.1, this implies that the total acoustic pressure of tube 1 must be the same as that in tube 2, over the infinitesimal time of interaction. The same concept applies for the continuity of net flow at the junction. Therefore,

$$F_1 + B_1 = F_2 + B_2 \quad \text{, and} \tag{3.8}$$

$$\frac{1}{Z_1}\left(F_1 - B_1\right) = \frac{1}{Z_2}\left(F_2 - B_2\right) \quad , \tag{3.9}$$

where F_1 and B_2 are incident acoustic pressures into the junction, and B_1 and F_2 are reflected pressures that propagate back into each respective tube. In this sense, F_1 and B_2 can be considered as input pressures and B_1 and F_2 as output pressures for the scattering equations. As stated before, the goal is to compute the output acoustic pressure in term of the input acoustic pressure. This concept is illustrated in Figure 3.2, where the junction is considered as a system that affects the input.

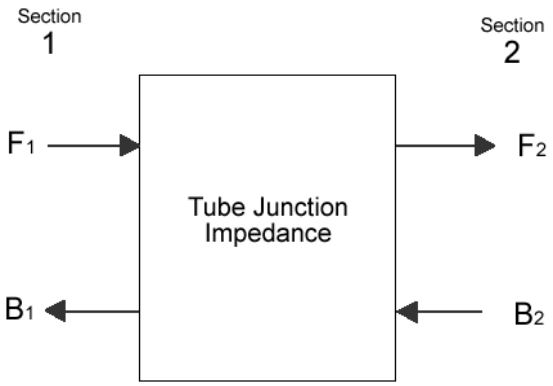


Fig. 3.2 Input/Output junction representation

Solution of equations (3.8) and (3.9) yields an explicit solution for B_1 and F_2. Since there are two unknowns and two equations, this is reduced to

$$F_2 = F_1\left(\frac{2A_1}{A_1 + A_2}\right) + B_2\left(\frac{A_2 - A_1}{A_1 + A_2}\right) \quad \text{, and} \quad (3.10)$$

$$B_1 = F_1\left(\frac{A_1 - A_2}{A_1 + A_2}\right) + B_2\left(\frac{2A_2}{A_1 + A_2}\right) \quad , \quad (3.11)$$

where, A_1 and A_2 are the cross-sectional areas of tubes 1 and 2. It is convenient to define a dimensionless reflection coefficient to simplify these equations, defined as

$$r_1 = \frac{A_1 - A_2}{A_1 + A_2} \quad . \quad (3.12)$$

The scattering equations are then

$$F_2 = F_1 + (F_1 - B_2)r_1 \quad , \quad (3.13)$$

$$B_1 = B_2 + (F_1 - B_2)r_1 \quad . \quad (3.14)$$

Equations (3.13) and (3.14) completely describe the propagation for ideal, lossless, hard-walled tubes. Concurrent solution of these equations at all junctions yields an updated set of acoustic pressures for the next time interval. Every propagating wave component requires marching in time and in space. Since the scattering equations are only valid at tube junctions, "half time" computations must be performed to update every variable at each full time step. This means that the time required for a wave component to pass through a tube section is one half the sample interval. This time-space relation is illustrated in Figure 3.3. This relation between time and space also defines the length for every tube section

$$l = \frac{c}{2F_s} \quad , \quad (3.15)$$

where l is the constant length in [m] of every tube section, c is the speed of sound in [m/s] and F_s is the sampling frequency in [Hz]. Note from equation (3.15) that F_s determines the resolution for both time and space. In the time domain, F_s is defined by the Nyquist criterion. Aliasing can be avoided if this frequency is at least twice the value

of the highest frequency component of the signal being sampled. For audible signals, a sampling frequency of 44.1 kHz is normally used. In the spatial domain, a large sampling frequency implies a large number of tube sections, and a fine spatial resolution. This is particularly useful when realistic shapes of area functions of sub and supraglottal tracts are designed. Therefore, for realistic shapes of the tracts and/or synthesized output signals, a sampling frequency of 44.1 kHz was used. For idealized tube configurations, the sampling frequency was decreased to 10 kHz to be consistent with the bandwidth of interest (up to 5 kHz). This increases efficiency without compromising accuracy.

When coupling with a sound source is involved, the acoustic pressures are supplied only in the first "half time" (where two "half times" make one sampling interval). The computational expense may be reduced, as proposed by Liljencrants [26], by solving the scattering equations at all even junctions during the first "half time"; then perform the same operation for all odd junctions during the second "half time". This procedure alleviates the need for a large number of trivial steps.

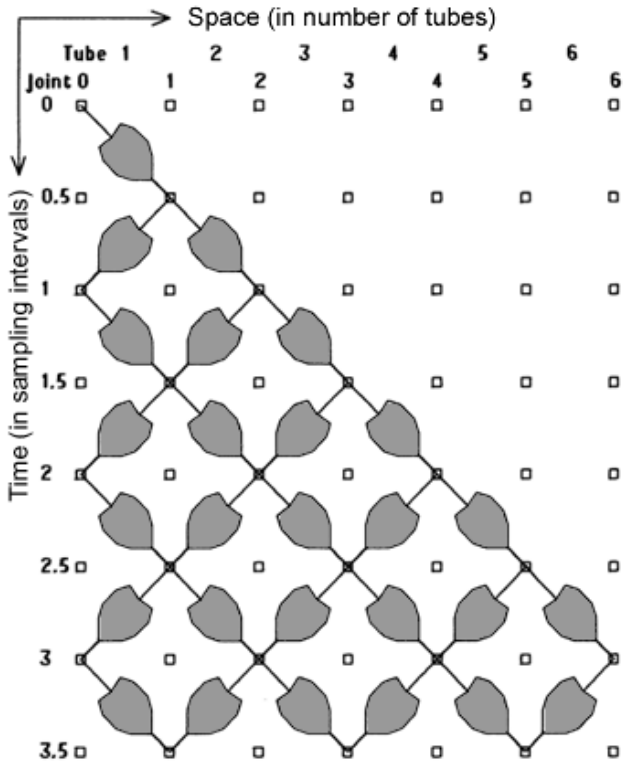


Fig. 3.3 Time-space representation of the wave reflection analog technique (adapted from Liljencrants [26])

3.3 Radiation Impedance

When acoustic waves traveling through the vocal tract reach the terminal section, (mouth opening), reflections back into the tract occur. This is due to the radiation impedance imposed by the air in free space in addition to the tract configuration. With respect to the scattering equations, it is convenient to estimate the reflected and radiated pressures using only the forward acoustic pressure in the terminal section as the input. Since radiation occurs only in the last section, no losses are considered for this tube section. This simplification makes it possible to develop scattering equations for the radiation using equations (3.13) and (3.14). Compared with the propagation inside the tube described by these equations, two differences at the mouth opening can be noted. First, since no acoustic pressure is considered to come from the outside into the vocal tract, no backward acoustic pressure is required in the radiation at the mouth opening. Secondly, the reflection coefficient needs to be expressed in terms of a radiation impedance instead of area relations. The radiation impedance of the mouth opening was originally modeled by Fant [10] as a circular piston in a sphere, the sphere representing the effect of the head surface. Flanagan showed [12] that a simpler expression can be used with similar results, that have been widely utilized for speech synthesis. The model consists of a radiation impedance load using a resistance in parallel with an inertance, which is equivalent to that of a circular piston in an infinite baffle:

$$\hat{Z}_{rad}(\omega) \approx \frac{j\omega RL}{R + j\omega L} \quad , \tag{3.16}$$

where Z_{rad} is the radiation impedance as a function of frequency, R is the equivalent resistance given by

$$R = \frac{128\, Z_M}{9\pi^2} \quad , \tag{3.17}$$

and L is the equivalent inertance given by

$$L = \frac{8a\, Z_M}{3\pi c} \quad , \tag{3.18}$$

where a is the radius of the piston, and Z_M is the impedance of the terminal tube section given by equation (3.8). One can link the expression for the radiation impedance with the reflection coefficient given by equation (3.12). Using equation (3.8), the reflection coefficient can be rewritten as,

$$r_1 = \frac{A_1 - A_2}{A_1 + A_2} = \frac{Z_2 - Z_1}{Z_2 + Z_1} \tag{3.19}$$

The terminal reflection coefficient, r_M , may be expressed in terms of the impedances Z_M and Z_{rad} using

$$\hat{r}_M = \frac{\hat{Z}_{rad} - Z_M}{\hat{Z}_{rad} + Z_M} \quad . \tag{3.20}$$

In order to make the relation between the reflection coefficient and the radiation impedance compatible, the latter must be transformed from the frequency domain into the time domain. An efficient approach is to use first the bilinear transformation and make use of the z-domain as an intermediate step [41]. The inverse z-transform then yields the desired time domain expression. The derivation is presented here since it can be useful for future developments of this technique. Only the final expression is used in the algorithm. The bilinear transformation

$$j\omega = 2F_s \frac{(1 - z^{-1})}{(1 + z^{-1})} \quad , \tag{3.21}$$

in which F_s is the sampling frequency and z^{-1} is a unit delay in the z-domain is used. Substituting equation (3.21) into (3.16), an expression for Z_{rad} in the z-domain is obtained. Using this in equation (3.20), the equivalent expression for r_M is simplified to obtain

$$r_M = \frac{a_2 + z^{-1} a_1}{b_2 - z^{-1} b_1} \quad , \tag{3.22}$$

where the coefficients are given by

$$a_1 = -R^* + L^* - R^* L^* \quad , \tag{3.23}$$

$$a_2 = -R^* - L^* + R^* L^* \quad , \tag{3.24}$$

$$b_1 = -R^* + L^* + R^* L^* \quad , \tag{3.25}$$

$$b_2 = R^* + L^* + R^* L^* \quad , \tag{3.26}$$

and the modified resistance and inertance are

$$R^* = \frac{128}{9\pi^2} \quad \text{, and} \tag{3.27}$$

$$L^* = 2Fs\frac{8a}{3\pi c} \tag{3.28}$$

As noted previously, the basic scattering equations (3.13) and (3.14) can be used to compute the backward and the radiated pressure. Since no acoustic pressure is coming from the outside into the vocal tract, no backward acoustic pressure (B_2) is considered. Thus, the radiation scattering equations are

$$B_1 = F_1 r_M \quad \text{, and} \tag{3.29}$$

$$F_2 = \left(F_1 + 1\right) r_M \quad . \tag{3.30}$$

Finally, by substituting equation (3.22) into equations (3.29) and (3.30), and by computing the inverse z-transform for each expression, the time domain formulas for the reflected (B_1) and radiated (F_2) acoustic pressures due to the radiation impedance are obtained. These are

$$B_1[n] = \frac{1}{b_2}\left(a_2 F_1[n] + a_1 F_1[n-1] + b_1 B_1[n-1]\right) \quad \text{, and} \tag{3.31}$$

$$F_2[n] = \frac{1}{b_2}\left((a_2 + b_2) F_1[n] + (a_1 - b_1) F_1[n-1] + b_1 F_2[n-1]\right) \quad . \tag{3.32}$$

Note that theses expressions reveal the inertive and resistive components of the radiation impedance through time delays. The stencil involves only two time steps, which makes the scheme efficient. The algorithm to compute reflected and transmitted acoustic pressures in presence of a termination impedance can be used not only for the mouth and nose openings, but also in the subglottal tract if a “lung opening” is considered. A similar

procedure may be used for different impedance expressions. Story [41] and Liljencrants [26] have discussed radiation from the nose, the neck and cheeks. The incorporation of these factors may be the object of future work.

3.4 Yielding Walls, Viscous Effects and Heat Conduction Losses

Yielding walls, viscous effects and heat conduction losses are the main sources of sound energy dissipation within the tubes. Yielding wall losses refer to the energy that is extracted from the acoustic waves to excite vibrations of the tracts walls. Heat conduction losses are associated with temperature gradients within the acoustic boundary layers resulting in heat conducted into the wall. Viscous losses refer to friction-generated heat in the fluid boundary layer adjacent to the tube wall. The main effects of these losses are in the bandwidth and peak values of the formants. A high damping results in larger formant bandwidths and reduced formant amplitudes.

In [26], Liljencrants designed a scheme to account for these losses that modified the scattering equations in the wave analog. As discussed by Story [41], part of this scheme has proven to be unsuccessful. Story modified Liljencrants' original version by adding a lumped element circuit to the scattering equations. All reactive terms in the circuit are associated with delays in the time domain. Buffers are required to achieve this goal. The scattering equations become complex in comparison with the lossless case. As a result, this derivation affects not only the bandwidth of the formants but also the frequency of the first formant, raising it by about 10% (30-40 Hz). Although Story's loss factor appears to work in the right direction, its implementation is rather problematic. Having an overall effect over all frequencies, a simplified approach for the losses in a wave reflection analog scheme was developed by Rahim [33], [34]. This expression was based on previous work by Flanagan [12], where heat conduction losses were included as an attenuation factor that is a function of length. Rahim added yielding walls and viscous effects in the attenuation factor. One coefficient is then sufficient to account for losses in the original lossless scattering equations. This propagation loss factor is inserted at each section of the tracts to attenuate the forward and backward traveling waves in a wave reflection analog scheme. This approach was selected here due to its simplicity and acceptable accuracy. The attenuation factor, α_k , is defined as

$$\alpha_k = \frac{R_v A_k}{2\rho c} + \frac{\rho c}{2A_k}(G_t + G_w) \tag{3.33}$$

where R_v , G_t and G_w are the contribution to the losses due to viscous friction, heat conduction and yielding walls vibrations, A_k is the cross-sectional area of the k-th section, ρ is the air density, and c is the speed of sound. In [12], Flanagan derives numerical expressions for R_v , G_t and G_w, which are presented in the following equations, using cgs units.

$$R_v = 2.89 \text{ x } 10^{-3} \frac{\zeta f^{1/2}}{A_k^{3/2}} \tag{3.34}$$

$$G_t = 8.07 \text{ x } 10^{-7} \zeta f^{1/2} A_k^{1/2} \tag{3.35}$$

$$G_w = \frac{2 r_s \zeta \sqrt{\pi A_k}}{r_s^2 + x_s^2} \tag{3.36}$$

In these equations, ζ is a shape factor of the sectional area, f is frequency, r_s and x_s are the resistance and reactance for the wall vibration losses. Values for theses variables are given in Table 3.1 for different cases.

Table 3.1
Suggested resistance and reactance values for yielding wall losses from [34]

Tract Section	r_s **(cgs Ω)**	$\frac{x_s}{\omega}$ **(cgs Ω/cps)**
Relaxed cheek	800	2.1
Tensed cheek	1060	1.5
General	1600	1.5
Neck	2320	2.4
Stomach surface – Subglottal tract	6500	0.4
Vocal Tract	1000	2.0

Using equations (3.33) through (3.36) and the values in Table 3.1 for the vocal tract, Rahim simplified the attenuation factor α_k to obtain

$$\alpha_k = \zeta(5.23 \text{ x } 10^{-5} f^{\frac{1}{2}} + 439 f^{-2}) A_k^{-\frac{1}{2}} \tag{3.37}$$

The attenuation factor is plotted as a function of frequency in Figure 3.4 for a shape factor ζ=1.08, a typical tract shape, and a unit area. As shown in Figure 3.4, α_k is nearly constant at a value near 0.004 per [cm] above 400 Hz. In [34], Rahim suggests that it is acceptable to assume an average effect over frequency, selecting then a choice of this factor that provides a reasonable compromise over the frequency for the supraglottal tract. This value as a function of the area is

$$\alpha_{k(\text{supraglottal})} \approx 3.8 \text{ x } 10^{-3} A_k^{-\frac{1}{2}} \quad . \qquad (3.38)$$

Although all the tools were available, no suggestions for the losses in the subglottal tract are present in [34]. Assuming that viscous friction and heat conduction are the same compared with the vocal tract, the only difference is observed in the contribution of the yielding walls to the loss factor. Following the same procedure, equations (3.33) through (3.36) and the values suggested in Table 3.1 for the subglottal tract case, yield the attenuation factor

$$\alpha_k = \zeta(5.23 \text{ x } 10^{-5} f^{\frac{1}{2}} + \frac{23041}{6500^2 + 6.31 f^2}) A_k^{-\frac{1}{2}} \qquad (3.39)$$

The values for wall reactance and resistance used to compute this expression have been obtained from [12] and [15]. Figure 3.4 shows this function of frequency, using the same unit area and shape factor as before (ζ=1.08). This assumption is important, since studies performed by Dang [7] have shown that the shape factor could take values of up to ζ =3 for other tract shapes, as in the case for the nasal tract. Since no special information on the shape factor for the subglottal tract has been considered, it is assumed the same as in the vocal tract. Figure 3.4 illustrates the differences in the losses between the two tracts. Although the supraglottal tract attenuation peaks at low frequency, it decreases rapidly with frequency. The subglottal tract, on the other hand, has significant attenuation over the entire frequency range. Assuming that the main effects of the subglottal are observed below 3000 Hz, an average value for this attenuation factor per cm was selected as

$$\alpha_{k(\text{subglottal})} \approx 11.2 \text{ x } 10^{-3} A_k^{-\frac{1}{2}} \quad . \qquad (3.40)$$

Now, defining γ_k as the propagation loss factor for the k-th section as an exponential decay, this can be approximated as

$$\gamma_k = e^{-\alpha l} \approx 1 - \alpha l \quad , \tag{3.41}$$

where l is the length of each tract section defined in (3.15) and α_k the coefficients defined in equations (3.38) and (3.40). Finally, this loss factor is combined with the scattering equations given in equations (3.13) and (3.14) simply by multiplying it with the incident waves at the junction prior to compute the scattering equation, then

$$F_1 = \gamma_k F_1 \tag{3.42}$$
$$B_2 = \gamma_k B_2 \tag{3.43}$$

Note that the difference between the propagation loss factors in the subglottal and supraglottal tract is a function of the distance and the area of the section. The best approach to compare them is through the selected attenuation factors, where the subglottal tract value is larger by a roughly a factor of 3.

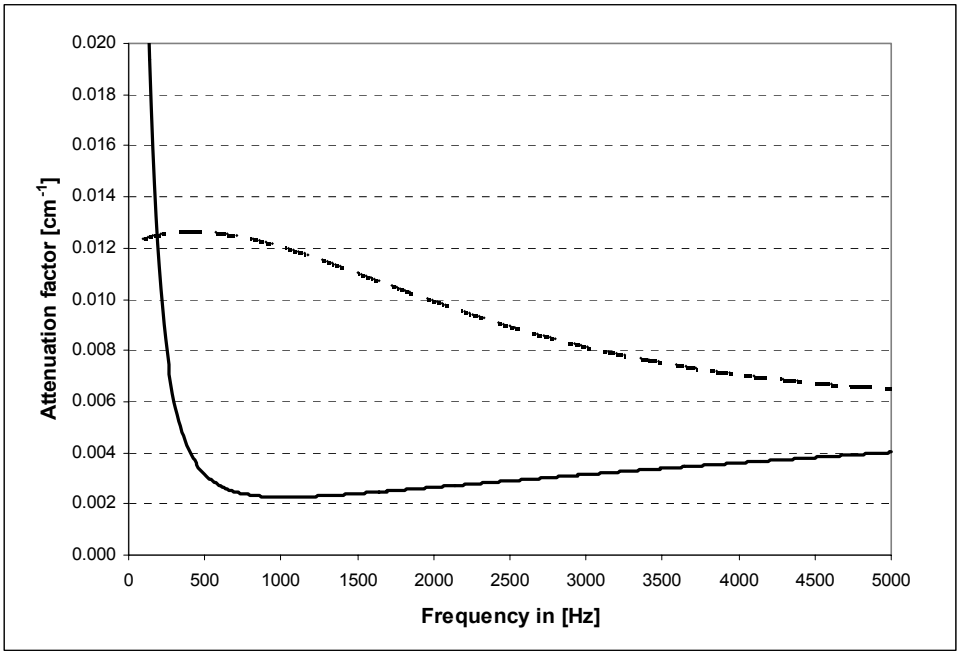


Fig. 3.4 Attenuation factor for the sub and supra glottal tracts.
-- : subglottal tract, —: supraglottal tract.

3.5 Connection with the Voice Source

The wave reflection analog scheme requires a sound source model at the glottis location. The source can be either an ideal generator, which means that no interaction is allowed between the source and tracts, or an interactive model, where the reflected acoustic pressures affect the behavior of the source. The latter correspond to the source model developed in Chapter 4. In either case, it is convenient to express the source as a volumetric flow source. The connection with the wave reflection scheme is given by

$$F_{\text{supra}} = B_{\text{supra}} + \left(\frac{\rho c}{A_{\text{supra}}} \right) Q \quad \text{, and} \tag{3.44}$$

$$B_{\text{sub}} = F_{\text{sub}} - \left(\frac{\rho c}{A_{\text{sub}}} \right) Q \quad , \tag{3.45}$$

where F_{supra}, F_{sub}, B_{supra} and B_{sub} represent the forward and backward acoustic pressures in the supra and subglottal tracts, A_{supra} and A_{sub} are the areas of the first sections of the supra and subglottal tracts, ρ is the air density and c the speed of sound. This scheme is illustrated in Figure 3.5.

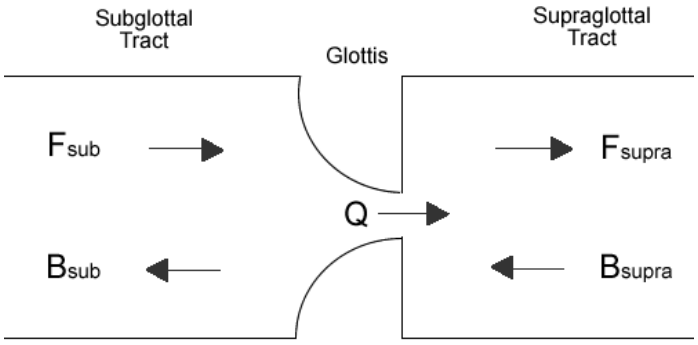


Fig. 3.5 Incident and reflected acoustic pressures due to the connection with the voice source

3.6 Acoustic Interaction between Subglottal and Supraglottal Tracts

Acoustic coupling between supra and subglottal tracts has been discussed in [39] and [15], where it is stated that such coupling exists only when the glottis is open. For normal vibrations of the folds, coupling is observed only during the opening phase of the cycle. In some voiceless sounds, this coupling is continuous and more evident since the glottis remains open. Stevens [39] shows that the transfer function of coupled sub- and supra-glottal tracts is roughly the same as adding the independent transfer functions obtained with a closed glottis. More specifically, a set of poles (resonances) and zeros (anti-resonances) that are close to each other, are added to the transfer function of the supraglottal tract. These poles and zeros are located around the resonance frequencies of the subglottal tract, in a closed glottis configuration. Since some degree of cancellation between poles and zeros occurs, the global effect can be simply seen as the addition of small resonances in the subglottal resonant frequencies [39]. This effect has been also noticed by Harper [15]. These studies have used a transmission line approach to model this phenomenon. The problem is now to identify how this coupling works in the wave analog scheme.

In [41], Story followed the guidelines described by Titze [47], where the interaction between tracts is given by the volume velocity source, where volumetric flow rate is computed directly based on reflected and incident acoustic pressures. Note that this approach uses the non-linear acoustic theory for orifices described by Ingard [18]. This author demonstrated that for high acoustic pressures and small orifices, the relation between acoustic pressure and particle velocity (hence the impedance) is quadratic, instead of linear. Ingard states that the relation between these two parameter is better described using a Bernoulli flow equation in this case. As detailed in chapter 4, a Bernoulli flow equation is used to compute the volumetric flow rate at the glottal opening. Therefore, the connection between the tracts and the source described in equations (3.44) and (3.45) already includes a nonlinear coupling between the two tracts.

Using a different approach, Liljencrants [26] described the glottis as another tube section in his wave analog approach. Traveling waves in the glottal area are required to obtain the transmitted and reflected waves in the supra- and sub-glottal tracts. This introduces a delay when pressure waves are transmitted between tracts. Note that in the nonlinear approach, it is assumed that the thickness of the folds (length in flow direction) does not produce a significant delay in the waves, so the glottal source is considered as a

junction instead of a tube. In general, Liljencrants' technique defines a linear coupling, with none of the effects mentioned by Ingard.

Two challenges relative to the acoustic coupling emerge at this point. First, it is of interest to test if the wave analog scheme exhibits the coupling effects observed by previous authors. Secondly, a comparison between the linear and nonlinear approaches should be established to clarify their differences. These questions are answered through case studies performed and discussed in chapter 5. This helps to better understand the acoustic coupling in this model of voice production. Note that by using an interactive source model and acoustic coupling between the subglottal and supraglottal tracts, the system reflects more accurately the high levels of interaction that occur in phonation.

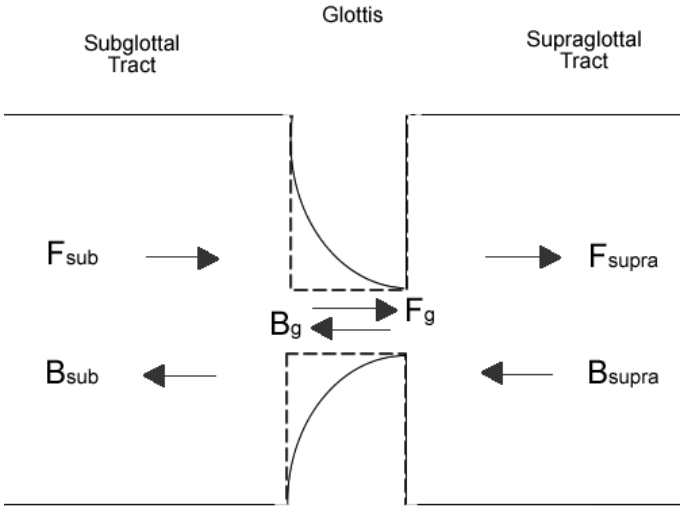


Fig. 3.6 Linear acoustic coupling between subglottal and supraglottal tracts.

3.7 Subglottal and Supraglottal Tracts Design

The tracts are designed as a collection of cylindrical tubes. A particular tract shape is imposed by specifying only an array of cross-sectional areas for each section called the "area function". For both tracts, it is assumed that the first term of the array corresponds to the section next to the glottis, and the last section corresponds either to one next to the mouth or lung openings. For a complete system design, including subglottal and supraglottal tracts, the glottis is the connection between the two.

The area functions of the supraglottal tract used in this research were taken from the vocal tract imaging study performed by Story and Hoffman [43]. A set of 18 vocal tract shapes (3-D) for one male subject, who vocalized while being scanned for 12 vowels, 3 nasals and 3 plosives, was collected. The 3-D images were acquired using magnetic resonance imaging (MRI). The images were analyzed to find the area functions within planes chosen to be perpendicular to the centerline of the tract. Noteworthy limitations are due to the nature of the MRI approach. The acquisition time is slow, so no motion of the articulatory system can be observed. In addition, neither the vocal folds nor the subglottal tract are seen very clearly. Although the approach has limitations, the database was shown to be useful in speech synthesis and in other studies of acoustic phonetics. In [41] these area functions were supplied into an articulatory synthesizer using a wave reflection analog approach, showing reasonable similarity compared with the natural speech of the imaged subject. Further details of that study as well as the resulting area function database can be found in [43].

For the subglottal area functions, two designs were tested. The first one is an MRI shape from [41], which was obtained using the imaging procedure described before. Only one section of the subglottal tract was imaged, from the glottis to the carina. Due to the difficulties to obtain a high-quality image in this case, an interpolation was suggested to reduce irregularities in the area function. Note that since the tract is not complete, an equivalent radiation impedance is required. Since no expression for this subglottal impedance is available, Flanagan's radiation impedance described before for the mouth opening was incorporated. Modifications to obtain a more realistic shape are discussed in chapter 5. The second approach used to define the area function of the subglottal tract was based on Weibel [54]. The total cross-sectional area of the airways as a function of the distance from the larynx was measured using silicone rubber casts and excised adult lungs. Using only this data, an equivalent area function with no tube branching is defined, as showed in Figure 3.7. Note that since the equivalent cross-section at the end of the tract is significantly larger than other tract sections, the radiation impedance does not affect propagation inside the duct. Most of the reflected waves are generated upstream of this opening. Therefore, no radiation term is included in this design.

Although tube branching can be included using a wave reflection analog approach, none has been included here for simplicity. A more detailed study of the acoustics of the respiratory tract by Harper [15] has illustrated the importance of branching and yielding walls features in the subglottal tract. Since the model used in this thesis has a simplified version of these features (i.e. a global loss factor for all

frequencies), some deviations are expected. To compensate for this error, some corrections in the original tract design are suggested. These corrections are defined and discussed in Chapter 5. Proper improvements are to be included in further development of this model.

In addition to the tract designs explained above, simplistic shapes have also been included. Tract shapes such as simple tubes, infinitely long tubes, two tube versions of the vocal tract and Helmholtz resonators are useful to test the general wave reflection analog algorithm and their effect in laryngeal self-oscillations.

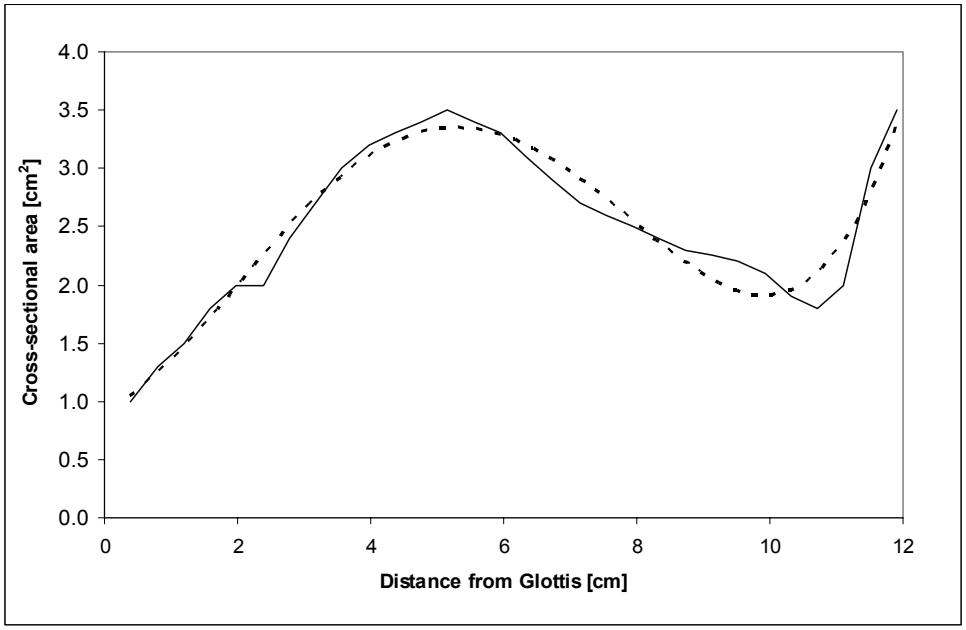


Fig. 3.7 Subglottal tract from Story [41] and polynomial regression.
-- : MRI data, —: Polynomial regression .

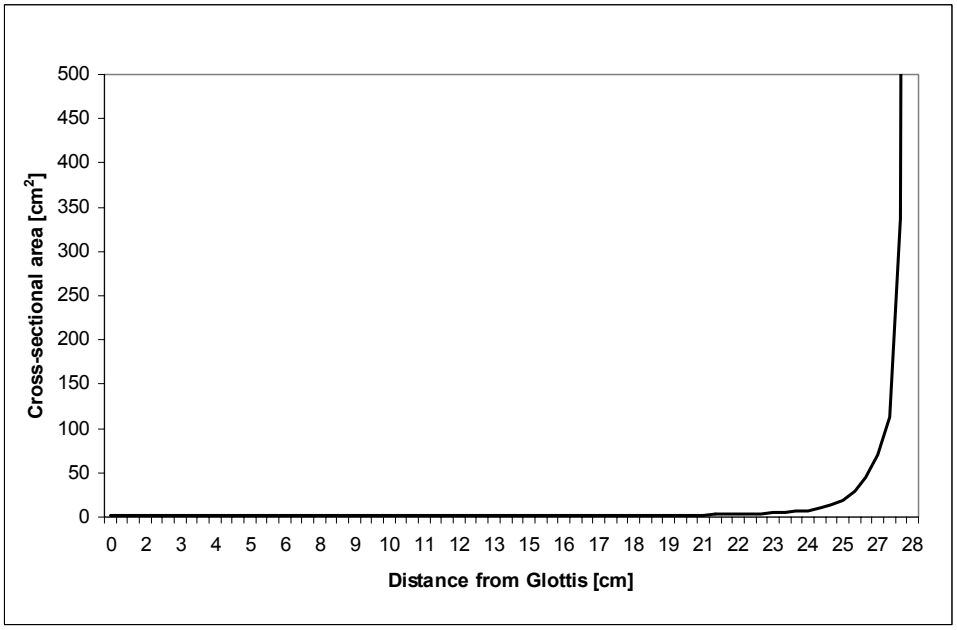


Fig. 3.8 Subglottal tract from Weibel [54].

4. EFFECTIVE SINGLE DEGREE OF FREEDOM SELF OSCILLATING MODEL

This chapter describes the source model. A simplified model of the vocal folds to test the wave analog scheme described in chapter 3 was designed. The single degree-of-freedom model achieved self-sustained oscillations without acoustic loading. Preliminary results obtained using the source model are discussed in chapter 6.

4.1 Preliminary Work

The source model developed in this chapter is an updated version of the "effective one-mass model" of the vocal folds developed by Fulcher *et al.* [14]. In this work, a one-mass model of the vocal folds is excited by a negative Coulomb damping force. This effect is assumed equivalent to the action of aerodynamic forces on the vocal folds. This one single degree of freedom model performed self-sustained oscillations without the need for other forces. As discussed previously, several authors have discussed the unability of one-mass models to self-oscillate unless other factors are included, such as acoustic loading. Due to its reduced order, Fulcher's one-mass model serves as starting point of the model developed in this thesis.

Using the fact that forces in phase with velocity are favorable to phonation [48], the negative Coulomb damping is applied as a fixed switching force that depends on the velocity of the folds. The switch is associated with a change in the glottal shape that keeps aerodynamic forces acting on the folds always in phase with velocity. Using this idea, the equation of motion for this oscillator is given by

$$m\ddot{y} + b\dot{y} + ky = F_p \quad , \tag{4.1}$$

where

$$F_p = \begin{cases} F_o, & \dot{y} > o \\ -\dfrac{F_o}{2}, & \dot{y} < o \end{cases} \tag{4.2}$$

The "dot" notation used in equation (4.1) and for the rest of this document is the differentiation operator. In addition, m is the mass of one fold, y is the displacement, b is the viscous damping constant, and k is the spring constant. The magnitude of the force (F_o) is given by the lung pressure multiplied by the area of the medial surface of the glottis. The mean pressure in the vocal tract is atmospheric. Typical results for this model are shown in chapter 6.

Important simplifications should be noted. The driving force was manufactured to resemble pressure forces that have been modeled or measured in previous work [37]. A weaker pressure force is expected for the inward, closing phase portion of the cycle than for the outward opening phase. This reasoning led to the assumption that there was a factor two difference between aerodynamic wall pressures during these two stages. Obviously, this is a gross oversimplification that does not accurately represent the aerodynamic forces. It ignores geometry of the vocal folds and coupling issues. In addition, the force magnitude remains constant over each half-cycle. In a more realistic scenario, the force that drives the folds varies continuously in magnitude over one cycle. Finally, Fulcher's model does not include collision effects, although this is the object of an ongoing work.

4.2 Basic Oscillation Mechanism

Fulcher's one-mass model was modified by defining the driving force associated with the negative Coulomb damping based on the aerodynamic forces acting on the vocal folds. Flow-structure interactions, flow-sound interactions, and collision effects were included in this modified version of the model. The switching property was included using a time-varying orifice discharge coefficient that modifies the flow equations and driving forces. The variation of this parameter due to glottal shape changes has been discussed and measured [32], [58] ,[37], and [8]. Experimental data for discharge coefficients during the closing phase (diverging glottal shape) and the opening phase (converging glottal shape) is available. A smooth transition of the aerodynamic force is

enforced, in place of a step function. An impact force is also considered during collision between vocal folds. These modifications make the model more realistic.

Figure 4.1 shows a schematic of the vocal folds model. It is assumed that the movement of the folds is bilaterally symmetric, with identical displacements. For each fold, the equation of motion is given by

$$m\ddot{y} + b\dot{y} + k(y - y_o) = F_p + F_H \quad , \tag{4.3}$$

where y is the displacement of the mass at the glottal end, y_o is the equilibrium position of the mass without the action of external forces, b is the viscous damping constant , k is the spring constant. The expressions for F_p and F_H are derived in section 4.3 and 4.4.

4.3 Pressure Force

To find an expression for the pressure force, F_p, only half of the symmetric system is considered. The values of the pressure force acting on one fold are assumed to be the same as for the other half of the system. Figure 4.2 shows details of the orifice cross-section and the nomenclature.

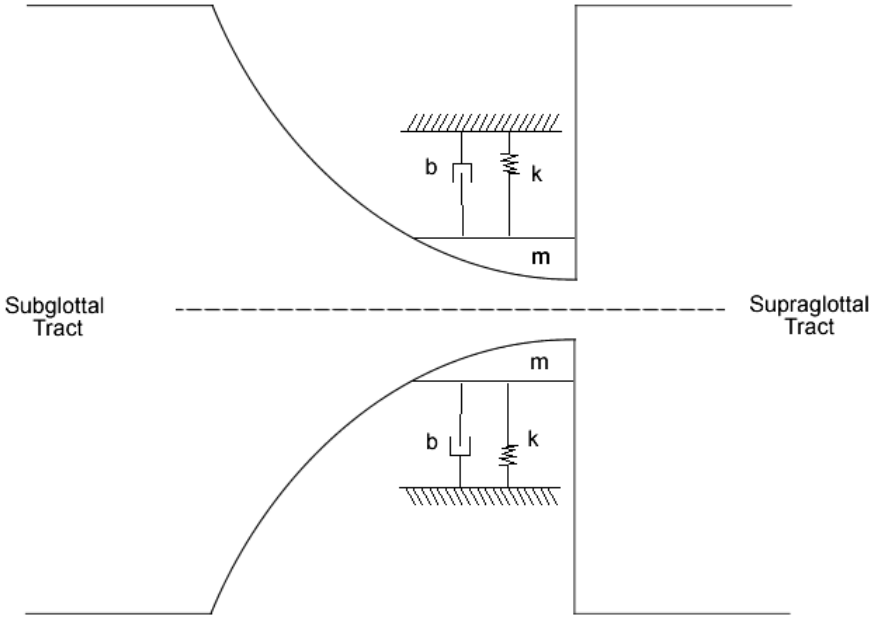


Fig. 4.1 Basic diagram of the single-degree-of-freedom, self-oscillating model

Note that the values for A_s and A_e , as illustrated in Figure 4.2, correspond to half the values of the subglottal and supraglottal areas.

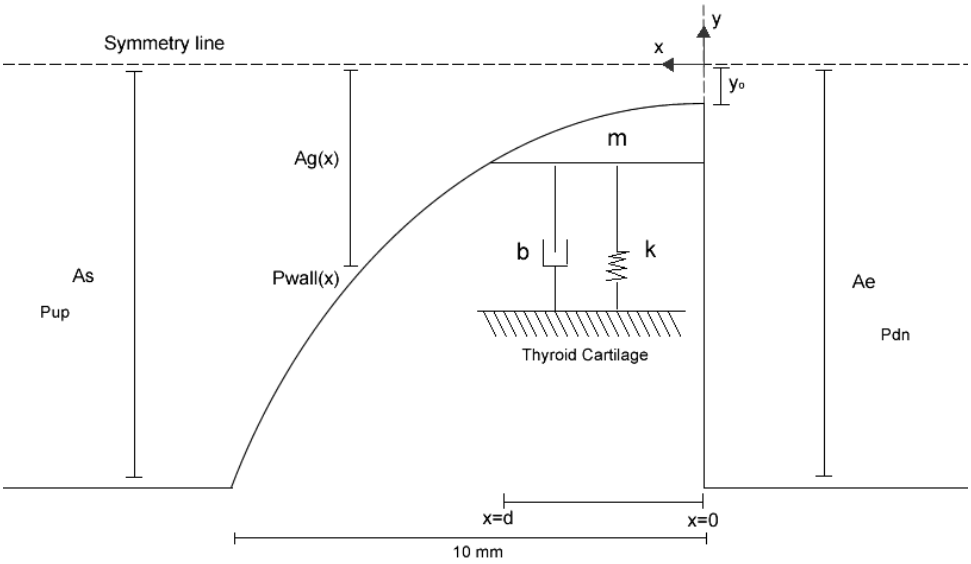


Fig. 4.2 Half of the symmetric representation of the model. Details of the areas and pressures considered to compute the pressure force.

The flow is assumed to be inviscid, irotational, and approximately incompressible. The "quasi-steady approximation" is made, i.e. the time-varying flow is treated as a sequence of stationary flows for similar geometries and boundary conditions at each time step. Under these conditions, Bernoulli's equation is valid along a streamline in the region from the subglottal to the glottal opening, where the flow is assumed to detach from the wall. Aerodynamic effects after/at the glottal opening (pressure recovery, dynamic head loss, momentum diffusion, etc) are included through the use of an orifice discharge coefficient. Note also that the wave reflection analog supplies the acoustic pressures (subglottal and supraglottal) that are used in the calculations. From conservation of mass, the relation between subglottal and glottal flow velocities is

$$A_s u_{up} = A_g(x) u_{cl}(x) \quad . \tag{4.4}$$

The function that describes the glottal area $A_g(x)$ along the depth of the folds (or length in the flow direction) is assumed to be quadratic. This area is obtained using a shape function $f(x)$, such that $A_g(x)$ is equal to $f(x)$ times the length of the folds. The

effects of the geometry are discussed in section 4.2.1. To relate pressure and velocity, Bernoulli's equation is used along the centerline (or symmetry line) between the subglottal section ($x=d$ in Figure 4.2) and the opening of the vocal folds ($x=0$ in Figure 4.2). Assuming the total head at $x=d$ equals the subglottal head, and that the centerline pressure at $x=0$ is equal to the supraglottal value, then

$$\frac{1}{\rho}p_{cl}(d)+\frac{1}{2}u_{cl}^2(d)=\frac{1}{\rho}p_{cl}(0)+\frac{1}{2}u_{cl}^2(0) \quad , \tag{4.5a}$$

$$\frac{1}{\rho}(p_{up}+p_{up}^{'})+\frac{1}{2}u_{up}^2=\frac{1}{\rho}(p_{dn}+p_{dn}^{'})+\frac{1}{2}u_{cl}^2(0) \quad . \tag{4.5b}$$

Using equation (4.4) to relate centerline velocities between the opening and the subglottal section, and defining β as the area ratio, this relation is expressed as

$$u_{cl}=u_{cl}(0)=u_{up}\frac{A_s}{A_g}=u_{dn}\beta \quad . \tag{4.6}$$

Using equation (4.6) in (4.5b) yields the ideal velocity at the glottal opening

$$u_{cl(ideal)}=\sqrt{\frac{2\left(\Delta p+(p_{up}^{'}-p_{dn}^{'})\right)}{\rho(1-\beta^2)}} \quad . \tag{4.7}$$

Since the volumetric flow rate Q must be constant across the system, it can be computed at the glottal opening as

$$Q_{ideal}=u_{cl}A_g \quad . \tag{4.8}$$

Note that all equations up to this point correspond to an ideal flow case. It is important to introduce aerodynamic effect that happens after/at the glottal opening, such as pressure recovery, dynamic head loss, momentum diffusion and so on. Using the definition of Park *et al.* in [32], the volumetric flow rate can be modified through the use of an empirical orifice discharge coefficient; the actual volumetric flow rate is then obtained from

$$Q = c_d(t) u_{cl} A_g \quad , \tag{4.9}$$

where $c_d(t)$ is the time orifice varying discharge coefficient (ODC). This coefficient varies during the oscillation cycle. It constitutes a key factor to obtain self-sustained oscillation. Details of this coefficient are discussed in section 4.2.2. Using equation (4.9) the flow velocity at the opening from equation (4.7) is

$$u_{cl} = c_d(t) \sqrt{\frac{2\left(\Delta p + (p_{up}^{'} - p_{dn}^{'})\right)}{\rho(1-\beta^2)}} \quad . \tag{4.10}$$

Using this modified version of the flow velocity, downstream values are also corrected to obtain

$$u_{up} = u_{cl} \frac{A_g}{A_s} \quad \text{, and} \tag{4.11a}$$

$$u_{cl}(x) = u_{cl} \frac{A_g}{A_g(x)} \tag{4.11b}$$

Using Bernoulli's flow equation again between the subglottal section and any point in the glottis along the centerline, the centerline pressure as a function of the position in the glottis is computed as

$$p_{cl}(x) = p_{up} + p_{up}^{'} + \frac{\rho}{2}(u_{up}^2 - u_{cl}^2(x)) \quad . \tag{4.12}$$

An approximate pressure force value may be obtained by integrating this function along the surface of the vocal folds. A more accurate value for the force can be obtained without adding much complexity. The axial flow velocity is approximately constant along *x*, which means that the x-component of the flow velocity near the wall is the same as along the centerline, as illustrated in Figure 4.3.

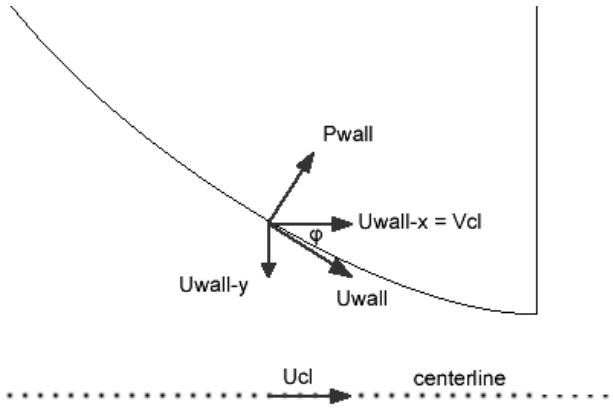


Fig. 4.3 Pressure and velocity along the wall of the vocal folds.

The relation between centerline velocity and the wall velocity can be obtained geometrically from Figure 4.3, where

$$u_{cl}(x) = u_{wall}(x)\cos\varphi \quad , \quad \text{where} \tag{4.13}$$

$$\tan\varphi = \frac{d}{dx}(f(x)) = f'(x) \quad , \quad \text{then} \tag{4.14}$$

$$u_{wall}(x) = \frac{u_{cl}(x)}{\cos(\tan^{-1}(f'(x)))} \quad . \tag{4.15}$$

Using Bernoulli's flow equation again for any pair of points in the glottis, the pressure along the wall is obtained as

$$p_{wall}(x) = p_{up} + p'_{up} + \frac{\rho}{2}(u_{up}^2 - u_{wall}^2(x)) \quad . \tag{4.16}$$

Finally, the pressure force acting on the folds in equation (4.3) is the integral of this function along the surface of the folds. To obtain the total force that it is applied along the surface it is necessary to consider both depth and length of the fold. Therefore

$$F_p = L\int_0^d p_{wall}(x)dx \tag{4.17}$$

4.3.1 Geometrical Considerations

As expected, the geometry of the folds assumed for the vocal fold affects the force that is applied to drive self-sustained oscillation. It can be observed from Figure 4.3 and equation (4.13) that the flow velocity at the wall is always larger than the centerline velocity. In addition, equation (4.4) shows that the glottal area influences the centerline velocity, and therefore the wall pressure force.

The aim of the one-mass model is to keep the design as simple as possible, in order to promote efficiency over complexity. An appropriate balance between the two factors is required to ensure the accuracy of the model. Complex geometries like M5 or others used in different models of the vocal folds must be appropriately simplified to meet the design goals.

Flanagan [12] used a linear geometry to describe the surface that is in direct contact with the flow for the one-mass model. The main problem with that geometry is a sharp discontinuity at the entrance of the glottis from the subglottal section. This leads to problems in the flow behavior (flow displacement, head loss, etc) that his model neglects. To keep the aerodynamic equations consistent, a geometry that reaches the subglottal area without discontinuities is preferred. A better flow profile is obtained by including a rigid structure that serves as a transition element to reach the folds. This structure has been included in other models of the folds [42] [28] [27], and it is embodied in the M5 shape. However, since the model has only one degree of freedom, minor discontinuities are observed when the system oscillates. This produces fewer problems in the flow behavior than in Flanagan's one-mass model. The discontinuity is given by the magnitude of the displacement, which has peak values normally less than 0.5 mm. It is better to account for the error produced by this discontinuity than to modify the fold geometry to keep the junction with the subglottal tract fixed. In that case, when the fold vibrates, the vocal fold would squeeze to keep the juncture fixed. This is not desired since it violates the single degree-of-freedom assumption.

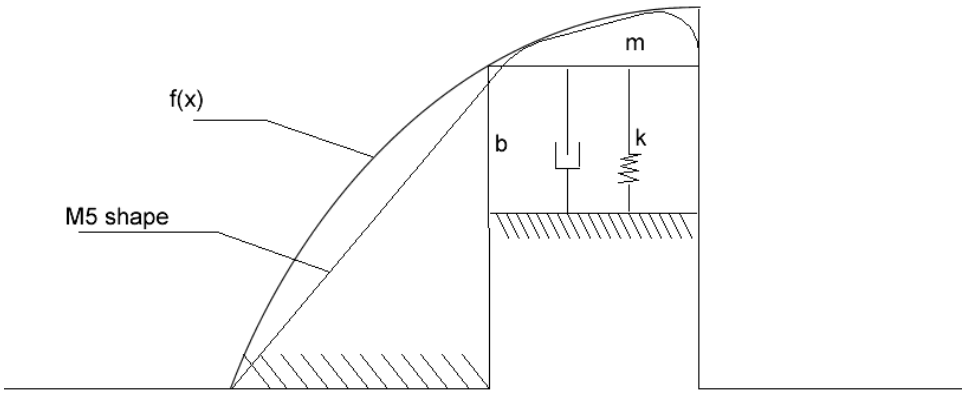


Fig. 4.4 Comparison between the glottis profile and Scherer's M5 model geometry.

Shown in Figure 4.4, a quadratic geometry was selected. The height of the vocal fold surface with respect to the fixed wall is

$$f(x) = \frac{A_s}{L}\left(\frac{x}{d}\right)^2 + y \quad \text{, and} \tag{4.18a}$$

$$A(x) = L \cdot f(x) = A_s\left(\frac{x}{d}\right)^2 + yd \quad , \tag{4.18b}$$

where the values for the length of the folds L, the depth d and the subglottal area A_s are defined later in this chapter.

Figure 4.5 illustrates three different instants during the oscillation of the vocal fold model. These instants are: where the vocal fold is initiating (or finishing) collision ($y=0$), where the vocal fold is in a fully opened state ($y=y_{max}$), and where the vocal fold is in fully closed state ($y=-y_{max\ collision}$). Note that the displacement of the mass affects the complete structure, meeting the one-degree of freedom assumption. The discontinuity at $x=d$ can be also observed.

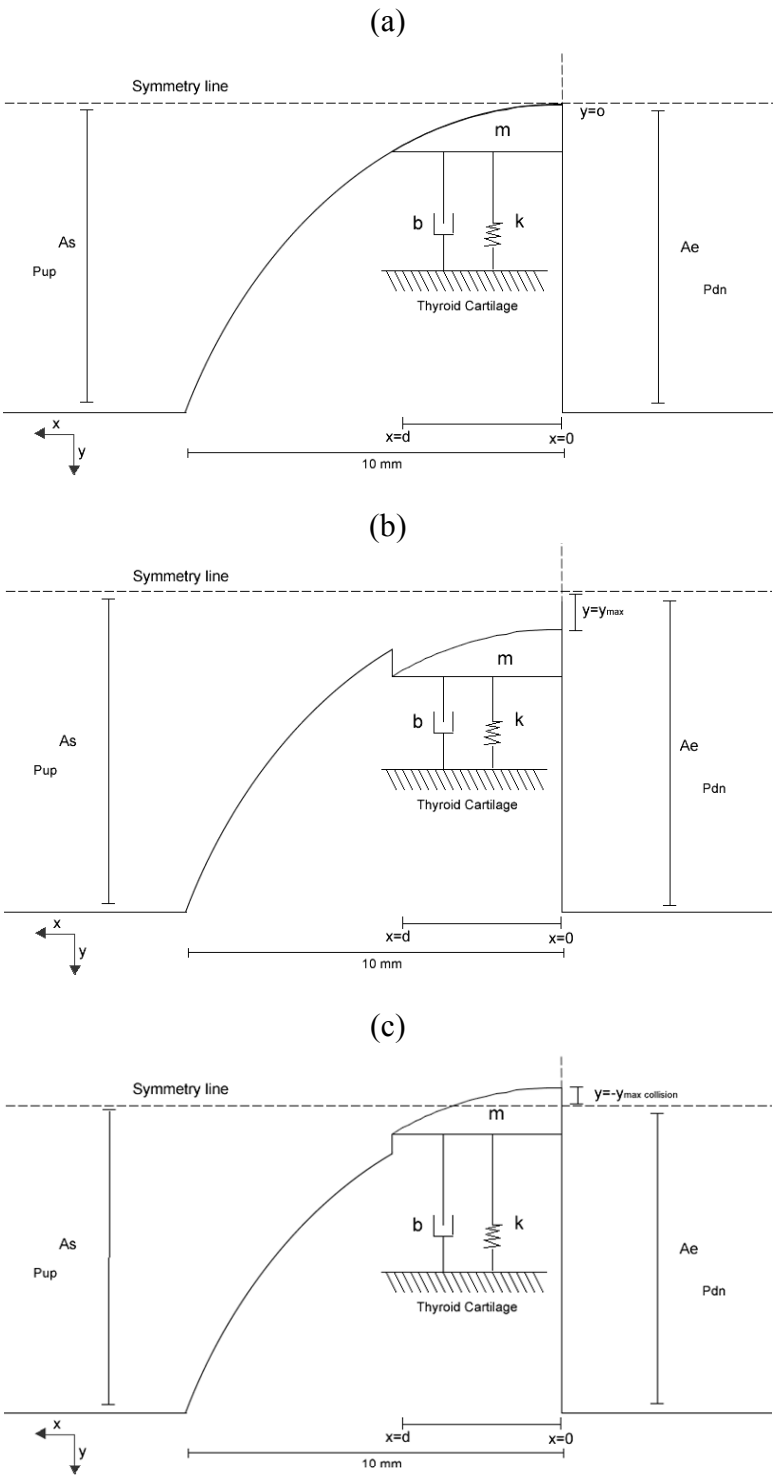


Fig. 4.5 Schematic of the model at three instants during one oscillation period (a): at $y=0$ (b): at $y=y_{max}$ (c): at $y=-y_{max\ collision}$

4.3.2 Orifice Discharge Coefficient

Experiments have been made to measure the orifice discharge coefficient in typical glottal orifices [32], [37], [58], [30], [36]. In [32], Park measured the orifice discharge coefficient to test the quasi-steady approximation using a physical driven model of the human larynx. Figure 4.6 shows the values for the orifice discharge coefficient obtained in this study.

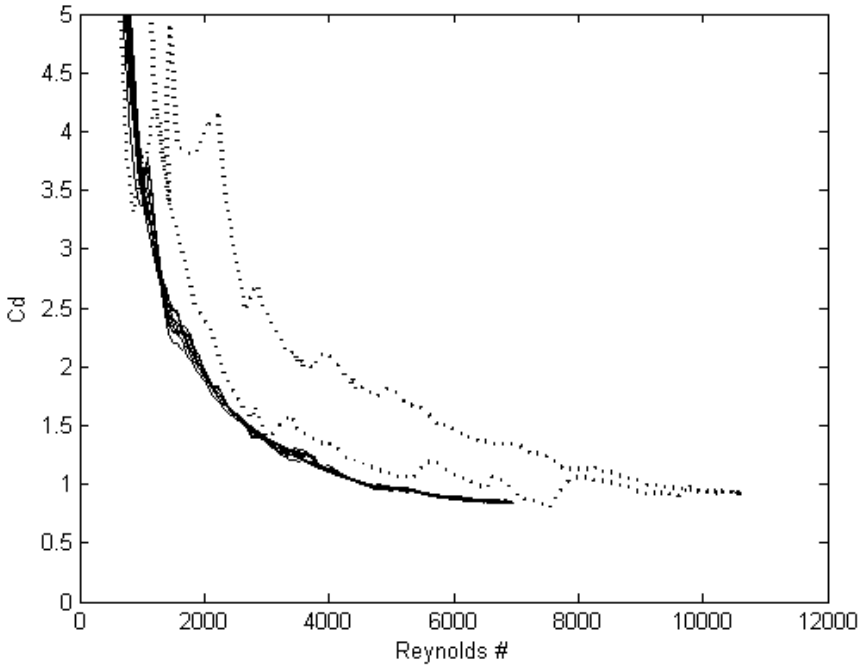


Fig. 4.6 Experimental orifice discharge coefficient data from Park [32].
—: Converging; --: Diverging.

In [32], convergent and divergent glottal geometries were tested separately, for both the opening and closing phases. Note that the orifice discharge coefficient values are greater for the divergent shape during the closing phase (dotted line). On the other hand, in the convergent shape, the values do not differ much on each phase. Values for the discharge coefficient were selected from this experimental data for an opening converging case and a closing divergent case. These two values were used to modify the ideal volumetric flow rate in equation (4.9) to yield the pressure forces, F_P. Since for higher Reynolds numbers the quasi-steady approximation is better supported, the values are selected at the highest Reynolds number (close to Re=7000) where the orifice

discharge coefficient is available for both geometries. Therefore, the values of the discharge coefficient are c_d =0.85 (opening phase) and c_d =1.34 (closing phase). In terms of the oscillation, it is rewritten as

$$c_d(t) = \begin{cases} 0.85, & \dot{y}(t) \geq o \\ 1.34, & \dot{y}(t) < o \end{cases} \tag{4.19}$$

This step-like function creates the difference between the opening and closing phase that changes F_p as in Fulcher model. However, these jumps also produce discontinuities in the volumetric flow rate that make it unrealistic. A smooth function was used to smooth the evolution of the orifice discharge coefficient as

$$c_d(t) = 1.095 - 0.355\frac{\dot{y}}{|\dot{y}_{\max}|} \quad , \dot{y}_{\max} \neq 0 \tag{4.20}$$

Note the slight difference in magnitude between equations (4.19) and (4.20). The latter has peak values of 0.74 and 1.45. This difference is the consequence of a minor increment in equation (4.20) to compensate the reduced average over each cycle. This increment only modifies by about 10% the original $c_d(t)$ values. Numerical results for the pressure force and the volumetric flow rate between these two functions are discussed in chapter 6.

4.3.3 Alternative Volumetric Flow Rate Equation

An alternative formulation for the volumetric flow rate was proposed by Titze in [47] and [50]. This formulation comes from the same basic principles and assumptions previously described, but it was designed for a different model of the vocal folds. Its application to the effective one-mass model cannot reflect all the potential of this equation, but is still valid and comparable with the initial approach. An attractive component of this formulation is the use of a wave reflection analog scheme to supply the feedback from the acoustic pressures. The expression for the volumetric flow rate is then given by,

$$Q_{titze} = \frac{a_d c}{(1-k_e)}\left\{-\frac{a_d}{A^*} \pm \left[\left(\frac{a_d}{A^*}\right)^2 + \frac{4(1-k_e)}{\rho c^2}(p_{up}^{+} - p_{dn}^{-})\right]^{1/2}\right\} \quad , \tag{4.21}$$

where

$$A^* = \frac{A_s A_o}{A_s + A_e} \quad , \tag{4.22}$$

is an effective tract area that include both the tracheal tube and the supraglottal area, p_{up}^{+} and p_{dn}^{-} are the sub and supraglottal acoustic pressures incident into the glottis (obtained from a wave reflection analog approach), a_d is the glottal area at the flow detachment point (in the model developed in this thesis the flow is detached at the glottal opening or *x=0*), and k_e is the pressure recovery coefficient from Ishizaka and Flanagan [12] written as

$$k_e = 2a_d\left(1 - \frac{a_d}{A_e}\right) \quad . \tag{4.23}$$

Since this model includes a moving flow separation point and a pressure recovery, some differences are observed when comparing it with the volumetric flow rate model proposed in (4.9).

4.4 Collision Forces

When the vocal folds surface cross the symmetry line, collision between the folds occurs. Different assumptions have been made to estimate the behavior of the oscillator during collision. A Hertz impact force described by Stronge [40] was used to describe the force collision in F_H that is used in equation (4.2) that describes the motion of the vocal folds. This force is considered for each fold and it is given by

$$F_H = k_H \delta^{3/2}(1 + b_H \dot{\delta}) \quad \text{, and} \tag{4.24}$$

$$k_H = \frac{4}{3}\frac{E\sqrt{r}}{1-\mu^2} \tag{4.25}$$

The penetration of each vocal fold through the contact plane is defined as δ. It corresponds to the absolute value of the position of the mass when it crosses the centerline $y=0$. Similarly, $\dot{\delta}$ is the magnitude of the velocity of the mass during collision. The vocal fold material properties are given by the damping factor b_H, the Young modulus E and the Poisson ratio μ. Finally, r is the radius of curvature of the colliding surface. Due to the geometrical considerations explained in section 4.3, this parameter is approximately the depth of the folds, d. The values for all the constants required to compute F_H are presented in Table 4.1.

In addition to the Hertz force, two other modifications in the system occur during collision. Following the guidelines suggested by Titze and Story in [42], an increment in the damping ratio η_H of the oscillator is included to reflect the effect of the impact in the oscillation. This damping ratio η_H is incremented by 40% with respect to its normal value. Note that the damping ratio defines the damping factor, as expressed in equation (4.27). Finally, the pressure force is not completely suppressed during collision, since it sill applied onto one part of the folds surface. However, the pressure force expression differs from the one derived before because the glottis is closed. The pressure force during collision was assumed to be given by the product of the upstream pressure and the effective surface in contact with the fluid.

4.5 Parameter Values

The values for the parameters used in the model were taken from previous studies. These values have been also used in previous models of the vocal folds, so the results can be compared. The main references are Titze [50], Story and Titze [42], and Berry et al. [3]. These values are presented in Table 4.1. The values for the areas of the subglottal tract and the supraglottal tract are from Story [43]. Since a symmetric model is used, the values supplied into the model are only half those for a complete model.

Table 4.1
Parameters and constants used in the vocal folds model for the numerical simulations

Parameter	Symbol	Default value
Length of the vocal folds	L	0.015 [m]
Depth of the vocal folds (flow direction)	d	0.003 [m]
Equilibrium position of the mass	y_o	0.0001 [m]
Mass of one fold	m	2 x 10^{-4} [kg]
Stiffness (spring constant)	k	200 [N/m]
Damping ratio	ς	0.1
Damping ratio (collision)	ς_H	0.5
Subglottal lung pressure	p_{up}	800 [Pa]
Supraglottal pressure	p_{dn}	0 [Pa]
Density of air	ρ	1.15 [kg/m^3]
Speed of sound	C	350 [m/s]
Young modulus of the folds	E	8 [kPa]
Poisson Ratio	μ	0.4

Using these constants, the theoretical frequency of oscillation and the damping factor are computed. The fundamental frequency is given by

$$f_o = \frac{1}{2\pi}\sqrt{\frac{k}{m}} \quad , \tag{4.26}$$

and the damping factor by

$$b = 2\varsigma\sqrt{mk} \quad . \tag{4.27}$$

Using these formulas, the fundamental frequency for the linear spring-mass system is approximately 160 Hz. The damping factor is b = 0.04 [kg/s] during dynamic oscillations and b_H = 0.2 [kg/s] during collision.

4.6 Flow Diagram of the Source Model

A simplified flow diagram of the source model is shown in Figure 4.7. The forces acting on the model, the fluid-structure interaction and the fluid-sound interaction were computed according to the sequence described.

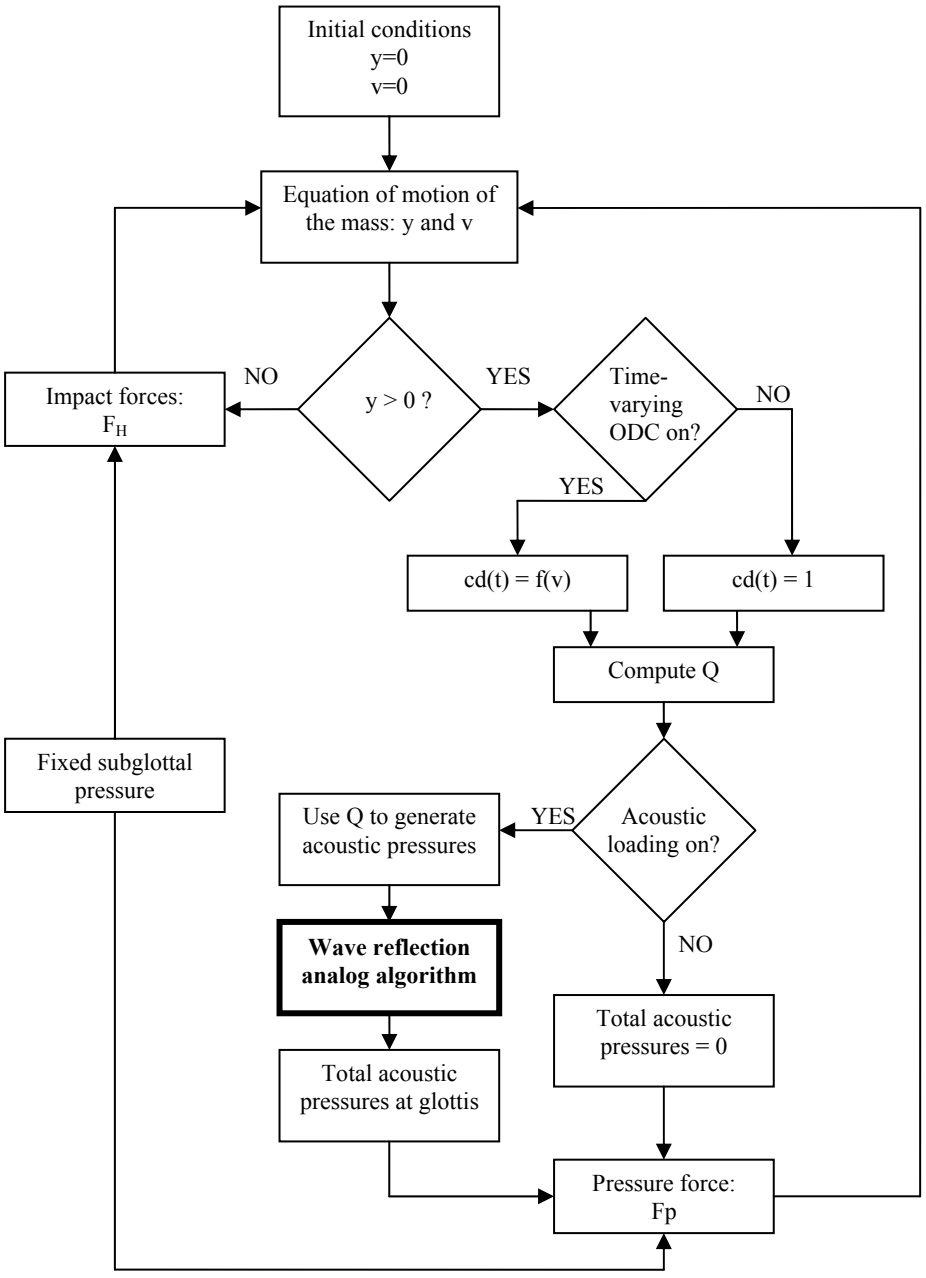


Fig. 4.7 Simplified flow diagram of the source model

5. RESULTS OF THE WAVE REFLECTION ANALOG MODEL

In this chapter, the wave reflection analog model for the acoustics of the subglottal and supraglottal tracts is evaluated. Several tests were performed to verify the accuracy of the model.

5.1 Testing Approach

To verify the wave reflection analog model, the impulse response of the tracts and their respective transfer functions were calculated. Assuming that it was generated by a volumetric flow velocity source, an impulsive pressure input with finite amplitude of 1 [Pa] or 94 dB (ref:$2x10^{-5}$ [Pa]) was supplied at the glottis ($x=0$ in Figure 5.2) to the acoustic load under evaluation. This impulse propagates in the tract, where multiple reflections with the junctions or ends occurred. The total acoustic pressure is shown vs. time at $x=L$ in Figure 5.2. This corresponds to the impulse response of the system. The frequency response of the system is the Fourier transform of the impulse response. In practice, only the magnitude of the FFT (Fast Fourier Transformation) of the radiated acoustic pressured is reported. This general approach is illustrated in Figure 5.1.

The time evolution of the waves was visualized using Matlab-based movies. This feature was particularly useful to evaluate the performance of the model for different boundary conditions.

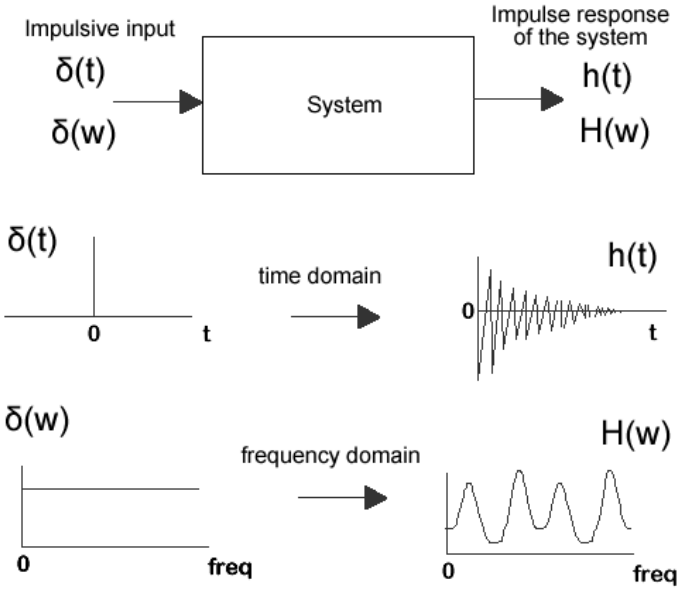


Fig. 5.1 Ideal impulse response analysis for a system in the time and frequency domains.

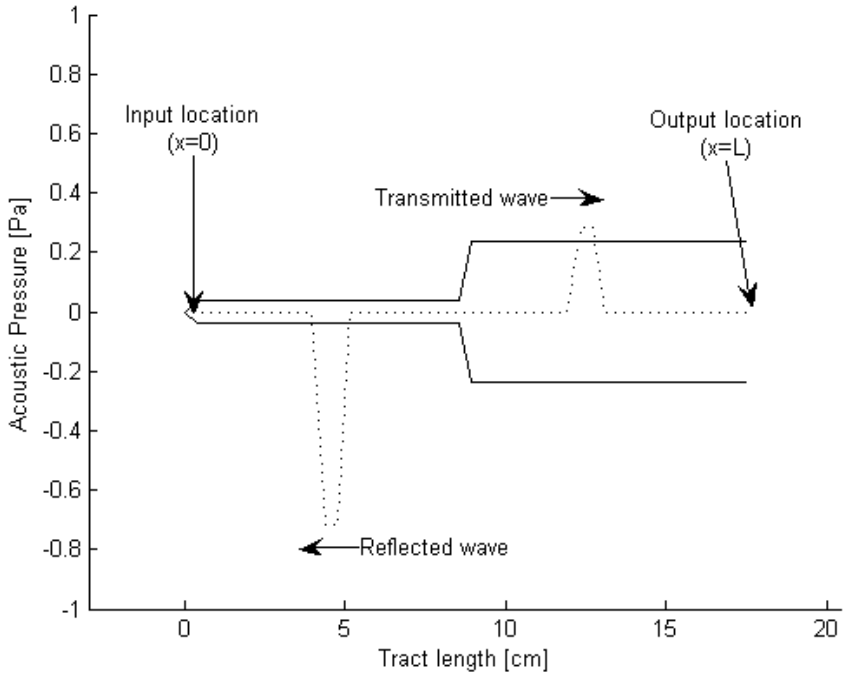


Fig. 5.2 Schematic view of a Matlab-based movie for a two-tube model.

For all the cases and tests (unless otherwise stated) using the wave reflection analog in this chapter, the parameters presented in Table 5.1 were used.

Table 5.1
Parameters and constants used in the wave reflection analog for preliminary numerical simulations

Parameter	Symbol	Default value
sampling frequency	F_s	44100 [Hz]
Cross-sectional area – only for uniform tubes	A	4 [cm^2]
Total tube length	L	17.46 [cm]
Tube sections	*N_tubes*	44
total time frame of each test	*t_test*	0.5 [s]
tube section where the impulse is supplied	*in_location*	2
tube section where the response is measured	*out_location*	*N_tubes*

5.2 Boundary Conditions in a Uniform Tube

To evaluate the wave reflection analog scheme's handling of boundary conditions, two sections were added at the ends of the domain for each boundary region. For a closed boundary, a small section of area 10^{-8} [m^2] was included. For an open boundary, a section of area 100 [m^2] was added. This is sufficiently large to be considered as an infinitely baffled tube. Ideally, no losses or radiation impedance should be introduced in these tests. However, some degree of loss is required to have a decaying impulse response. Thus, minor losses were included at boundaries. The following equations indicate the resonance frequencies for the normal modes of the system, also called formants, for different boundary conditions [38]:

For a open-closed tube $$f_{n\ (open-closed)} = \frac{(2n+1)c}{4L} \tag{5.1a}$$

For a closed-closed tube $$f_{n\ (closed-closed)} = \frac{nc}{2L} \tag{5.1b}$$

For a open-open tube $$f_{n\ (open-open)} = \frac{nc}{2L} \tag{5.1c}$$

The values for f_n are in Hz. Here L is the total tube length, c the speed of sound and n = 0, 1, 2, 3, … is the mode number. Using equations (5.1a), (5.1b) and (5.1c), the theoretical resonant frequencies were computed and are presented in Table 5.2.

Table 5.2
Calculated resonance frequencies for different boundary conditions for a uniform tube using the constants from Table 5.1

Tube configuration	F1 [Hz]	F2 [Hz]	F3 [Hz]	F4 [Hz]	F5 [Hz]
Open- Closed	501	1503	2506	3508	4510
Closed – Closed	0	1002	2005	3007	4009
Open - Open	0	1002	2005	3007	4009

As shown in Figures 5.3, 5.4 and 5.5, the values computed using the wave algorithm match very closely the values presented in Table 5.2. This supports the accuracy of the model. Note that since only minor losses were introduced in these tests, the bandwidths of the formants are very narrow. In addition, the different peak values in the magnitudes are product of the boundary conditions imposed during the tests.

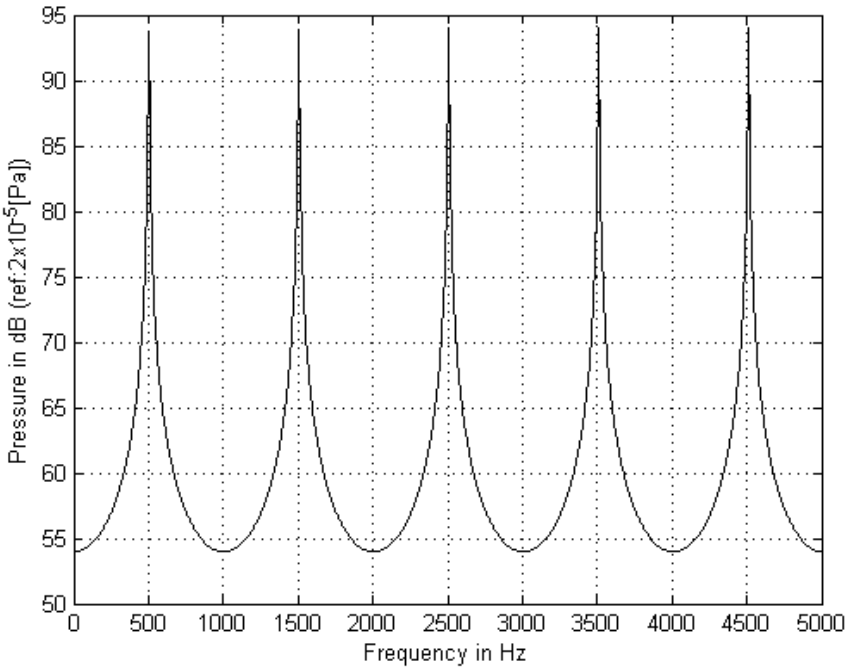


Fig. 5.3 Spectrum of the acoustic pressure at $x=L$. Tube closed at $x=0$ and open at $x=L$.

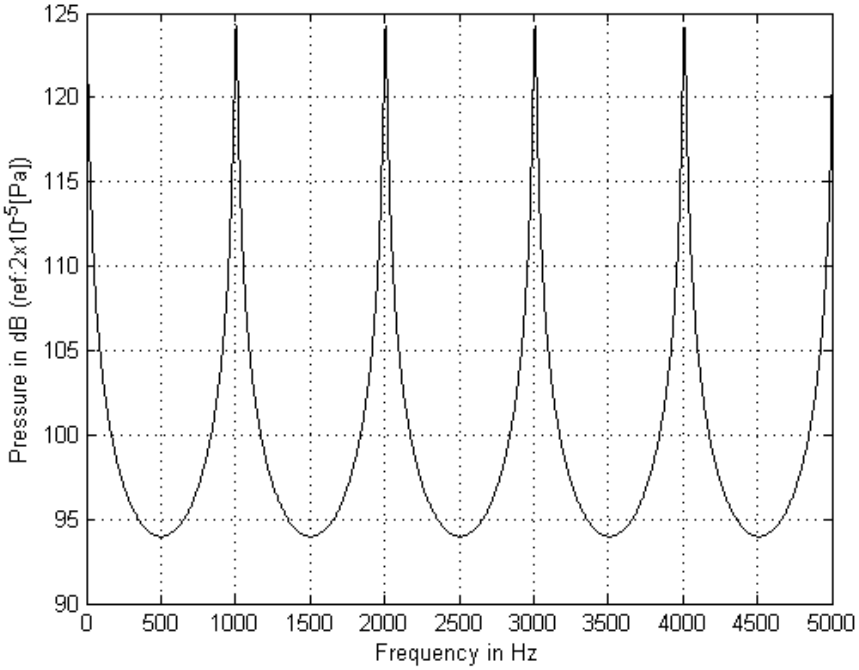


Fig. 5.4 Spectrum of the acoustic pressure at $x=L$. Tube closed at $x=0$ and at $x=L$.

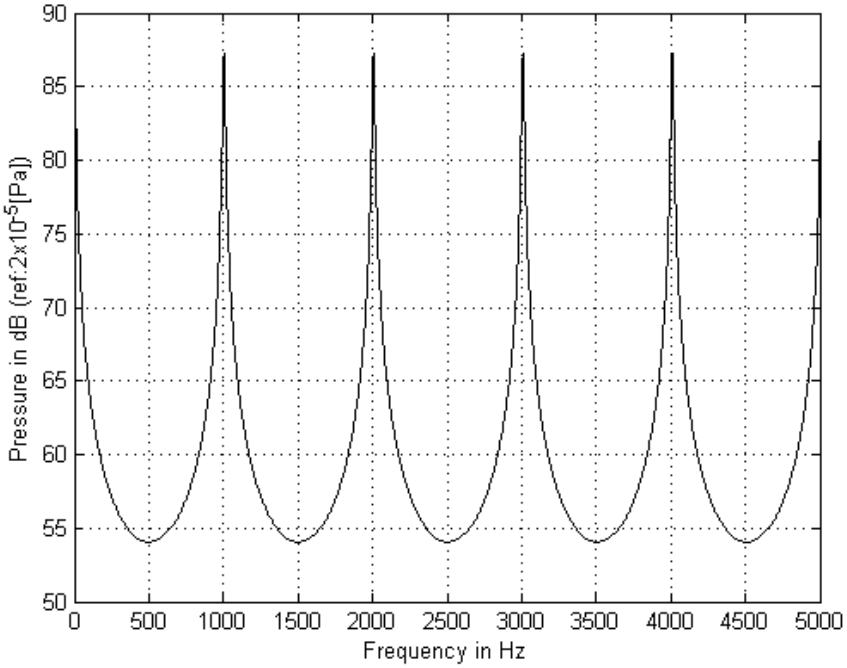


Fig. 5.5 Spectrum of the acoustic pressure at $x=L$. Tube open at $x=0$ and at $x=L$.

5.3 Radiation Impedance

The effects of a radiation impedance were investigated for the case of a uniform tube closed at *x=0* and open *x=L*. The boundary created by adding a section with area 10^{-8} [m^2] resembles a closed glottis. When evaluating the radiation impedance, no extra section is needed at *x=L*. Only the losses observed at the boundary are included in this analysis. The spectra of the acoustic pressure at the termination with and without a radiation impedance are shown in Figure 5.6. The differences in amplitude are due to the more realistic reflections and transmissions at the boundary *x=L* for the case with a termination impedance. For the case with no impedance, the open-ended boundary was created by adding an arbitrarily large section to resemble an infinite baffle. This produced an artificial boundary condition, designed for testing purposes. The adjustments in amplitude are due to the resistive part of the impedance. The shift in the formant frequencies is due to the effects of the radiation reactance.

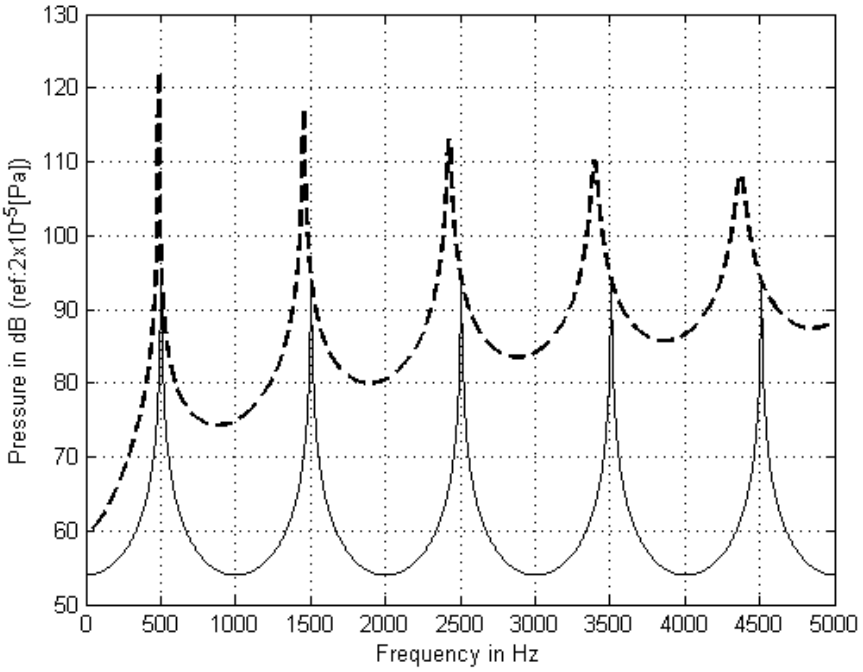


Fig. 5.6 Spectrum of the radiated pressure for a tube closed at *x=0* and with termination impedance at *x=L*. -: No radiation impedance; --:With radiation impedance.

Two approaches were used to verify the accuracy of the results. The effective length of the tube was modified through the addition of a correction factor given by Fant [10]:

$$L_{effective} = L + 0.8\sqrt{\frac{A}{\pi}} \quad . \tag{5.2}$$

This correction was used along with equation (5.1a) to compute the modified resonance frequencies with termination impedance. The theoretical values of these formants are presented in Table 5.3. In addition, a more complete evaluation was made. The acoustic pressure at the termination impedance was obtained using the complex solution from the planar wave equation [23]. The relation between acoustic pressure and frequency is given by

$$\hat{p}(x) = \hat{A}e^{j\hat{k}(L-x)} + \hat{B}e^{-j\hat{k}(L-x)} \quad , \tag{5.3}$$

where $\hat{k}$ is the complex wave number, adjusted to be

$$\hat{k} = k + j\alpha \quad , \tag{5.4}$$

where α was set to 0.19 to match the wave reflection analog losses for this case. $\hat{A}$ is the incident acoustic pressure given by an ideal velocity source at $x=0$, and $\hat{B}$ is the reflected pressure given by the reflection coefficient at $x=L$

$$\hat{R} = \frac{\hat{Z}_{mL} - 1}{\hat{Z}_{mL} + 1} \quad , \tag{5.5}$$

where the termination impedance was modeled as

$$\hat{Z}_{mL} = \frac{1}{4}ka + j0.6\ ka \quad , \tag{5.6}$$

where a is the radius of the terminal cylindrical section. Using this theoretical complex solution approach, the acoustic pressure at $x=L$ is given by

$$\hat{p}(L) = \hat{A} + \hat{B} = \hat{A}(1 + \hat{R}) \tag{5.7}$$

As shown in Table 5.3, the theoretical values of the formant frequencies computed using both the effective length correction and the complex solution of the wave equation are in good agreement with those from the wave reflection analog. The inclusion of radiation impedance yields nearly the same resonance frequency values as those obtained analytically. The reactance also cause a slight shift in the upper formants; they are no longer exact integer multiples of the first formant.

Table 5.3
Calculated and measured resonance frequencies with the effect of a radiation impedance for a uniform tube using the constants from Table 5.1

Tube configuration	**F1 [Hz]**	**F2 [Hz]**	**F3 [Hz]**	**F4 [Hz]**	**F5 [Hz]**
No length correction	513	1538	2564	3590	4615
With effective length correction	487	1461	2435	3409	4383
Theoretical complex solution	488	1463	2438	3415	4392
Modeled with wave analog	487	1461	2433	3409	4385

A comparison between the magnitudes of the radiated acoustic pressures from the theoretical complex solution and the wave reflection analog is shown in Figure 5.7. The effects of losses are included for both cases. In the theoretical complex solution, the complex wave number from equation (5.4) defines the losses. In the wave reflection analog algorithm, the losses are based on equation (3.38). The wave reflection analog scheme yields accurate frequency response amplitude predictions. Note that for both cases, the radiation impedance increases the bandwidth of the upper formants. Minor differences between the curves are due to the slightly different impedance termination.

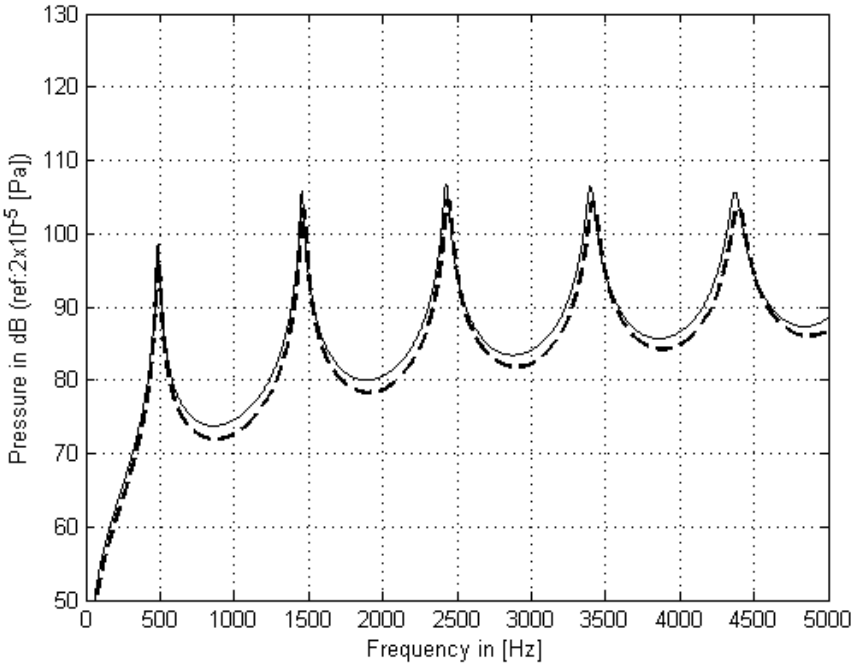


Fig. 5.7 Spectrum of the radiated pressure for a tube closed at *x=0* and with termination impedance at *x=L*. -: Wave reflection analog; --: Theoretical frequency domain solution.

5.4 Global Loss Factor

To investigate the effects of yielding walls, viscous effects and heat conduction losses, the same uniform tube with closed-open ends was used. The boundary conditions were defined as in the previous tests. No radiation impedance was initially used in this test. Figure 5.8 shows the comparison between the frequency responses with and without the effects of the losses. Note that in this test case all formant have the same bandwidth (7 Hz). A reduction of 13 dB in magnitude and an increase of 21 Hz in bandwidth is observed for all formants. The influence of the loss factor was investigated for the case of the same uniform tube with radiation impedance. Figure 5.9 shows the results. The narrow formant around 500 Hz is reduced by about 26 dB. The broader formant near 4500 Hz is reduced by only 3.5 dB. Thus, the loss factor reduces marked tonal components in the system.

These comparisons show that the loss factor influences all frequency components of the waves of interest. Its effects match theoretical expectations.

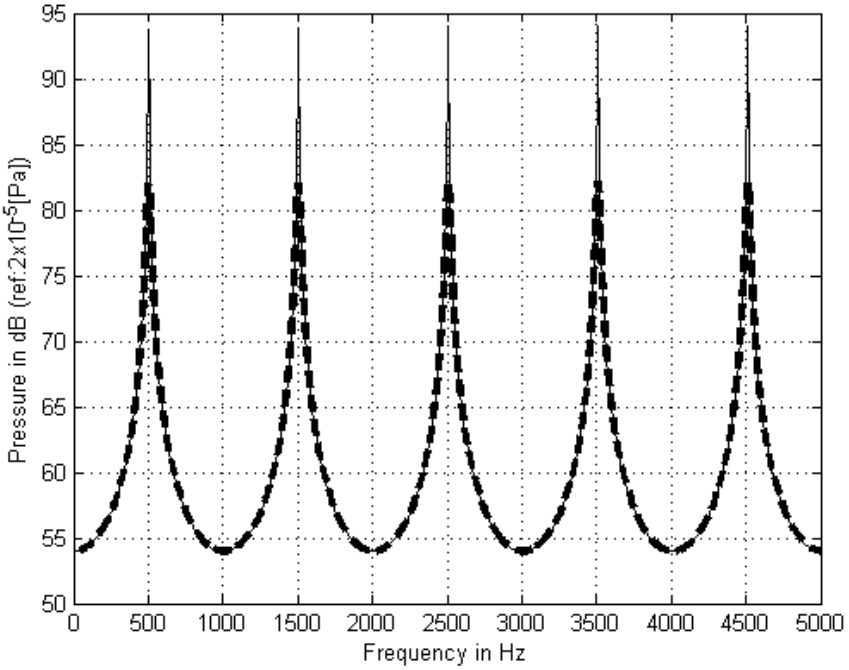


Fig. 5.8 Effects of losses in the frequency response of a tube closed at *x=0* and open at *x=L* -: No losses; --:With losses.

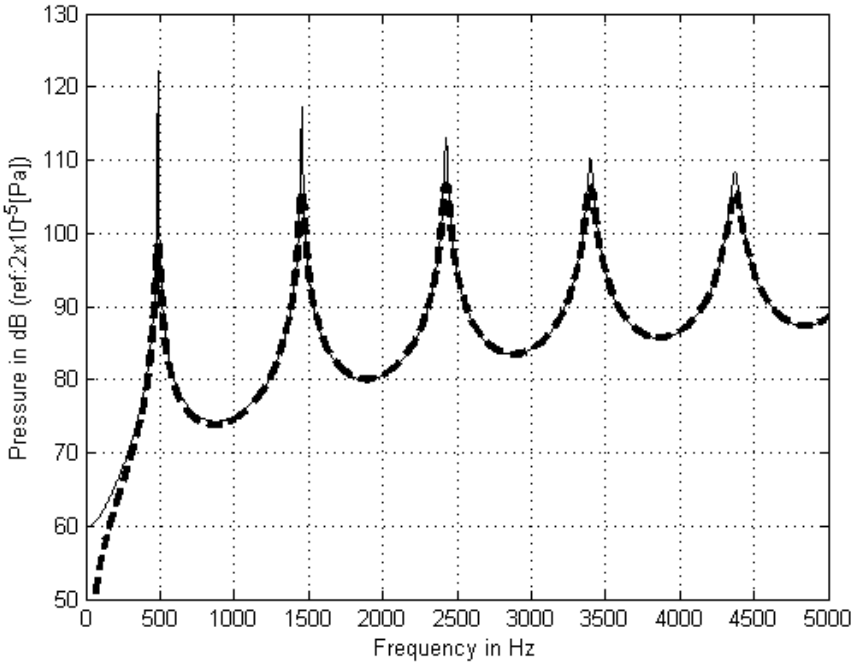


Fig. 5.9 Effects of losses and radiation impedance in the frequency response of a tube closed at *x=0* and open at *x=L* -: No losses; --:With losses.

5.5 Comprehensive Evaluation: Tract Configuration

Three shapes resembling the supraglottal tract were tested. These shapes were previously evaluated by Story in [41], therefore a direct verification of the transfer functions can be made. Note that no exact values of the formant are reported for this test in [41], which makes the verification qualitative more than quantitative. The test cases correspond to a uniform tube, an idealized vowel /a/ and an idealized vowel /i/. The glottis is considered closed and the sections are defined as shown in Figure 5.10. The radiation impedance and the loss factor were included for these results.

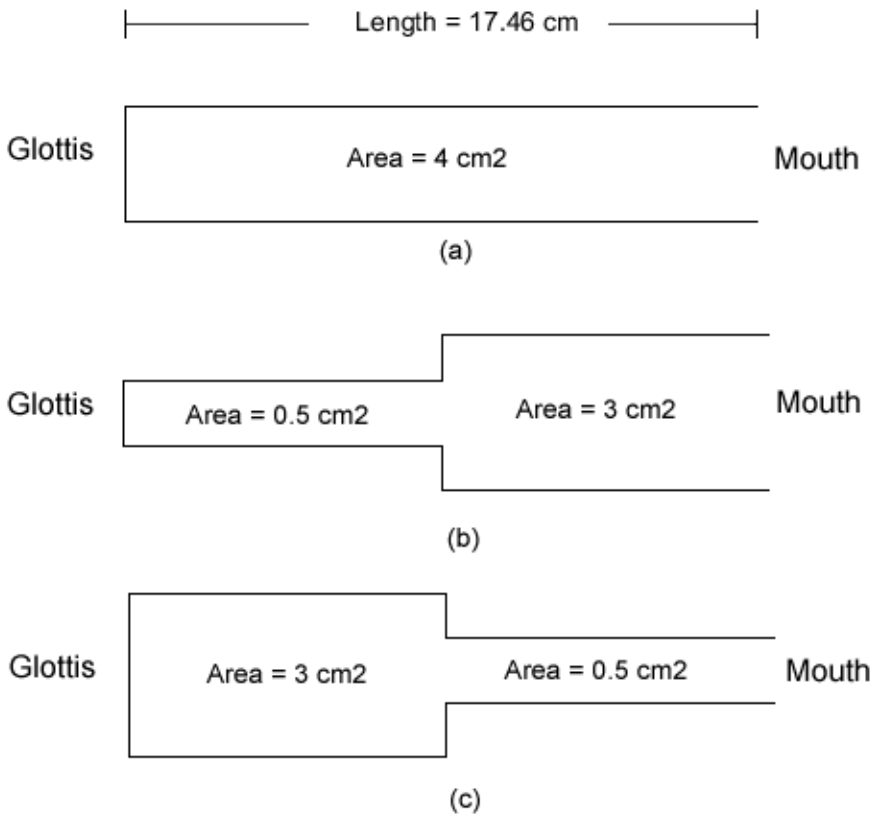


Fig. 5.10 Test cases for the complete scheme and vowel evaluation. All the tubes have a length of 17.46 cm (a) Uniform tube, (b) idealized /a/, (c) idealized /i/

Figure 5.9 (from last section) presents the frequency response of a uniform tube when termination impedance and losses are considered. This is a neutral case because no constrictions are present. The effects of radiation impedance and global losses are combined, creating a slightly different result than that reported by Story [41]. Losses are handled differently in that study, producing only a noticeable change in the amplitude of the first formant. No notoworthy differences in the frequency of the formants are observed.

Using perturbation theory [39], it is possible to predict relative movements of the formants frequency with respect to constrictions based on a uniform tube. Using this idea, when a uniform tube is constricted in its first half (close to the source), the first and third formant frequencies are raised and the second and fourth formant frequencies are reduced. This behavior is clearly observed in Figure 5.11. Using the same principle, when the second half (close to the opening) of a uniform tube is constricted, the first and third formant frequencies are reduced, and the second and third formant frequencies are increased. This behavior is observed in Figure 5.12.

The complete design of the wave reflection scheme behaves as expected for different tract configurations.

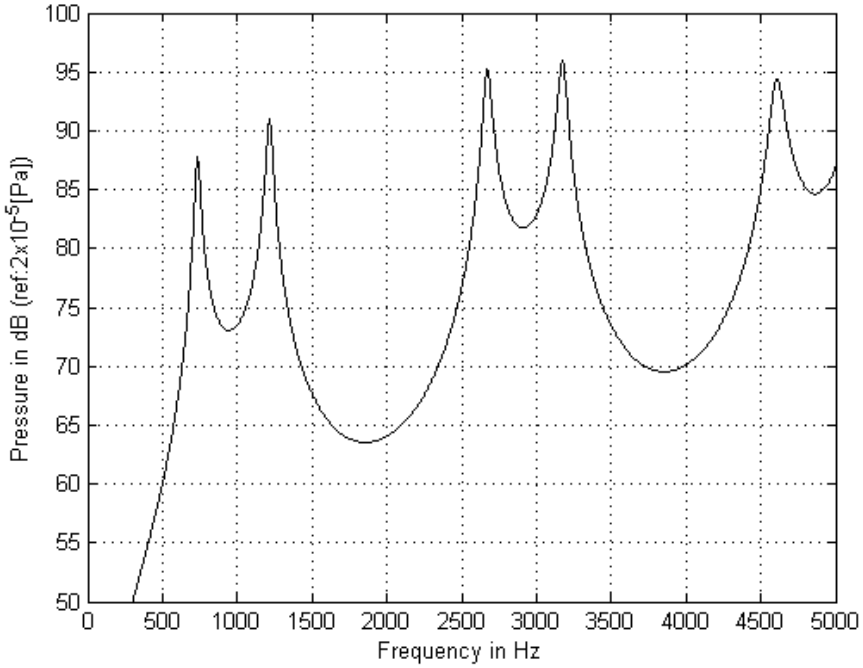


Fig. 5.11 Spectrum of the radiated pressure at $x=L$ for an idealized vowel /a/.

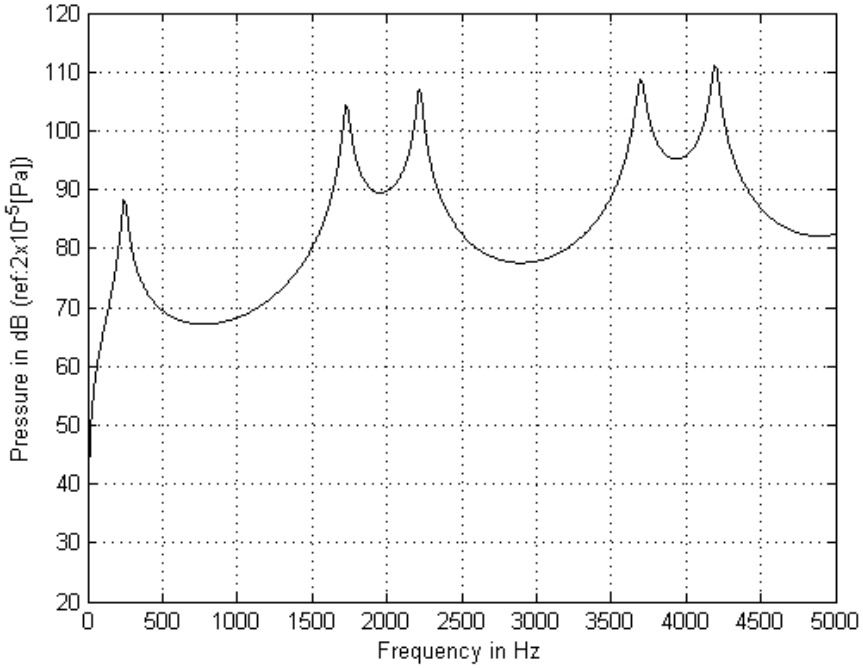


Fig. 5.12 Spectrum of the radiated pressure at $x=L$ for an idealized vowel /i/.

5.6 Subglottal Tract Design Considerations

As mentioned in section 3.7, two subglottal tract designs were investigated. The first design was an incomplete MRI based shape that requires a termination impedance at the location of the carina to simulate the complete tract. The second was an area function of the subglottal tract based on lung studies, which includes a significant increase in cross-sectional area towards the end of the tract. No radiation impedance is required in the latter case.

Table 5.4 summarizes results from several experimental and analytical studies on the acoustics of the subglottal tract [39], [10], [15]. The formant frequencies are the most relevant parameters to describe the acoustic behavior of the subglottal tract. The bandwidth for the subglottal resonances are between 200Hz and 400Hz [39] [21].

Table 5.4
Subglottal resonance frequencies from previous studies.

Study	Details	F1 [Hz]	F2 [Hz]	F3 [Hz]	F4 [Hz]
Stevens' model [39]	Male	600	1550	2200	-
Stevens' model [39]	Female	700	1650	2350	-
Ishizaka's model [21]	-	615	1335	2110	2379
Harper's model [15]	-	569	1360	1980	2500
Harper's measurements [15]	Subject 1	606	1279	2256	2919
Harper's measurements [15]	Subject 2	529	1092	1729	2343
Harper's measurements [15]	Subject 3	743	1191	2024	2617
Harper's measurements [15]	Subject 4	699	1318	2228	2748
Harper's measurements [15]	Mean	644	1220	2059	2656

To simplify notation, the tract design from the MRI data of Story [41] with radiation impedance is referred as subglottal tract design #1. The tract design based on the area function of Weibel [54] is referred as subglottal tract design #2. The transfer functions were computed using the impulse response method described before. The amplitude for the impulse is 1 [Pa] or 94 dB (ref:$2x10^{-5}$[Pa]). It is convenient to show the frequency responses of the tracts not only at $x=L$ but also at $x=0$, since the acoustic pressure in these subglottal tracts interact with the rest of the system through the glottis. The results for both cases are presented in Figures 5.13 and 5.14.

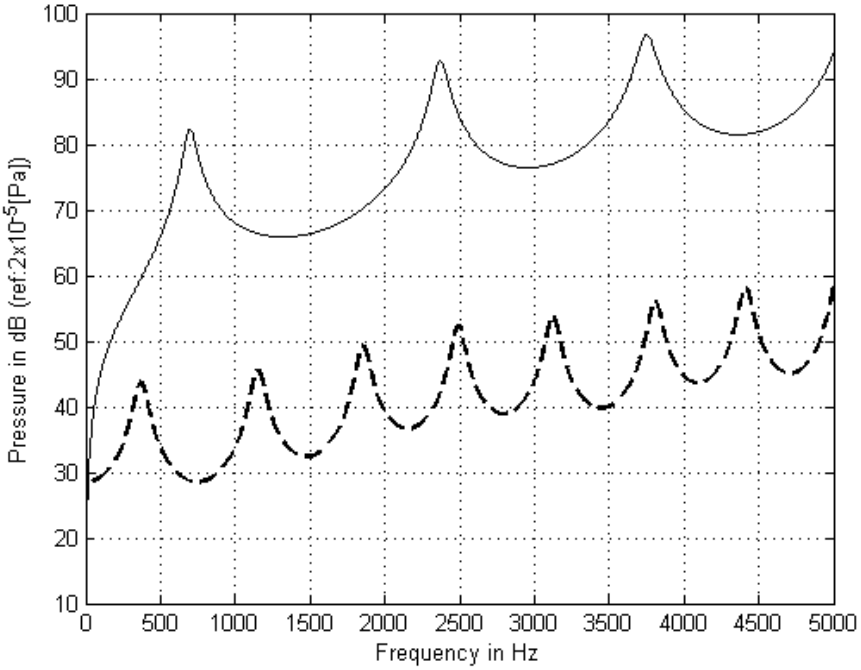


Fig. 5.13 Spectrum of the acoustic pressure at $x=L$. -: Design#1; --:Design #2.

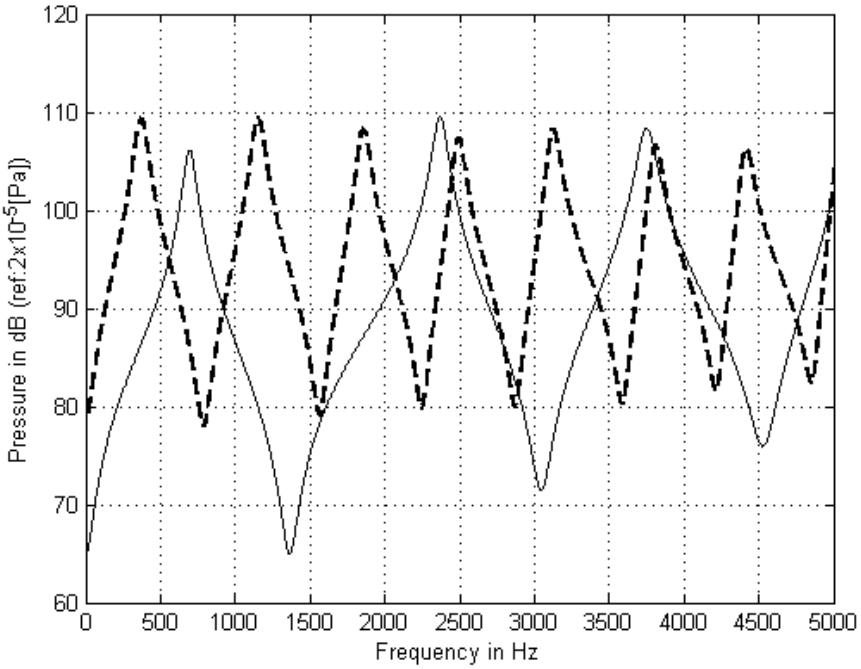


Fig. 5.14 Spectrum of the acoustic pressure at $x=0$. -: Design#1; --:Design #2.

Note that design #1 uses Flanagan's original radiation impedance derived in section 3.3. This expression is for to the radiation with an infinite baffle, which it is not directly applicable to the subglottal case, but serves as a first cut model.

Table 5.5 summarizes the results of Figures 5.13 and 5.14. Note that a modified version of the subglottal tract design #2 is also included, as discussed below. A comparison between the values presented in Table 5.4 and Table 5.5 highlights problems for both tract designs.

Table 5.5
Subglottal resonances obtained using the wave reflection analog scheme.

Tract Design	Details	F1 [Hz]	F2 [Hz]	F3 [Hz]	F4 [Hz]
#1	With radiation impedance	620	2220	3500	4800
#2	No radiation impedance	369	1151	1855	2490
#2	Modified version	613	1341	2100	2896

Tract design #1 partially matches the results of Stevens' model [39], but it does not exhibit a second formant. This is because in this design the tract is not complete. Altering the radiation impedance could modify to some degree the bandwidth and frequency of the formants. More work needs to be done to identify an appropriate termination impedance for this design. Tract design #2 yield values in the range of the model and measurements performed by Harper [15], except for the first formant. The lower frequency of the first formant is likely to be due to the global loss factor. This factor is averaged over frequency, and tends to overestimate yielding wall effects at low frequency. Other studies using wave reflection analog [41] have shown a first formant shift of less than 15% compared to the hard wall case. This shift is still insufficient to produce a first formant at the desired frequency, which is approximately 70% above the hard walled case. A more accurate expression for yielding walls is needed to improve the wave reflection analog model of the subglottal tract.

A modifed design # 2 was created using profile of Weibel [54] up to the point where the area increases significantly (introducing a reduction of 3.6cm in the tract length). The subglottal tract is terminated at this point by a closed boundary with a reflection coefficient of 0.8. This increases the first formant without altering significantly the others and produces formant bandwidths between 200Hz and 400Hz. This design is based on Titze [47]. The frequency response of the modified design#2 is shown in

Figures 5.15 and 5.16. The formant frequencies vary between 200 Hz and 300 Hz for all formants. Since noticeable differences are observed between the relative magnitudes of the original and modified transfer functions due to the different boundaries, it is convenient to compare them at the glottis location, where the impulse is injected. Figure 5.16 shows that there are no important differences in magnitude between the two cases. A shift in frequency is clearly observed in this Figure. The new formant frequencies are shown in Table 5.5. Note that these values fall within the range of the values presented in Table 5.4.

The sign of the reflected wave at the boundary was investigated. Acoustic theory in tubes states that for an open ending, the acoustic wave is reflected out of phase. For a closed ending, the acoustic wave is reflected in phase. Although the boundary is considered open, the reflection observed at the open boundary in the original design #2 appears to be in phase with the incident wave. This is product of the rapidly increasing opening towards the end of the tract, producing multiple reflections that are combined to create an in phase reflection. Several reflections are observed prior to this main component. Simulations showed that the original reflection has reduced amplitude and it is in fact out of phase with respect to the incident wave pulse. However, the multiple reflections that are produced during the expansion create an apparent reflection in phase, having a larger amplitude. This phenomenon can be observed in animations of the wave motion, as illustrated in Figures 5.17 and 5.18. The incident pulse had positive amplitude. Although the sign of the reflected wave at the boundary in the modified design #2 is consistent with the original design #2, "in phase" reflections at the terminal section are not expected to occur. This has been noticed by Wodicka (personal communication) in measurements of acoustic pulses in the human respiratory system. This inconsistency is a product of the area function used to construct the acoustical model. More work needs to be done to develop a more accurate representation of the subglottal tract using the wave reflection analog technique. However, since the modified version of the subglottal tract design #2 yields the expected formant frequencies and bandwidths, this tract design serves as an initial approach to the problem**.** Thus, this design was used in subsequent analysis.

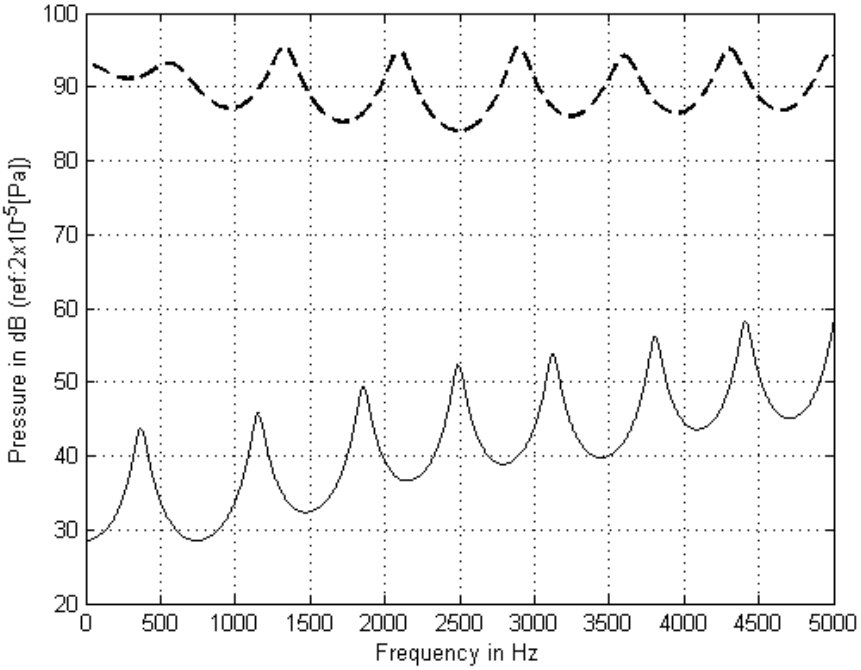


Fig. 5.15 Spectrum of the acoustic pressure at $x=L$. -: Design#2; --:Modified design #2.

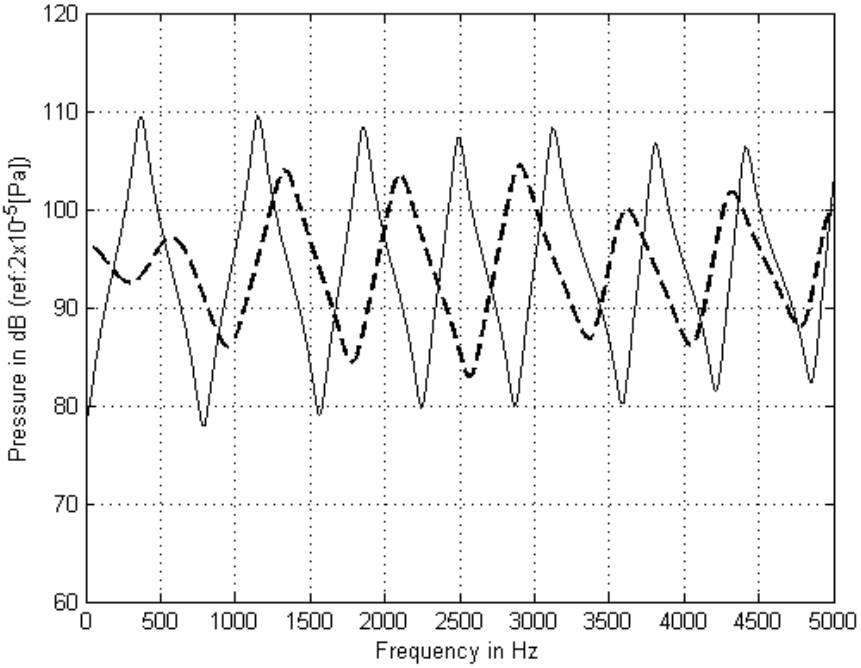


Fig. 5.16 Spectrum of the acoustic pressure at $x=0$. -: Design#2; --:Modified design #2.

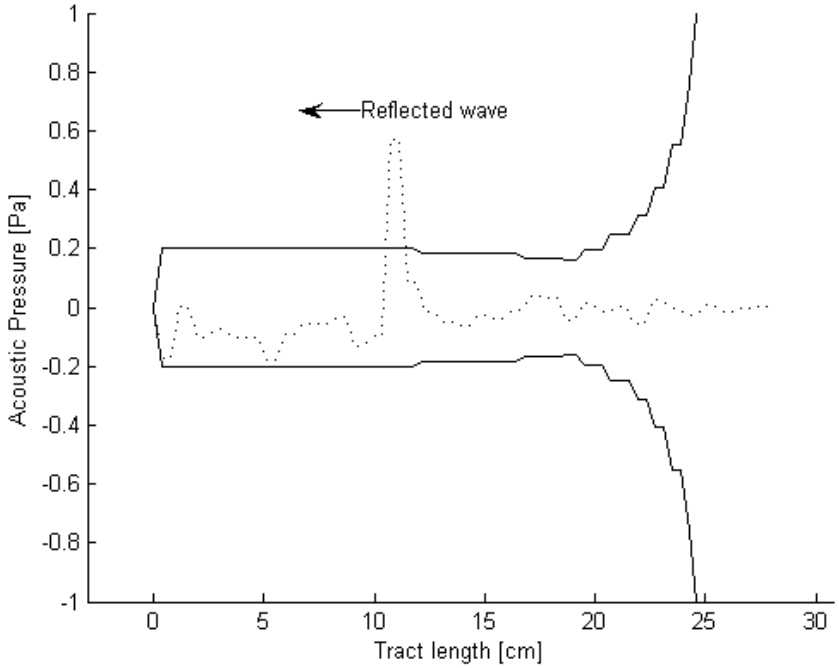


Fig. 5.17 Snapshot of the sound field in the subglottal tract. Reflected wave from a positive incident wave. Original case from [54].

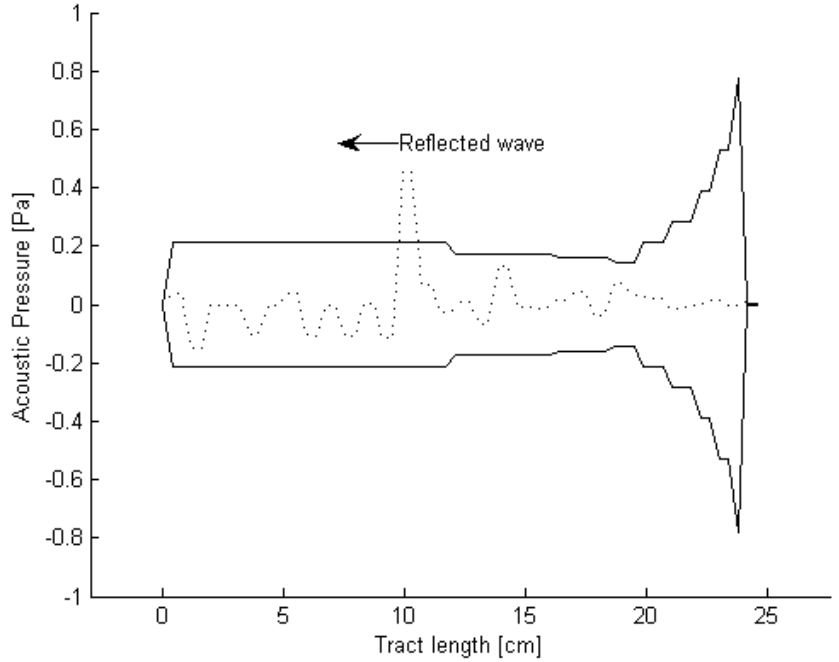


Fig. 5.18 Snapshot of the sound field in the subglottal tract. Reflected wave from a positive incident wave. Modified design #2.

5.7 Acoustic Coupling between Tracts

Acoustic coupling between tracts occurs due to the influence of the subglottal tract on the sound field of the supraglottal tract. Due to its small area during phonation, the glottis is normally considered a closed boundary. Thus, a reasonable approximation of the transfer function at the mouth opening may be obtained by considering only the acoustics of the supraglottal tract. However, as discussed in chapter 3, previous studies [39] [15] have suggested that some interaction between the sub- and supra-glottal tracts occurs. If the glottal opening is considered a connection between both tracts, poles and zeros appear close to each other in the supraglottal transfer function. As previously discussed, the location of the zeros is obtained from the resonant frequencies of the subglottal tract when the glottis is closed. Poles are located close to these zeros, with no more than a 200 Hz difference in frequency. Note from Table 5.5 that the first four subglottal formants (with a closed glottis) are at 613 Hz, 1341 Hz, 2100 Hz and 2896 Hz. Higher frequency formants can be observed in Figure 5.15.

In [39], Stevens obtained the spectrum of a vowel /i/ produced by an adult female speaker. A small prominence around 1600 Hz was assumed to be caused by the second subglottal resonance. To illustrate the same point, the vowel /i/ was selected for the vocal tract. To test the acoustic coupling, both tracts were joined together through the glottal opening. Figure 5.19 shows a schematic of the system.

Figure 5.20 presents the frequency response of the complete system (both tracts) for the case of a total glottal closure. Obviously, this case matches the spectral features of the vowel /i/, given only by the supraglottal tract considering the glottis as a close boundary, and no coupling is observed due to the closure.

Two approaches were used to join the tracts and test the presence of acoustic coupling: linear and non-linear acoustic coupling. As discussed in section 3.6, the linear case treats the glottis as an extra tube section, keeping the same scattering equations to compute the transmitted pressure wave through the orifice. In the non-linear case, a Bernoulli equation is used to compute the transmitted and reflected pressures from the flow rate, as originally suggested by Ingard [18]. It is of interest to evaluate how the acoustic coupling between the two tracts behaves and to compare these two approaches.

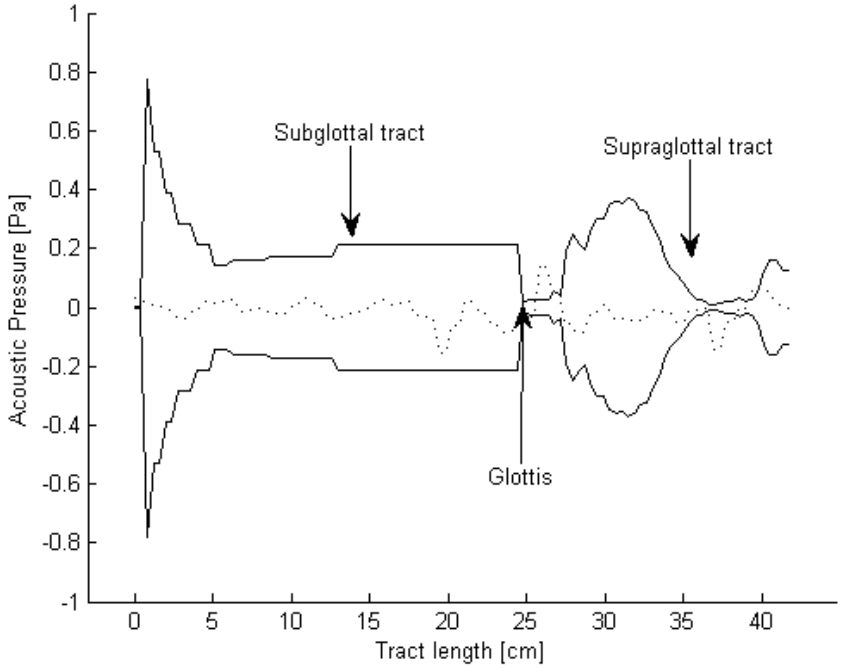


Fig. 5.19 Coupling between the subglottal tract, the glottis and the vocal tract. For these tests, an MRI based /i/ vowel was selected.

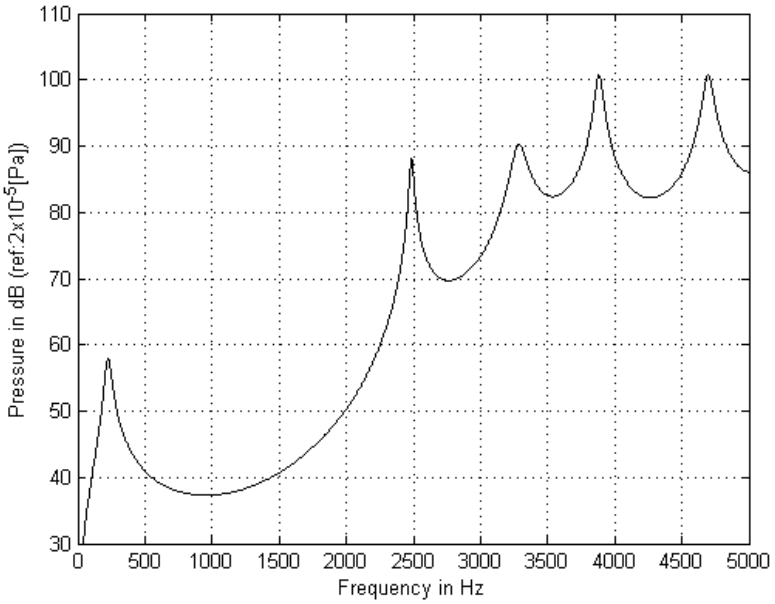


Fig. 5.20 Frequency response of the complete system with a complete glottis closure.

In these tests, an impulse was supplied into both tracts at the glottis, and the transfer function at the mouth opening was measured. A fixed glottal opening of 1 mm was used. For the linear coupling case an extra tube section was added. Note that the sampling frequency of 44.1 KHz, defined a tube length of approximately 0.39[cm], which was similar to the thickness of the vocal folds. In the non-linear approach, the glottal source was considered a junction instead of a tube. A volumetric flow rate source was used as defined in equation (4.9), without the action of lung pressure to simplify the problem. The connection between this source and the tracts was achieved using equations (3.44) and (3.45).

Figures 5.21 and 5.22 show the frequency response for the linear and non-linear coupling. Both figures suggest that when a glottal opening joins the tracts, poles and zeros at the resonant frequencies of the subglottal tract appears when using the wave reflection analog scheme. In both cases, the poles and zeros are close to each other around the closed-glottis subglottal formant frequencies, matching observations by Stevens. The main difference is the larger amount of irregularities observed in the non-linear coupling case. Additionally, different levels of displacement on the location of the original formants (compare with Figure 5.20) were observed on each case. With respect to the frequency shift, the non-linear case resembles more closely the effects observed by Stevens in [39]. Although the presence of pole/zero pairs is more pronounced in the linear case in this test, it is believed that the non-linear approach is more realistic for the acoustic coupling between the tracts. Interestingly, the irregularities observed in the non-linear case make the resulting signal sound more realistic when played back through a digital sound card for non-standardized auditory tests.

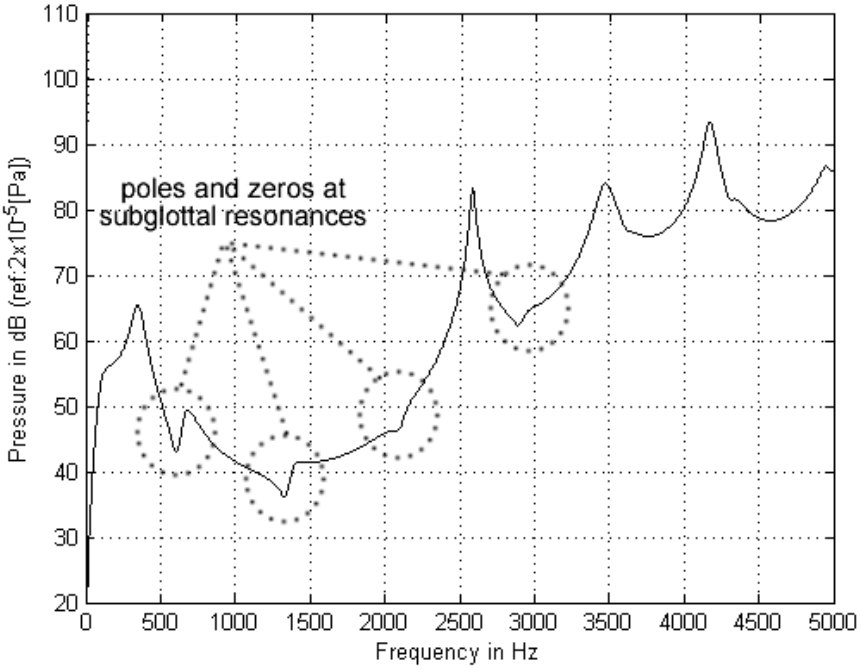


Fig. 5.21 Frequency response of the linearly coupled system for a fixed glottal opening.

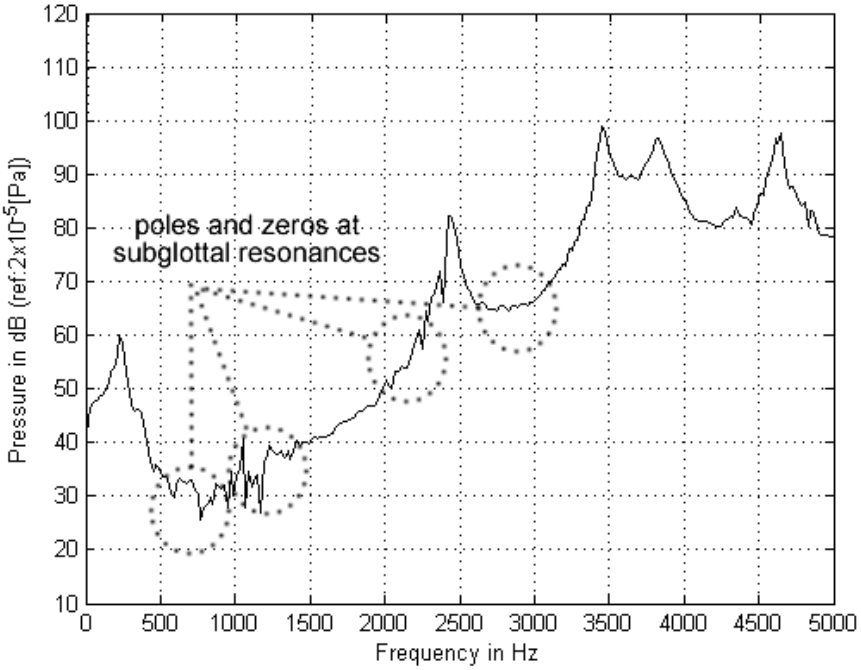


Fig. 5.22 Frequency response of the non-linearly coupled system for a fixed glottal opening.

6. RESULTS OF THE SOURCE MODEL

In this chapter, the main features of the optimized one-degree of freedom self-oscillating source model are evaluated. Several tests were performed to verify the accuracy of the model. No acoustic loading was included for these cases.

6.1 Effects of the Discharge Coefficient

As described in section 4.3, the orifice discharge coefficient c_d accounts for the changes observed in the volumetric flow rate (and hence flow velocity and wall pressure) between converging and shapes of the vocal folds. The orifice discharge coefficient takes different values during the opening and closing portions of the cycle. The changes in value of this coefficient are expected to modify the pressure force sufficiently to drive the mass into self-sustained oscillations. To illustrate the effect of the discharge coefficient, three cases are presented: a discontinuous $c_d(t)$, a continuous $c_d(t)$ and a fixed c_d.

Figure 6.1 shows a time record of 50 [ms] when two types of orifice discharge coefficients are used. The dotted lines of different variables show parameters of for a discontinuous discharge coefficient defined in equation (4.19). The solid line represents the same parameters for the continuous discharge coefficient case as in equation (4.20). Note that negative glottal areas indicate closure. In the computations, the glottal areas are set to zero upon closure. Although some differences in the pressure force are observed, they do not produce noticeable effects in the oscillatory behavior of the mass in these two cases. The frequency of oscillation is 180 Hz instead of the theoretical 160 Hz from the linear spring-mass system. This difference is assumed to be due to the effects of collision.

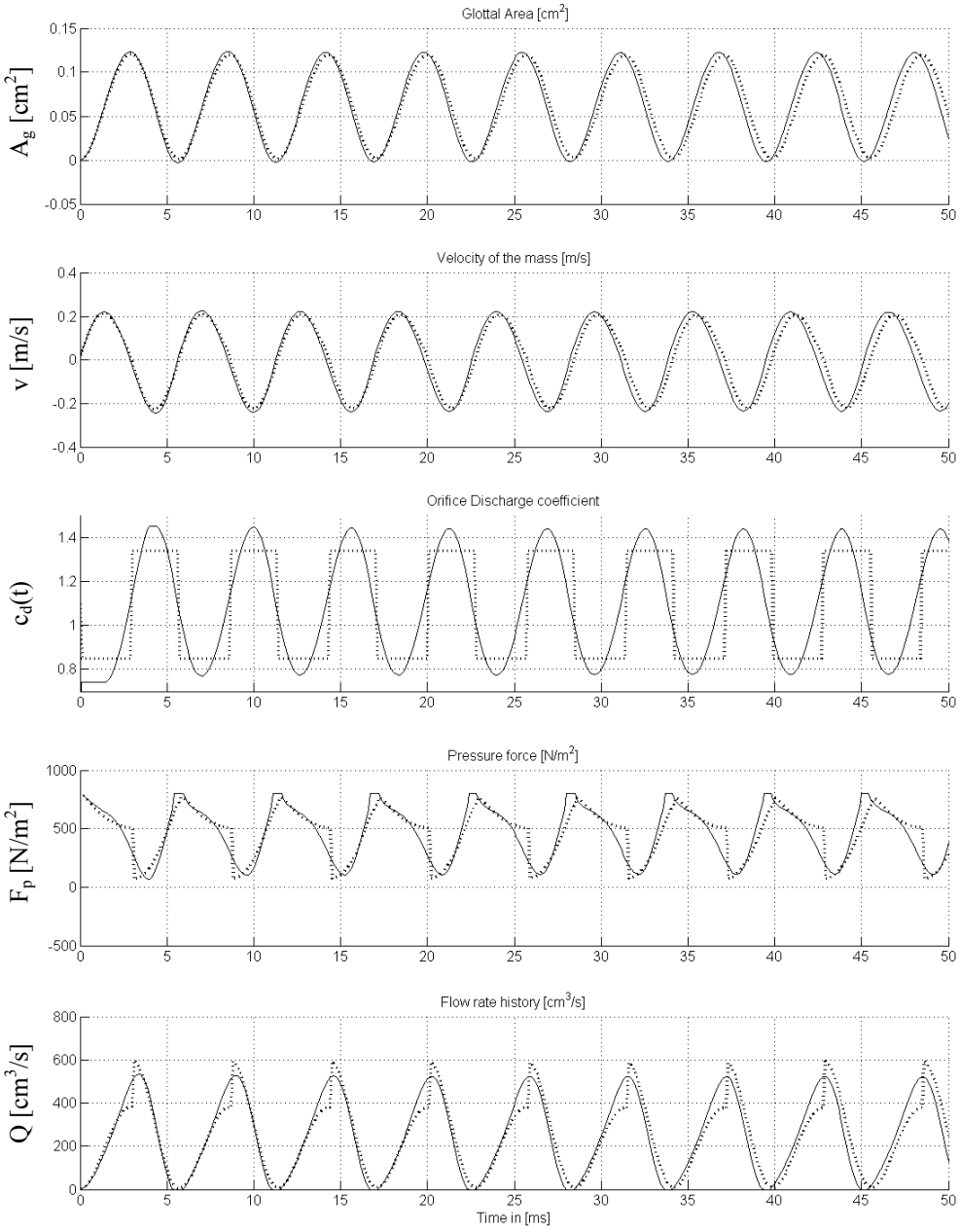


Fig. 6.1 Effects of the discharge coefficient on different variables.
-: Continuous $c_d(t)$ from equation (4.19); --: discontinuous $c_d(t)$ from equation (4.20).

The most significant variation is observed in the volumetric flow rate signal. As can be seen from equation (4.9), the orifice discharge coefficient affects directly this variable, making it continuous or discontinuous depending on the approach used. It is believed that a smooth variation of the discharge coefficient provides a more realistic transition between the opening (convergent) and closing phase (divergent). As explained in chapter 2, this natural transition is also observed in the "mucosal wave" propagated in the axial flow direction (streamwise). To reproduce this effect in a vocal fold model, more degrees of freedom are required. As explained by Titze in [48], the action of the mucosal wave introduces a skewing in the volumetric flow rate shape. This is normally evaluated through a direct comparison between the glottal area and the flow rate. As observed in Figure 6.2, the "smooth" discharge coefficient produces the same effect. The discontinuous $c_d(t)$ case does lead to skewing in the flow signal, as observed in Figure 6.1. For both time varying discharge coefficients, a steady state is reached after the first cycle. This is confirmed by the phase plane plot shown in Figure 6.3, where the mass velocity is plotted with respect to its displacement. A steady state is reached when the trajectory defines a closed curve, normally a circle. No differences are observed between the two different orifice discharge coefficient functions, since their displacement and velocity profiles are essentially the same. The results are comparable in order and structure with those from other studies [19] [42] [44]. Thus, the smooth function was used in all subsequent cases (unless otherwise stated).

A constant orifice discharge coefficient value of unity was also tested. As shown in Figure 6.4, the absence of a proper time varying c_d profile results in the absence of self-sustained oscillations. This behavior has been confirmed by several authors [10], [12], [48], where in absence of either a "mucosal wave" (or equivalent) or acoustic loading, the one-mass model is not able to exhibit self-sustained oscillations. In this case, the model behaves like a damped oscillator, where displacement and velocity of the mass reach their equilibrium states and the volumetric flow rate reaches its residual DC component. This is illustrated in Figure 6.5 where the decay is shown in a phase plot of the system.

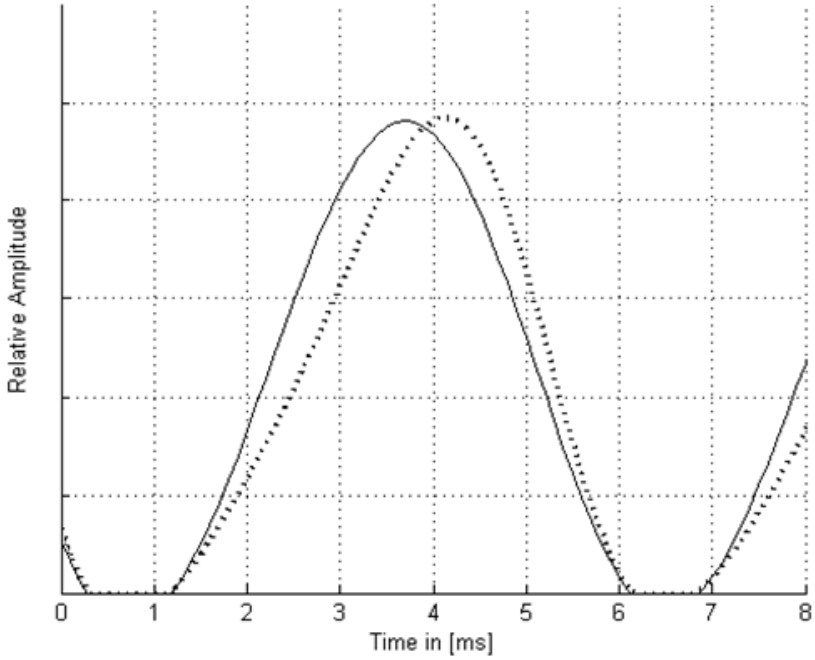


Fig. 6.2 Skewing introduced by the smooth variation of $c_d(t)$
-: Glottal area; --:Volumetric flow rate

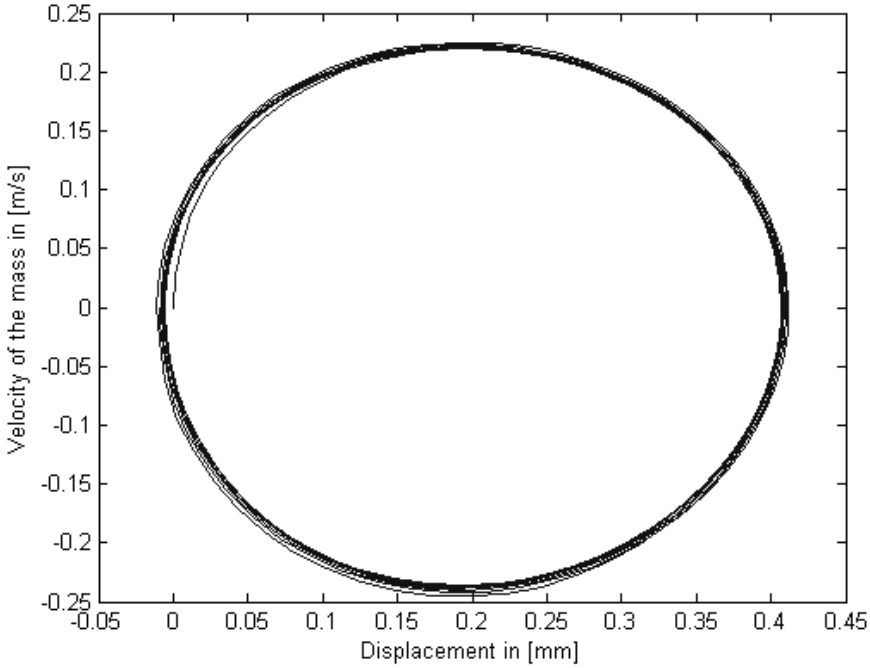


Fig. 6.3 Phase plane diagram of the oscillator.

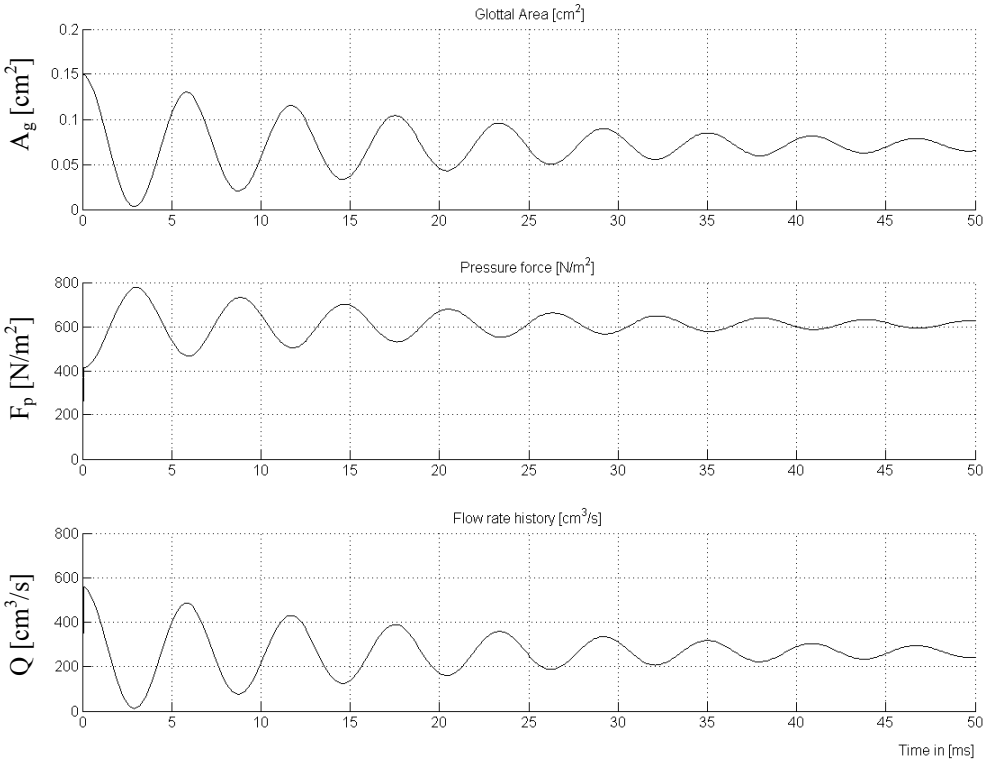


Fig. 6.4 Time history of different variables for $c_d(t) = 1$.

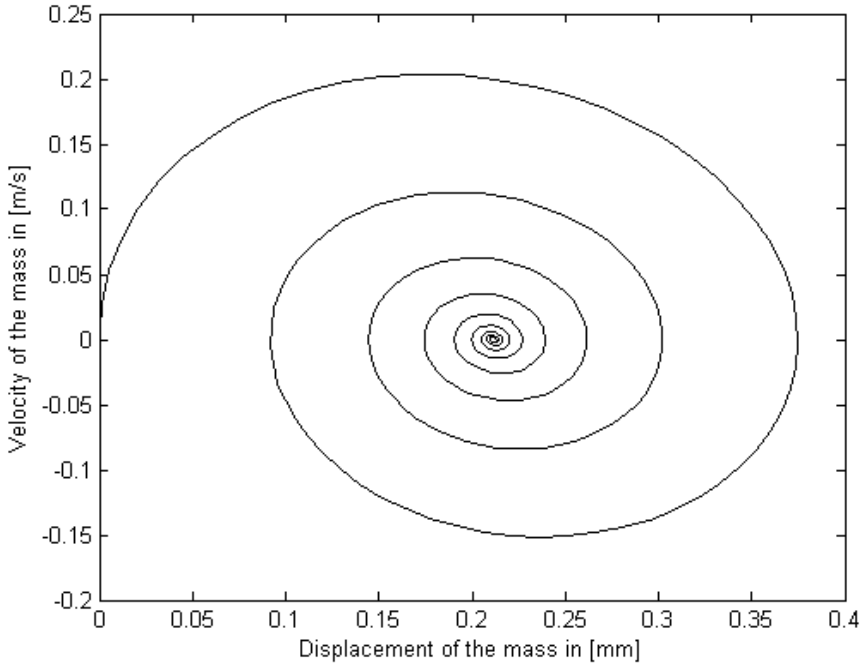


Fig. 6.5 Phase plot for $c_d(t) = 1$.

6.2 Comparisons with Fulcher's Model

The differences between this model and that of Fulcher [14] were investigated. As stated in equation (4.1), Fulcher's excitation is a fixed switching force that depends on the velocity of the folds. No collision forces are considered. An ideal volumetric flow rate with no discharge coefficient effects has been included in Fulcher's model. Figure 6.6 shows that there are significant differences between the two versions of the model. The increased driving forces and the absence of collision forces result in larger magnitudes and differences in the frequency of the oscillations.

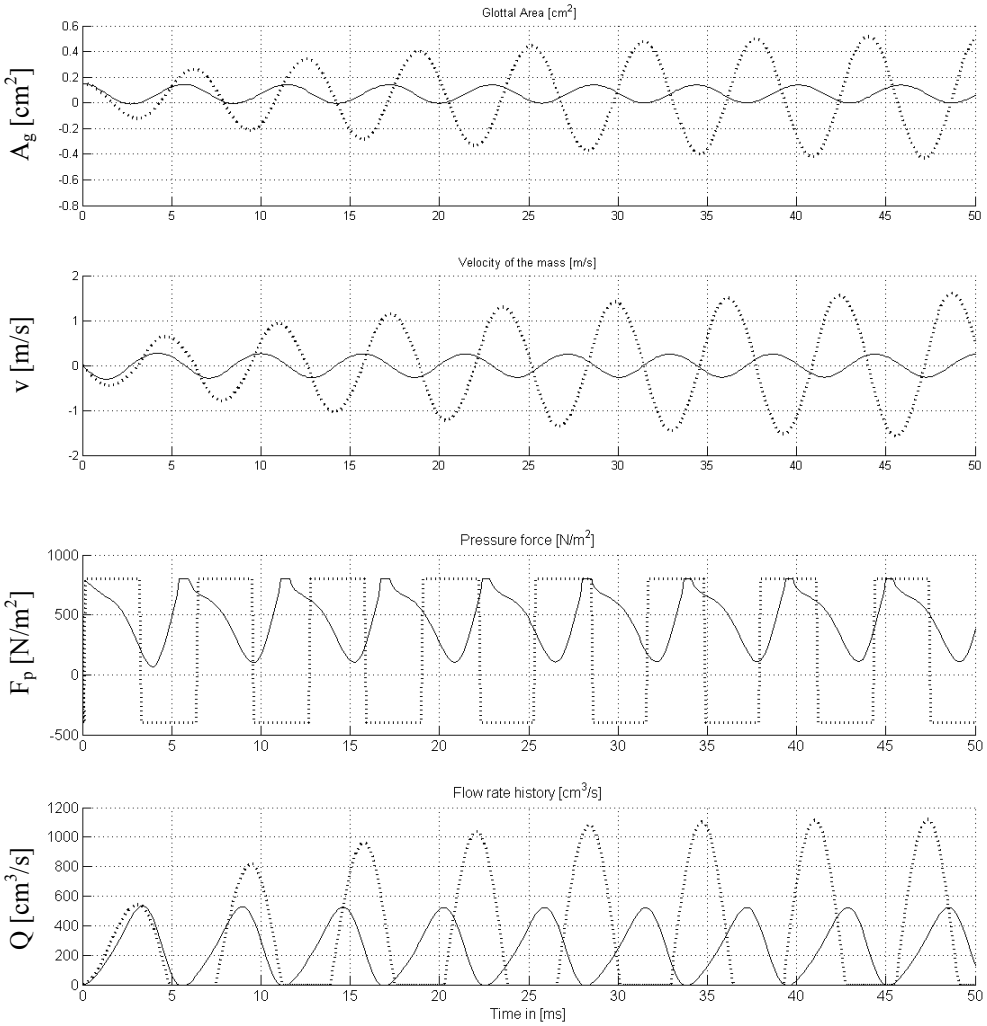


Fig. 6.6 Comparison with Fulcher's original model
--: Fulcher's model; -:Current model

6.3 Effects of Collision Forces

The effects of collision are significant. Fulcher's original model was used since it exacerbated collision effects, facilitating their study. Three forces act during collision: Hertz force, pressure force and a modified damping force. These forces were tested separately. Figure 6.7 illustrates the effects of collision forces. The increase in the damping ratio of the system (from 0.1 to 0.5) during impact is believed to be the most important factor. The amplitude of oscillation of the displacement and velocity of the mass are reduced and the pattern of oscillation is changed. An increase in the fundamental frequency and skewing of the velocity waveform are observed. This influences the behavior of other related variables, for example the volumetric flow rate. The effect of the pressure force applied to the part of the vocal folds that is not colliding is also significant. It affects the pattern of the driving force, which is a function of both velocity and displacement. This is seen in the case where the mass is in the closing phase (with negative velocity) during collision, where the pressure force is positive due to the closure in the glottal opening produced during collision. This change reduces the period of negative pressure force applied in each cycle, causing a reduction in magnitude for both displacement and velocity. Finally, the Hertzian force described in equation (4.24) is the least significant factor during collision. Its magnitude is approximately 4% of the added pressure force during collision. This force can be ignored without noticeable changes.

Figure 6.7 shows the change in several variables when collision forces are included. An ideal volumetric flow rate is also included for comparison. The inclusion of collision effects allows Fulcher's model to achieve results that appear more in line with previous studies in the field.

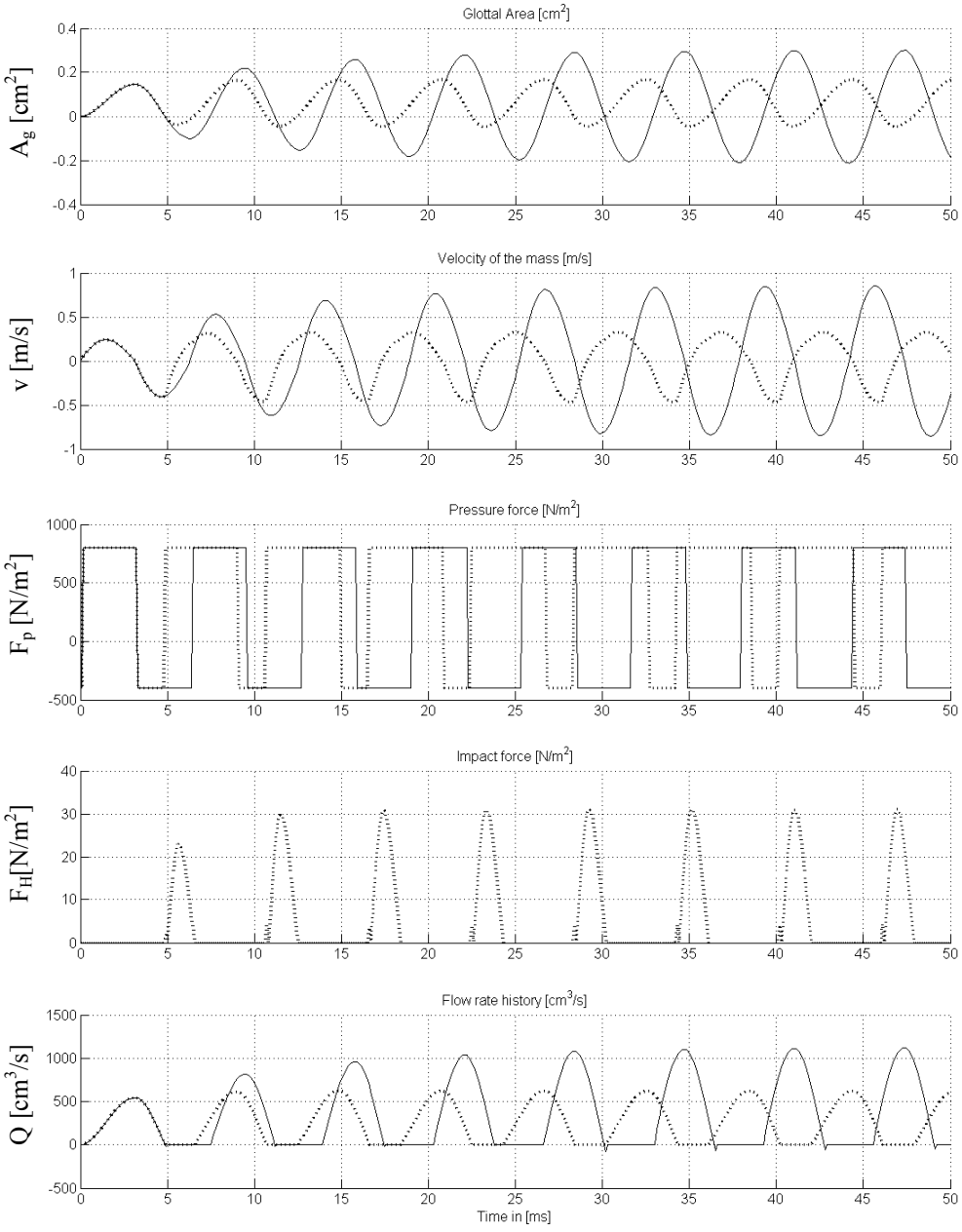


Fig. 6.7 Effects of impact forces on Fulcher's model.
--: With collision forces; -: Without collision forces

6.4 Effects of the Discontinuity at the Glottal Entrance

During oscillations, the profile of the model of the vocal folds presents minor discontinuities at the juncture with the structure that defines the transition with the subglottal tract. These discontinuities are shown in Figure 4.5 during oscillation. To evaluate the associated error, a comparison with a non discontinuous geometry was made. During oscillation, a continuous model changes its shape, which is more realistic but violates the one-degree of freedom assumption. Comparisons of the flow and the pressure force are shown in Figure 6.8. No significant differences are observed in the pressure force between the two cases. For the continuous orifice discharge coefficient, the differences are of the same order as for the discontinuous case, shown in Figure 6.8. Therefore, the error introduced by the discontinuous profile in the shape of the vocal folds does not introduce significant differences in the driving force of the mass.

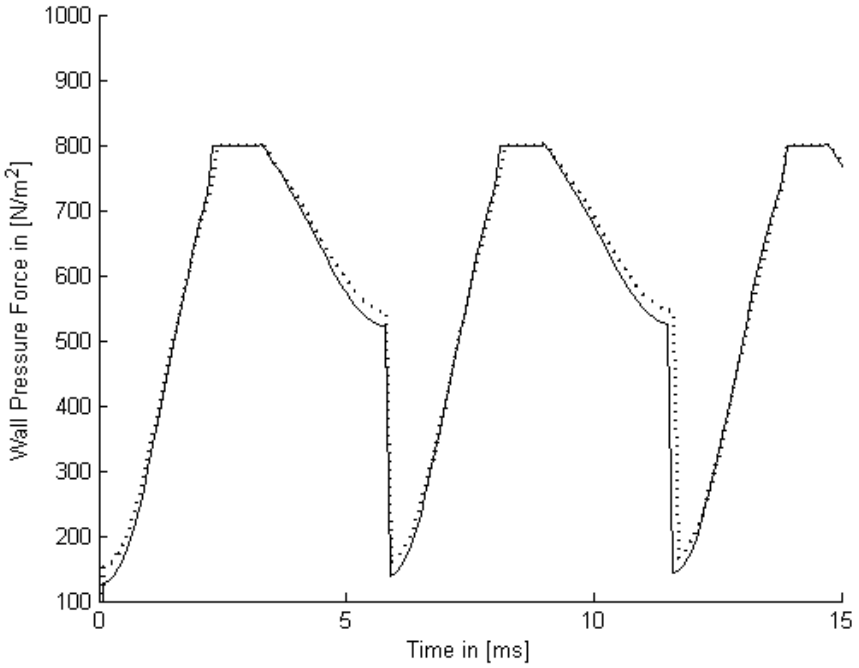


Fig. 6.8 Effects of geometry of the vocal folds model on the pressure force.
-: streamlined profile; --: discontinuous profile.

6.5 Pressure Distribution along the Wall

The pressure distribution along the wall regulates the pressure force that drives the vocal folds. In normal phonation, this distribution changes significantly between the opening (converging shape) and the closing (diverging shape) portion of the cycles. The total instantaneous pressure force is obtained by integrating the pressure distribution along the axial axis (flow direction). The pressure distribution shows the transition between the upstream pressure and the downstream pressure. This transition is clearly seen in Figure 6.9, for a fixed orifice discharge coefficient $c_d = 1$. It can be also observed that the centerline pressure is a good approximation of the wall pressure. Figure 6.10 shows a different situation, where convergence and divergence in the geometry of the folds are introduced by setting different orifice discharge coefficient values, as it happens in the model during oscillation. The convergent mode is associated to the upper lines for both centerline and wall pressures, where $c_d = 0.85$, and the divergent mode is associated to the lower lines, where $c_d = 1.34$. The centerline pressure is approximately equal to the wall pressure in each mode. Most importantly, discontinuities are observed at the glottal opening. These discontinuities are artifacts, due to the fact that the discharge coefficient is imposed without a change in orifice geometry. The ideal volumetric flow rate at the outlet is multiplied by the discharge coefficient to yield a more realistic profile between opening and closing. However, following Bernoulli's flow equation this also introduces a discontinuity in the pressure at the outlet. Hence, an inconsistency arises from the simultaneous use of Bernoulli's obstruction theory and Bernoulli's flow equation in this case. The pressure discontinuity is observed analytically from the wall pressure described in equation (4.16). It is noticed from that derivation that, at the entrance of the folds, the wall velocity is almost the same as the upstream pressure, independently of the discharge coefficient. However, at the outlet the effect of the discharge coefficient modifies significantly the value of the wall pressure, producing the discontinuities in pressure that appear in Figure 6.10. Although discontinuities are observed in the distribution, the integration along the axial direction yields values that are comparable with other models of phonation [19] [44] [49]. It can be argued that pressure adjustments are assumed to happen after the glottal opening, which justifies the presence of discontinuities. This type of pressure adjustment after the glottal opening has been noted in different experimental studies [52],[25],[58],[36].

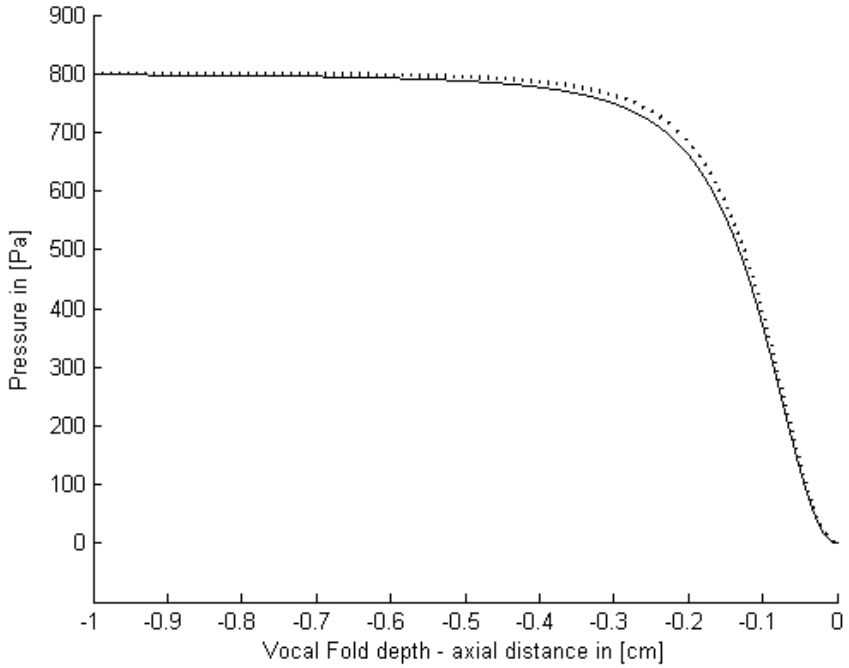


Fig. 6.9 Pressure distributions for wall and centerline pressures. Ideal case ($c_d(t)$=1).

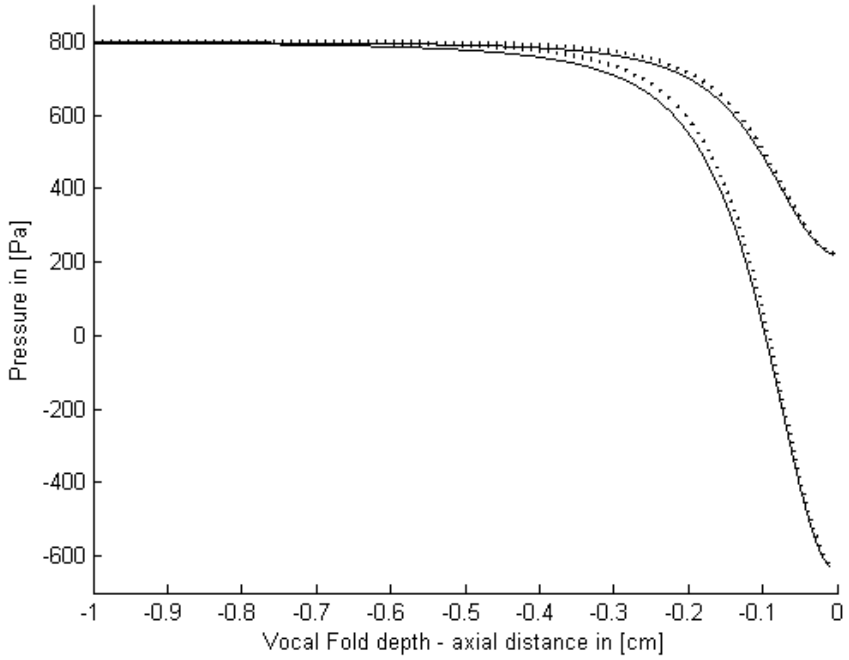


Fig. 6.10 Pressure distributions for wall and centerline pressures. Convergent orifice case (upper lines) and divergent orifice case (lower lines). -: p_{wall}; --:p_{cl}

Recent measurements of pressure distributions performed by Li and Scherer [36] confirm the typical patterns observed in phonation for a number of convergent and divergent shapes. It can be inferred from that data that for a fully diverging shape, the total pressure applied to the folds is nearly zero, which is in fact the value of the supraglottal pressure. For a fully converging shape, the total pressure force obeys Bernoulli's equation, with values that are around a 65% of the subglottal pressure. Similar trends were also observed in the "cover model" [44]. Here, the convergent mode uses a Bernoulli flow to compute the driving force, and the divergent mode uses the supraglottal pressure as the driving force. The same pattern is present in the pressure force of the model developed in this thesis. This is shown in Figure 6.11, where divergent and convergent modes are clearly observed when a discontinuous orifice discharge coefficient is used. Comparison of this pressure force with other reported results [19] [44] [49], yields a good match. A reconstruction of the pressure force from the cover model for a non-loading case, as reported in [44], is shown in Figure 6.12. Comparing Figure 6.11 and 6.12, it can be noticed that the converging mode yields the correct decay slope. The divergent mode introduces a significant reduction, but not a null. This confirms, as several authors have noticed, that the pressure force does not have to be negative for self-sustained oscillations to occur. Note that if the discharge coefficient is slightly increased, negative pressures could be achieved. In order to keep the model consistent with the experimental data reported by Park [32], this was avoided. Figure 6.12 shows that not only the forces have similar structures, but also that different assumptions were made in that model. Since the cover model has more degrees of freedom, the pressure force during closure was set to the average of subglottal and supraglottal pressures. For the particular case shown, a larger collision time was observed in the cover model.

Finally, it can be argued that although a discontinuity in the pressure distribution appears in the glottal opening, the total force supplied to the folds is realistic in the present model. A pressure recovery is assumed following glottal opening. However, this recovery does not affect the pressure force that drives the folds.

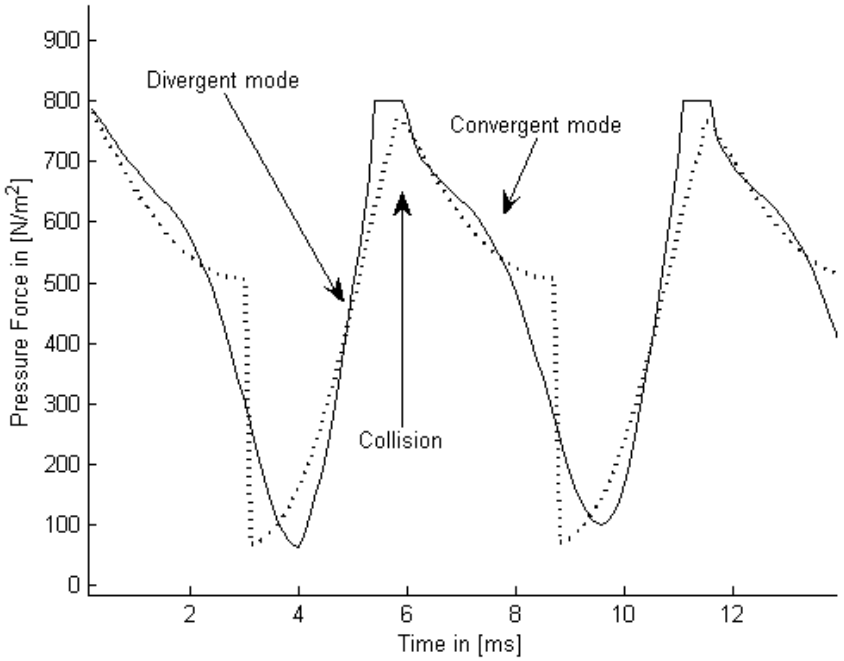


Fig. 6.11 Pressure time history of the driving force for convergent and divergent modes. No acoustic loading. -:Continuous $c_d(t)$; --:Discontinuous $c_d(t)$

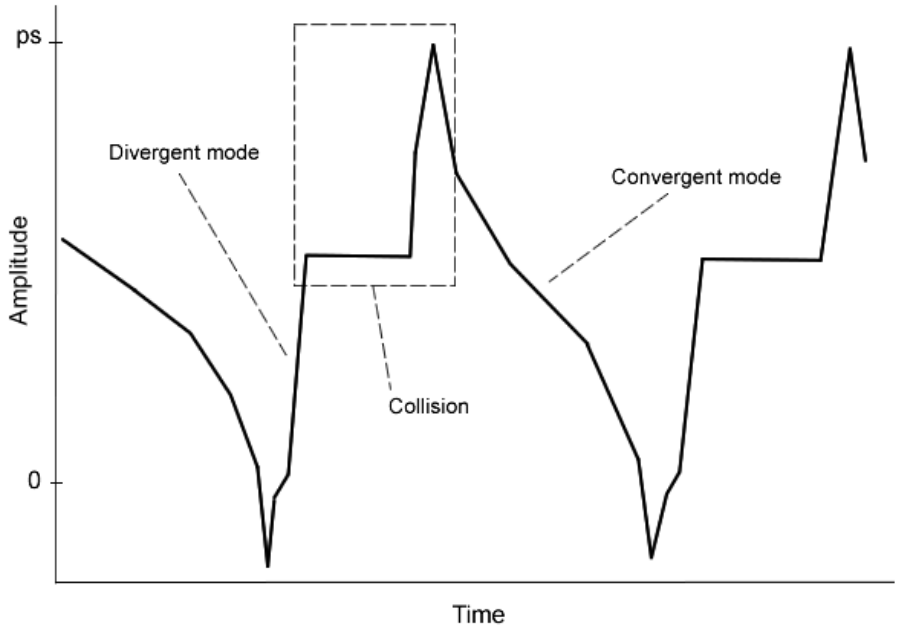


Fig. 6.12 Simplified time history of the driving force for the cover model [44] in absence of acoustic loading.

6.6 Derivative of the Volumetric Flow Rate

The volumetric flow rate is one of the most important parameters in the source model. It constitutes the input signal that excites the vocal tract. In addition, the derivative of the volumetric flow rate signal is useful to predict the behavior of acoustic loading. As explained in detail in [10] and [39], for a simplified inertive supraglottal loading, the radiated pressure at the mouth opening can be approximated by the derivative of the glottal waveform (using a monopole source concept). When a fully coupled system is tested, the pressure immediately downstream of the vocal folds should also have a similar shape. This constitutes a good baseline. Figure 6.13 shows that the derivative of the volumetric flow rate vs. time is similar to that reported by Stevens in [39]. It is relatively constant during closure, reaching a peak value and then becoming negative after closure. Since the flow rate is controlled by the displacement of the folds, its derivative is in phase with the velocity of the mass, hence the pressure force. If the downstream pressure has the same behavior as the derivative of the flow, it contributes to get self-oscillation in the system. The downstream and upstream pressures are presented in chapter 7, where a fully coupled system is tested.

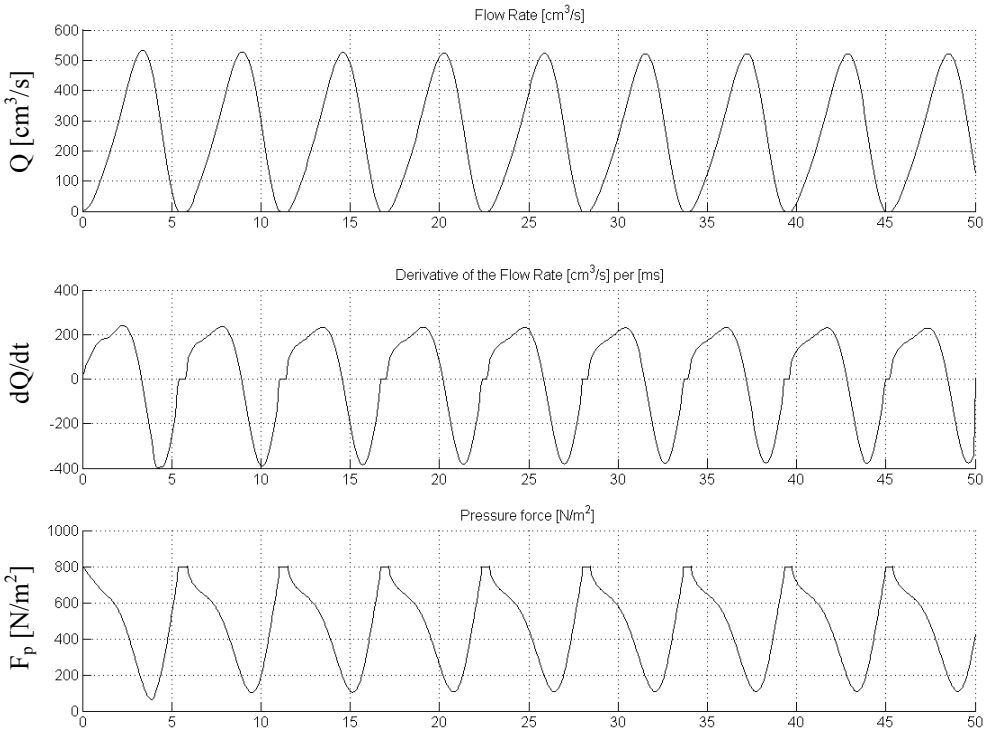


Fig. 6.13 Volumetric flow rate, derivative of the flow rate and pressure force.

6.7 Stability of Volumetric Flow Rate for the Coupled System

The behavior of the source model in presence of acoustic loading was investigated. When the sections of the tracts connected with the source model are narrow, they have impedances comparable with that of the source. This produces a favorable environment for noticeable coupling between the source and the tracts. In practice, narrow sections lead to large acoustic pressures. These large acoustic pressures can in fact produce problems in the model. Note that the flow velocity and volumetric flow rate are originally computed from the quadratic equation (4.5b), where only one solution was considered. Due the large acoustic pressures observed in some coupled cases, the actual flow velocity (hence, volumetric flow rate and wall pressure) cannot be computed directly using equation (4.7) since the root square of the flow velocity produces complex solutions that have no physical interpretation. Errors in the volumetric flow rate lead to errors in the radiated acoustic pressures, hence the complete oscillator. Unfortunately, most of loading profiles in this thesis produce those undesired effects on the source model. Figure 6.14(a) shows how an MRI vowel /A/ drives the flow rate defined in equation (4.9) numerically unstable. In this case, no variation of the discharge coefficient was considered. A different behavior in Figure 6.14(b) is observed, where a different technique is used to compute volumetric flow rate in presence of acoustic coupling.

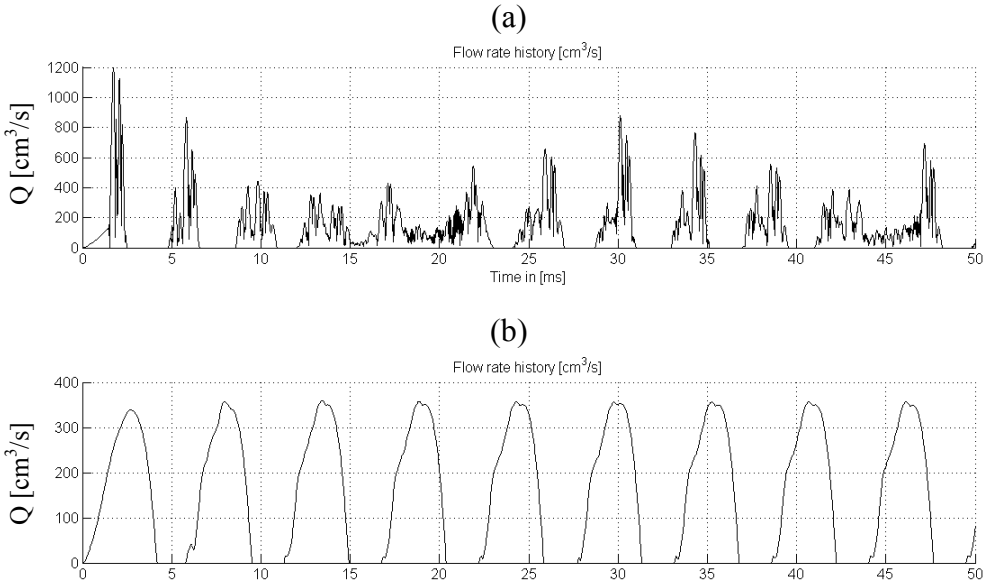


Fig. 6.14 Flow rate computation when acoustic loading is present.
(a): Original case ; (b): Titze's volumetric flow rate

Note that the numerical stability problem arises when the root square of the flow velocity is taken from equation (4.5b). Fortunately, this is not a problem in the pressure force, since that magnitude is later squared, eliminating the above problems. Therefore, a new expression for volumetric flow rate is required to compute the acoustic pressures when the system is coupled. Earlier studies [12], [50], [47], [42] have suggested different approaches to compute volumetric flow rate. Flanagan [12] designed a technique to compute flow rate using glottal flow resistances and inertances, but it is not directly applicable for irregular geometries. The best approach is given by Titze [47]. Note that this volumetric flow rate uses both subglottal and supraglottal tract areas and a pressure recovery factor in its derivation. These produce some differences in the peak values compared with the simplified expression from equation (4.9). For the non-loading case, these differences are shown in Figure 6.15. To force Titze's volumetric flow rate in agreement with the effective source model, the orifice discharge coefficient was multiplied to Titze's flow expression, reflecting the effects of converging and diverging geometries, as in equation (4.9). Using this idea, the expression used to compute the volumetric flow rate is

$$Q = c_d(t) Q_{titze} \quad , \tag{6.1}$$

where $c_d(t)$ is the orifice discharge coefficient as it was proposed in equation (4.20) and Q_{titze} is the flow rate described in equation (4.21). Figure 6.14 shows the differences between the original and Titze's approach for a non-loading case. The flow rate obtained using Titze's scheme is slightly larger and does not exhibit skewing. By introducing the effects of the discharge coefficient, the skewing becomes noticeable and the differences in magnitude are larger. The same occurs when the ideal flow is multiplied with the orifice discharge coefficient, as in the original development of the flow rate. It is important to note that in Titze's flow rate, the non-loading case assumes that the areas of the subglottal and supraglottal tract are infinite. This increases the peak values, as it can be inferred from equation (4.21). When real tract areas are supplied into this equation, the volumetric flow rate normally presents values that are lower than those for the non-loading case.

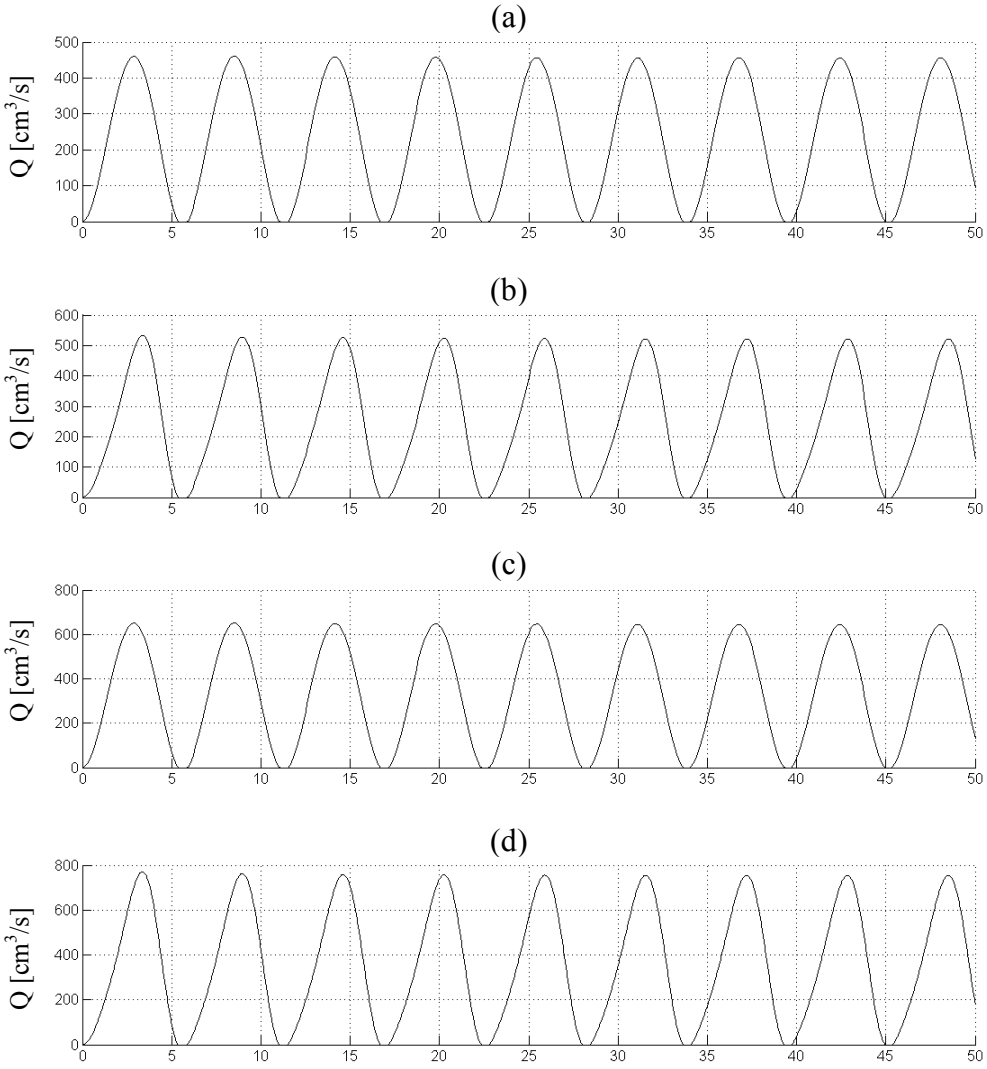


Fig. 6.15 Time history for different flow rates in absence of acoustic loading. (a): Ideal flow; (b): Original case: Ideal flow times $c_d(t)$; (c): Titze's original flow; (d): Titze modified: Titze's original flow times $c_d(t)$

7. RESULTS OF THE COUPLED MODELS

This chapter contains the main results of this study, aimed at quantifying the factors that govern the self-oscillations of the vocal folds. Different acoustic loading profiles were tested to evaluate their influence on the source model for a simplified one-mass model, unable to achieve self-sustained oscillations without acoustic loading, and the effective one-mass model, able to reach self-sustained oscillations without acoustic loading.

7.1 Single Degree of Freedom Model with Constant Orifice Discharge Coefficient

It has been suggested by Flanagan [13], Rothenberg [35], Titze [48], Lucero [27] and others, that single degree of freedom models of the vocal folds are not amenable to self-oscillation without acoustic loading. As discussed in chapter 2, these authors modeled the vocal tract as an inertive load, assumption that over-simplifies the acoustic properties of the vocal tract. In addition, the effects of the subglottal tract were simplified or suppressed. More recent studies [49], [44] have shown the effects of more accurate loading profiles in a "cover model" (three mass model). Although these studies provide insights on the importance of acoustic loading, they are based on multi-mass models that are able to perform self-sustained oscillation without acoustic loading. Since Flanagan's one-mass model failed to perform self-sustained oscillations without acoustic loading, little attention has been paid to one-mass models of the vocal folds. No clear guidelines on what type of acoustic loading could produce self-sustained oscillations have been suggested since.

The interaction between the source model and realistic shapes of the subglottal and supraglottal tracts was studied. To perform these tests, the source model developed in chapter 4 and discussed in chapter 6 was used. No variation of the orifice discharge coefficient was included and it is set to have a value of one. As shown in Figure 6.4, for a no loading condition, the model was unable to achieve self-sustained oscillation. The results obtained in this section similar to those reported by Flanagan in [13] and [12], but

use different pressure force and acoustic pressures. Moreover, the technique used in this thesis to compute the wall pressure (driving force) and the acoustic pressures (loading) are considered an improvement over Flanagan's original version.

7.1.1 Coupling using Realistic Loading Profiles

In sections 7.1.2 and 7.1.3, the effects of acoustic coupling on laryngeal self-sustained oscillations were investigated. Two realistic and complete loading profiles for the subglottal and supraglottal tracts were used. All the features of the wave analog were included in these tests. The vocal tract was set to the MRI vowels /i/ (close front unrounded vowel – IPA 301) and /A/ (open back unrounded vowel - IPA 305). These shapes were selected since they represent two completely different acoustic loads. In addition, comparable loads were used in previous studies [12] [13] [44] [50]. The time varying orifice discharge coefficient was set to unity for all times. The volumetric flow rate used to compute the radiated acoustic pressure was based on Titze's approach, as described in equation (4.21).

7.1.2 Coupling with the Vowel /i/

As explained in chapter 2, the inertive behavior of the vocal tract is based on the assumption that the fundamental frequency is considerably less than the first formant. The acoustic loading defined by the MRI vowel /i/ has a first formant at 225 Hz and a second at 2486 Hz. The fact that the first formant is relatively close to the fundamental frequency (in theory 160 Hz) makes this case particularly interesting. The vowel /i/ presents a loading that appears to be closer to a resistive loading than an inertive one. On the other hand, the section connected to the glottis is 0.33 cm^2, which is considered very narrow. This produces a high vocal tract impedance and strong coupling. In practice, this also implies that the acoustic pressure in the tracts is large. On the other hand, the subglottal tract has its first two resonances at 613 Hz and 1341 Hz and a cross-sectional area of 2.7 cm^2 at the glottis. As discussed in section 7.1.2, it is not clear that the subglottal tract impedance can be considered a purely inertive loading. However, as discussed by Titze in [49], both the supraglottal and subglottal impedances act together in a fully coupled system. In this case, their combined effects could be neutralized.

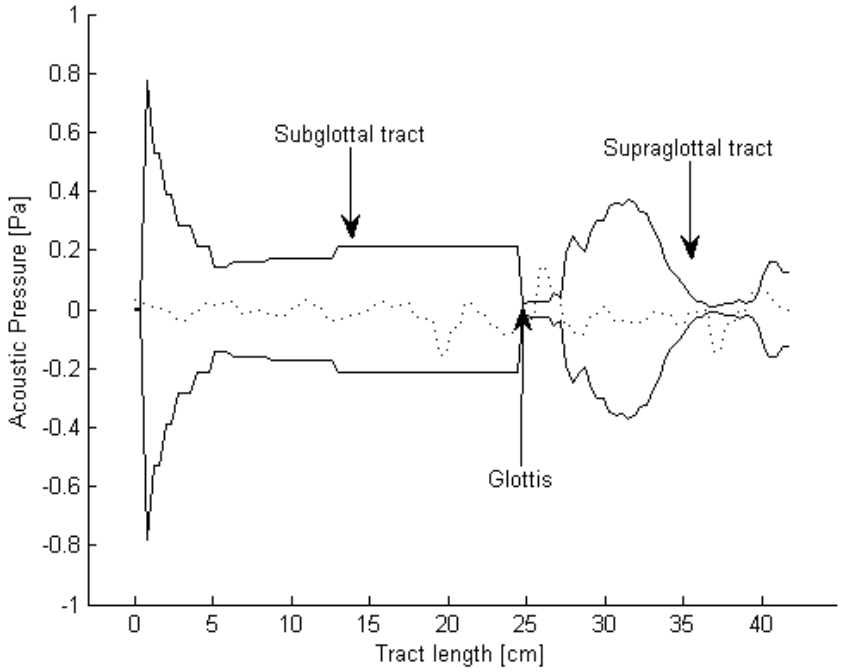


Fig. 7.1 Complete system using the MRI vowel /i/.
Snapshot of the acoustic pressure amplitude vs. position at t=5ms.

The results of the first simulation are presented in Figure 7.2. The system was able to perform self-sustained oscillations with the presence of a realistic loading. This matches the observations of Flanagan in his original one-mass model [13].

The oscillation reaches a steady state after the second cycle. Collisions are larger than for the non-loading case. The glottal opening is the least affected of all variables. A decrease in fundamental frequency from 180 Hz in the non-loading case to 170 Hz in this case was observed. This trend has been observed in other models [42]. Note that the fundamental frequency is now closer to the theoretical value of 160 Hz for the linear spring-mass system. The velocity does not change with respect to the non-loading case, being again almost sinusoidal. Note that the volumetric flow rate changes, although the glottal opening has the same structure as in the non-loading case. Compared with the non-loading case, the amplitude of the flow rate is reduced, reaching peak values similar to the ideal flow (see Figure 6.15a).

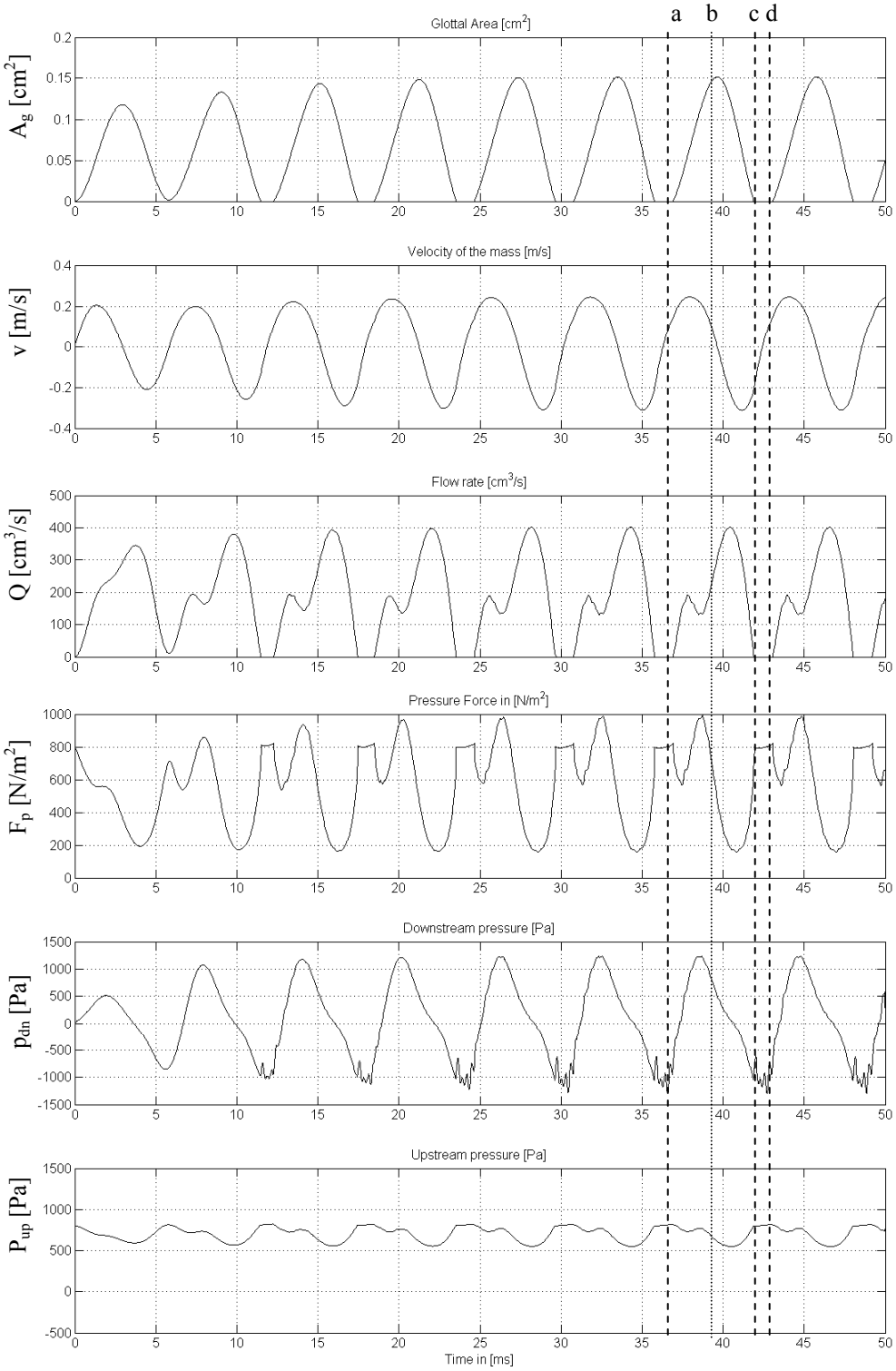


Fig. 7.2 Non ODC coupled system: MRI /i/. (a): Initial glottal opening; (b): Maximum displacement; (c): Initial collision; (d): Initial glottal opening for a new cycle.

The waveform of the volumetric flow rate is slightly skewed rightward in time, which has been explained in [48] and [35] as an effect of the inertia of the combined tract impedances. A depression is also observed in the flow rate during the opening phase. This depression has been also noted by Flanagan in [12] and [13]. Note that the volumetric flow rate is a function of the transglottal pressure (p_{up}-p_{dn}). Thus, the large peak in the downstream pressure causes the depression in the opening phase of the flow rate. The peak in the downstream pressure is likely to be associated with the pressure oscillation produced by the low frequency of the first supraglottal formant. Since the variation of the upstream pressure is considerably less than the downstream pressure, the subglottal tract plays a minor role in the production of this depression. Cases without the subglottal tract, however, have shown an increased depression in the volumetric flow, suggesting that the subglottal tract plays a role in the alteration of the flow rate.

As shown in Figure 7.2, the pressure force is in phase with velocity, especially during the open phase. This positive energy transfer is the main factor that maintains self-sustained oscillations. Note that the presence of acoustic loading did not make the pressure force negative. Overall, the pressure force remains similar to that in the non-loading case. To evaluate the degree of coupling between the source and the tracts, it is necessary to compare the pressure force (the main driving force) with the downstream and upstream pressures. Note that these pressures are similar in form but opposite in phase. During the opening phase, the downstream pressure reaches a peak value (which explains the depression in the flow rate) and then decays almost in phase with the pressure force and the mass velocity, with a slight time lead. During collision, faster oscillations in the downstream pressure occur, mostly associated with the high second formant frequency of the supraglottal tract. The upstream pressure exhibits less variation over each cycle. A clear coupling between source and subglottal tract can be observed during collision. This was expected since, during this phase, the force is governed by the upstream pressure. It is necessary to note that the upstream pressure is slightly reduced compared with the lung pressure. The pressure oscillates around a value that is below the lung pressure. Overall, the behavior of the variables aforementioned match the ones reported in other studies [42] [44] [1].

7.1.3 Coupling with the Vowel /A/

The second case of interest is the MRI vowel /A/ load, which has a first formant at 786 Hz and a second at 1147 Hz, and a cross-sectional area of 0.45 cm^2 at the glottis. The first formant is much larger than the frequency of oscillation, and so the load impedance is inertive. The subglottal tract is the same as in the previous case. A representation of the system and its sound field is shown in Figure 7.3.

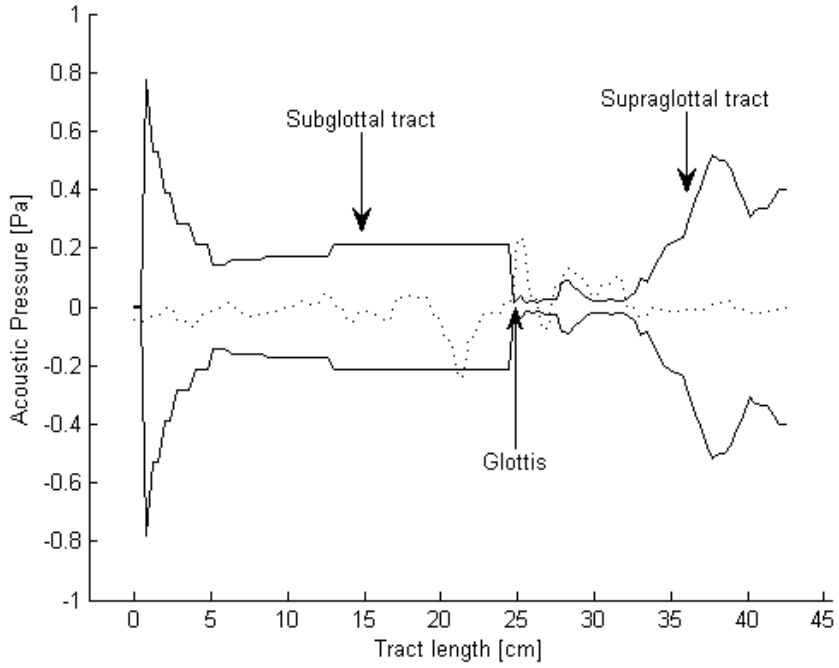


Fig. 7.3 Complete system using the MRI vowel /A/.
Snapshot of the acoustic pressure amplitude vs. position at t=5ms.

Results of this simulation are presented in Figure 7.4, As for the previous case, the oscillation reaches a steady state after the second cycle. Collisions are strong compared with the non-loading case. The glottal area was again the least affected of all variables. The fundamental frequency increased from 180 Hz in the non-loading case to 190 Hz in the case with loading. This difference is attributed to the more pronounced effects of the supraglottal loading and collision.

The main differences with respect to the previous case are observed in the volumetric flow rate signal, and the coupling between the pressure force and the downstream pressure. The volumetric flow rate reaches similar peak values, but is more

skewed rightward in time. This is explained by the increased inertance in the supraglottal tract due to the higher first formant frequency. A ripple during the opening phase is observed, in place of the large depression observed in the previous case. As before, this is produced by fluctuations observed in the downstream pressure (zooming in). These fluctuations are likely to be associated with the pressure oscillation produced by the higher frequency of the first supraglottal formant. As for the vowel /i/, the reduced variation in the upstream pressure suggests that the subglottal tract plays a minor role in the production of this ripple. Tests without subglottal, however, have shown that this ripple is reduced by the effects of the subglottal tract. This means that although it does not produce the ripple, it influences its magnitude. Previous studies have reported a slightly increased flow rate ripple for this vowel. Yet, the volumetric flow rate closely resembles the ones reported in [1] [13] [19] and [44], in structure and magnitude.

It can be observed from Figure 7.4 that the pressure force is again clearly in phase with velocity, except during the collision phase, where other forces also contribute. Compared with the previous vowel, a higher degree of coupling with the supraglottal tract is observed. During the open phase, the downstream pressure has almost the same structure than the pressure force, for which the system is highly coupled with the tract. As before, the downstream pressure drives the force during collision.

The upstream pressure appears to be less coupled with the source than in the previous case, showing also less fluctuations. Nevertheless, a small ripple from the effects of the supraglottal loading on the flow rate can be observed. The upstream pressure remains as before, with small oscillations and out of phase with the downstream pressure. Note again that the upstream pressure oscillates slightly below the lung pressure.

It can be concluded from sections 7.1.2 and 7.1.3 that the coupling between the source model and the tracts is qualitatively similar to that reported in previous studies of non-linear models of phonation. Minor differences are believed to be due to different flow models, loss factors and tract designs.

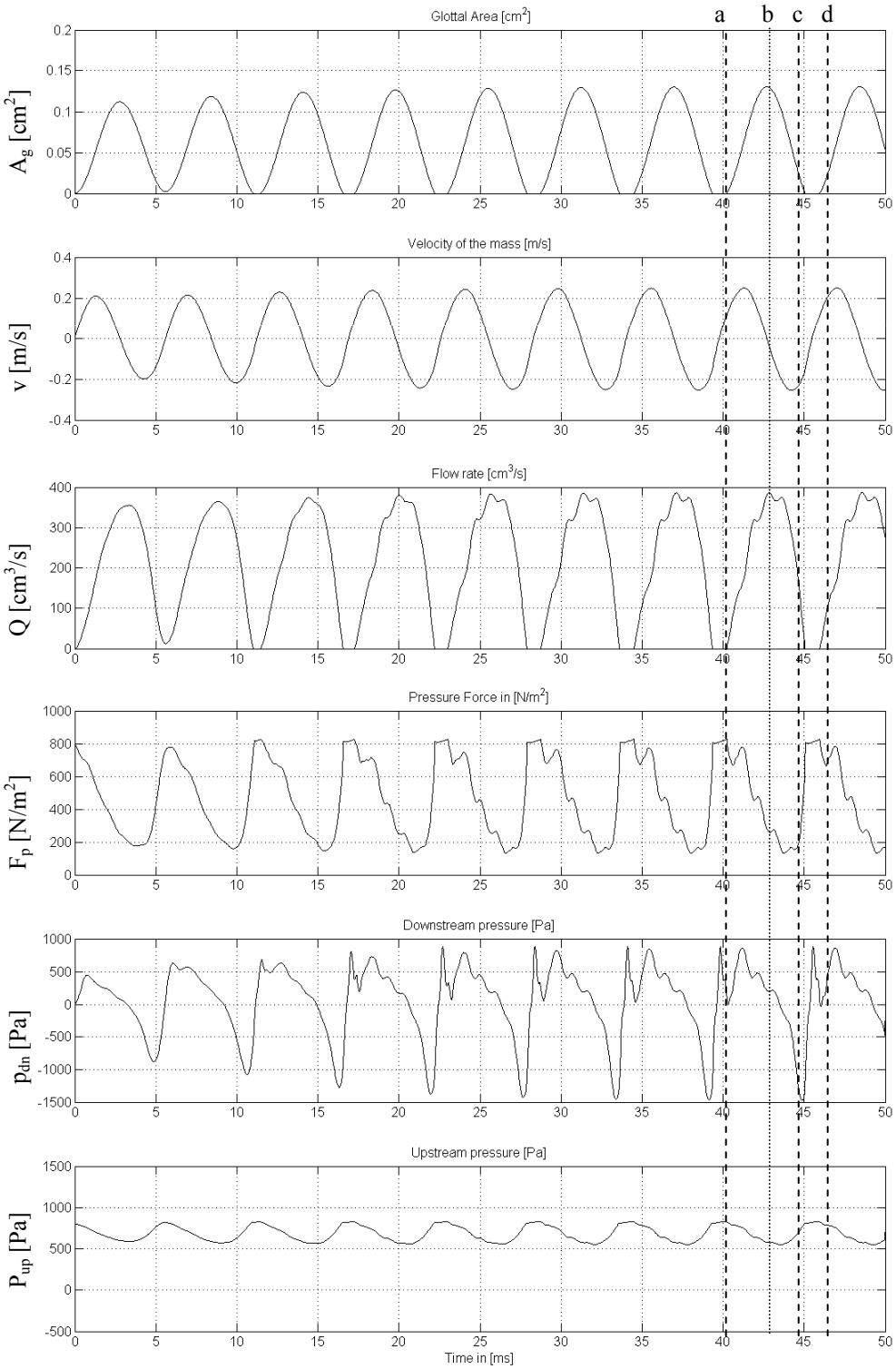


Fig. 7.4 Non ODC coupled system: MRI /A/. (a): Initial glottal opening; (b): Maximum displacement; (c): Initial collision; (d): Initial glottal opening for a new cycle.

7.1.4 Effects of the Tracts on Oscillatory Conditions

Several other loading profiles were used to evaluate the conditions that yield self-sustained oscillations in the one-mass model. An infinitely long tube, a simple tube, subglottal tract designs and the same two MRI vowels used before for the vocal tract. For simplicity, only phase plots are presented in each case. Figures 7.5, 7.6 and 7.7 show phase plots for different loading profiles. Initially supraglottal and subglottal tracts were tested separately and later together. Recall that the effects of the time varying discharge coefficient are not considered here. Also note that the acoustics of the supra and subglottal tracts differ in the loss factor.

Figure 7.6 shows the phase plots for different supraglottal loadings. Infinitely long tubes and uniform long tubes (with low formant frequencies) do reach a steady state. For a uniform tube of length L=17 cm (or less) and an area of A=1 cm^2 (or less), self-sustained oscillations are reached in the model. The cross-sectional area is more sensitive in this case, i.e. short tubes needed the same area to reach a steady state. Length plays a different role in this case. For the same area, larger tubes did not lead to self-oscillation. Based on the inertance theory from Titze [48], these data suggest that when the first formant frequency of the load is considerably higher than the natural frequency of the oscillator, the impedance of the tract is inertive and the acoustic pressure is in phase with the velocity of the mass. This leads to a favorable condition for self-sustained oscillations. Note that, in an infinitely long tube, the downstream pressure is in phase with displacement and not velocity, which explains the inability to reach a steady state in that case. For different supraglottal loading profiles, the downstream pressure varies in phase from that of the displacement to that of the velocity. The data confirms that not only the tube length is important to achieve this goal, but also the tract area. Smaller tract areas produce a better impedance match between the source and the tract. Titze and Story [49] described that the impedance of the vocal tract is mainly controlled by the ratio between the length and the area of the first portion of the tract, the epilarynx.

Finally, the data presented in Figure 7.6 confirm the importance of the inertance of the vocal tract in self-sustained oscillation. However, this theory cannot be extended to the subglottal tract. It is interesting to note that no steady state is reached if a supraglottal load that produces self-sustained oscillation is moved at an upstream location. The reflections that are fed into the source model are now opposite in phase with velocity. This stresses the different character of the subglottal impedance.

Figure 7.7 shows the phase plots for different subglottal loadings. The infinitely long tube does not reach a steady state and decays as rapidly as for the supraglottal and the non-loading cases. Note again that a load that produces self-sustained oscillations as supraglottal load, does not allow a steady state to be reached when used as subglottal load. This is the case of the uniform tube of 17 cm of length and an area of $1cm^2$. Setting the subglottal tract as a simple uniform tube with radiation impedance according to Flanagan [12], neither the area nor the length appears to contribute significantly to self-sustained oscillations. The length, however, appears to be correlated with the steady state or phonation threshold. Note that when the boundary conditions are modified and a closed boundary is used instead of an open one (with radiation impedance), a steady state is easily reached using the same loading profiles as for the supraglottal case. This is expected since boundary conditions affect the behavior of the pressure in the tract, hence, its impedance. It can be argued that when a closed boundary is considered in the subglottal tract, its impedance behaves like the open tube impedance in the supraglottal tract. That explains the fact that longer lengths and a close boundary do not contribute much to the establishment of a steady state in this case.

To simulate the acoustics of subglottal tract in models of phonation, either simple or shaped open tubes with some radiation impedance, closed tubes with a specific reflection coefficient, or tubes with an expansion chamber or "semi-lung" termination have been used. A case resembling a "semi-lung" was also tested. The behavior resembled more closely that of open-ended tubes. This is believed to be due to both the boundary conditions and the large acoustic attenuation in the subglottal tract. Special care must be used in experiments involving excised larynges or synthetic models to account for the effects of the subglottal tract. The lack of an appropriate attenuation factor or tract design can modify the boundary and propagation conditions, distorting the measurements. More work is needed to obtain accurate sub-glottal impedance data, and determine its influence in phonation.

Note that the subglottal tract designed by Story in [41] has an open boundary. An evaluation of its transfer function was performed in chapter 5, where it was noticed that it did not include the second formant. Figure 7.6(e) shows that this type of loading does not lead to self-sustained oscillations. Although its behavior is similar, the tract design used in this thesis is slightly more favorable than Story's tract design. Recent studies [31] have reported that the subglottal tract is in fact favorable and required to reach self-sustained oscillations. This supports the tract design used in this thesis. More work upon on an

appropriate subglottal tract design using a wave reflection analog is suggested for future research.

Figure 7.8 shows the phase plots for different supraglottal loads when the subglottal tract design used in this thesis is also present. The cases resemble closely the behavior of the supraglottal loading observed in Figure 7.5. The effects of both tracts appear to be combined. Although the patterns are clearly dominated by the supraglottal tract, the presence of the subglottal tract can be observed in the slight reduction of the oscillation compared with the theoretical non-subglottal case. This implies that the two cases are comparable. This observation can be confirmed from the development of the source model in chapter 4. Note that to make Bernoulli's equations consistent, a default subglottal cross-sectional area of $3cm^2$ was used. Thus, the "non-loading" case is in fact an infinitely long tube with, which explains the aforementioned similitude.

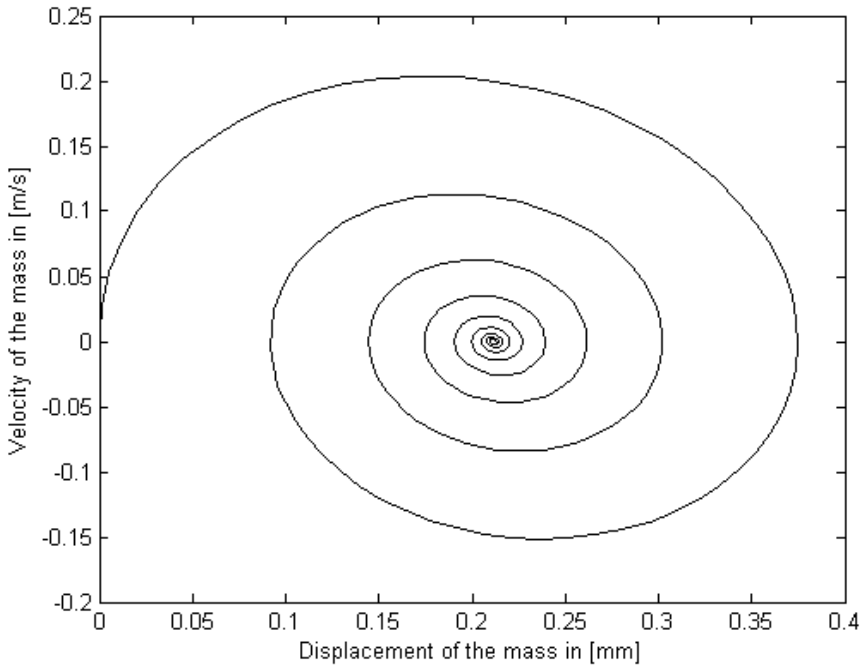
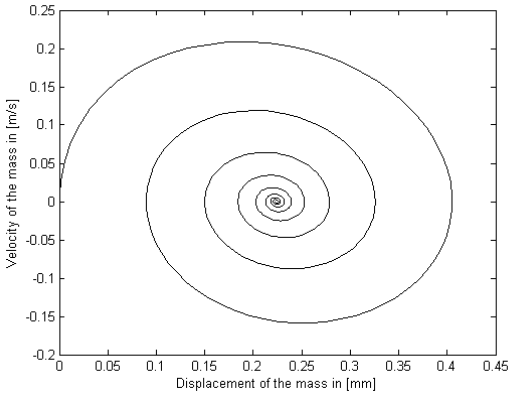
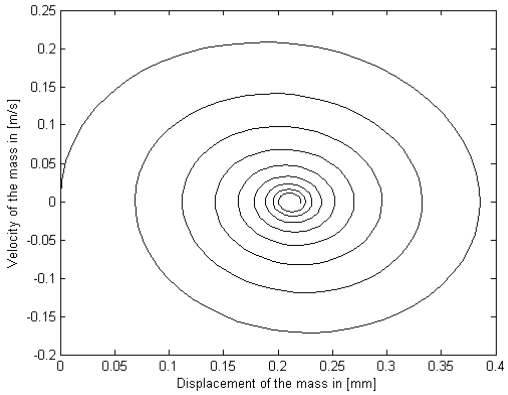


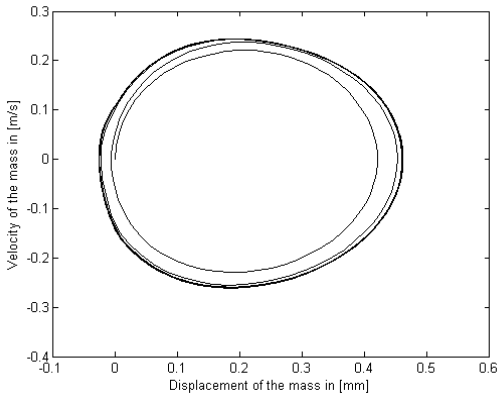
Fig. 7.5 Phase plot for a non-loading case.



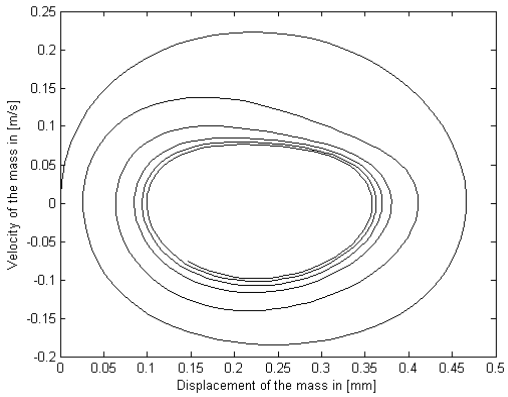
(a): Infinitely long tube, 4 cm^2



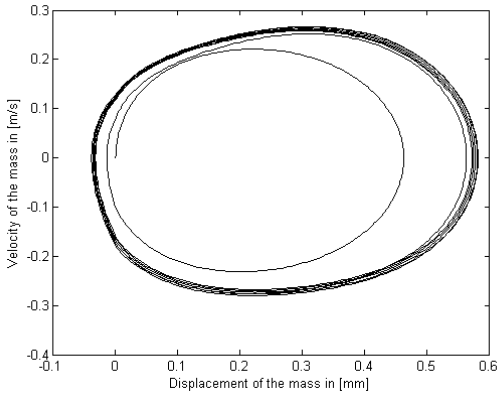
(b): Simple tube 17 cm, 4 cm^2



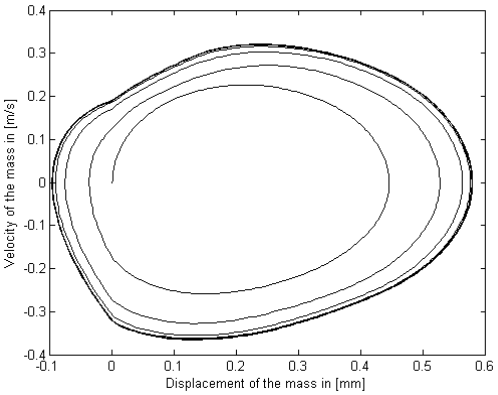
(c): Simple tube 17 cm, 1 cm^2



(d): Simple tube 70 cm, 1 cm^2

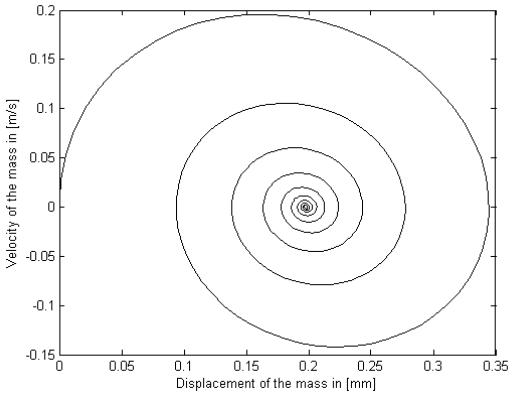


(e): MRI vowel /i/

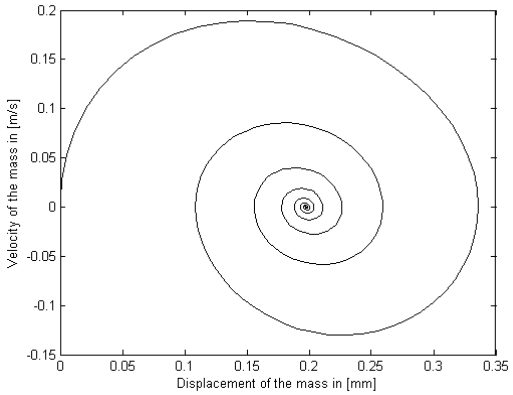


(f): MRI vowel /A/

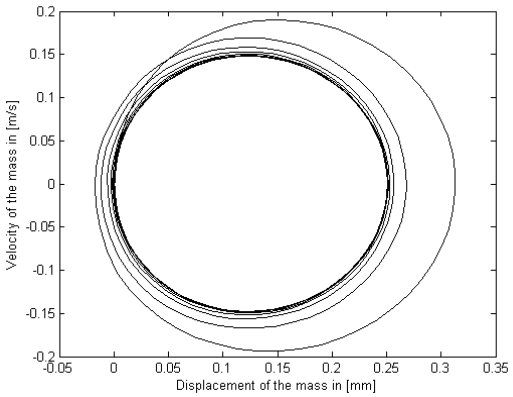
Fig. 7.6 Simplified stability tests for different supraglottal tracts.



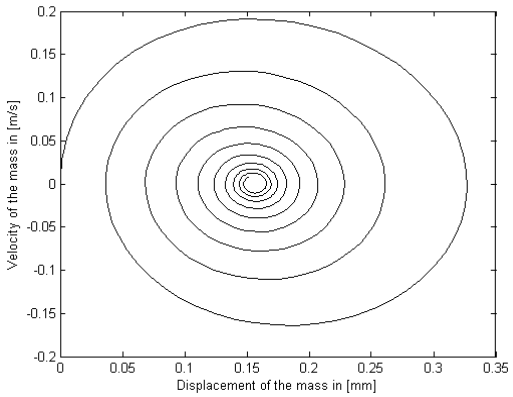
(a): Infinitely long tube, 4 cm^2



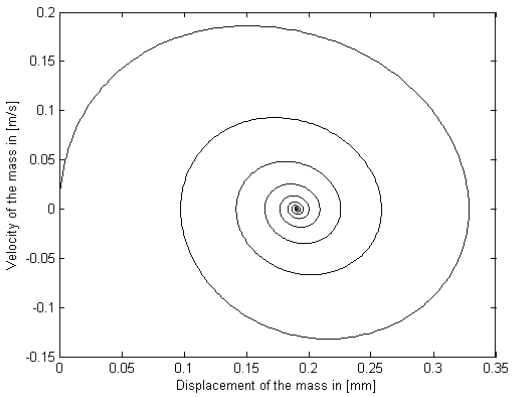
(b): Uniform tube 17cm, 1 cm^2



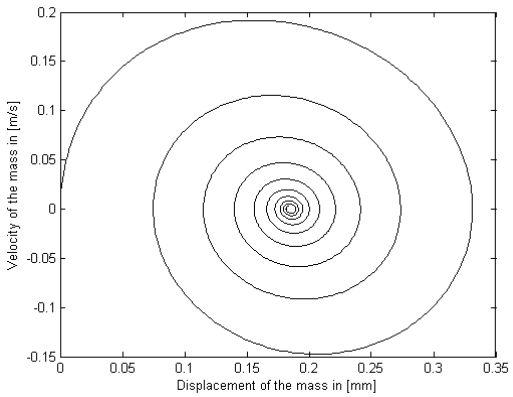
(c): Closed tube 12 cm, 2.3 cm^2



(d): Closed tube 26 cm, 2.3 cm^2

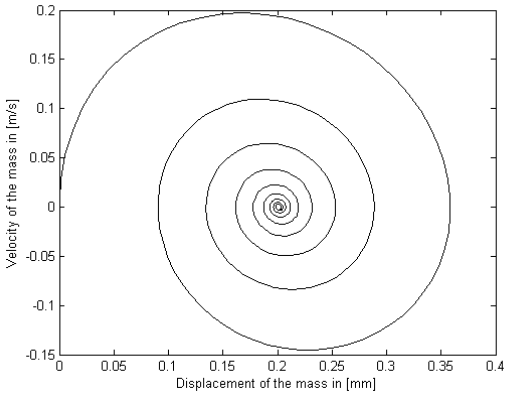


(e): Story's subglottal design

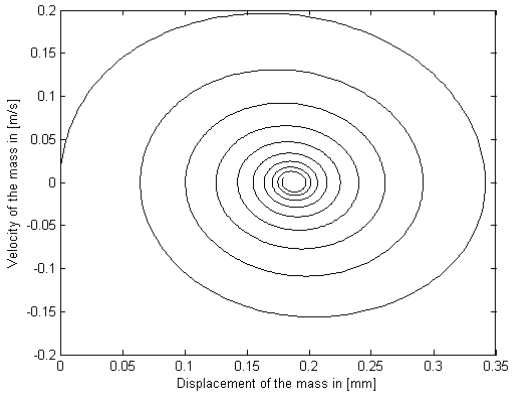


(f): Subglottal tract design in this thesis

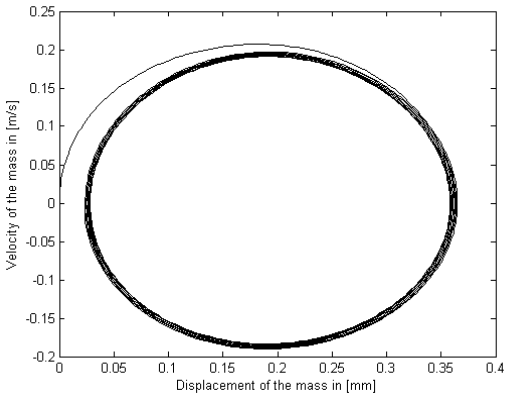
Fig. 7.7 Simplified stability tests for different subglottal tracts.



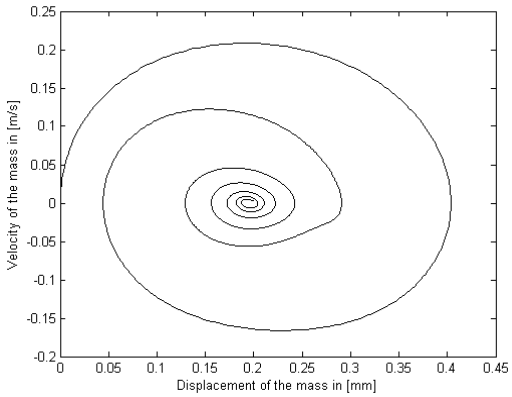
(a): Infinitely long tube, 4 cm^2



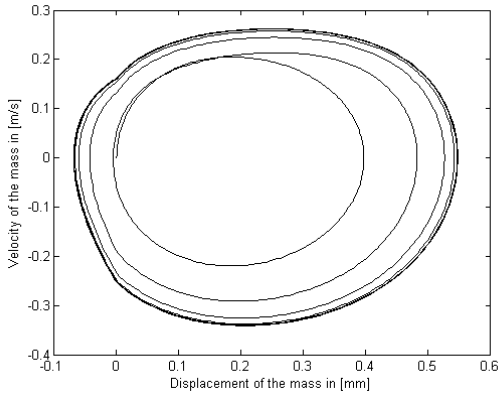
(b): Simple tube 17 cm, 4 cm^2



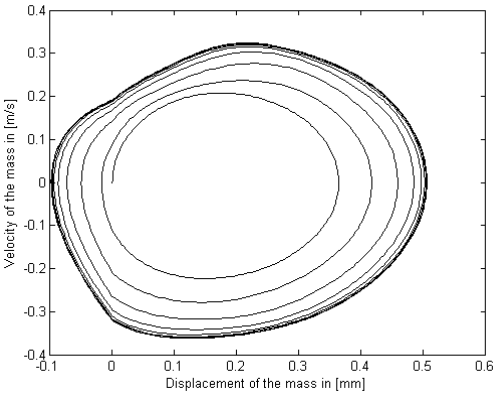
(c): Simple tube 17 cm, 2 cm^2



(d): Simple tube 60 cm, 2 cm^2



(e): MRI vowel /i/



(f): MRI vowel /A/

Fig. 7.8 Simplified stability tests for both tracts.

7.2 Effective Single Degree-of-Freedom Model

In section 7.1, tests used a fixed unit value for the time varying orifice discharge coefficient. That condition made the model behave as a simple one-mass model. As explained in chapters 4 and 6, the function that describes the orifice discharge coefficient modifies the volumetric flow rate and the pressure force supplied into the vocal folds, allowing them to perform self-sustained oscillation is absence of acoustic loading. The effects of acoustic loading on this "effective one-mass model" were investigated in this section.

The results presented here can be compared with the ones reported in [19] [42] [44] [1], without disregarding the differences between the models. The same MRI vowels /i/ and /A/ as in the previous section were used in this case. All the features of the wave reflection analog for both tracts are also included in these tests. As before, the volumetric flow rate was used to compute the radiated acoustic pressure based on Titze's approach as described in equation (4.21).

7.2.1 Coupling with the Vowel /i/

A similar behavior than the one reported is section 7.1.2 was expected to occur, since the acoustic loads are the same. This vowel represents a case with a non-inertive impedance. Figure 7.8 shows that the amplitude is increased with respect to Figure 7.2. Note that the frequency of oscillation is reduced to 150 Hz in this case. The velocity waveform is also comparable in magnitude to that in Figure 7.2, with a drop when the closure phase begins. This is a result of the oscillation observed in the pressure force in the transition from opening to closure. This variation in the force is due to variations observed in the downstream pressure.

Perhaps the most important difference between this case and the one presented in Figure 7.2 is the volumetric flow rate. Its amplitude is significantly increased. This affects the upstream and downstream pressure amplitudes. Due these larger acoustic pressures, the patterns observed in the previous case are emphasized. The depression at the beginning of the opening phase in the volumetric flow rate is more pronounced.

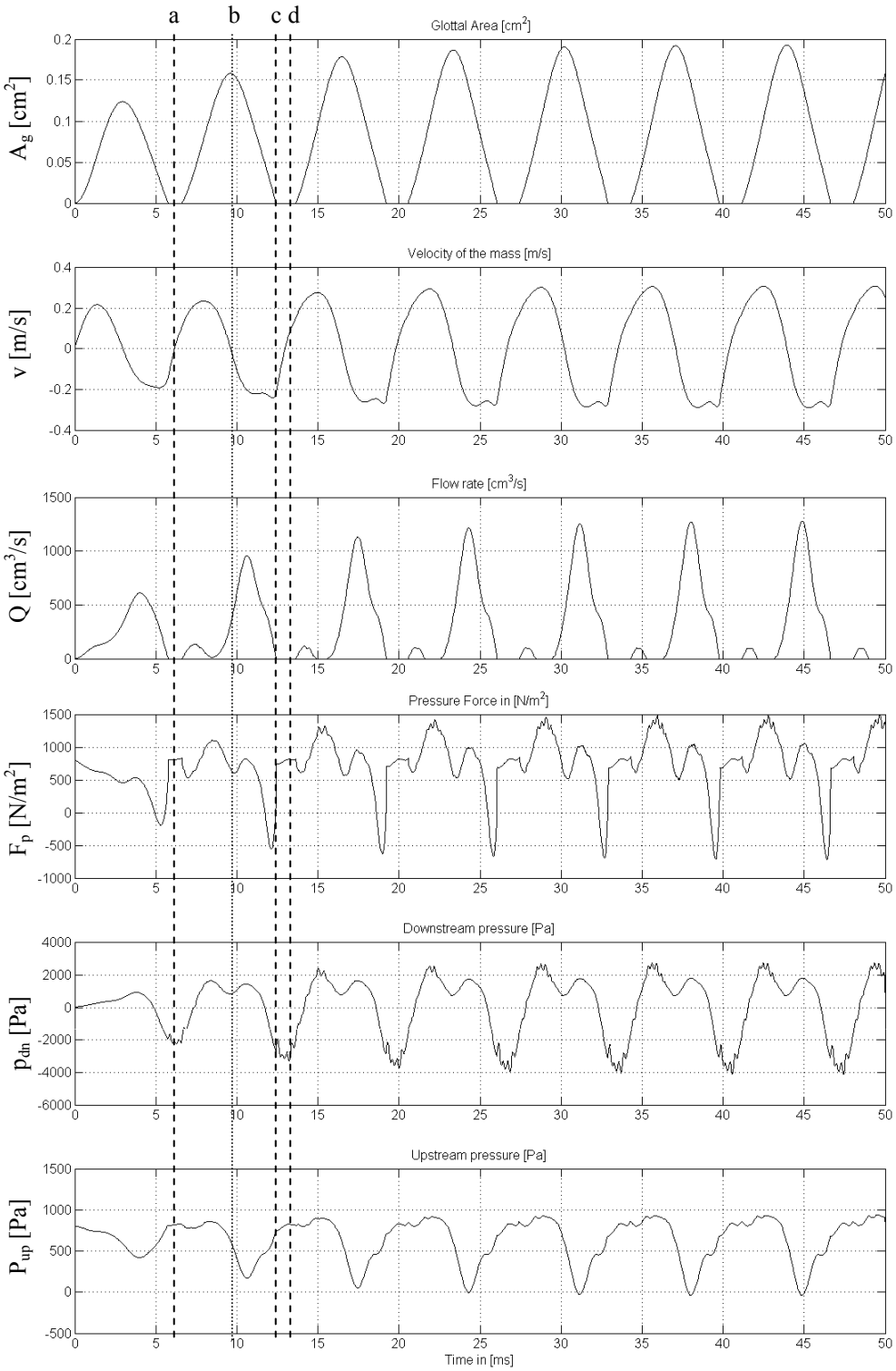


Fig. 7.9 Complete coupled system: MRI /i/. (a): Initial glottal opening; (b): Maximum displacement; (c): Initial collision; (d): Initial glottal opening for a new cycle.

A minor drop is observed at the beginning of the closing phase, which did not appear in the previous case. This drop is due to the increased downstream pressure, which includes a depression that triggers the one in the volumetric flow rate. The oscillation in pressure is attributed to the marked presence of the first formant oscillation of the supraglottal tract. The pressure force is clearly controlled by the downstream pressure during the open phase. This indicates that a larger coupling with the vocal tract is present in this case. Note that the orifice discharge coefficient also modifies the force by making it reach large negative peak values. Similar peaks are observed in both downstream and upstream pressures. Both pressures exhibit new oscillations that alter the behavior of the complete system. These oscillations are caused by the increment in energy introduced by the volumetric flow when the discharge coefficient is used in the flow.

These observations contrast previously reported data [19] [42] [44] [1] in the magnitude and the extra ripple in the pressures, and the variations that they produce in the other variables. These studies match more closely the case presented in Figure 7.2. This suggests that the inclusion of the time varying orifice discharge coefficient over estimates the coupling effect.

7.2.2 Coupling with the Vowel /A/

A similar behavior than the one reported is section 7.1.3 is expected to occur. This vowel represents a case where a more inertive impedance is present. As shown in Figure 7.10, most of the variables do not vary as they did for the vowel /i/. The frequency of oscillation was reduced to 150 Hz. The glottal area was comparable in magnitude and structure to that for Figure 7.3. For velocity, the same range with a different oscillation pattern is observed. This change is attributed to the increased effects of collision.

Showing less variation than the previous case, the volumetric flow rate has the same trends but with an incremented amplitude and ripple in the opening phase. As before, this ripple is due to oscillations observed in the acoustic pressure of the supraglottal tract. The degree of skewing in the volumetric flow rate is similar. As in section 7.2.1, the pressure force features larger negative peak values. These peaks are directly produced by the discharge coefficient and the increased acoustic pressures.

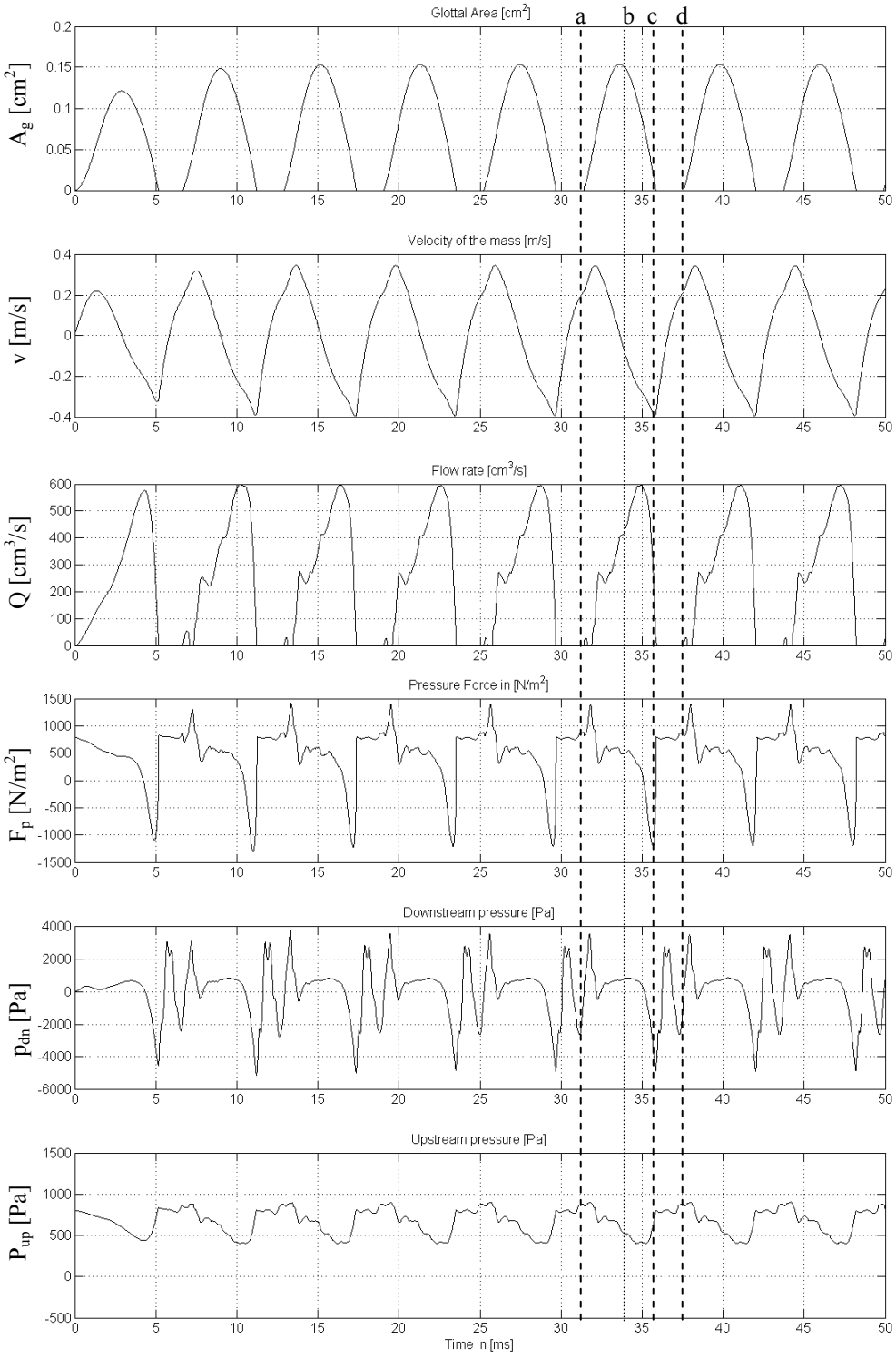


Fig. 7.10 Complete coupled system: MRI /A/. (a): Initial glottal opening; (b): Maximum displacement; (c): Initial collision; (d): Initial glottal opening for a new cycle.

The degree of coupling with the tracts is observed through the close similitude between this force and the downstream pressure during the open phase. As in the case of the MRI vowel /i/, larger amplitudes are observed for both downstream and upstream pressures. Extra oscillations in both pressures are again observed. These two variations are related with the effect of the time varying orifice discharge coefficient in the system.

As before, previously reported data resemble more closely the case where no discharge coefficient was included. This suggests that the inclusion of the time varying orifice discharge coefficient in the model over estimates the effect of the coupling between source and loads. However, this appears to be less critical for this vowel. Note that the two-mass model from Ishizaka and Flanagan [19] and the cover model used by Story [44] include the effects of the mucosal wave. The fact that such models resemble more closely the data reported in sections 7.1.2 and 7.1.3, suggests that the orifice discharge coefficient, representing the effects of the mucosal wave in the effective one-mass model, might be used incorrectly. Note that in these two last test cases, the effects of the orifice discharge coefficient are affecting the acoustic pressures due to its direct application to the volumetric flow rate. It can be argued that the time varying orifice discharge coefficient was originally designed to account for the effects of converging and diverging modes in the oscillation of the folds. In that sense, it was designed to compute a time-varying pressure force. The results presented in this section have shown that its direct application in the volumetric flow rate to compute the acoustic pressure is questionable. This may not be the case if the time varying orifice discharge coefficient were used only to estimate only the pressure force. This case is presented in section 7.3.

7.3 Modification of the Effects of the Orifice Discharge Coefficient

In this section, the effects of the orifice discharge coefficient were included only in the wall pressure that drives the vocal folds. The objective was to illustrate a different usage of this parameter to, perhaps, better account for its effects. The only difference with respect to the tests performed in section 7.2 is that the variation of the orifice discharge coefficient is not considered in the volumetric flow rate that produces the acoustic pressures (Titze's volumetric flow rate). This assumption is based on the suggestion that the effect of the mucosal wave appears to be less significant than presumed in section 7.2 shows. This is supported by the fact that in the source model, the orifice discharge coefficient (ODC) was originally designed to account for a more precise pressure force and not to affect the volumetric flow rate used to compute the acoustic pressures. This implies that its direct application to compute the acoustic pressures might be reason of the overestimation. Although this approach introduces an inconsistent application of the orifice discharge coefficient, it is useful for diagnostics of the model.

The same MRI vowels were tested. The orifice discharge coefficient corresponds to the smooth function described in equation (4.20). Note from Figures 7.11 and 7.12 that only minor variations are observed with respect to Figures 7.2 and 7.3. The frequency of oscillation remains 170 Hz and 190Hz, as in section 7.1. The only variable that changes is the pressure force, which in these tests is able to reach larger negative peak values. This variation is not sufficiently large to produce significant alteration in the other variables. However, this data is the one that more closely resembles the ones reported in more elaborated non-linear models of voice production [19] [44] [42] [1]. This suggests that the effects introduced by orifice discharge coefficient are likely to be in the range described by this approach. A revision of the application of the orifice discharge coefficient should be considered in further developments of the model.

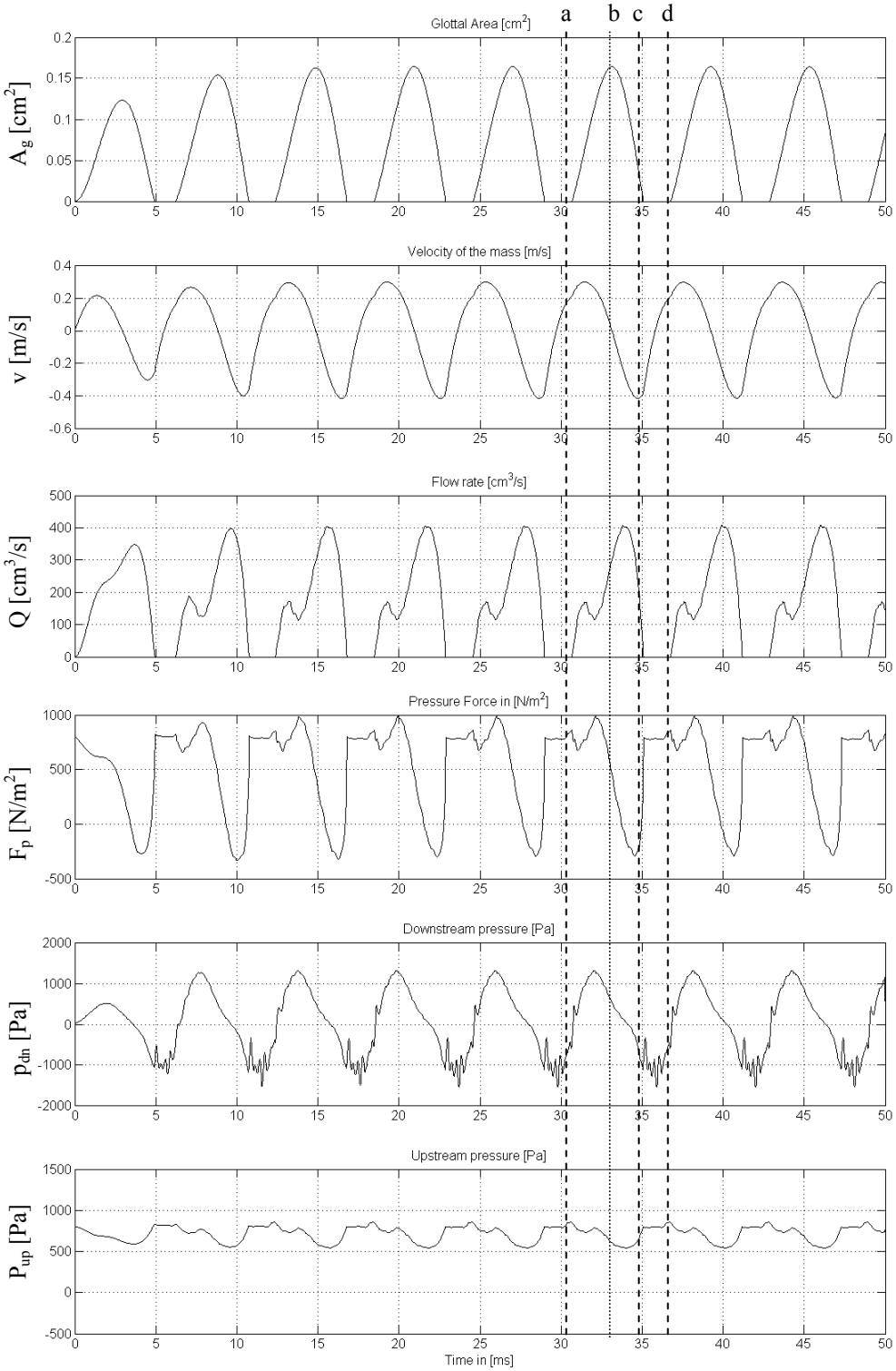


Fig. 7.11 Modified effect of ODC: MRI /i/. (a): Initial glottal opening; (b): Maximum displacement; (c): Initial collision; (d): Initial glottal opening for a new cycle.

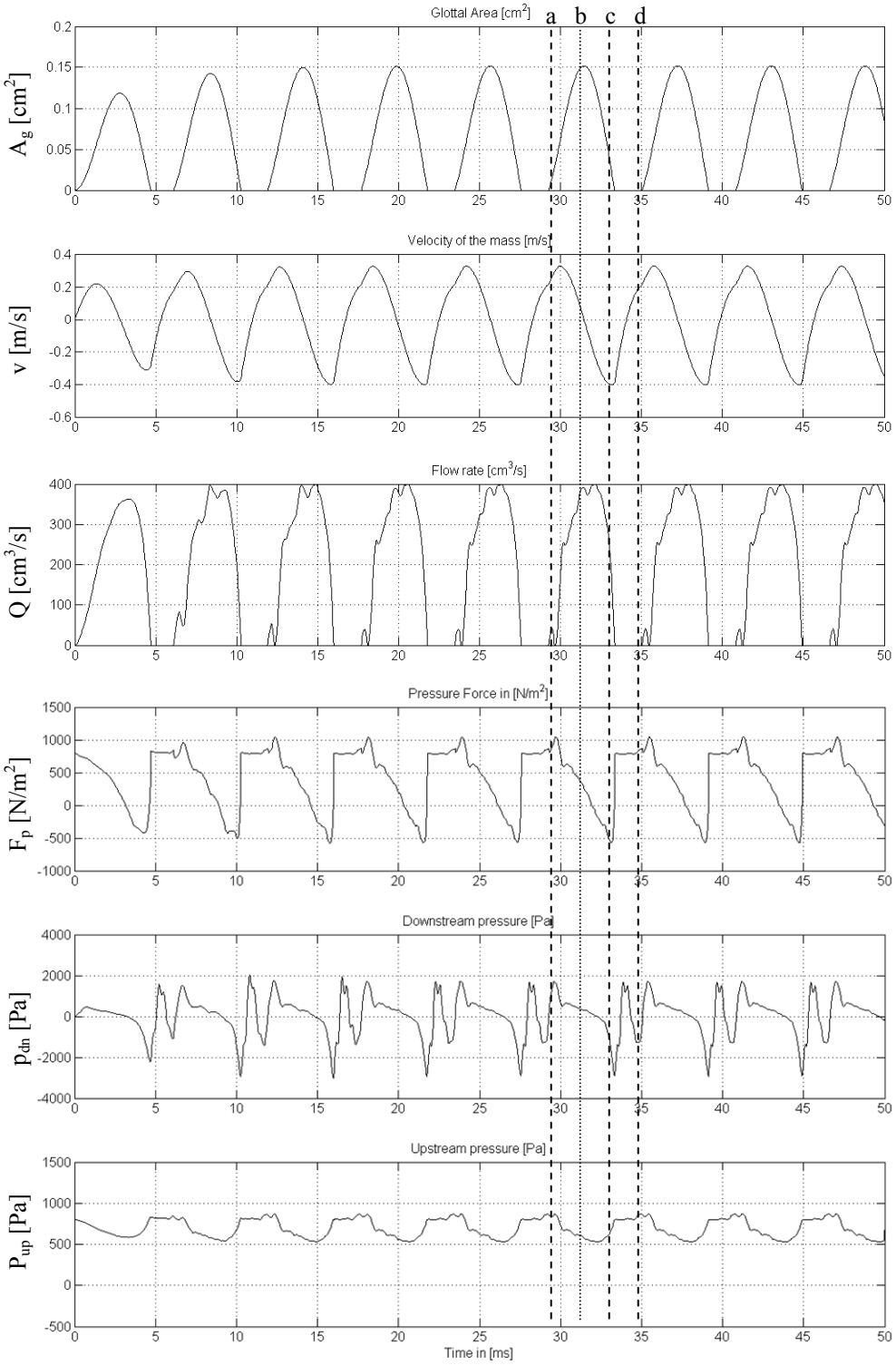


Fig. 7.12 Modified effect of ODC: MRI /A/. (a): Initial glottal opening; (b): Maximum displacement; (c): Initial collision; (d): Initial glottal opening for a new cycle.

8. CONCLUSIONS

8.1 Conclusions

Numerical models for sound wave propagation in the subglottal tract and the vocal tract were developed along with models for the vocal folds vibrations and the airflow through the larynx. For the acoustics of the tracts, a flexible time domain scheme was implemented that allowed investigations of realistic tract geometries. The technique is based on a wave reflection analog principle, which was originally designed to produce articulatory synthesis. It includes different sources of energy dissipation and radiation. Special attention was paid to the subglottal tract design, to ensure consistency with previous studies. The scheme allowed the coupling between tracts to be modeled using either a linear acoustic model, or through non-linear pressure-flow relations. The technique was shown to be useful to produce synthesized speech sounds, and to supply the proper acoustic feedback into the source model.

A simplified source model, amenable to feedback, was designed based on previous studies. A one-mass model was used along with a Bernoulli type flow model and a time-varying orifice discharge coefficient. This so-called "effective one-mass model" accounts for the effects of the mucosal wave on the flow resistance of the orifice, observed in higher order models of the vocal folds. The relative importance of acoustic loading and time varying flow resistance for fluid-structure energy transfer was compared for various configurations. The premise is that both acoustic coupling and aerodynamic adjustments are factors that lead to self-sustained oscillation during normal phonation.

By initially neglecting the effects of the time varying discharge coefficient, the model allowed the effects of the inertive impedance of the vocal tract to be verified. The results were consistent with those of previous studies. The volumetric flow rate waveform was affected by acoustic loading, and sensitive to variations in the vocal tract. The presence of acoustic loading was manifested by ripples or depressions in the opening phase of the volumetric flow rate, skewing rightward in time, and variations in the

fundamental oscillation frequency. These effects have been noted in previous studies, and are associated with instantaneous variations in the upstream and downstream pressures. The ripple in the volumetric flow rate excites high-frequency harmonics, causing more spectral variety in the source.

The vocal tract loading dominated that of the subglottal load in most cases. The subglottal tract design used appeared favorable to self-oscillation, but not sufficiently to induce self-sustained in absence of a vocal tract. The effects of the subglottal tract may be critical in experiments involving excised larynx and synthetic models of the vocal folds, where the subglottal tract used to supply the flow may play a comparatively large role.

The use of a time-varying orifice discharge coefficient may have led to an over-estimation of the effects of the acoustic coupling. One possible solution to this problem is to use more accurate flow models and glottal geometries. The orifice discharge coefficient could be used to modify the pressure force only, and not the volumetric flow rate used to compute the acoustic pressure. The results so far suggest that acoustic loading may be more significant than time-varying flow resistance effects for the maintenance of self-oscillations.

8.2 Suggestions for Future Research

More work is needed to clarify the interactions between source and acoustic loadings. The limitations of the effective one-mass model in terms of wall pressure force accuracy and mechanical behavior might result in a bias. Different loading profiles can alter the source-filter interactions. Only a few specific acoustic loads were tested, assuming that their formant frequencies and impedances represent extreme cases. However, during normal speech or even during singing, transitions between formants and high fundamental frequencies can present completely different challenges.

Although the model developed for the acoustics of the tracts is fairly accurate, several enhancements can be introduced. Some efforts are needed to describe more accurately the losses in the tracts, with special attention to the subglottal tract. A revision of the expression for the yielding wall losses, the introduction of tube branching, and other sources of radiation are suggested. This would allow the study of coupling with the nasal tract and the piriform sinus section, neglected in this thesis. More importantly,

these improvements would lead to a more accurate acoustic model of the subglottal tract, which would be an important contribution.

Upgrades in the source model are needed. The objective in its design was to account for acoustic interaction and aerodynamic adjustments in the simplest possible model. Since the assumptions made for the model make it difficult to expand the results obtained with this design, it appears that a re-design or a new model must be considered. The effective single degree of freedom model could be enhanced by introducing a codebook to accurately compute the pressure force. This codebook could be based on measured or modeled pressure distributions for different configuration that involve mainly position and geometry of the folds. Downstream and upstream pressures would need to be variables in the codebook. This not only would require extensive work to create an appropriate database, but also would significantly increase the complexity of the model. It is believed that a different source model is a better choice. It is suggested to initially test known multi-mass models to later develop a high order (finite element) model that includes fluid-structure and acoustic interactions. In addition, it is suggested to include perturbation analysis to test the performance of the source statistically. This would allow computing other descriptors such as jitter, shimmer, SHR among others, that are used in voice diagnosis by speech pathologists. The results of the model could be compared with normal and pathological phonation data.

From a theoretical perspective, it is recommended to develop analytical expressions that explain the interactions between the subsystems, particularly including the effects of the subglottal tract. This can be done through a complete impedance analysis. An interactive model can be also used to perform stability, observability and controllability analysis. These analyses are currently unavailable in the literature and they should be examined in future research.

Finally, experimental work is suggested to verify the behavior observed in the models. The experimental setup might consider acoustic coupling in self-oscillating synthetic physical models of the vocal folds. Other simplified sources could be measured first, to test the procedure. In any case, measurements of the acoustic pressures downstream and upstream, and volumetric flow rate would be required. The objective would be to measure variations in the volumetric flow rate, oscillation threshold pressures and others variables, as functions of the loading. Digital image correlation (DIC) measurements of the physical models of the vocal folds might be included. However, this type of measurement adds constrains that may cause difficulties when performing measurements.

LIST OF REFERENCES

LIST OF REFERENCES

[1] F. Alipour, D. A. Berry and I. R. Titze, "A finite-element model of vocal-fold vibration," *J. Acoust. Soc. Am.,* vol. 108(6), pp. 3003-3012, Dec. 2000.

[2] S. F. Austin and I. R. Titze, "The effect of subglottal resonance upon vocal fold vibration", *Journal of Voice*, vol. 11(4), pp. 391-402, Dec. 1997.

[3] D. A. Berry and I. R. Titze, "Normal modes in a continuum model of vocal fold tissues," *J. Acoust. Soc. Am.,* vol. 100(5), pp. 3345-3354, Nov. 1996.

[4] R. W. Chan and I. R. Titze "Dependence of phonation threshold pressure on vocal tract acoustics and vocal fold tissue mechanics," *J. Acoust. Soc. Am*, vol. 119(4), pp. 2351-2362, Apr. 2006.

[5] R. W. Chan "Estimation of viscoelastic shear properties of vocal-fold tissues based on time–temperature superposition," *J. Acoust. Soc. Am*, vol. 110(3), pp. 1548-1561, Sep. 2001.

[6] D. D. Cook, *Computational Models of Fluid Flow, Structural Vibration and Fluid-Structure Interactions of Human Phonation*, MS Thesis, School of Mechanical Engineering, Purdue University, 2005.

[7] J. Dang, K. Honda, and H. Suzuki, "Morphological and acoustical analysis of the nasal and the paranasal cavities," *J. Acoust. Soc. Am.*, vol. 96(4), pp. 2088-2100, Oct. 1994.

[8] B. D. Erath, *An Experimental Investigation of Velocity Fields in Divergent Glottal Models of the Human Vocal Folds*, MS Thesis, School of Mechanical Engineering, Purdue University, 2005.

[9] L. J. Eriksson, "Higher order mode effects in circular ducts and expansion chambers," *J. Acoust. Soc. Am.*, vol. 67(2), pp. 545-550, Aug. 1980.

[10] G. Fant, *The Acoustic Theory of Speech Production*, First edition, Mouton & Co. N.V. Publishers, The Hague , 1960.

[11] G. Fant and Q. Lin "Glottal source - vocal tract acoustic interaction", *Speech Transmission Laboratory Quarterly Progress and Status Report*, Royal Institute of Technology, Stockholm, pp. 013-027, 1987.

[12] J. L. Flanagan, *Speech Analysis, Synthesis and Perception*, Second edition, Springer-Verlag, New York, 1972.

[13] J. L. Flanagan and L. L. Landgraf, "Self-oscillating source for vocal tract synthesizers," *IEEE Trans. Audio Electroacoust*., vol. AU-16, 57-64, Mar. 1968.

[14] L. P. Fulcher , R. C. Scherer, A. Melnykov and V. Gateva, "Negative Coulomb damping, limiting cycles, and self-oscillation of the vocal folds," *Phys. Rev. APS,* in Press Oct. 2005.

[15] V. P. Harper, *Respiratory Tract Acoustical Modeling and Measurements*, Ph.D. Dissertation, School of Electrical and Computer Engineering, Purdue University, 2000.

[16] V. P. Harper, S. S. Kraman, H. Pasterkamp, and G. R. Wodicka, "An acoustic model of the respiratory tract," *IEEE Trans. Biomed. Eng*., vol. 48, pp. 543-550, May 2001.

[17] E. J. Hunter, I. R. Titze and F. Alipour , "A three-dimensional model of vocal fold abduction/adduction", *J. Acoust. Soc. Am*, vol. 115(4), pp. 1747-1759, Apr. 2004.

[18] U. Ingard and H. Ising "Acoustic nonlinearity of an orifice," *J. Acoust. Soc. Am.,* vol. 42(1), pp. 6-17, Jul. 1967.

[19] K. Ishizaka and J. L. Flanagan, "Synthesis of voiced sounds from a two mass model of the vocal cords," *Bell Syst. Tech*. J.51, pp. 1233-1268, Jul. 1972.

[20] K. Ishizaka, J. C. French and J. L. Flanagan, ``Direct determination of vocal-tract wall impedance," *IEEE Trans. on Acoust., Speech, and Sig. Processing*, vol. ASSP-23, pp. 370-373, Aug. 1975.

[21] K. Ishizaka, M. Matsudaira and T. Kaneko, "Input acoustic-impedance measurement of the subglottal system," *J. Acoust. Soc. Am*, vol. 60(1), pp. 190-197, Jul. 1976.

[22] J. L. Kelly and C. C. Lochbaum, "Speech synthesis," *Proceedings of the Fourth International Congress on Acoustics*, Copenhagen, pp. 1-4, Sep. 1962.

[23] L. E. Kinsler, A. R. Frey, A. B. Coppens and J. V. Sanders, *Fundamental of Acoustics*, Fourth Edition, New York, Springer, 2000.

[24] D. H. Klatt, "Software for a cascade/parallel formant synthesizer," *J. Acoust. Soc. Am*, vol. 67(3), pp. 971-995, Mar. 1980.

[25] S. Li, R. C. Scherer, M. Wan, S. Wang and H. Wu, "The effect of glottal angle on intraglottal pressure," *J. Acoust. Soc. Am.,* vol. 119(1), pp. 539-548, Jan. 2006.

[26] J. Liljencrants, *Speech Synthesis with a Reflection-type Line Analog, DS Dissertation*, Dept. of Speech Commun. and Music Acoust., Royal Inst. of Tech., Stockholm, Sweden, 1985.

[27] J. C. Lucero, "Dynamics of the two-mass model of the vocal folds: Equilibria, bifurcations, and oscillation region," *J. Acoust. Soc. Am*, vol. 94(6), pp. 3104-3111, Dec. 1993.

[28] J. C. Lucero and L. L. Koenig, "The Phonation thresholds as a function of laryngeal size in a two-mass model of the vocal folds (L)," *J. Acoust. Soc. Am.,* vol. 118(5), pp. 2798-2801, Nov. 2005.

[29] P. Mergell, H. Herzel and I. R. Titze, "Irregular vocal-fold vibration—High-speed observation and modeling," *J. Acoust. Soc. Am.,* vol. 108(6), pp.2996-3002, Dec. 2000.

[30] L. Mongeau, N. Franchek, C. H. Cocker and R. A. Kubli, "Characteristics of a pulsating jet through a small modulated orifice, with application to voice production," *J. Acoust. Soc. Am.,* vol. 102(4), pp. 1121-1133, Oct. 1997.

[31] J. Neubauer, Z. Zhang, and D. Berry, "Effects of subglottal acoustics on phonation onset," *149th Meeting: Acoustical Society of America*, *J. Acoust. Soc. Am*, vol. 117(4), p. 2542, Apr. 2005.

[32] J. B. Park and L. Mongeau, "Instantaneous orifice discharge coefficient measurements of a physical, driven model of the human larynx," *J. Acoust. Soc. Am.*, submitted in Oct. 2005.

[33] M. G. Rahim, C. C. Goodyear, W. B. Kleijn, J. Schroeter, and M. M. Sondhi, "On the use of neural networks in articulatory speech synthesis," *J. Acoust. Soc. Am.,* vol. 93(2), pp. 1109-1121, Feb. 1993.

[34] M. G. Rahim, *Artificial Neural Networks for Speech Analysis/Synthesis*, First edition, Kluwer Academic Publishers, New York, 1994.

[35] M. Rothenberg, "Acoustic interaction between the glottal source and the vocal tract," in *Vocal Fold Physiology*, K. N. Stevens and M. Hirano Eds, University of Tokyo Press, Tokyo, 1981.

[36] R. C. Scherer, I. R. Titze, and J. F. Curtis, "Pressure-flow relationships in two models of the larynx having rectangular glottal shapes," *J. Acoust. Soc. Am.,* vol. 73(2), pp. 668-676, Feb. 1983.

[37] R. C. Scherer, D. Shinwari, K. J. De Witt, C. Zhang, B. R. Kucinschi,, and A. A. Afjeh, "Intraglottal pressure profiles for a symmetric and oblique glottis with a divergence angle of 10 degrees," *J. Acoust. Soc. Am.,* vol. 109(4), pp. 1616-1630, Apr. 2001.

[38] K. N. Stevens and A. S. House, "Development of a quantitative description of vowel articulation," *J. Acoust. Soc. Am*, vol. 27(3), pp. 484-493, May 1955.

[39] K. N. Stevens, *Acoustic Phonetics*, First edition, The MIT Press, Massachusetts, 2000.

[40] W. J. Stronge, *Impact Mechanics*, First Edition, New York, Cambridge University Press, 2000.

[41] B. H. Story, *Physiologically-Based Speech Simulation using an Enhanced Wave-Reflection Model of the Vocal Tract*, Ph.D. Dissertation, University of Iowa, 1995.

[42] B. H. Story and I. R. Titze, "Voice simulation with a body-cover model of the vocal folds," *J. Acoust. Soc. Am.,* vol. 97(2), pp. 1249-1260, Feb. 1995.

[43] B. H. Story, I. R. Titze and E. A. Hoffman, "Vocal tract are functions from magnetic resonance imaging," *J. Acoust. Soc. Am.,* vol. 100(1), pp. 537-554, Jul. 1996.

[44] B. H. Story, "An overview of the physiology, physics and modeling of the sound source for vowels," *Acoust. Sci. & Tech.* vol. 23(4), pp. 195-206, 2002.

[45] S. L. Thomson, *Fluid-Structure Interactions within the Human Larynx*, PhD Dissertation, School of Mechanical Engineering, Purdue University, 2004.

[46] I. R. Titze, *Principles of Voice Production*, Second Edition, Prentice-Hall, New Jersey, 1994.

[47] I. R. Titze, "Parameterization of glottal area, glottal flow and vocal fold contact area," *J. Acoust. Soc. Am.,* vol. 75(2), pp. 570-580, Feb. 1984.

[48] I. R. Titze, "The physics of small-amplitude oscillation of the vocal folds," *J. Acoust. Soc. Am.,* vol. 83(4), pp. 1536-1552, Apr. 1988.

[49] I. R. Titze and B. H. Story, "Acoustic interactions of the voice source with the lower vocal tract," *J. Acoust. Soc. Am*, vol. 101(a), pp. 2234-2243, Apr. 1997.

[50] I. R. Titze, “Regulating glottal airflow in phonation: Application of the maximum power transfer theorem to a low dimensional phonation model,” *J. Acoust. Soc. Am.*, vol. 111(1), pp. 367-376, Jan. 2002.

[51] I. R. Titze, “Nonlinear source-filter interaction in singing”, *2nd International Conference on the Physiology and Acoustics of Singing*, Denver, Colorado, U.S.A. October 7th - 9th, 2004.

[52] J. W. Van den Berg, J. T. Zantema and P. Doornenbal, “On the Air Resistance and the Bernoulli Effect of the Human Larynx,” *J. Acoust. Soc. Am*, vol. 29, 626-631, May 1957.

[53] J. W. Van den Berg, “Myoelastic-aerodynamic theory of voice production,” *J. Speech Hear. Res.*, vol. 1. 227–244, 1958.

[54] E. R. Weibel, *Morphometry of the Human Lung*, First Edition, New York, Springer, 1963.

[55] G. R. Wodicka, K. N. Stevens, H. L. Golub, E. G. Cravalho and D. C. Shannon, “A model of acoustic transmission in the respiratory system” *IEEE Trans. Biomed. Eng*., vol. 36(9), pp. 925-934, Sep. 1989.

[56] G. R. Wodicka, K. N. Stevens, H. L. Golub and D. C. Shannon, “Spectral characteristics of sound transmission in the human respiratory system” *IEEE Trans. Biomed. Eng*., vol. 37(12), pp. 1130-1135, Dec. 1990.

[57] K. Zhang and T. Siegmund and R. W. Chan, “A constitutive model of the human vocal fold cover for fundamental frequency regulation,” *J. Acoust. Soc. Am*, vol. 119(2), pp. 1050-1062, Feb. 2006.

[58] Z. Zhang, L. Mongeau, and S. H. Frankel, “Experimental verification of the quasi-steady approximation for aerodynamic sound generation by pulsating jets in tubes,” *J. Acoust. Soc. Am.,* vol. 112(4), pp. 1652-1663, Oct. 2002.

[59] W. Zhao, *A Numerical Investigation of Sound Radiated from Subsonic Jets with Application to Human Phonation*, PhD dissertation, School of Mechanical Engineering, Purdue University, 2000.

[60] W. Zhao, S. H. Frankel, and L. Mongeau, "Numerical Simulations of Sound from Confined Pulsating Axisymmetric Jets", *AIAA J.*, vol. 39(10), pp.1868-1874, 2001.