A NUMERICAL AND ANALYTICAL STUDY OF PHONATION THRESHOLD PRESSURE AND EXPERIMENTS WITH A PHYSICAL MODEL OF THE VOCAL FOLD MUCOSA

Chen Liu

A Thesis
Submitted to the Graduate College of Bowling Green State University in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE

December 2009

Committee:
Lewis P. Fulcher, Advisor
Ronald C. Scherer
Haowen Xi
ABSTRACT

Lewis P. Fulcher, Advisor

Titze (1988) used the surface wave model to derive an analytical expression for threshold pressure in terms of the glottal geometry and biomechanical parameters of the larynx. This formula was tested in a series of experiments in 1995 and 1997. Since the membrane of the physical model used in the experiments becomes rounded when a fluid flows beneath it, the effects of glottal curvature were investigated. Because physical model used for the experiments could be adjusted to give a divergent prephonatory glottal geometry, an angle $\theta$ was also introduced in addition to the curvature correction. Including the curvature coefficient $b$ and the prephonatory glottal angle $\theta$ do not seem likely candidates for discrepancies observed in the experiments because such effects are hard to distinguish from changes in the effective values of the glottal halfwidth and the damping coefficient.

Nardone’s mathematical model was built from the classic, lumped element, two-mass model of Ishizaka and Flanagan. The mathematical model is based on ten, second-order, nonlinear, coupled, ordinary differential equations that are solved simultaneously using the software Mathematica. Nardone’s model was employed to study the role of vocal tract parameters and viscous damping constants in determining the threshold pressure. Calculated results were compared with Chan and Titze’s (2006) experimental data. The results indicate that the threshold pressure is consistently lowered when the vocal tract is included, which follows the same trend as the experimental results of Chan and Titze (2006). Increasing the vocal tract area in the mathematical model achieved a larger difference between the calculated results with vocal tract and with a larger vocal tract area. We have also shown that an increased viscous damping constant leads to a bigger threshold pressure differences. Comparison of the calculated results with Chan and Titze’s (2006) experiments were not able to generate a consistent fit over the entire range of glottal widths studied. Successes were recorded over parts of the range with different sets of parameters.
ACKNOWLEDGMENTS

I am very grateful to Dr. Lewis Fulcher for his encouragement and support in my research and for his perfect advising of my thesis. I am also very thankful to Nardone Marco for his help and Dr. Ronald Scherer for his interest in this project.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Chapter</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>INTRODUCTION</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>REVIEW OF NARDONE'S MODEL</td>
<td>17</td>
</tr>
<tr>
<td>3</td>
<td>CURVATURE CORRECTION</td>
<td>30</td>
</tr>
<tr>
<td>4</td>
<td>RESULTS WITH NARDONE'S MODEL</td>
<td>37</td>
</tr>
<tr>
<td>5</td>
<td>CONCLUSIONS</td>
<td>45</td>
</tr>
<tr>
<td></td>
<td>REFERENCES</td>
<td>47</td>
</tr>
</tbody>
</table>
LIST OF FIGURES

Figure Page
1.1 Schematic of the upper part of the human body, which shows the vocal airways (Rosistem, 2003) ................................ 1
1.2 Phases of the vocal fold oscillation cycle (Hanavan, 2006) ......................... 3
1.3 Frontal section of body-cover model used for small-oscillation analysis (Titze, 1988) ........................................ 4
1.4 Schematic of the physical model of the vocal fold and the glottal airway (Titze et al., 1995) .................................... 6
1.5 Phonation threshold pressure for ascending subglottal pressure (open circles for onset) and descending subglottal pressure (filled circles for offset) as a function of glottal halfwidth (Titze et al., 1995) ..................... 7
1.6 Phonation threshold pressure for phonation onset (open circles) and phonation offset (filled circles) for fluids with different viscosities (Titze et al., 1995) . 8
1.7 Schematic of the physical model with a divergent prephonatory glottal geometry (Titze et al., 1997) .............................. 9
1.8 Phonation threshold pressure as a function of vocal-fold thickness for different membrane thicknesses and mucosal fluid viscosities. In each case, both the onset and the offset pressures are shown (Titze et al., 1997) ................... 11
1.9 Offset phonation threshold pressure as a function of prephonatory glottal half-width and glottal angle (Titze et al., 1997) .............................. 12
1.10 Offset and onset phonation threshold pressures as a function of prephonatory glottal half-width, for two concentrations of hyaluronic acid implanted under the membrane used to model the vocal fold cover (Chan and Titze, 2006) .... 13
1.11 Schematic of the physical model of the larynx with coupling to a rectangular uniform-tube supraglottal vocal tract (Chan and Titze, 2006) .......................... 14
1.12 Phonation threshold pressure of the physical model as function of prephonatory
glottal half-width, with and without a uniform-tube vocal tract (Chan and
Titze, 2006) .................................................. 15

2.1 Schematic of the IF72 model of the vocal folds (Nardone, 2007) .............. 17

2.2 The wind tunnel of model M5 (Scherer et al., 2001) ............................. 20

2.3 Schematic of pressure taps on one of the M5 vocal fold inserts (Scherer et al.,
2001) ................................................................. 21

2.4 Samples of M5 experimental data for a rectangular glottis ($\phi = 0^\circ$), a con-
verging angle of $10^\circ$, and a diverging angle of $10^\circ$, when the minimal glottal
diameter is 0.08 cm. Locations of the pressure taps are indicated by numbers
from 1 to 16 (Scherer, 2009). ........................................ 22

2.5 Driving forces on mass $m_{R1}$ for the original IF72 model (upper), the modified
IF72 force (center), and the M5 experimental data (lower). The transglottal
pressure is 8 cm $H_2O$ and the glottal area is 0.05 $cm^2$ (Nardone, 2007) .... 26

2.6 IF72 network representation of the glottis and vocal tract (Ishizaka and Flana-
gan, 1972) .......................................................... 28

3.1 Schematic of the medial surface geometry with curvature .......................... 30

3.2 Comparison of the calculated threshold pressures with the experimental results
of Titze et al. (1997) .................................................. 36

4.1 Comparison of threshold pressures when the parameter set IF1 (Table 4.1) is
used and the results of Chan and Titze (2006) .................................. 41

4.2 Comparison of threshold pressures when the parameter set IF1 is used and the
results of Chan and Titze (2006). The vocal tract area was set to 8 $cm^2$ to
simulate the case of no vocal tract. .............................................. 42

4.3 Comparison of the threshold pressures obtained with the parameters IF2 and
IF3 with Chan and Titze (2006) .................................................. 43
4.4 Comparison of the threshold Pressures as the damping parameters get smaller
# LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1</td>
<td>Four variations of membrane thickness and fluid viscosity</td>
<td>10</td>
</tr>
<tr>
<td>2.1</td>
<td>Force conditions if the modified IF72 model pressures are used</td>
<td>27</td>
</tr>
<tr>
<td>2.2</td>
<td>Force conditions if the M5 pressure data are used</td>
<td>27</td>
</tr>
<tr>
<td>4.1</td>
<td>List of parameters and their values for the two-mass model</td>
<td>39</td>
</tr>
</tbody>
</table>
CHAPTER 1
INTRODUCTION

Phonation is a process of producing speech sounds that begins with vocal fold vibrations. Before the sounds are produced, a stream of air from the lungs passes into the trachea and then into the larynx. The larynx is the principal structure that converts the flow of air from the trachea into a more complicated time-dependent air stream. The vocal folds are the active parts of the larynx that chop the air stream into flow pulses. The vocal folds periodically interrupt the air stream by rapidly opening and closing. Different acoustical properties of vocal tract transform the flow pulses into meaningful speech sounds. A schematic of the anatomy surrounding the larynx and its associated airways is presented in Figure 1.1.

Figure 1.1: Schematic of the upper part of the human body, which shows the vocal airways (Rosistem, 2003)
Phonation threshold pressure is defined as the minimum lung pressure required to produce a sustained oscillation of the vocal folds (Titze, 1988). It has been studied in both theoretical and experimental approaches. Phonation threshold pressure has clinical significance in that it determines the ease with which phonation can be achieved and sustained. It is an important measure of vocal health after surgical intervention and when attempting to diagnose vocal fold pathologies.

To understand the mechanics of phonation in mathematical terms, Titze (1988) assumed that energy was transferred from the glottal airflow to the motion of the vocal folds when a surface wave propagated along the cover of the vocal folds with a small amplitude. He based this assumption on careful experimental studies (Titze, 2006, Hanson et al., 1988, Berke et al., 1989) of upper and lower lips of the vocal folds with high speed glottographs. A schematic representation of these experimental studies is presented in Figure 1.2. As shown in frame 1 of Figure 1.2, the vocal folds are closed at the beginning of the cycle that leads to phonation. In frame 2, the lower portions of the glottal walls are separated by the increased air pressure from trachea, and eventually the left and right vocal folds are separated by the air flow, as shown in frames 3 and 4. As the vocal folds separate, the lower lips of vocal fold are always leading the upper lips of vocal fold, and a convergent shape is produced, as shown in frame 4. In frame 5, the lower portions of the vocal fold start to return to the midline, after they reach the maximum excursion from the midline, while the upper portions of the vocal fold still move outward. By the time the top portions of the vocal fold reach their maximum extent, the lower portions of the vocal fold are already returning and a divergent shape is achieved, as shown in frame 6. Both top portions and lower portions of the vocal fold keep moving towards each other, and they collide at the midline at the end of the cycle, as shown in frame 7 and 8. The behavior of the vertical phase difference between the lower and upper glottal surfaces of the vocal folds described in these frames resembles that of a surface wave propagating from the bottom lip of the vocal fold to the top lip.
Figure 1.2: Phases of the vocal fold oscillation cycle (Hanavan, 2006)
Titze’s surface wave model views the vocal fold as composed of two structures: a body and a cover, as shown in Figure 1.3. The body is assumed to be stationary and consists of the muscle and deep layers of the vocal fold ligament. The cover is assumed to provide a medium for a surface wave to respond to the airflow. Its wave velocity is $c$. It is free to oscillate and to move back and forth in response to a subglottal pressure from the trachea and the lungs.

![Figure 1.3: Frontal section of body-cover model used for small-oscillation analysis (Titze, 1988)](image)

In Figure 1.3, the parameters $\xi_{01}$ and $\xi_{02}$ describe the initial position of the lower and the upper lips of the vocal fold. The subglottal pressure is described by $P_s$, the pressure within the glottis by $P_g$, and the pressure in the vocal tract by $P_1$. Based on the surface wave model, Titze (1988) derived an analytical expression for threshold pressure $P_{th}$ in terms of
the glottal geometry and biomechanical parameters of the larynx, which is given by

\[ P_{th} = \frac{B}{l_g T^2} c \xi_0 (k_{ent} - k_{ex}), \]  

(1.1)

where \( B \) is the oscillator damping coefficient, \( c \) is the surface wave speed, \( l_g \) is the glottal length, \( T \) is the glottal thickness, \( \xi_0 \) is the glottal halfwidth for the rectangular glottis (where \( \xi_{01} = \xi_{02} = \xi_0 \)), \( k_{ent} \) is the entrance coefficient, and \( k_{ex} \) is the exit coefficient. According to the analytical expression, the phonation threshold pressure is reduced by reducing the mucosal wave velocity, and by bringing the vocal folds closer together. Increasing the resistance of the vocal fold oscillator increases the phonation threshold pressure.

The validity of this analytical expression was tested in experimental measurements by Titze, Schmidt, and Titze (1995) on a physical model of the larynx, which is shown in Figure 1.4. The physical model included a glottal airway, which was constructed to allow control of the glottal width by adjusting the micrometer screw. This physical model of the larynx includes a silicone membrane of thickness 200 \( \mu m \) which encloses a region through which a fluid may flow. This membrane and the flowing fluid represent the vocal fold mucosa, or the outer layer of the vocal fold. By adding materials to the fluid, one can change the viscous properties of the membrane and fluid, and thus examine the dependence of phonation threshold pressure on these properties. As seen from Figure 1.4, only a hemilarynx can be simulated by this apparatus instead of a full larynx. The hemilarynx geometry was chosen over the full larynx because it is easier to construct than the full larynx. If one attempted to construct a full larynx with the left-right symmetry used in most models of normal phonation, it is very difficult to ensure that small differences between the membranes with encapsulated fluids on both sides would not introduce spurious effects. The glottal length, which is perpendicular to the plane depicted in Figure 1.4, is 2.3 cm.

The onset threshold pressure for a given glottal width and fluid was recorded as that pressure at which the membrane oscillation began as the air pressure was increased from a
small value. The pressure was increased above the threshold value, and then decreased until
the membrane no longer oscillated. This was identified as the phonation offset pressure. As
shown below in Figure 1.5, the onset pressure for a given fluid and a given glottal halfwidth
was always higher than the offset pressure. This difference between the onset of oscillations
and the offset of oscillations suggests that a history of the material is important. Such effects
are often described as hysteresis phenomena (Lucero 1999).

Figure 1.5 shows $P_{th}$ as a function of the prephonatory glottal half-width $\xi_0$. The fluid
was pure water which had a viscosity of 1.0 cP (centipoise). The most noteworthy feature of
this result is that $P_{th}$ has a minimum value near $\xi_o = 1.0 \, mm$, which was not predicted by
Eq. (1.1). However, $P_{th}$ increases almost linearly with $\xi_0$ on the right side of the minimum,
Figure 1.5: Phonation threshold pressure for ascending subglottal pressure (open circles for onset) and descending subglottal pressure (filled circles for offset) as a function of glottal halfwidth (Titze et al., 1995)

as predicted by Eq. (1.1), and thus it is apparent that the physical model of Figure 1.4 contains some physical effects not described by Titze's (1988) surface wave model.

Figure 1.6 shows that there is an approximate linear relationship between $P_{th}$ and viscosity of the fluid inside the membrane. This would be expected from Eq. (1.1), since the viscosity of the fluid should affect the damping parameter $B$. The viscosity of the fluid flowing through the membrane was changed by mixing graduated weights of sodium carboxymethyl cellulose powder with the water. The phonation threshold pressure of Figure 1.6 does not approach zero in the limit that the viscosity of the flowing fluid approaches zero. Presumably, this reflects a residual damping from the membrane itself, in addition to
Figure 1.6: Phonation threshold pressure for phonation onset (open circles) and phonation offset (filled circles) for fluids with different viscosities (Titze et al., 1995)

...the energy loss from the viscous flow of the fluid. However, the magnitudes of the viscosities in Figure 1.6 are a bit of surprise, since they are much lower than the viscosity of vibrating vocal fold tissues (600 cP by Titze and Talkin (1979)), although the corresponding threshold pressures is already higher than that of typical human phonation (about 0.3 kPa at 100 Hz). In fact, Titze et al. found that increasing the fluid viscosity an order of magnitude beyond the values of Figure 1.6 produced phonation threshold pressures that were completely unrealistic, which presents a dilemma in interpreting the data of Figure 1.6.

Chan, Titze, and Titze (1997) conducted further experiments with the physical model of Figure 1.7 in order to explore some additional aspects of the surface wave model that pertain to a nonrectangular prephonatory glottal geometry, with an oblique angle $\theta$. As seen
in Figure 1.7, the physical model used by Chan, Titze, and Titze (1997) was very similar to the one described in their previous experiment, but the angle could be adjusted by tilting the micrometer assembly.

![Figure 1.7: Schematic of the physical model with a divergent prephonatory glottal geometry (Titze et al., 1997)](image)

The angle $\theta$ is positive for the diverging glottis, according to our conventions (Scherer et al., 2001). The analytical expression for the nonrectangular prephonatory glottal geometry with an angle $\theta$ derived by Titze (1988) is

$$P_{th} = \frac{2 k_t B c \xi^2_{01}}{l_g T^2 (\xi_{01} + \xi_{02})}, \quad (1.2)$$

where $k_t (= k_{ent} - k_{ex})$ is the transglottal pressure coefficient. These quantities $\xi_{01}$ and $\xi_{02}$ are connected with the angle $\theta$ by

$$\tan \theta = \frac{\xi_{02} - \xi_{01}}{T}. \quad (1.3)$$
Because of the square of \( \xi_{01} \) in the numerator, Eq. (1.2) predicts that the convergent glottal geometry \((\xi_{01} > \xi_{02})\) requires a higher \( P_{th} \) than divergent glottis \((\xi_{01} < \xi_{02})\).

Experiments were done on this physical model for several sets of parameters. In the first set, onset and offset \( P_{th} \) were measured for two different vocal-fold thicknesses (7.5 and 11.0 mm). For each vocal fold thickness the viscosity of the encapsulated fluid was changed, and thus there were four combinations of parameters of membrane thickness and mucosal fluid viscosity, as listed in Table 1.1.

| Table 1.1: Four variations of membrane thickness and fluid viscosity |
|--------------------|---|---|---|---|
|                   | Set A | Set B | Set C | Set D |
| Membrane thickness (\( \mu m) \) | 70 | 210 | 70 | 210 |
| Fluid viscosity (cP) | 1.0 | 1.0 | 471 | 471 |

A glottal halfwidth of 2.0 mm was maintained throughout these experiments. Membrane thickness was either 70 \( \mu m \) (thin) or 210 \( \mu m \) (thick), and fluid viscosity was either 1.0 cP (low) or 471 cP (high). The combinations of vocal fold thickness, membrane thickness, and fluid viscosity required 8 different sets of measurements. Since onset and offset pressures were collected for each set, the data of Figure 1.8 comprise 16 different points. Figure 1.8 shows that offset \( P_{th} \) is consistently lower than the corresponding onset \( P_{th} \) and that the thicker vocal fold has a lower \( P_{th} \) than the thinner one, as one might expect from the damping factor of Eq. (1.2).

In another set of experiments, both the prephonatory glottal half-width and the glottal orientation angle were varied. The results are shown in Figure 1.9. As seen from Figure 1.9, for small values of \( \xi_{0} \), \( P_{th} \) increases with prephonatory glottal half-width, in accord with the 1995 experiments. But, when \( \xi_{0} \) rises above 3.0 mm, the relationship between prephonatory glottal half-width and \( P_{th} \) becomes more complicated and \( P_{th} \) does not increase any further,
Figure 1.8: Phonation threshold pressure as a function of vocal-fold thickness for different membrane thicknesses and mucosal fluid viscosities. In each case, both the onset and the offset pressures are shown (Titze et al., 1997)

a kind of saturation effect. However, Eq. (1.2) does not predict a minimum near $\theta = 0^\circ$ as most of the data in Figure 1.9 seems to indicate. Instead, Eq. (1.2) predicts a uniform decrease with values of $\xi_02 > \xi_01$, when the shape of the glottis is divergent. As in the earlier case regarding the data of Figure 1.5 and Eq. (1.1), the data of Figure 1.9 point out short-comings of the surface wave model and Eq. (1.2).

Chan and Titze (2006) extended the surface wave model to include a vocal tract, and thus they had to revise the analytical expression of Eq. (1.1). This revised theory was tested by several experiments. In order to study the effect of vocal fold tissue viscoelasticity on $P_{th}$, the fluid flow assembly was replaced so that different biological materials could be implanted
into the vocal fold cover. Also a uniform-tube supraglottal vocal tract was added to the previous physical model to investigate the effect of vocal tract inertance on $P_{th}$.

Chan and Titze’s revised analytical expression is given by

$$P_{th} = \frac{B c \xi_0}{l_g T^2} - \frac{2 I l_g v c \xi_0}{T}. \quad (1.4)$$

The air velocity at the glottal exit is represented by $v$, and the inertance of the vocal tract is given by $I = \rho L / A$, where $\rho$ is density of air in the vocal tract, $L$ is the vocal tract length, and $A$ is the vocal tract cross-sectional area. According to Eq. (1.4), the threshold pressure can be lowered by the following conditions: a decrease in the tissue damping constant, a decrease in the velocity of the mucosal wave $c$, a decrease in the prephonatory glottal halfwidth $\xi_0$, a increase in the vocal fold length $l_g$, an increase in the vocal fold vertical thickness $T$, and a
increase in the vocal tract inertance (Chan and Titze, 2006).

One major improvement in the experiments conducted by Chan and Titze was that the superficial layer of the lamina propria was simulated by implanting viscoelastic biomaterials in between the silicone epithelial membrane and the stainless steel vocal fold body instead of using viscous fluids. These biomaterials were fat, hyaluronic acid (HA), and fibronectin. The results of the hyaluronic acid experiments are shown in Figure 1.10, where the vocal tract was not present. The results of Figure 1.10 are consistent with two of the predictions carried by Eq. (1.4): an approximately linear relationship between $P_{th}$ and $\xi_0$ and an increase in $P_{th}$ as the viscous properties of the implant (with the higher concentration of HA) are increased.

Figure 1.10: Offset and onset phonation threshold pressures as a function of prephonatory glottal half-width, for two concentrations of hyaluronic acid implanted under the membrane used to model the vocal fold cover (Chan and Titze, 2006)
The other part of the experiment was to test the effect of vocal tract inertance by coupling a rectangular Plexiglas tube to the original model. The transverse dimension of the Plexiglas tube was 2.22 cm, about the same as the glottal length, but its dimension in the plane of the diagram was 1.27 cm instead of 2.7 cm, which can be seen in Figure 1.11. The addition of the tube allowed Chan and Titze to determine if the inertance of the vocal tract would lower the phonation threshold pressure in accord with Eq. (1.4).

![Figure 1.11: Schematic of the physical model of the larynx with coupling to a rectangular uniform-tube supraglottal vocal tract (Chan and Titze, 2006)](image)

Keeping all other key geometrical and biomechanical variables of the physical model constant, Figure 1.12 shows that $P_{th}$ was consistently lower when the physical model had a vocal tract than when the model did not have a vocal tract.

The experiments of Figures 1.4, 1.7, and 1.11 were designed to test specific predications
Figure 1.12: Phonation threshold pressure of the physical model as function of prephonatory glottal half-width, with and without a uniform-tube vocal tract (Chan and Titze, 2006)

of the surface wave model. These experiments record agreement with some of the trends predicted by Eqs. (1.1) and (1.2). For example, data in Figure 1.5 show a region of almost linear behavior as $\xi_0$ is increased (beyond 1 mm), and similar behavior is seen in Figures 1.10 and 1.12 for the 2006 experiments. In Figure 1.6 the phonation threshold pressure increases as the fluid is made more viscous and this trend is also seen in Figure 1.10. Clearly, the vocal tract lowers the phonation threshold pressure (Figure 1.12), as prescribed by the surface wave result of Eq. (1.4).

However, the experiments have effects that are not consistent with the predictions of the surface wave model. For example, the data of Figure 1.9 do not suggest that the divergent glottis has a lower phonation threshold pressure than a convergent shape, which is
predicted by Eq.(1.2). Further, it is not clear that the experiments are all consistent with each other, since the results of Figure 1.5 for the rectangular glottis have a clear minimum near $\xi_0 = 0.5 \text{ mm}$, which is not seen in the data of Figures 1.10 and 1.12.

This picture suggests that certain effects left out of Titze's (1988) original formulation of the surface wave model should be carefully examined. In Chapter 3, the question of the curvature of the medial surface of the vocal fold is analyzed, in order to isolate the leading corrections to the formulas for the threshold pressure given in Eqs. (1.1) and (1.2).

A more comprehensive analysis of some possible corrections is carried out in Chapter 4, where the lumped element model of phonation developed at Bowling Green by Nardone (2007) is applied to the 2006 experiments of Chan and Titze. Nardone considered the vocal folds as coupled oscillators with values of the stiffness, mass, and damping not far from those used in the original work of Ishizaka and Flanagan (1972) and experimental input for the introglottal pressures. Models of this sort have been used to address many aspects of both normal and pathological phonation. Our goal in Chapter 4 is to see if this lumped element model can be adapted to give a reasonable account of the data collected in the 2006 experiments of Chan and Titze.

The deeper theme running through our calculations is an attempt to clarify the role of physical models in the process of understanding human phonation. What properties do these models share with human phonation? What aspects of their data are consistent with data collected from human subjects? To us it seems that a reasonable beginning of such a project is to evaluate the success of mathematical models often used to account for certain aspects of human phonation in describing the results of experiments with physical models, such as the three done by Titze and his colleagues.

The immediate goal of this thesis is to ask such questions in the context of the surface wave model and Nardone's adaptation of Ishizaka and Flanagan's classic two-mass model.
A mathematical model was developed to investigate possible causes of jitter and shimmer by Marco Nardone (2007). The model builds on the two-mass model which was formulated by Ishizaka and Flanagan (1972), hereafter referred to as the IF72 model. In the IF72 model, each vocal fold is divided into top and bottom parts. These oscillators are anchored to the stationary parts of the larynx with nonlinear springs and dampers, and they are coupled to each other by linear springs. The oscillators are driven by aerodynamic forces and are allowed only lateral displacement. A schematic of the IF72 model is provided in Figure 2.1.

![Schematic of the IF72 model of the vocal folds (Nardone, 2007)](image)

Since Nardone was interested in jitter and shimmer, it was natural for him to explore asymmetrical oscillations with different sets of biomechanical parameters for the oscillators.
on the opposing sides of the glottis. But much of the theoretical work on phonation threshold pressure (Titze, 1988, Lucero, 1999) has been done with the symmetric glottis. Nardone’s equations can be applied to the symmetric glottis by choosing the parameters describing the left vocal fold to be the same as those of the right vocal fold. In this case, the left and right parts of Figure 2.1 oscillate as mirror images of each other. On the other hand, it is possible to adapt Nardone’s equations to the hemilaryngeal geometry of Figures 1.4, 1.7, and 1.11 by choosing stiffness parameters for one side that are much larger than the other side. This approach is used in Chapter 4.

The set of equations describing the motion of both vocal folds (Nardone, 2007) is given by,

\[
\begin{align*}
    m_{R1} \ddot{x}_{R1} + r_{R1} \dot{x}_{R1} + s_{R1}(x_{R1}) + k_{Re}[(x_{R1} - x_{R10}) - (x_{R2} - x_{R20})] &= F_{R1}, \\
    m_{R2} \ddot{x}_{R2} + r_{R2} \dot{x}_{R2} + s_{R2}(x_{R2}) + k_{Re}[(x_{R2} - x_{R20}) - (x_{R1} - x_{R10})] &= F_{R2}, \\
    m_{L1} \ddot{x}_{L1} + r_{L1} \dot{x}_{L1} + s_{L1}(x_{L1}) + k_{Le}[(x_{L1} - x_{L10}) - (x_{L2} - x_{L20})] &= F_{L1}, \\
    m_{L2} \ddot{x}_{L2} + r_{L2} \dot{x}_{L2} + s_{L2}(x_{L2}) + k_{Le}[(x_{L2} - x_{L20}) - (x_{L1} - x_{L10})] &= F_{L2}.
\end{align*}
\] (2.1)

In these equations, the subscripts “R” and “L” represent the right and left vocal folds, respectively and the subscripts “1” and “2” represent the lower and upper masses, respectively. The quantities \(m_{ij}\) represent the masses of the oscillators. The quantities \(x_{ij}\) represent the displacements of the oscillators away from the midline, while \(\dot{x}_{ij}\) and \(\ddot{x}_{ij}\) denote the velocities and accelerations of the oscillators, respectively. The quantities \(r_{ij}\) and \(s_{ij}\) represent viscous damping and restoring force functions. The spring constants \(k_{ic}\) describe the coupling of the upper and lower oscillators. For the open glottis, the restoring force for each oscillator includes linear and nonlinear terms, that is,

\[
s_{ij}(x_{ij}) = k_{ij}[(x_{ij} - x_{ij0}) + \eta_{kij}(x_{ij} - x_{ij0})^3],
\] (2.2)

where \(k_{ij}\) and \(\eta_{kij}\) represent the spring constants and nonlinear coefficients when the glottis is open, and \(x_{ij0}\) is the initial position of the oscillators relative to the glottal midline. When
the glottis closes, it is necessary to add an additional term to Eq. (2.2) to consider the elastic effects of the collision. The damping coefficients are also increased when collisions occur.

The driving forces applied in Nardone’s model were based on a combination of a modified version of the IF72 driving forces and data collected from the experimental apparatus M5 (Scherer et al., 2001). The modified IF72 driving forces were derived from IF72 model by limiting the strength of the partial vacuum that occurs during the diverging part of the glottal cycle. Intraglottal pressures collected with diverging angles (Scherer et al., 2001) suggest that the IF72 approach exaggerates the effects of the partial vacuum for diverging angles greater than 5°. Model M5 is a Plexiglas, static model of the normal human male larynx that is scaled up by a factor of 7.5. The apparatus is comprised of a rectangular wind tunnel, removable vocal fold pieces to consider varying geometrical shapes of the glottis, and sixteen pressure taps which are used to measure pressure distributions (Scherer et al., 2001). A three-dimensional schematic of M5 is presented in Figure 2.2. Air enters from the right and passes through the narrow gap between the Plexiglas inserts, one of which contains 14 pressure taps, as shown in Figure 2.3. After the air leaves the gap between the inserts, it enters a downstream volume with much larger cross section, where taps 15 and 16 are located. This larger volume simulates the vocal tract.
Details of the M5 apparatus, methodology, and some results are provided in the published literature (Scherer et al., 2001; Scherer et al., 2002; Thapa, 2005). However, most of the information obtained from M5 is in the form of an extensive collection of EXCEL spreadsheets located in the Department of Communication Sciences and Disorders at Bowling Green State University (Scherer, 2009).

For each subglottal pressure, the pressure along the axial distance $z$ of the glottis for different minimal diameters $d_{\text{min}}$ and included angles $\phi$ was measured. Examples of such pressure distributions are shown in Figure 2.4 for $d_{\text{min}} = 0.08$, a rectangular glottis ($\phi = 0^\circ$), a converging angle of $10^\circ$, and a diverging angle $10^\circ$. For each angle pressures were taken for 4 subglottal pressures, 3, 5, 10, and 15 cm of $H_2O$ ($1 \text{ cm } H_2O = 980 \text{ dynes} = 98 \text{ Pascals}$). All of the pressure distributions show a decrease in the subglottal region (taps 1 to 5). In Figure 4(A), the intraglottal pressure decreases in an approximately linear fashion as the airflow moves through the glottis (taps 6 to 11). In Figure 4(B), A large pressure gradient
within glottis can be seen as a common feature for converging angles. For the diverging angle of Figure 4(C), the pressure increases along the glottis after the airflow passes a minimum of the pressure distribution at the glottal entrance. Because of the formation of a flow jet which favors one vocal fold over the other, two pressure distributions were collected for each of the angles and subglottal pressure. When the air jet flows along a wall, this wall is called the flow wall (FW), and the other wall is called the non-flow wall (NFW).

The M5 data was collected for both the FW and NFW sides with symmetric vocal fold configurations for minimal glottal diameters, 0.02, 0.04, 0.08, 0.16, and 0.32 cm. At the smaller diameters, 0.005, 0.0075, and 0.01 cm, it was not necessary to consider differences between FW and NFW, since these diameters were too small to support bistable flow. For
Figure 2.4: Samples of M5 experimental data for a rectangular glottis ($\phi = 0^\circ$), a converging angle of $10^\circ$, and a diverging angle of $10^\circ$, when the minimal glottal diameter is 0.08 cm. Locations of the pressure taps are indicated by numbers from 1 to 16 (Scherer, 2009).
all of the diameters except 0.0075 cm, data were collected at four converging angles (-40, -20, -10, and -5 degrees), the uniform case (0 degree), and four diverging angles (5, 10, 20, and 40 degrees). At d=0.0075 cm, only the uniform case was considered. When the minimal diameter was 0.005, 0.01 cm, and 0.02 cm, data were taken for transglottal pressures of 3, 5, 10, 15, and 25 cm H$_2$O, and when the minimal diameter was 0.04, 0.08, and 0.16 cm, data were taken for transglottal pressures of 3, 5, 10, and 15 cm H$_2$O. When the minimal diameter was 0.32 cm, data were taken for transglottal pressures of 1, 3, and 5 cm H$_2$O.

In Nardone’s adaptation of the IF72 model, the total M5 data set was truncated to a rectangular grid bounded by diameters of 0.005 cm and 0.16 cm and subglottal pressures of 3 cm and 15 cm H$_2$O. Data outside these limits were too sparse to give reliable results. Linear interpolations were used to provide the appropriate pressure distributions at intermediate glottal diameters and transglottal pressures. When the subglottal pressure is outside of the range of the M5 data, a modified version of IF72 approach to the driving forces was used to supplement the M5 data.

Now we briefly review the fluid mechanical considerations underlying the IF72 approach to the intraglottal pressures. The IF72 analysis began with a consideration of the Bernoulli work-energy equation, which describes effects due to flow channels of different sizes. This equation may be expressed

$$P + \frac{\rho}{2} (\frac{U_g}{A})^2 = constant,$$

where $P$ is the static pressure, and $A$ is the area of the channel. The quantities $\rho$ and $U_g$ are the density of the air and the rate of airflow, or volume velocity, through the glottis, respectively. Empirical studies indicate that the pressure drop in the contraction region near the glottal entrance is increased by an average additional factor of 0.37 (van den Berg et al., 1957). This loss factor was included in the IF72 model. The Bernoulli equation was also used to consider the discrete area change between masses $m_1$ and $m_2$ and thus the pressure
change there is given by

\[ \Delta P_{\text{junction}} = \frac{1}{2} \rho U_g^2 \left( \frac{1}{A_{g1}^2} - \frac{1}{A_{g2}^2} \right), \]  

(2.4)

where \( A_{g1} \) and \( A_{g2} \) represent the cross-sectional areas between the opposing oscillators of Figure 2.1. Based on the Poiseuille equation (Ishizaka and Flanagan, 1972), the resistance to airflow in the first section of the glottis is given by

\[ R_{v1} = \frac{12 \mu l_g d_1}{A_{g1}^3}, \]  

(2.5)

where the quantity \( \mu \) is the viscosity and the areas of the medial surfaces of masses \( m_1 \) and \( m_2 \) are given by \( l_g d_1 \) and \( l_g d_2 \), respectively. From an analysis of momentum conservation, Ishizaka and Flanagan (1972) estimated the pressure recovery at the glottal exit to be equal to

\[ \Delta P_{\text{rec}} = \rho \frac{U_g^2}{A_1 A_{g2}} (1 - \frac{A_{g2}}{A_1}), \]  

(2.6)

where \( A_1 \) is the cross-sectional area of the first section of the vocal tract.

This sequence of equations may be used to find the total (or transglottal) pressure drop as the air flows through the entire glottis. Using these equations, it would only be possible to find the instantaneous pressure-flow relations for a fixed geometry. To include all dynamical effects requires that one consider the force required to accelerate the air due to its inertance, that is \( \Delta P_{\text{inert}} = I \frac{dU}{dt} \), where \( I = \rho \frac{d}{A} \) for each part of the glottis. Based on the above approach to aerodynamics gives the pressures on the medial surfaces of the vocal folds. These are inversely related to the glottal areas. The mean pressures acting on the lower and upper oscillators are given by,

\[ P_{m1} = P_s - 1.37 \frac{\rho}{2} \left( \frac{U_g}{A_{g1}} \right)^2 - \frac{1}{2} \left( \frac{12 \mu l_g d_1}{A_{g1}^3} U_g + \frac{\rho d_1}{A_{g1}} \frac{dU_g}{dt} \right), \]

\[ P_{m2} = P_{m1} - \frac{1}{2} \left[ \left( \frac{12 \mu l_g d_1}{A_{g1}^3} \right) U_g + \left( \frac{\rho d_1}{A_{g1}} + \frac{\rho d_2}{A_{g2}} \right) \frac{dU_g}{dt} \right] - \frac{1}{2} \rho U_g^2 \left( \frac{1}{A_{g2}^2} - \frac{1}{A_{g1}^2} \right). \]  

(2.7)

Although the pressures given by the IF72 model are reasonable during the opening part of the glottal cycle, they are not accurate for the closing, or divergent, part of the cycle. A
consistent feature identified in the M5 data is that during the divergent part of the cycle, the intraglottal pressure is typically a negative value about 20% of the subglottal pressure. This observation is used as a basis to modify the IF72 pressures by truncating the magnitude of the negative pressure during the divergent part of the cycle to 20% of the subglottal pressure. Figure 2.5 provides a comparison for the original IF72 force, the modified IF72 force, and the M5 force acting on mass $m_{R1}$. As seen from Figure 2.5, it is clear that the modified IF72 driving forces are a better approximation to the M5 data than the original IF72 forces, especially during the closing part of the glottal cycle.
Figure 2.5: Driving forces on mass $m_{R1}$ for the original IF72 model (upper), the modified IF72 force (center), and the M5 experimental data (lower). The transglottal pressure is 8 cm $H_2O$ and the glottal area is 0.05 cm$^2$ (Nardone, 2007)
Logical controls are necessary to describe the driving forces of Eq. (2.1) at different times during the glottal cycle since different glottal shapes (converging, rectangular, or diverging) occur at different times. A further consideration is whether the glottis is open or closed. The logical conditions that govern the application of the driving forces are given in following tables. Table 2.1 pertains when the modified IF72 approach is used for the pressures, and Table 2.2 if the M5 pressures are used.

<table>
<thead>
<tr>
<th>(A_{g1})</th>
<th>(A_{g2})</th>
<th>(F_{R1})</th>
<th>(F_{R2})</th>
<th>(F_{L1})</th>
<th>(F_{L2})</th>
</tr>
</thead>
<tbody>
<tr>
<td>(A_{g1} &lt; A_{g2})</td>
<td>(A_{g2} &gt; A_{g1})</td>
<td>(-0.20P_s l_g d_{R1})</td>
<td>0</td>
<td>(-0.20P_s l_g d_{L1})</td>
<td>0</td>
</tr>
<tr>
<td>(A_{g1} &lt; A_{g2})</td>
<td>(A_{g2} &gt; A_{g1})</td>
<td>(P_{m1} l_g d_{R1})</td>
<td>(P_{m2} l_g d_{R2})</td>
<td>(P_{m1} l_g d_{L1})</td>
<td>(P_{m2} l_g d_{L2})</td>
</tr>
<tr>
<td>(A_{g1} &lt; 0)</td>
<td>(A_{g2} &gt; 0)</td>
<td>(P_s l_g d_{R1})</td>
<td>0</td>
<td>(P_s l_g d_{L1})</td>
<td>0</td>
</tr>
<tr>
<td>(A_{g1} &lt; 0)</td>
<td>(A_{g2} &lt; 0)</td>
<td>(P_s l_g d_{R1})</td>
<td>0</td>
<td>(P_s l_g d_{L1})</td>
<td>0</td>
</tr>
<tr>
<td>(A_{g1} &gt; 0)</td>
<td>(A_{g2} &lt; 0)</td>
<td>(P_s l_g d_{R1})</td>
<td>(P_s l_g d_{R2})</td>
<td>(P_s l_g d_{L1})</td>
<td>(P_s l_g d_{L2})</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>(A_{g1})</th>
<th>(A_{g2})</th>
<th>(F_{R1})</th>
<th>(F_{R2})</th>
<th>(F_{L1})</th>
<th>(F_{L2})</th>
</tr>
</thead>
<tbody>
<tr>
<td>(A_{g1} &gt; 0)</td>
<td>(A_{g2} &gt; 0)</td>
<td>(\langle P_{M5FW1}\rangle l_g d_{R1})</td>
<td>(\langle P_{M5FW2}\rangle l_g d_{R2})</td>
<td>(\langle P_{M5FW1}\rangle l_g d_{L1})</td>
<td>(\langle P_{M5FW2}\rangle l_g d_{L2})</td>
</tr>
<tr>
<td>(A_{g1} &gt; 0)</td>
<td>(A_{g2} &lt; 0)</td>
<td>(P_s l_g d_{R1})</td>
<td>0</td>
<td>(P_s l_g d_{L1})</td>
<td>0</td>
</tr>
<tr>
<td>(A_{g1} &lt; 0)</td>
<td>(A_{g2} &gt; 0)</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>(A_{g1} &lt; 0)</td>
<td>(A_{g2} &lt; 0)</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

Determination of the glottal volume velocity \(U_g\) depends upon its downstream interaction with the vocal tract. Nardone (2007) chose the same form for his representation of the vocal tract as IF72. The IF72 model employs an equivalent network representation of the
vocal system to combine the movement of air through the glottis and the vocal tract. In this method, each part of the vocal tract is divided into segments with an inductance $L_i$, a resistance $R_i$, and a volume capacitance $C_i$. The inductance plays the role of the inertance, the resistance describes energy losses due to viscous effects and absorption of sound waves by the walls of the vocal tract, and the capacitance describes pressure changes due to the changing amount of air in the given segment. The acoustic properties of each segment are represented by an equivalent resistor-capacitor-inductor circuit. The network representation of the IF72 model is depicted in Figure 2.6.

![Network representation of the IF72 model](image)

Figure 2.6: IF72 network representation of the glottis and vocal tract (Ishizaka and Flanagan, 1972)

The network representation of Figure 2.6 allows one to apply the acoustical analogue of Kirchoff’s electromotive force law. This leads to a set of differential equations governing the
airflow through the system. These coupled equations take the form,

\[ (R_{K1} + R_{K2})|\dot{q}_g|\dot{q}_g + (R_{v1} + R_{v2})\dot{q}_g + (L_{g1} + L_{g2})\ddot{q}_g + L_1\ddot{q}_g + R_1\dot{q}_g + \frac{1}{c_1}(q_g - q_1) - P_s = 0; \]

\[ (L_1 + L_2)\ddot{q}_1 + (R_1 + R_2)\dot{q}_1 + \frac{1}{c_2}(q_1 - q_2) + \frac{1}{c_1}(q_1 - q_g) = 0; \]

\[ (L_2 + L_3)\ddot{q}_2 + (R_2 + R_3)\dot{q}_2 + \frac{1}{c_3}(q_2 - q_3) + \frac{1}{c_2}(q_2 - q_1) = 0; \]

\[ (L_3 + L_4)\ddot{q}_3 + (R_3 + R_4)\dot{q}_3 + \frac{1}{c_4}(q_3 - q_4) + \frac{1}{c_3}(q_3 - q_2) = 0; \]

\[ (L_4 + L_r)\ddot{q}_4 + R_4\dot{q}_4 - L_r\ddot{q}_r + \frac{1}{c_4}(q_4 - q_3) = 0; \]

\[ L_r(\ddot{q}_r - \ddot{q}_4) + R_r\dot{q}_r = 0, \]

(2.8)

where all the flow rates \( U_i \) have been replaced by their integrals over time \( q_i \), that is \( U_i = \dot{q}_i \).

This step allows one to write the equations above as second-order differential equations instead of the integro-differential equations used in the original IF72 work.

Nardone’s model is based on the ten nonlinear, ordinary, differential equations given by equations Eqs. (2.1) and Eqs. (2.8). They were solved simultaneously using the software Mathematica. In our study, the solutions will be analyzed graphically and numerically to identify the threshold pressure as a function of the requisite input parameters.
CHAPTER 3  
CURVATURE CORRECTION

Titze’s (1988) analytical expression for threshold pressure of rectangular glottis is given in Eq. (1.1). Since the membranes of Figures 1.4 and 1.7 become rounded when a fluid is flowing or when materials are inserted beneath it, it is appropriate to investigate the role of curvature of the medial surface. Then the simplest possible form for this curvature is a parabolic form, as shown in Figure 3.1.

![Figure 3.1: Schematic of the medial surface geometry with curvature](image)

The initial configuration of the glottis is given by

\[
\xi_0(z) = \frac{(\xi_{01} + \xi_{02})}{2} + \frac{\xi_{02} - \xi_{01}}{T} z + b (z^2 - \frac{T^2}{4}),
\]  

(3.1)

where the first term gives the average of the center \((z=0)\) of the vocal fold surface if curvature is not considered, the second term allows for a slant to the vocal fold wall, and the third
term introduces curvature of the vocal fold surface in the simplest possible way. Titze’s 1988 treatment of the surface wave model included only the first two terms. In Figure 3.1 and Eq. (3.1), $\xi_0(z)$ is the prephonatory glottal halfwidth, $\xi_{01}$ is the inferior glottal halfwidth, $\xi_{02}$ is the superior glottal halfwidth, $T$ is the vocal fold thickness, and $b$ is the glottal curvature coefficient.

The simplest case is that of the uniform glottis where $\xi_{02} = \xi_{01} = \xi_0$, and the second term disappears. This yields

$$\xi_0(z) = \xi_0 + b \left(z^2 - \frac{T^2}{4}\right).$$

(3.2)

Thus $\xi_0(0) = \xi_0 - b \frac{T^2}{4}, \xi_0(\pm T/2) = \xi_0$, and we can see that for positive $b$, $\xi_0(0) < \xi_0$, and the surface is convex. For negative $b$, the surface is concave.

Since the experiments done by Chan, Titze, and Titze (1997) used a physical model with a divergent prephonatory glottal geometry, it will be convenient to introduce the angle $\theta$, which is shown in Figure 3.1,

$$\tan \theta = \frac{\xi_{02} - \xi_{01}}{T},$$

(3.3)

and to introduce the coordinate of the center of the medial surface $\bar{\xi}$ in the absence of curvature, which is given by

$$\bar{\xi} = \frac{\xi_{01} + \xi_{02}}{2}.$$ (3.4)

After these modifications, the initial configuration for the glottis can be expressed by

$$\xi_0(z) = \bar{\xi} + \tan \theta \ z + b \left(z^2 - \frac{T^2}{4}\right).$$

(3.5)

The glottal area $A(z, t)$ depends on the initial shape function $\xi_0(z)$, a term to describe the time-dependence of the surface $\xi_1(z, t)$, and the glottal length $l_g$,

$$A(z, t) = 2 \ l_g \ [\xi_0(z) + \xi_1(z, t)].$$

(3.6)

The factor 2 occurs since each vocal fold is assumed to undergo a symmetric displacement relative to the glottal midline.
According to Titze’s (1988) model, the function $\xi_1(z, t)$ represents a surface wave and thus satisfies a wave equation. Therefore, it has the form of a general solution of a one-dimensional wave equation with wave velocity $c$, that is,

$$\xi_1(z, t) = \xi_1(t - z/c), \quad (3.7)$$

where the minus sign is required since the wave travels in the direction of the positive $z$-axis. This gives a time delay between the motion of the bottom and top lips of the vocal fold. By expanding Eq. (3.7) in a Taylor series around the midpoint ($z = 0$) the displacement of the cover at points along the medial surface can be approximated by

$$\xi_1 \left( t - \frac{z}{c} \right) \approx \xi(t) - \frac{z}{c} \dot{\xi}(t), \quad (3.8)$$

where $\xi$ and $\dot{\xi}$ are the midpoint displacement and midpoint velocity, respectively. This expansion can be carried to higher orders, but keeping only the first two terms suffices as long as the phase difference between the top and bottom surfaces is not too large.

The glottal areas at entry and exit are defined by setting $z = \pm T/2$ and introducing a time delay $\tau = T/2c$. Then,

$$A_1(t) = 2l_g[\xi_01 + \xi(t) + \tau \dot{\xi}(t)], \quad (3.9)$$

$$A_2(t) = 2l_g[\xi_02 + \xi(t) - \tau \dot{\xi}(t)]. \quad (3.10)$$

In terms of Bernoulli’s work-energy equation, the subglottal pressure $P_{\text{sub}}$ and the intraglottal pressure at $z$ are connected with the pressure at the glottal exit by

$$P_{\text{sub}} = P(z) + \frac{1}{2} \rho \, v^2(z) = P_2 + \frac{1}{2} \rho \, v_2^2. \quad (3.11)$$

Using the expression $v_i = U_g/A_i$ and setting $P_2 = 0$, the intraglottal pressure can be expressed as

$$P(z) = P_{\text{sub}} \left[ 1 - \frac{A_2(t)^2}{A(z)^2} \right]. \quad (3.12)$$
Defining the average glottal pressure by
\[ P_g(\xi, \dot{\xi}) = \frac{1}{T} \int_{-T/2}^{+T/2} P(z)dz. \] (3.13)

It may be expressed as
\[ P_g(\xi, \dot{\xi}) = \frac{P_{sub}}{T} \int_{-T/2}^{+T/2} \left[ 1 - \frac{A_2^2(t)}{A^2(z,t)} \right]dz. \] (3.14)

By substituting Eq. (3.6) and Eq. (3.8) into Eq. (3.14), and carrying out an expansion in powers of the ratio of \( \xi_1(z,t) \) to \( \xi_0(z) \), which assumes that the surface wave oscillations are small in comparison with the glottal halfwidth, the expression for the average pressure on the glottal surface is found to be
\[ P_g(\xi, \dot{\xi}) \approx \frac{P_{sub}}{T} \int_{-T/2}^{T/2} dz \left[ 1 - \frac{\xi_0^2(T/2)}{\xi_0^2(z)} - \frac{2 \xi_0(T/2) \xi_1(T/2, t)}{\xi_0^2(z)} + \frac{2 \xi_0^3(T/2) \xi_1(z, t)}{\xi_0^2(z)} \right]. \] (3.15)

After the \( z \) integration, the result is
\[ P_g(\xi, \dot{\xi}) \approx P_{sub} \left[ 1 + \left[ 2 \frac{T^2}{\xi_0^2(z)} \xi_0^2(t) \xi_1(t) \xi_0^2 - 2 \xi_0^2 \xi_1(t) - \xi_0^2 \right] \frac{I_1}{T} + 2 \xi_0^2 \xi_1(t) \frac{I_2}{T} - \frac{2 \xi_0^2 \xi_1(t)}{c} \frac{I_3}{T} \right]. \] (3.16)

The integrals in Eq. (3.16) are related to inverse powers of \( \xi_0(z) \), and they may be evaluated analytically for the form of \( \xi_0(z) \) given in Eq. (3.1). For our purposes it suffices to carry out an expansion in powers of certain dimensionless parameters, whose sizes can be controlled in the experiments. These parameters are \( \tan\theta \) and \( b \frac{T^2}{\xi} \), which gives the ratio of curvature effects to those associated with initial glottal geometry. For the present we keep terms through the second order in both parameters,
\[ I_1 = \int_{-T/2}^{T/2} \frac{dz}{\xi_0^2(z)} = \frac{T}{\xi^2} \left[ 1 + \frac{b T^2}{3 \xi} + \frac{b^2 T^4}{10 \xi^2} + \frac{T^2}{4 \xi^2} \tan^2\theta \left( 1 + \frac{b T^2}{\xi} \right) \right] + \cdots, \]
\[ I_2 = \int_{-T/2}^{T/2} \frac{z \, dz}{\xi_0^3(z)} = \frac{T^3}{\xi^4} \left[ 1 + \frac{2 b T^2}{5 \xi} + \frac{b^2 T^4}{2 \xi^2} + \cdots \right]. \] (3.17)

The expression for \( I_2 \) is not given, since it does not multiply \( \dot{\xi} \) and therefore is not directly involved in the transfer of energy to the vocal folds. Such terms determine the threshold pressure.
Applying Newton’s law to the vocal fold cover yields

\[ l_g T P_g = M \ddot{\xi} + B \dot{\xi} + K \xi. \] (3.18)

In this equation, \( M, B, \) and \( K \) are oscillator mass, damping coefficient, and stiffness of the oscillating tissue, respectively.

By substituting the result of Eq. (3.16) into the equation of motion (Eq. (3.18)), the effective damping \( B^* \) can be calculated by examining all of the coefficients of the linear velocity term, which yields

\[ B^* = B - l_g T P_{sub}(2\tau \xi_0^2 I_1 T - 2 \frac{\xi_0^2}{c} \frac{I_3}{T}). \] (3.19)

It is evident that oscillations will grow (or be sustained) whenever \( B^* \leq 0 \), which means a negative (or zero) damping. Therefore, setting the effective damping to zero will give an expression for the subglottal pressure at which the oscillation begins to grow, instead of being damped out. This pressure is of course the threshold pressure. It depends upon the glottal curvature coefficient and the prephonatory glottal angle as follows:

\[ P_{th} \approx \frac{B c \bar{\xi}}{l_g T^2} \left[ 1 - \frac{b T^2}{3 \xi} - \frac{T}{\xi} \tan \theta \left( 1 - \frac{3 b T^2}{10 \xi} \right) + \frac{T^2}{4 \xi^2} \tan^2 \theta + \cdots \right]. \] (3.20)

Considering that the expansion of Eq. (3.15) is only valid when the surface wave displacements are small in comparison with \( \xi_0 \), it is expedient to carry out an expansion in powers of \( \tan \theta \) and the curvature parameter \( b \). Working to the second order in \( \tan \theta \) and to the first order in \( b \) yields

\[ P_{th} = \frac{B c \bar{\xi}}{l_g T^2} \left[ 1 - \frac{b T^2}{3 \xi} - \frac{T}{\xi} \tan \theta \left( 1 - \frac{3 b T^2}{10 \xi} \right) + \frac{T^2}{4 \xi^2} \tan^2 \theta + \cdots \right]. \] (3.21)

From Eq. (3.21) it is easy to isolate the leading curvature correction for the rectangular case (where \( \theta = 0 \)). This is

\[ P_{th} = \frac{B c \xi_0}{l_g T^2} \left( 1 - \frac{b T^2}{3 \xi_0} \right). \] (3.22)

The first term, of course, gives Titze’s result of Eq.(1.1), and the second term shows that the curvature lowers the phonation pressure when the medial surface is convex (bent towards
the midline, positive b) and raises it when the medial surface is concave (bent away from the midline, negative b). However, the form of Eq. (3.22) does not suggest that the curvature dependence of the threshold pressure for the rectangular case will be easy to detect, since the curvature correction is equivalent to a small change of $\xi_0$, or $B$. Both $\xi_0$ and $B$ must be determined from experiment, and uncertainties of these parameters are likely to be indistinguishable from small effects due to nonzero $b$. One possible interpretation of Eq. (3.22) is to regard the curvature correction as determining an effective half width.

Keeping both the leading curvature correction and the leading angle correction yields

$$P_{th} = \frac{B c \tilde{\xi}}{l_g T^2} \left[ (1 - \frac{b T^2}{3 \tilde{\xi}}) - \frac{T^2}{\tilde{\xi}} \theta (1 - \frac{3 b T^2}{10 \tilde{\xi}}) \right].$$

This equation shows that when $b$ is small, the threshold pressure for positive angles (diverging) is lower than the threshold pressure for negative angles (converging), in accord with the behavior predicted by Eq. (1.2). Equation 3.23 shows that the slope of the threshold pressure curve (as a function of angle) is sensitive to the parameter $b$. However, Eq. (3.23) has a feature which makes the observation of the curvature correction difficult. The two factors in the parenthesis are nearly identical since the coefficients of their second terms (0.333 and 0.300) are almost the same. Thus, as in the case of Eq. (3.21), the curvature correction may be absorbed into either an effective glottal width or an effective damping term.

To examine the validity of Eq. (3.23), the first factor there can be determined from the experimental results of Figure 1.9. Choosing a prephonatory glottal halfwidth of 0.25 cm and glottal angle of 0° yields a value of 150 pa (1500 dynes/cm²) for the first factor of Eq. (3.23), if curvature effects are ignored. Setting $b = 0$ in Eq. (3.23), one obtains

$$P_{th} = \frac{B c \tilde{\xi}}{l_g T^2} (1 - \frac{T^2}{\tilde{\xi}}),$$

which can be used to calculate the threshold pressures at 0°, 5°, and -5° without the curvature correction, as shown in Figure 3.2. For comparison, Titze et al.’s (1997) measurements from Figure 1.9 are also included. Making a definite choice for the size of the curvature correction,
say $b \frac{T^2}{\bar{\xi}} = 0.5$ yields a value of 1800 dynes/cm$^2$ for the first factor of Eq. (3.23). Then one can use Eq. (3.23) to calculate the threshold pressure at small angles. As one can see in Figure 3.2, the curvature correction gives very small differences from the case $b = 0$. Figure 1.9 shows clearly that the calculated results indicate a uniform decrease from negative angle (converging) to positive angle (diverging), which is different from the experimental results which present a minimum near $\theta = 0^\circ$.

Figure 3.2: Comparison of the calculated threshold pressures with the experimental results of Titze et al. (1997)
CHAPTER 4
RESULTS WITH NARDONE’S MODEL

Nardone’s mathematical model is comprised of the four motion equations [Eqs. (2.1)], the six vocal tract equations [Eqs.(2.8)], and the force conditions listed in Tables 2.1 and 2.2. These ten nonlinear, ordinary, differential equations were solved by NDSolve function in Mathematica version 7.0 (32 bit). NDSolve is a general numerical differential equation solver. It can handle a wide range of ordinary differential equation and partial differential equations. The solutions of the equations produced by the NDSolve algorithm are numerical interpolating functions which can be studied by methods of calculus, statistics, and graphical analysis, and thus the ten differential equations of Nardone’s mathematical model were solved by obtaining the solutions in the form of interpolating functions of time. The solution set included the displacement equations of the four oscillators and the airflow equations for the glottis, the four segments of the vocal tract, and the mouth output segment (Nardone 2007).

The general procedure for determining the threshold pressure for a given set of conditions requires the following steps:

1. Set the initial condition including glottal areas, spring constants, viscous damping constants, and vocal tract parameters to specific values, such as those shown in Table 4.1 below.

2. Run the model at some subglottal pressure \( P_s \) and see if oscillations occur. Under most circumstances, running the model for about 0.100 s is enough to determine whether sustained oscillations occurred. The solutions of the model were presented graphically to observe the time dependence of the oscillator displacements.

3. If oscillations occur, then lower pressure \( P_s \) until the pattern of oscillations becomes a decaying pattern. The threshold pressure is the minimum subglottal pressure that gives sustained oscillations.
Selecting the set of parameters required considerable experimentation. The new set of parameters is called IF1 and the full suite of IF1 parameters is given in Table 4.1. They are quite different from the set Nardone (2007) used for his studies of jitter and shimmer, which were close to the set of “typical glottal parameters” chosen by Ishizaka and Flanagan (1972). It is clear from the Figure 1.4 that the size of the physical model is large compared to the human larynx, so that one would expect larger masses than those of the IF72 “typical” set. Some guidance in this choice is given by considering the ratio of the glottal thickness $T$ (1.1 cm) to that of the human case (0.3 cm), which is almost a factor of 4. Further, the glottal length of the physical model was 2.3 cm, almost twice that of the IF72 set (1.4 cm). Thus, it was decided to choose a total mass of 1.2 g, about 8 times that of the IF72 set (0.15 g). To keep the frequencies realistic, the stiffness parameters were also scaled up by factors of 5 or 6. The masses and thickness of the upper and lower masses have same ratio as IF72. The total thickness is 1.1 cm, the dimension of the Chan and Titze’s apparatus. Some of the preliminary runs were done with the symmetrical glottal geometry, where the parameters for the left fold were the same as those for the right fold. This was done to establish the regions of parameter space that would give rise to reasonable physical oscillations.

Because only a hemilarynx was present in Chan and Titze’s experiments instead of a full larynx, the left vocal fold parameters for the spring constants were multiplied by 1000 in order to immobilize this oscillator and the prephonatory glottal areas of the left side were set to zero. In this case, only the right vocal fold was allowed to move. In addition, to obtain the threshold pressure for the case of no vocal tract, the vocal tract differential equations in Nardone’s mathematical model were modified to keep only one segment as the vocal tract. The vocal tract area was changed from 2.82 $cm^2$ (Chan and Titze’s value) to 6 $cm^2$ and then to larger values. Since the inertance of the vocal tract depends inversely on its cross-sectional area, this should be a reasonable procedure to minimize the effects of the vocal tract. The threshold pressures were found at the glottal halfwidths of 0.025, 0.05, 0.075, 0.1, 0.15, 0.2,
Table 4.1: List of parameters and their values for the two-mass model

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Description</th>
<th>Typical Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>( m_{R1} )</td>
<td>mass of lower right oscillator</td>
<td>0.97 g</td>
</tr>
<tr>
<td>( m_{R2} )</td>
<td>mass of upper right oscillator</td>
<td>0.19 g</td>
</tr>
<tr>
<td>( d_{g1} )</td>
<td>thickness of lower mass(inferior/superior)</td>
<td>0.916 cm</td>
</tr>
<tr>
<td>( d_{g2} )</td>
<td>thickness of upper mass(inferior/superior)</td>
<td>0.184 cm</td>
</tr>
<tr>
<td>( l_g )</td>
<td>length of vocal folds(anterior/posterior)</td>
<td>2.22 cm</td>
</tr>
<tr>
<td>( k_{R1} )</td>
<td>spring constant, lower right mass, open glottis</td>
<td>440000 dyne/cm</td>
</tr>
<tr>
<td>( k_{R2} )</td>
<td>spring constant, upper right mass, open glottis</td>
<td>75000 dyne/cm</td>
</tr>
<tr>
<td>( h_{R1} )</td>
<td>spring constant, lower right mass, closed glottis</td>
<td>240000 dyne/cm</td>
</tr>
<tr>
<td>( h_{R2} )</td>
<td>spring constant, upper right mass, closed glottis</td>
<td>24000 dyne/cm</td>
</tr>
<tr>
<td>( k_{Rc} )</td>
<td>coupling spring constant, right side</td>
<td>20000 dyne/cm</td>
</tr>
<tr>
<td>( \eta_{kR1} )</td>
<td>nonlinear spring coefficient, upper right mass, open glottis</td>
<td>100</td>
</tr>
<tr>
<td>( \eta_{kR2} )</td>
<td>nonlinear spring coefficient, lower right mass, open glottis</td>
<td>100</td>
</tr>
<tr>
<td>( \eta_{hR1} )</td>
<td>nonlinear spring coefficient, upper right mass, closed glottis</td>
<td>500</td>
</tr>
<tr>
<td>( \eta_{hR2} )</td>
<td>nonlinear spring coefficient, lower right mass, closed glottis</td>
<td>500</td>
</tr>
<tr>
<td>( r_{R1} )</td>
<td>viscous damping constant, lower right mass, open glottis</td>
<td>68 g/s</td>
</tr>
<tr>
<td>( r_{R2} )</td>
<td>viscous damping constant, upper right mass, open glottis</td>
<td>20 g/s</td>
</tr>
<tr>
<td>( r_{R1c} )</td>
<td>viscous damping constant, upper right mass, closed glottis</td>
<td>220 g/s</td>
</tr>
<tr>
<td>( r_{R2c} )</td>
<td>viscous damping constant, upper right mass, closed glottis</td>
<td>45 g/s</td>
</tr>
<tr>
<td>( \rho )</td>
<td>density of warm, moist air</td>
<td>( 1.14 \times 10^{-3} ) g/cm(^3)</td>
</tr>
<tr>
<td>( \mu )</td>
<td>shear viscosity of air</td>
<td>( 1.8 \times 10^{-4} ) dyne s/cm(^2)</td>
</tr>
<tr>
<td>( c )</td>
<td>speed of sound in warm air</td>
<td>35000 cm/s</td>
</tr>
<tr>
<td>( l_i )</td>
<td>length of vocal tract</td>
<td>16.4 cm</td>
</tr>
<tr>
<td>( A_i )</td>
<td>area of vocal tract</td>
<td>2.82 cm(^2)</td>
</tr>
</tbody>
</table>
0.25, and 0.3 cm, which are those used in Chan and Titze 2006 experiments. The results of the threshold pressures for the conditions described by the IF1 parameters of Table 4.1 with the vocal tract and without the vocal tract are presented in Figure 4.1. For comparison, Chan and Titze's measurements from Figure 1.12 are also included. Figure 4.1 shows that the threshold pressure was consistently lower when Nardone’s model included the vocal tract with Chan and Titze’s area than when the model used the large area. Thus the calculations follow the same trend as the experimental results in Figure 1.12. In addition, the differences of the threshold pressures calculated in the situations with and without vocal tract become smaller when the prephonatory halfwidths are smaller, which is also the same trend as the experiment results in Figure 1.12. However, the magnitudes are clearly different, since the calculated differences are much smaller than the measured differences at small halfwidths. At larger halfwidths the size of the difference between the two calculations is closer to the measured difference.

So one is forced to conclude that there is some kind of incompatibility of the mathematical model and the physical model used by Chan and Titze. One possibility for a difference would arise if the glottal halfwidths were not determined in the same way. Another worrisome difference is the appearance of a minimum in the calculated results for both cases, which does not show up in Chan and Titze’s 2006 data. It is tempting to suggest that the calculated minimum is an example of the effect recorded in Figure 1.5.

The threshold pressure was 700 \( \text{dyne/cm}^2 \) lower for the physical model with vocal tract than for situation without the vocal tract when the prephonatory halfwidth was 0.3 cm in Chan and Titze’s experiments. Enlarging the area of vocal tract to simulate the effect of no vocal tract given larger differences between the results with and without the vocal tract. When the area of vocal tract is 8 \( \text{cm}^2 \), the threshold pressure with the vocal tract is 700 \( \text{dyne/cm}^2 \) lower than the threshold pressure without the vocal tract at 0.3 cm prephonatory halfwidth, as shown in Figure 4.2. However, the situation at smaller half widths is not really
Figure 4.1: Comparison of threshold pressures when the parameter set IF1 (Table 4.1) is used and the results of Chan and Titze (2006) improved and the minimum still appears in the calculated results. Increasing the area of the large vocal tract to 12 cm$^2$ increased the separation between the two calculated curves, so that the differences between the vocal tract and the no vocal tract at 0.15 cm and 0.20 cm halfwidths were closer to those measured by Chan and Titze. However, the separation of the two calculated curves was too large at 0.30 cm and again too small when the halfwidth were 0.1 cm or less. Thus, the strategy of simulating a decreasing effect of the vocal tract by increasing its cross sectional area gives a reasonable account of some of the qualitative features of Chan and Titze’s experiments, but fails to give a quantitative account.

Several calculations were done in an attempt to isolate the nature of the difficulty encountered in fitting Chan and Titze’s data. Most of the IF1 parameters were kept the
Figure 4.2: Comparison of threshold pressures when the parameter set IF1 is used and the results of Chan and Titze (2006). The vocal tract area was set to 8 \(cm^2\) to simulate the case of no vocal tract.

same, but the damping constants were varied. Two different set of damping parameters are required to give reasonable partial fits Chan and Titze’s data, as shown in Figure 4.3. From this figure, one sees that Nardone’s model can only fit a part of the experimental results. The IF2 set of parameters with 40 and 20 \(g/s\) viscous damping constants gives a good fit with the threshold pressures when the prephonatory halfwidths are 0.15, 0.2, 0.25, 0.3 cm. However, the IF3 parameters with 40 and 35 \(g/s\) fit the threshold pressures at 0.025, 0.05, 0.075, and 0.1 cm prephonatory halfwidths. It is worth nothing that the calculated results have a tendency to give linear behavior which is only found in certain parts of the data. It has been difficult to isolate the factors responsible for the discrepancies between calculation and experiment in Figure 4.3.
To explore the role played by the damping coefficients in determining the threshold pressure, a set of modified IF1 parameters with different viscous damping constants was compared with IF1 in Figure 4.4. These parameters, designated IF4, were obtained by lowering the viscous damping constants $r_{R1}$ and $r_{R2}$ to 40 and 12 g/s, respectively. The results of Figure 4.4 show an increase in $P_{th}$ as the damping coefficient increases, and the differences between threshold pressures obtained with the two sets IF1 and IF4 tend to become smaller as the prephonatory halfwidths get smaller. Another feature of Figure 4.4, which suggests that lowering the damping constants is not a fruitful way to pursue an explanation of the Chan and Titze experiments, is that the effects of the vocal tract mostly disappear for the IF4 parameter set.
Figure 4.4: Comparison of the threshold Pressures as the damping parameters get smaller
CHAPTER 5
CONCLUSIONS

The calculations carried out in Chapter 3 to examine the curvature corrections were not conclusive. This was a consequence of an expansion in powers of a dimensionless parameter, and the final formulas only included curvature effects to the lowest order. The integrals that arose can be evaluated exactly in an analytic form and thus the question of possible manifestations of curvature corrections can be investigated more thoroughly. It would be interesting to explore more systematically a series of experiments where the curvature correction could be observed. If these higher-order effects reveal an interferences between curvature effect and angle effects, curvature may have a role to play in understanding the differences in the experiments of Figure 1.9 and the formula of Eq. (1.2).

Nardone’s adaptation of the IF72 model revealed that a lumped element model could account for several of the features observed in Chan and Titze’s 2006 experiments. These included lower threshold pressures as the effect of vocal tract inertance was decreased of the threshold pressure with glottal width. However, the calculations of Chapter 4 provided a result not seen in the 2006 experiments, namely, a minimum phonation pressure around 0.5 mm, in contrast to the monotonic behavior of Figures 1.10 and 1.12. Since such small values of glottal halfwidth may allow the vocal folds to collide, which would lead to more energy dissipation and thus increased threshold pressure, the effects of such collisions on threshold pressure should be investigated. It will be interesting to see if such an investigation leads to some insight into the minimum (Figure 1.5) in the threshold pressure observed in the 1995 experiments of Titze et al..

Another calculation that the research in this thesis has suggested is an improvement in the treatment of vocal tract inertance beyond the result of Eq. (1.4), where the time-dependence of the glottal flow was considered only in the leading order. A preliminary calculation has succeeded in a more accurate solution of the glottal flow equation (Fulcher
et al., 2009) and shows promise of giving a quantitative account of the data in Figure 1.12.
REFERENCES


