Evaluation of Contact Pressure in Human Vocal Folds During Phonation Using High-Speed Videoendoscopy

Jennifer Robb

Master of Engineering

Department of Mechanical Engineering

McGill University

Montreal, Quebec

2011-04-04

A thesis submitted to McGill University in partial fulfillment of the requirements of the degree of Master in Engineering

ACKNOWLEDGEMENTS

I am grateful to my supervisor, Prof. Luc Mongeau, for his guidance, direction, and support in pursuing this project. I would also like to extend my gratitude to Prof. Katherine Verdolini, Nicole Li, Leah Helou, and Dr. Clark Rosen at the University of Pittsburgh. I thank Dr. Karen Kost and Jackie Griguolo at the Montreal General Hospital for their patience when our prototypes did not work, for their creative suggestions, and for their enthusiasm for this research. For being a willing and resilient test subject, I express my gratitude to Alfred Chan.

I owe much to Tony Micozzi, Sam Minter, Andy Hofmann, and Roy Westgate from the Machine Tool Lab for their expertise in creating many tools used in this project. I also thank Elizabeth Price for her phenomenal enthusiasm and work ethic in performing valuable but sometimes tedious experimental tasks. Li-Jen Chen and Jong Beom Park provided me with guidance for which I am grateful.

I thank Jean Elder for supporting my education. I am grateful to my parents for encouraging me to pursue graduate studies in this field and for supporting me in doing so. I thank Philippe for his infinite patience and encouragement.

This study was financially supported by the National Institute for Deafness and other Communication Disorders (NIH-NIDCD) through grants R01 DC008290 (K. Verdolini, P.I.) and R01 DC005788 (L. Mongeau, P.I.).

ABSTRACT

It is generally believed that voice overuse and abuse may eventually lead to a number of voice disorders. The link between mechanical stresses in vocal fold tissue and vocal fold damage, however, has not yet clearly been established. One key factor is often hypothesized to be the contact pressure associated with impact between the vocal folds during voicing. Studies of the contact pressure on the medial surface of the vocal folds are scarce. The objective of the present study was to estimate contact pressure in human subjects using high-speed videoendoscopy. The edge velocity of the vocal folds was estimated from the analysis of consecutive digital images. The edge velocity before and after contact, along with ad hoc approximate values for contact area and vocal fold mass, allowed the contact pressure to be estimated from the impulse momentum form of Newton's second law. The results were verified through comparisons with directly measured values in a silicone model. Investigations were carried out in human subjects to compare the contact pressure from high-speed videoendoscopy with contact pressure measured directly using a miniature probe microphone. The contact pressure values estimated from high-speed video were between 600 Pa and 9200 Pa. The values from the probe microphone were between 250 Pa and 800 Pa. The probe microphone values were significantly lower than those from high-speed video. The discrepancies are believed to be mainly due to interference between the probe and the vocal fold motion. Further work is needed to assess the accuracy of quantitative contact pressure estimations from videoendoscopy.

RÉSUMÉ

La sur-utilisation et l'abus de la voix peuvent mener à des problèmes qui nécessitent des soins cliniques. Les forces dues à l'impact des cordes vocales durant leurs oscillations sont assumées être très importantes dans la formation et la propagation de lésions. Aucun lien direct n'a encore été établi entre les contraintes mécaniques associées à l'impact entre les cordes vocales et les problèmes pathologiques. Le but de cette étude est l'estimation des contraintes mécaniques dues à l'impact entre les cordes vocales lors de la phonation à l'aide d'une caméra vidéo-endoscopique. La vitesse en bordure des cordes vocales a été estimée à partir d'une analyse d'images. La vitesse des cordes vocales ainsi que des valeurs estimées de la masse et de la surface de contact ont permis de calculer la pression sur la surface. Les résultats ont été vérifiés à l'aide d'une comparaison avec les mesures obtenues à l'aide d'une sonde dans une modèle auto-oscillant en silicone. Une série d'expériences a été effectuée sur des sujets humains. Le mouvement des cordes vocales a été enregistré avec une caméra vidéo-endoscopique pour estimer les contraintes d'impacts. Une sonde a été insérée entre les cordes vocales des sujets humains durant la phonation pour une mesure directe. Les valeurs de contraintes d'impacts estimées par l'analyse des images étaient entre 600 Pa et 9200 Pa. Les valeurs obtenues par mesure directe étaient entre 250 Pa et 800 Pa. Cette sous-estimation est due à l'effet de la sonde sur l'oscillation des cordes vocales. De plus amples vérifications sont nécessaires afin d'évaluer le potentiel de cette méthode à des fins cliniques.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1
Introduction

1.1 Motivation

It is estimated that up to 6% of Canadians are affected by voice disorders [1]. Voice disorders are the result of injury to the vocal fold tissue associated with phonation (speaking or singing). During phonation, the vocal folds vibrate at rates from 100 Hz to 1000 Hz. During normal speaking or singing, in every oscillation cycle the vocal folds collide, resulting in mechanical stress induced by the contact stress on the medial surface of the vocal folds. It is hypothesized in the present study that the contact pressure on the vocal fold surface during phonation is a primary indicator of tissue damage [2]. In normal voice usage, the vocal fold tissue heals over time and the cumulative tissue damage is insignificant. In some cases, however, vocal fold injury is more severe and may cause chronic voice disorders. Communication disorders, including voice disorders, have a negative impact on quality of life. Voice disorders also have a detrimental impact on the careers of teachers, singers, or other professionals who require the use of their voices.

Vocal fold injury usually involves lesions, defined as abnormal tissue caused by disease or injury, around the centre of the medial edge of the vocal folds. Such lesions affect phonation and cause discomfort. Voice disorders resulting from lesions are normally first treated with voice therapy, including voice rest and phonatory exercises. In cases where voice therapy is not effective, lesions may be repaired surgically, although surgical intervention often induces scarring.

Improvements to voice therapy treatment plans could result in a reduction in the number of cases of voice disorders requiring surgery. One obstacle to developing effective voice disorder treatments is that no link between contact pressure, the hypothesized mechanism of vocal fold injury, and tissue damage has yet been clearly established. Additionally, individual patients respond differently to voice therapies, indicating that patient-specific treatments might be beneficial. Quantitative information on the vocal fold contact pressure in individual patients could help voice therapists optimize treatment plans for each patient.

Vocal fold contact pressure has been measured in human subjects [3], [4], [5], but high quality data is still only available for a very limited number of subjects. The reason for this scarcity is often attributed to difficulties in sensor positioning, and to subject intolerance for the presence of a measurement probe between the vocal folds [6]. Most subjects do not tolerate a foreign object within the glottis, even with the use of anesthetic. The vocal folds' primary function is to protect the airway. One natural reflex is to violently expel any foreign object from between the vocal folds.

A better understanding of contact pressure and mechanical stresses is needed to assess the role of mechanical factors in voice disorders. Optimized patient-specific treatments are not yet available due to the lack of technologies to quantify tissue damage and mechanical stresses. A minimally invasive method for determining the contact pressure could allow data to be collected from a wider range of subjects and facilitate patient-specific voice disorder treatment.

1.2 Research Objectives

The primary objective of this research was to quantify the pressure on the contacting portions of the medial surfaces of the vocal folds during phonation. The influence of voice type and sound pressure level was investigated. A minimally invasive method that involved no sensor between the vocal folds was used. Videoendoscopy is commonly used in the clinic to capture video images of the vocal folds. Rigid endoscopes require little or no anesthetic, and are relatively easily tolerated by subjects. High-speed videoendoscopy captured the vibration of the vocal folds with a spatial resolution of 512 x 512 pixels. A frame rate of 2000 frames per second was used to capture approximately nine images per oscillation cycle, for a subject phonating at 220 Hz.

The first part of this research was to verify a method of assessing vocal fold contact pressure based on image analysis from high-speed videoendoscopy. The recorded video was processed and vibration parameters were extracted and used to estimate the contact pressure. Verification experiments were performed using physical model vocal folds. The contact pressure in the physical model was assessed using the image analysis method and results were compared with values measured directly using a probe microphone.

The second part was to apply the image analysis method in human subjects. Again, image analysis results were compared to directly measured contact pressure values obtained in a previous study [6].

1.3 Organization of the Thesis

The remainder of this document is organized as follows. Background information on voice production, vocal fold damage, mechanical stress in the vocal folds, high-speed videoendoscopy, and image processing is reviewed in Chapter 2. In Chapter 3 the methodology used in this study is described for both physical model verification and human subject experiments. Results are presented in Chapter 4. In Chapter 5, the limitations and sources of error in this study are discussed. Conclusions and suggestions for future work are summarized in Chapter 6.

CHAPTER 2
Background

2.1 Voice Production

Voice is produced when the vocal folds, two lips of soft tissue located within the larynx, vibrate. Air flow is induced by the air pressure below the vocal folds, referred to as the subglottal pressure. The glottis is the orifice between the vocal folds. It defines the supraglottal system and the subglottal system in voice production. Figure 2–1 shows a diagram of the supraglottal system, the subglottal system, and the anatomy of the chest and neck, involved in voice production.

An image of human vocal folds obtained from a videoendoscopic recording is shown in Figure 2–2. The anterior, posterior, medial, and lateral directions are identified in the image. The anterior direction points towards the front of the body, the posterior direction points towards the back. The medial direction points towards the mid-line of the body and the lateral direction points towards the outside of the body. The superior and inferior directions are towards the top and bottom, respectively.

Vocal fold tension is controlled by muscles [7]. The tension and length of the vocal folds alter the voice pitch, defined as perceived frequency. The muscles in the larynx also control the type of voice. A “pressed” voice is obtained when the vocal folds are overly adducted and the voice is strained. This voice type is believed to be representative of some types of pathological voice [8]. A “breathy” voice type is produced when the vocal folds are less adducted than in normal voice. This voice type is thought to be representative of a healing voice. Breathy voice often involves low amplitude oscillations with no contact.

2.2 Phonotrauma

Voice disorders may be caused by surgical complications, or by phonotrauma, defined as voice overuse, misuse, and abuse. Voice disorders caused by surgery include vocal fold paralysis and vocal fold scarring. Vocal fold paralysis prevents the vocal folds from closing properly at the mid-line [7]. This results in a perceived breathy voice and an increase in airflow required to produce phonation.

Vocal fold scarring increases the stiffness of the vocal fold tissue locally [7]. This change in material properties of the tissue hampers vocal fold vibrations and affects phonation.

Voice disorders caused by phonotrauma include nodules, cysts, and polyps. All three result from different types of lesions on the vocal folds. A general description of each is outlined below. For a thorough review of different vocal fold lesions see reference [9].

Nodules are small, round, hard, lesions located bilaterally in the centre of the medial edge of the vocal folds. They are differentiated from cysts and polyps by their white, opaque colour [10]. Nodules prevent the vocal folds from closing completely, and cause the voice to be hoarse [11, 2].

Polyps are larger than nodules, and usually occur unilaterally. They are translucent and red in colour and may be pedunculated, or a stalk-shaped mass of tissue.

Like nodules, polyps are formed in the centre of the medial edge of the vocal folds and cause hoarseness and vocal fatigue [10, 12].

Cysts vary in size, location, and appearance. They appear translucent and yellow and occur on the superior surface of the vocal fold. In addition to phonotrauma, cysts may also be caused by obstructed mucous glands or congenital defects [10, 12].

Nodules, polyps, and in some cases cysts are presumably formed where the contact pressure is the greatest. Voice disorders occur more commonly in people who use loud voice frequently, such as teachers or singers. Women and children also suffer from voice disorders more frequently than men. One hypothesis to explain this is that women have higher pitched voices, and therefore experience a greater rate of vocal fold collisions in their normal speech. These factors support the hypothesis of a relationship between contact pressure and voice disorders [7].

2.3 Measurement of Contact Pressure

The mechanical stresses on contacting surfaces include a normal component, called pressure, acting in the direction perpendicular to the contact area, and a shear stress acting along the transverse direction. Efforts to quantify contact stresses on vocal folds have been made using numerical models, physical models, excised canine larynges, and human subjects.

An aeroelastic model of voice production was used by Horacek et al. to investigate vocal fold contact stresses [13]. The model had two degrees of freedom, with three masses and two springs. A Hertzian impact model was used to estimate the contact pressure. The Hertzian impact model assumes a collision between two deformable bodies with curved surfaces at the location of contact. The maximum contact pressure values were found to be between 2000 Pa and 3000 Pa. The contact pressure was found to increase with sound pressure level (SPL) and lung pressure. Similar numerical studies reported contact pressure values up to 3000 Pa or 4000 Pa [14], [15].

The stress on the superior surface of a physical model of the vocal folds was investigated by Spencer et al. [16]. Digital image correlation (DIC) was used to measure the strain on a silicone physical model. The strain was measured using a series of digital images of the deformed surface covered with a random speckle pattern. A Hertzian impact model was used to obtain the contact pressure during collision from the measured strain and known material properties. Contact pressure values were reported to be slightly lower than 1000 kPa. Attempts have been made to use DIC to measure the strain on the superior surface of an excised porcine larynx in a flow supply [17]. DIC has not yet been used in human subjects due to the toxicity of the dyes required to produce a random speckle pattern.

Excised larynges have been used to investigate vocal fold contact pressure. Jiang and Titze [18] used excised canine larynges in a hemilarynx configuration, where just one vocal fold is used. A miniature pressure sensor was inserted into the port within a Plexiglas wall to measure the contact pressure on the medial surface of a canine vocal fold. Five different larynges were tested for different subglottal pressures while all other parameters were kept constant. The frequency of oscillation was not reported. For a subglottal pressure of 1000 Pa, the measured contact pressure values for the five larynges were between 500 Pa and 2000 Pa. For a subglottal pressure of 2000 Pa, the measured values were between 2000 Pa and 4500 Pa. For a subglottal pressure of

3000 Pa, the measured values were between 3500 Pa and 5000 Pa. An approximately linear relationship between subglottal pressure and contact pressure was found, but the authors cautioned that further studies are required to confirm the relationship.

Two complete excised canine larynges were used in a study by Verdolini et al. [19]. A small piezoresistive pressure sensor was used to measure the contact pressure at the anterior-posterior midpoint. Measurements were made for three different subglottal pressures, 800-900 Pa, 1400 Pa, and 2000 Pa. The tension in the vocal folds was adjusted to yield three different elongations: 80%, 100%, and 120% of the resting length, as well as three different glottal gaps: -1 mm (vocal folds overlapping), 0 mm (vocal folds just barely touching), and 3 mm (vocal folds apart). Data could not be collected for every combination of test parameters. Due to the sparse data set, no clear relationship between the test parameters and the contact pressure was observed. The reported contact pressure values were between 300 Pa and 5300 Pa.

In a similar excised larynx experiment, Jiang et al. [20] measured the contact pressure with a miniature pressure sensor while using photoglottography to measure the relative displacement of the medial edge of the vocal folds. The vocal fold acceleration was obtained from the second derivative of the displacement. A linear relationship between the acceleration and the contact pressure was found. The slope of the linear regression between contact pressure and acceleration was not consistent between larynges, however. The measured contact pressure values were between 300 Pa and 2600 Pa.

In two previous studies, a miniature pressure sensor with a diameter of 1.8 mm and a thickness of 0.4 mm was used to measure the contact pressure in human subjects. The first study was an exploratory study and for this reason the results were deemed “largely descriptive rather than inferential” [3]. Data were collected from twenty subjects with a variety of voice disorders and two healthy subjects. The contact pressure was measured at different locations between the arytenoids during a variety of phonatory and non-phonatory (i.e. throat clearing) tasks. The arytenoid cartiages are pyramid shaped and attach to the vocal folds at the posterior part of the larynx [7]. The contact pressure between the vocal folds was measured in some of the subjects. The contact pressure in the mid-membranous region of the contacting portion of the vocal folds was found to range from less than 1000 Pa to 4000 Pa.

Similarly, the contact pressure in human subjects was measured by Verdolini et al. [4]. The subjects produced three different voice types, three different pitches, and three different loudnesses. Quality data with consistent sensor placement was only obtained for seven of twenty subjects. Reported problems include the subjects' strong gag reflex and difficulties in sensor placement. The contact pressure values varied between 400 Pa and 3200 Pa for the seven subjects and the twenty-seven different phonation conditions.

In another study, a thin force sensor with a thickness of 0.29 mm was used to measure the collision force between the vocal folds [5]. Four subjects with one vibrating and one paralyzed vocal fold, and one subject with both vocal folds vibrating, participated in the study. Due to the subject's intolerance of the sensor, challenges in positioning the sensor in the centre of the vocal folds, and challenges in verifying sensor placement, only one segment of analyzable data could be obtained from the subject with two healthy vocal folds. Quality data were obtained for only two subjects of the four subjects with one vibrating vocal fold. The measured collision force values were reported to be between 5 mN and 210 mN. Since force and not pressure was measured, the results can not be compared directly to those of other studies. If the contact area of the vocal folds were known, the contact pressure could be determined by calculating the force divided by the contact area. Approximating the vocal fold contact area as a rectangle measuring 15 mm by 3 mm, the corresponding range of contact pressure measured in this study would be 100 Pa to 4700 Pa.

2.4 Stroboscopy and High-speed Videoendoscopy

The observation of vocal fold vibrations via either rigid or flexible endoscopy is important for the clinical evaluation of voice disorders [21, 22]. Stroboscopy is commonly used for clinical assessments. Video is recorded at a low frame rate (30 frames per second) and an asynchronous strobe light is used to capture complete glottal cycles through aliasing. One laryngeal contact microphone is used to measure phonation frequency. The main disadvantage of videostroboscopy is that it requires that the vocal fold oscillation be consistent from one glottal cycle to the next, which is often not the case in disordered voices.

High-speed videoendoscopy yields video recordings of vocal fold oscillation at frame rates between 2000 fps and 10000 fps. In most high-speed video systems, frame rates of 2000 fps and 4000 fps are achievable without the need to reduce spatial resolution. High-speed video allows for the complete glottal cycle to be recorded, with approximately ten images per cycle, depending on the frame rate and phonation frequency. Digital high-speed videoendoscopy has been used in voice research for nearly two decades. Prior to this, high-speed film had been used [23]. High-speed videoendoscopy is not widely used in standard clinical practice. In addition to higher cost, a lack of objective and quantitative assessment protocols and limited validation and repeatability measures have been cited as reasons for high-speed videoendoscopy sparsity in clinical settings [24].

2.4.1 Image Segmentation

Image processing is required to extract useful information from high-speed video recordings. The primary focus is normally on the glottal area. The glottal area is usually identified and segmented, and the glottal area waveform reconstructed. In each frame, the area and perimeter of the glottis is automatically identified as the region of interest, distinct from the background of the image. Different image segmentation methods have been used to automatically identify the glottal area.

Thresholding has been used for segmenting high-speed endoscopic video of the vocal folds [25, 26]. Pixels that belong to the glottal area are identified based on their intensity. Pixels with an intensity above a predetermined threshold value are considered background, and pixels with an intensity lower than the threshold value are considered part of the glottal area. The threshold value is normally determined from the histogram of the grayscale image. This method is efficient because it only requires each pixel to be examined once. Pixels that have a low intensity and are not part of the glottal area, however, may be improperly identified. Yan et al. [27] combined a thresholding method with a region growing segmentation algorithm to take advantage of the efficiency of the thresholding method without sacrificing accuracy.

Region growing algorithms have been used to segment the glottal area in high-speed endoscopic video of the vocal folds [23]. As for the thresholding method, the images are segmented based on pixel intensity, but with the requirement that the segmented pixels be connected in the same region [26]. User intervention is required to initiate the image segmentation process. Lohscheller et al. modified the basic region growing algorithm to reduce user intervention so that several glottal cycles could be analyzed without the need to re-initiate the segmentation process [28].

Level-set algorithms [29, 30] and active contours [31, 32] have also been used to segment the glottal area. Qin et al. used a level-set method to achieve a sub-pixel level detection of the vocal fold edge [29]. Active-contour methods generally yield good segmentation results without any user intervention [32] , but require more processing time [30].

2.4.2 The Kymogram and the Phonovibrogram

The kymogram displays information from a complete high-speed video sequence in a single image. Figure 2–3 shows one example of a kymogram. A sequence of images is obtained from high-speed endoscopic video. The pixel intensity along a fixed line in the medial-lateral direction is obtained from each frame in the sequence and reconstructed to form a single image. The kymogram is useful in examining asymmetries in vocal fold oscillation [33], identifying open and closed phases of the glottal cycle, and observing mucosal wave propagation [34].

The phonovibrogram is similar to the kymogram in that it also displays a complete high-speed endoscopic video sequence in one single figure. Phonovibrograms provide information over the entire length of the vocal fold as opposed to one specific location [21]. The phonovibrogram is formed by division of a segmented vocal fold image down the anterior-posterior axis. The right side of the image is rotated 180° around the posterior point. The distance between the vocal fold and the anterior-posterior axis is determined at a number of different points along the vocal fold. This distance is represented in the phonovibrogram by the colour intensity at that location. Specific patterns in the phonovibrogram may be identified and used to diagnose different voice types and voice disorders [35].

2.4.3 Quantification of Vocal Fold Dimensions and Kinematics

Data obtained from high-speed video endoscopy, such as the glottal area waveform, can not be quantified in absolute units unless the video image is calibrated. A number of different laser projection systems have been proposed to calibrate endoscopic video of the vocal folds in absolute units [36, 37, 38, 39]. Most laser projection systems project a laser beam along a second channel, attached to the side of a rigid endoscope. One single laser beam is split using two mirrors in series. One mirror reflects 50% of the laser beam, and the second mirror reflects 100% of the laser beam [36, 38]. Absolute dimensions are obtained assuming that the two laser dots are separated by a known distance, which is not accurate for angled surfaces. Another laser projection system projects a single laser dot onto the vocal folds [37]. The laser is projected at an angle relative to the optical axis of the endoscope. The position of the laser dot in the field of view of the endoscopic image can be used to determine the distance from the vocal fold to the endoscope. A second calibration is required to determine the size of the image at the specified distance from the lens. A flexible endoscope has also been used to project a laser grid pattern onto the vocal folds. This system transmits a laser beam through a fiber optic cable and displays a pattern of laser dots on the vocal folds using a miniature mask [39].

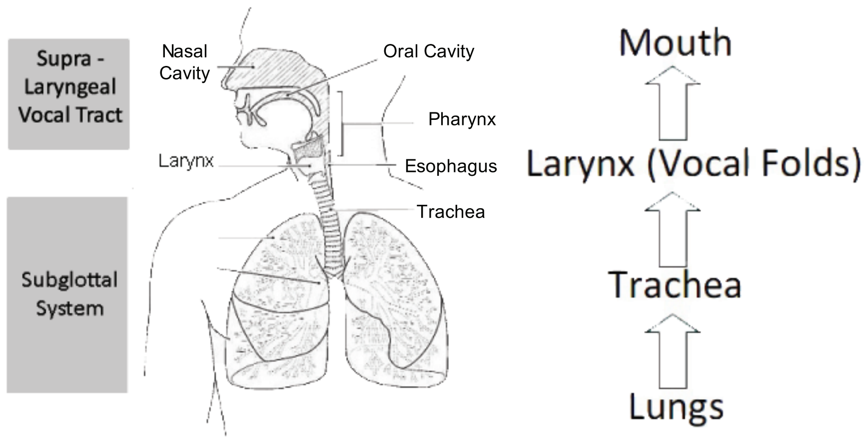


Figure 2–1: Diagram, adapted from [7], of the human anatomy involved in voice production, separated into the subglottal system and the supraglottal system. Air expelled from the lungs flows through the trachea, across the glottis, into the mouth and nasal cavity to produce sound.

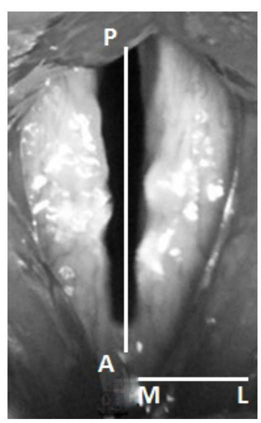


Figure 2–2: An endoscopic image of the human vocal folds. The medial-lateral direction is horizontal and the anterior-posterior direction is vertical. The superior direction is out of the plane of the image, and the inferior direction is into the plane of the image.

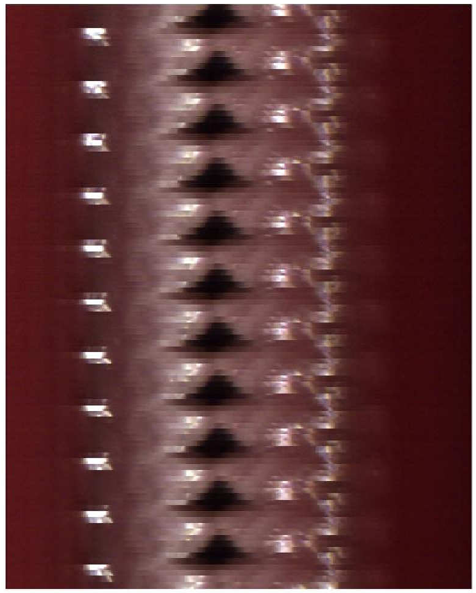


Figure 2–3: Example of a kymogram courtesy of Alfred Chan [40].

CHAPTER 3
Methodology

The experimental methods and materials used in this research are described. The chapter is divided into three main sections. The approach used to estimate vocal fold contact pressure using image analysis is described in the first section. In the second section, the verification studies using the physical model are described. The human subject experiments are described in the third section.

3.1 Basic Principles of Contact Pressure Estimation

Vocal fold impact may be approximately modeled as a collision between two rigid bodies. A diagram of the idealized collision problem is shown in Figure 3–1. The contact force between colliding bodies may be obtained from the change in linear momentum before and after collision. The force between the vocal folds during collision is determined from Newton's second law

$$F = m\frac{\Delta v}{\Delta t} \tag{3.1}$$

where F is the average collision force over the impact duration, m is the effective mass of the vocal fold, $\Delta v = v_2 - v_1$ is the change in velocity before and after collision, and $\Delta t = t_2 - t_1$ is the duration of the collision.

The contact pressure is obtained from the force divided by the average contact area

$$P_{contact} = \frac{F}{A_{avg}} = \frac{m(v_2 - v_1)}{A_{avg}(t_2 - t_1)} \tag{3.2}$$

where $P_{contact}$ is the contact pressure, v_2 is the vocal fold edge velocity after impact, v_1 is the edge velocity before impact, t_1 is the time at the start of the collision, t_2 is the time at the end of the collision, and A_{avg} is the average contact area during collision.

The duration of the vocal fold collision, $t_2 - t_1$, was obtained by analysis of the video recordings of the vocal fold oscillations. It was assumed that the vocal folds were in contact when the glottal area was at its minimum value. The time where the glottal area first reached zero is t_1 and the time point at which the glottal area began to increase is t_2. The procedure for determining the glottal area is described in more detail in Section 3.1.3.

The average contact area, A_{avg}, between the vocal folds is difficult to measure directly. The vocal fold contact region was assumed to be rectangular in shape, defined by its length along the anterior - posterior direction, and its depth along the inferior - superior direction. For verification studies using physical replicas, these quantities were measured directly. The length and the depth of the vocal folds could not be measured in human subjects. For that case, the depth was assumed to be one-fifth of the vocal fold length, L, which yields the approximate relation $A_{avg} = L(\frac{1}{5}L)$. This ratio is approximately the same as for the physical model, and it corresponds to values used in previous numerical studies [41]. Attempts were made to calibrate the endoscopic video to obtain the pixel size, but they were not entirely successful. Image calibration is described in more detail in Section 3.1.1.

Since it was not possible to measure the vocal fold length in the human subjects, the contact pressure was therefore determined from representative values of vocal fold length [42]. Lower bound estimates of the contact pressure were based on a length of 15 mm (L_l). The upper bound was based on a length of 10 mm (L_u).

The effective mass, m, of the vocal folds is also difficult to measure. In multi-mass models of phonation, a value of $m = 0.2$ g is often used and seems to be a reasonable approximation for the effective mass [41, 43, 14, 15]. This value was therefore used in the present study.

There are a number of limitations to this model. First, the vocal folds are modeled as rigid bodies and the deformation during collision is ignored. The collision is modeled as being perfectly elastic but in reality there is likely some energy loss when the vocal folds collide. The contact area is not measured directly. Length and depth of the human subject vocal folds were ad hoc values. The effective mass is also not precisely known, and the same value was assumed for all subjects. This obviously limits the accuracy of the analysis. The limitations and sources of error are discussed in greater detail in Chapter 5.

3.1.1 Image Calibration

Attempts were made to calibrate endoscopic images by focusing the endoscope on a grid pattern of known spacing. The variation in height between the endoscope and the vocal folds prevented accurate image calibration using this method.

Further calibration attempts were made using laser projections onto the vocal folds, as described in [36, 38]. A custom scope was fabricated to house a red laser at one end, and two mirrors in series at the other. The first mirror reflected 50%

of the light and the second mirror was 100% reflective. The mirrors were mounted such that the laser could project two dots onto the vocal folds at an angle of 70°, the same angle as the optical axis of the rigid endoscope used. The custom laser scope could not be used to accurately calibrate the endoscopic image to obtain the true dimensions of the vocal folds. The primary reasons were the following: 1) the laser dots were challenging to position on the vocal folds due to slight misalignment of the laser and mirror; 2) the laser dots had diffuse borders, making the location of the centre of the dot hard to determine; 3) the laser dots were too weak to be visible under the strong halogen lights used for videoendoscopy. Ad hoc values were used for the lengths of the human subject vocal folds because the absolute dimensions could not be obtained from the laser projection system.

3.1.2 Accuracy and Errors

The vocal fold edge velocity and collision duration are determined from the velocity time history obtained from image analysis. The error in the velocity time history, and hence the error in the edge velocity and collision duration, depends on the image size and the time resolution of the high-speed video.

The maximum image size for the camera used in high-speed videoendoscopy was 512 x 512 pixels. The anatomy surrounding the vocal folds was cropped out of the video to decrease the file size. The resulting images generally were between 200 x 200 pixels and 300 x 300 pixels. In the images, the glottis generally occupied approximately one-half of the height of the frame. Given that the length of the vocal folds is approximately 15 mm, this yields a resolution of 6.7 pixels per millimeter. In the image processing algorithm used to determine the vocal fold edge velocity, the position accuracy is to the nearest pixel. The accuracy of the vocal fold edge displacement is therefore estimated to be within ± 0.075 mm. For a measured displacement of 5 mm, this would yield an error of 1.5%. The image size for the camera used in the physical model experiments was much greater than that of the endoscopy camera (see Section 3.2.3). The displacement determined for the physical model vocal folds was more accurate.

The images were acquired at a rate of 2000 frames per second (F/s). The frequency of oscillation of the physical model vocal fold (described in more detail in 3.2) was between 70 Hz and 90 Hz, however, the phonation frequency of the human subjects was 220 Hz. This yields a time resolution of 0.5 ms, or approximately 9 frames per cycle. The average collision duration obtained from the velocity time history was generally between 0.5 ms and 1.5 ms. The error in this measurement was potentially between 17% and 50%. The error in the time measurement was considerably greater than the error in the displacement measurement. A camera with a frame rate greater than 20 000 frames per second is required to ensure temporal errors less than 5%. Using the high-speed videoendoscopy system presently available, the frame rate could be increased to 4000 frames per second and the resolution decreased to 512 x 256 pixels. This would reduce the temporal error by a factor of two and increase the spatial error by a factor of two. Future experiments using this system should be performed with increased frame rate.

3.1.3 Glottal Area Segmentation and Edge Velocity Measurement

The glottal area waveform was determined from the high-speed video images using a procedure similar to that described in previous studies [28] [23]. A digital video of the vocal fold oscillation was saved as a sequence of images. The images were then processed using MATLAB.

A region-growing algorithm was used to segment the glottal area [26]. This type of algorithm was selected because it is simple and relatively efficient. A threshold criterion was used to identify the pixels with a low intensity. These were assumed to be part of the glottis. A connectivity criterion was used to ensure that all pixels with a low intensity were included in one unique region.

One seed point was selected interactively by the user. The pixels forming the glottal area were then identified by comparing the intensity of the neighbouring pixels to that of the seed point. Pixels with an intensity less than a pre-defined threshold value were considered part of the glottal area. Other pixels were considered background. Only pixels connected to the seed point were considered part of the glottal area, even if their intensity was below the threshold value.

The segmented glottal area from a human subject (subject 11) is shown in Figure 3–2. A sample of the glottal area region was selected interactively by the user to define the first seed point. The glottal area and its perimeter were identified using the region growing method and custom functions from Gonzalez [26]. Occasionally the image segmentation algorithm did not properly identify the glottal area. An additional function was added to the algorithm to check if the identified glottal area was a realistic size. If the glottal area was greater than or less than the number of pixels that would be considered reasonable, the image was re-segmented using a new seed point. Aside from this occasional user intervention, the images were segmented automatically.

The edge velocity was determined from the perimeter of the glottal area. First, the anterior and posterior commissures of the glottis, the points where the right and left vocal folds appear to join, were automatically identified from the extremities of the glottal area perimeter, or the two points along the perimeter that are farthest from each other. Then the major axis of the glottal area was automatically determined by connecting the anterior and posterior commissures with a straight line. The minor axis was located at the midpoint and perpendicular to the major axis. The points where the minor axis intersected the perimeter of the glottal area were found. The distance between the intersection point and the major axis was determined for each image. The change in this distance over the time period between images was used to determine the vocal fold edge velocity. The edge velocity measurement from the images was completely automatic.

3.2 Physical Model Validation

A physical replica of the human vocal folds was used to verify the procedures to estimate contact pressure. The physical model and the validation experiments are described below.

3.2.1 Physical Model Hemilarynx Set-Up

The physical model vocal fold was fabricated following the procedures outlined in [45] and [6]. The physical model was created from a three-part silicone rubber solution. EcoFlex 10 (Smooth-On) was mixed using a 1:1 ratio of part A and part B. Approximately twice as much silicone thinner (Smooth-On) was added to the solution to obtain the desired onset pressure. The onset pressure is the subglottal pressure at which the physical model begins to oscillate and is shown in Figure 3–6.

The silicone rubber solution was de-gassed in a vacuum chamber for approximately five minutes before it was poured into a mold. It was then cured overnight at room temperature.

The shape of the physical model vocal fold contour was based on the "M5" model from [46], with an included glottal angle of 40°. An image of the physical model vocal fold, including its dimensions, is shown in Figure 3–3. A photograph of the physical model vocal fold is shown in Figure 3–4.

A hemilarynx model was used. The model consisted of one vocal fold vibrating against a rigid surface. Hemilarynx models and their use in excised larynx experiments have been described in [18]. A hemilarynx set-up allows the contact pressure to be measured on the medial surface of the physical model vocal fold with no interference with the vocal fold oscillation. This is achieved by recessing the pressure measurement probe into the rigid surface, facing the medial edge of the physical model vocal fold, thereby eliminating any possible probe interference.

The silicone model vocal fold was glued to an acrylic, C-shaped support using a silicone adhesive (Smooth-On), as shown in Figure 3–4 . The support was, in-turn, bolted to a custom made plastic surface. The surface facing the model vocal fold was perforated with fifteen holes, or ports, with a 1.6 mm internal diameter and spaced 2.5 mm apart from centre to centre. A miniature probe microphone was inserted into each port to measure the local pressure while the remaining ports were blocked. This configuration allowed the location of maximum pressure to be determined. Figure 3–5 shows the physical model and its supporting frame, including the contact surface.

3.2.2 Airflow Supply

Shop air was used to produce airflow. The flow rate was measured using a mass-flow meter (Type 558A, MKS). The static subglottal pressure was measured using a Baratron pressure transducer (Type 220D, MKS). Pressure taps were located at a distance of 3.5 cm upstream of the vocal folds. Both the static subglottal pressure and the flow rate were recorded using a digital readout display and a power supply (PR 4000F, MKS). One example of a typical pressure-flow curve is shown in Figure 3–6.

3.2.3 High-Speed Video

High-speed video images of the physical model were recorded using a Dantec NanoSense Mk III camera with a frame rate (F/s) of 2000 frames per second and an image size of 1280x1024 pixels. Two external light sources were used. The video recordings were saved as a series of image files before they were processed using the methods outlined in Section 3.1.3.

For the physical model, the edge velocity was determined at four locations along the anterior-posterior direction in addition to the intersection point of the minor axis. These locations were aligned with those of the pressure sensor ports. All edge velocity values were determined along the direction perpendicular to the major axis. The contact pressure was estimated from the edge velocity at each point, based on the assumptions outlined in Section 3.1. The estimated contact pressure values were then compared to directly measured pressures at the same locations.

3.2.4 Probe Microphone

For direct measurement, a probe microphone was fabricated as described in [6] from a 6.4 mm Brüel & Kjær condenser microphone, an acrylic coupler, and a 1.6

mm O.D. capillary tube. A schematic of the probe microphone assembly is shown in Figure 3–7. The acrylic coupler was made of two separate parts. The first part was bored to host the microphone and its preamplifier. An external thread allowed attachment to the second coupler. The capillary tube was inserted in the second part of the coupler and glued securely in place. The microphone holder was screwed to the capillary tube holder. The microphone cavity was sealed using a rubber O-ring.

The probe microphone was calibrated using a Brüel & Kjær sound intensity calibrator (Type 3541) following the procedure described in [6]. Two 6.4 mm diameter Bruel and Kjaer condenser microphones were insonified with white noise with a bandwidth of 10 kHz. The transfer function between the two microphones was determined. Then, the probe microphone was substituted for the first microphone, and the transfer function between the second microphone and the probe microphone was determined for the same input. The ratio of the two transfer functions then yielded the frequency response of the probe microphone. The Fourier transform of the probe microphone signal was thereafter compensated using the probe frequency response functions. The contact pressure was then calculated from the inverse Fourier transform of the result. More information on the validation and calibration of the probe microphone is included in reference [6].

3.3 Human Subject Experiments

The contact pressure estimation method was applied to human subjects. The protocol for these experiments was approved by the Research Ethics and Compliance Office both at McGill University and at the University of Pittsburgh, where the experiments were performed. A more detailed description of the human subject experiment protocol is available in reference [6]. Data will be reported on three subjects, referred to as subject 9, subject 10, and subject 11. Both high-speed video data and measured contact pressure data was available for these three subjects. All subjects were women between the ages of 18 and 40 with no history of voice disorders. They were asked to perform a series of phonation tasks while different measurements were made. Table 3–1 summarizes the tasks performed by the subjects. The subjects were required to phonate at a certain frequency using different voice types and different sound pressure levels. Each trial was only included if the voice type produced by the subject was judged by the experimenters to be accurate and if the sound pressure level produced by the subject was within 2 dB of the target.

A baseline endoscopic high-speed video was first recorded as the subjects performed the tasks outlined above. The subjects then repeated the same tasks while the contact pressure was measured with the probe microphone and high-speed endoscopic video was recorded simultaneously.

3.3.1 High-Speed Videoendoscopy

Endoscopic video was recorded using a Color High-Speed Video system from Kay-Pentax (model 9710). The system included a high-speed camera (Photron Fast-Cam MC2), a video processor, one light source, and a rigid endoscope. The video was recorded with a frame rate (Fs) of 2000 frames per second, with a maximum resolution of 512 x 512 pixels.

3.3.2 Contact Pressure Measurements

The contact pressure was measured in human vocal folds using the probe microphone described in 3.2.4 with some minor modifications. For the physical model experiments, the probe was oriented along the medial-lateral direction. For human subject measurements, the probe was oriented along the inferior-superior direction. A small hole was made in the side wall of the capillary tube, and the end of the capillary tube was then sealed.

Local anesthetic was required for the subjects to tolerate the presence of the probe. An otolaryngologist administered a solution of 4% atomized lidocaine and 0.25% atomized pontocaine to numb the vocal folds and larynx. The probe was inserted between the vocal folds by the otolaryngologists, using an Abraham's cannula, while the subject was breathing with an open glottis. The rigid endoscope was positioned trans-orally to simultaneously record the high-speed video. The vocal folds came into contact around the probe when the subject began phonating.

The high-speed video and probe signal were recorded simultaneously. Synchronization was performed using a trigger signal from the camera. The trigger signal was fed to the video recording system and the data acquisition system. This allowed synchronous data acquired independently to be examined.

3.4 Validation Criteria

Contact pressure estimations from image analysis were compared to directly measured values to validate the image analysis method. Criteria were established to determine what difference between the two contact pressure values would be considered acceptable. The contact pressure used for comparison was the average of five or more peak values. An acceptable margin of error may be determined from the standard deviation of the average value.

Given the significant variability in contact pressure values measured in human subjects in previous studies (see Section 2.3 and Section 5.1) verifications were performed. Contact pressure values that agree within one standard deviation are considered to be in “good” agreement. Contact pressure values within a factor of two are considered to be in “acceptable” agreement.

Table 3–1: Tasks for human subject experiments.

Frequency (Hz)	Voice Type	Sound Pressure Level (dB)
220	Breathy	60 70
220	Normal	60 70 80
220	Pressed	70 80

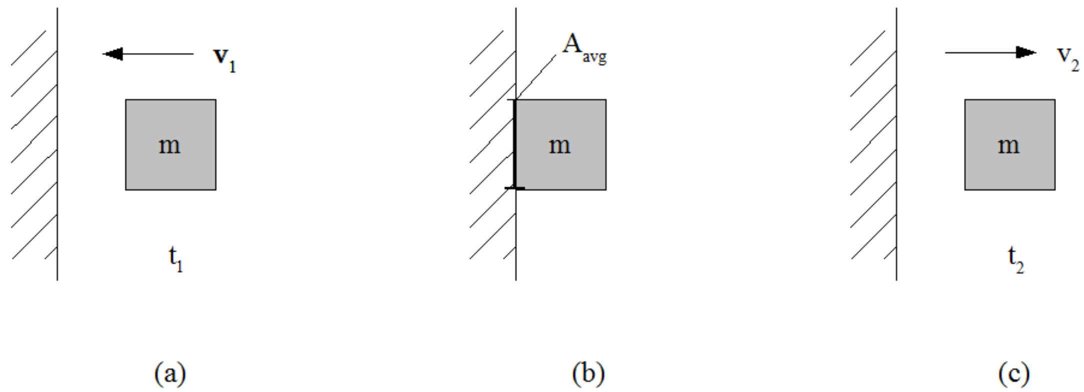


Figure 3–1: Diagram of the idealized vocal fold collision. A single vocal fold is shown colliding against a rigid surface, representative of a hemilarynx model. a) The vocal fold of mass m before impact at time t_1, traveling at velocity v_1, b) the collision with average contact area A_{avg}, c) the vocal fold after impact at time t_2, traveling at velocity v_2.

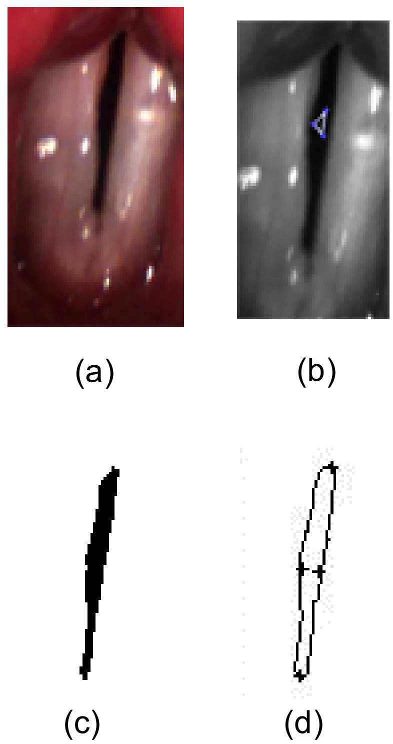


Figure 3–2: Glottal area segmentation for subject 11, normal voice type, 70 dB. a) The original frame from the high-speed video sequence, b) a sample of the glottal area is selected by the user to define the seed point, c) the glottal area is segmented, d) the vocal fold edges and major and minor axes are identified.

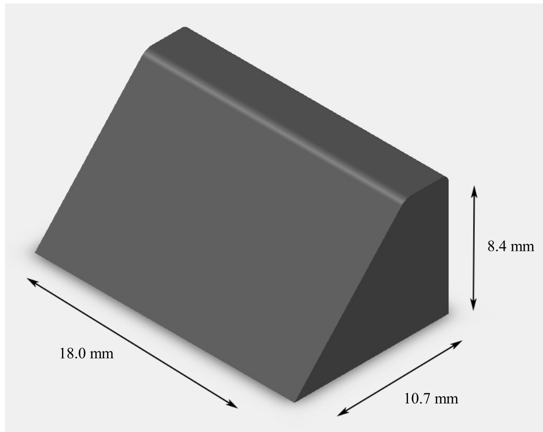


Figure 3–3: Dimensions of the physical model vocal fold.

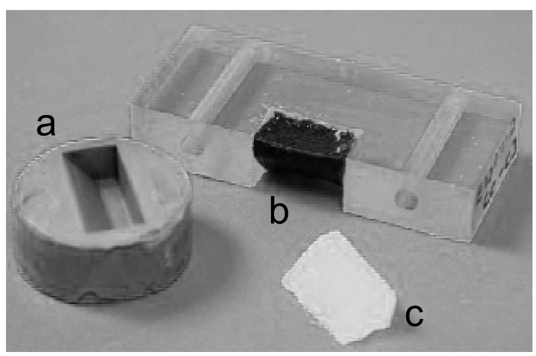


Figure 3–4: Pictures of a) the mold used to create the model vocal fold, b) the silicone model vocal fold fixed to the acrylic support, c) the silicone model vocal fold. Dye can be added to the silicone solution to create models in different colours.

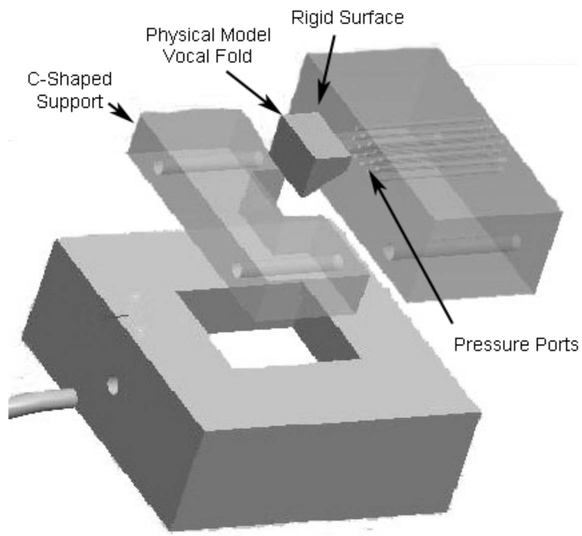


Figure 3–5: The hemilarynx physical model set-up with the C-shaped holder and plastic surface perforated with 15 holes, the pressure ports.

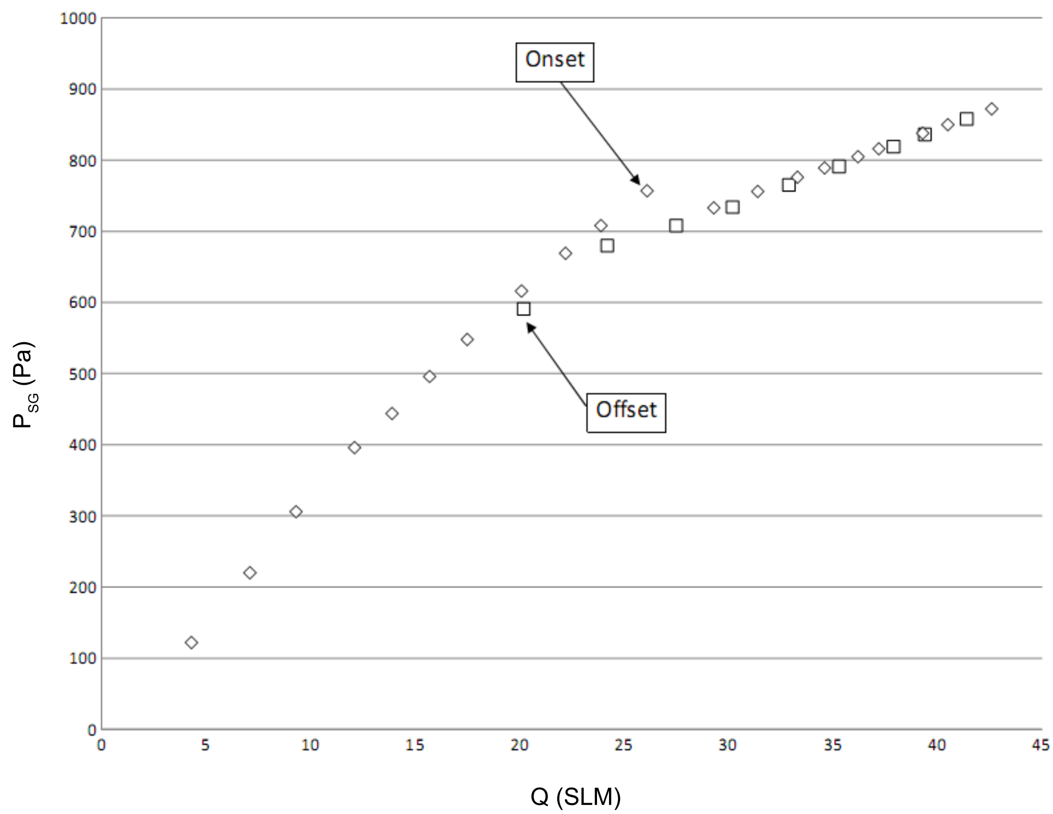


Figure 3–6: Typical relation between the static subglottal pressure (P_{SG}) and the air flow rate (Q) in physical model experiments. $\diamond$: increasing flow, $\square$: decreasing flow; the label "Onset" indicates the flow rate (27 SLM) and subglottal pressure (760 Pa) at which self-oscillation was initiated while the airflow rate was increased. The label "Offset" indicates the flow rate (20 SLM) and subglottal pressure (600 Pa) at which self-oscillation stopped.

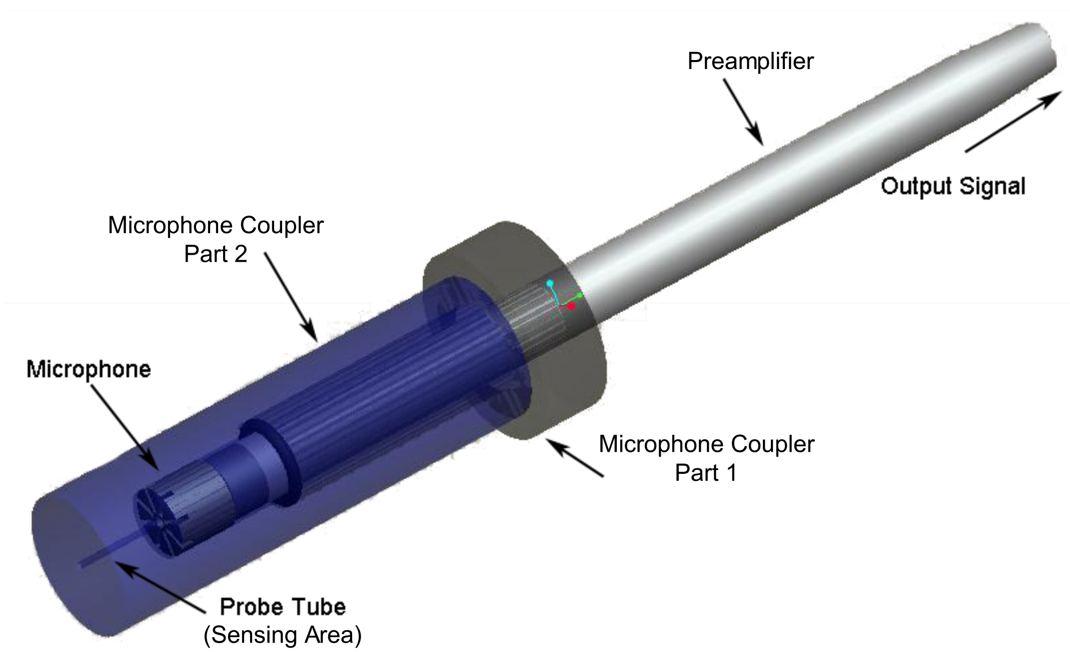


Figure 3–7: Schematic of the probe microphone assembly, including the condenser microphone and the plastic coupler. The capillary tube was glued into a small hole at the end of the plastic coupler.

CHAPTER 4
Results

4.1 Verification Using a Physical Model

A physical model was used to verify the contact pressure estimated from high speed video. Results from direct measurements using a probe microphone were used as a reference.

4.1.1 Probe Microphone Measurements

Figure 4–1 shows one example of the pressure waveform for the physical model measured using the probe microphone. The physical model oscillated at approximately 70 to 90 cycles per second. The peak pressure values were identified as the local maximum for each oscillation cycle. The average value of at least five peak pressure values was calculated and is reported.

Figure 3–5 shows a schematic of the physical model test set-up. The contact pressure was measured at fifteen different locations on the medial contact plane. The pressure ports were located along three columns along the inferior-superior direction, and five rows along the anterior-posterior direction. The locations of the pressure ports are shown in Figure 4–2.

The greatest contact pressure values were measured along the inferior edge of the contact plane (ports 3, 6, 9, 12, and 15). Table 4–1 shows the contact pressure distribution. The measured contact pressure values were different at each pressure port. The variations in the measured values were greater along the inferior-superior direction (i.e. between ports 7, 8, and 9) than in the anterior-posterior direction (i.e. between ports 6, 9, and 12). The greatest contact pressure values were measured at ports 6, 9, and 12.

4.1.2 High-Speed Video

Data from the high-speed video recordings are shown in Figures 4–3 and 4–4. Figure 4–3 shows the time history of the edge velocity at the centre of the medial edge of the vocal fold. The velocity is defined to be positive along the lateral direction, i.e. the direction of glottal opening. The velocity is negative along the medial direction, when the orifice is closing. The time history of the glottal area is shown on the same graph. The maximum vocal fold edge velocity occurs during the closing phase, approximately 3 ms before the instant of vocal fold collision.

The time history of the edge velocity of the physical model at different locations along the anterior-posterior direction is shown in Figure 4–4. As described in Section 3.2.3, the edge velocity was determined at points aligned with the pressure ports where the contact pressure was measured. Port 9 is at the centre of the vocal fold. Port 6 and port 12 are immediately posterior and anterior, respectively, to port 9. Port 3 is at the posterior end and port 15 is at the anterior end. As expected, the edge velocity is greater at locations near the centre region (port 9, port 12, and port 6), and smaller at locations close to the anterior and posterior ends (port 15 and port 3, respectively). The vocal folds have a greater displacement near the centre.

The contact pressure was estimated using the edge velocity and the assumptions outlined in Section 3.1. Table 4–2 shows the contact pressure values estimated from the image analysis and those measured directly with the probe microphone. The direct measurements were made at port 9 (at the mid-point) and the estimation used the edge velocity for the point aligned with port 9. Measurements were made for two different subglottal pressures, 730 Pa and 834 Pa. The directly measured values are in good agreement with those estimated from high-speed images.

Figure 4–5 shows directly measured contact pressure values and those estimated from the high-speed video images. Values measured at port 6, port 9, and port 12 are shown. The contact pressure from high-speed video images is in acceptable agreement with direct measurements from port 12 but not port 6. The best agreement was obtained at port 9. That location was used for further comparison between the two methods.

The validation experiments were carried out for a range of operating conditions (subglottal pressure and flow rate). Results for higher subglottal pressures are shown in Figure 4–6. The contact pressure values were taken from port 9. The agreement between the two methods is acceptable for the three subglottal pressures used.

4.2 Probe Microphone Installation Effects

The possible effects of the presence of the probe microphone on the vocal fold oscillation was investigated using a procedure similar to that described in [6]. A schematic of the experimental set-up is shown in Figure 4–7. The probe tip was mounted such that it protruded by 0.8 mm from the surface, instead of being mounted flush with the side wall. The offset value corresponds to one-half the outer diameter of the probe tube. In addition, a probe tube section was cut in half lengthwise and glued along the side wall of the test apparatus. The half-tube was extended in the superior direction to mimic the presence of the probe between actual vocal folds.

The contact pressure was first measured with the probe microphone in the "protruding" position. The protruding probe tube and the half-tube were removed while the same flow rate was maintained. The contact pressure was then measured with the probe microphone in the "flush" position.

The contact pressures measured with the probe microphone in the "flush" position and in the "protruding" position are shown in Figure 4–8 and Figure 4–9. The protruding probe microphone consistently underestimated the contact pressure by 16% to 36%. This indicated that the presence of the probe microphone affects the oscillation of the vocal folds. Similar errors induced in actual human subjects may be hard to quantify.

4.3 Human Subject Experiments

The contact pressure data from human subjects is shown in Figure 4–10 for subject 9, Figure 4–11, Figure 4–13, and Figure 4–14 for subject 10, and Figure 4–12 for subject 11. Contact pressure values are only available for a normal voice type for subjects 9 and 11. The limited contact pressure data is due to time constraints and to the subjects' inability to tolerate the probe. Upper and lower bound estimates for contact pressure from video data were obtained based on the assumed length of the subject's vocal folds. The greatest value is based on a vocal fold length L_u=10 mm. The lower value is based on a vocal fold length L_l=15 mm.

Contact pressure values from high-speed video images are in very poor agreement with directly measured data from the probe microphone. The high-speed image values are consistently greater than the directly measured values. The magnitudes of the discrepancies are not consistent. The differences range from a factor of over

500 for subject 11, normal voice type, at 60 dB, to a factor of 3, in the case of subject 11, normal voice type, at 70 dB. The directly measured contact pressure values or those from image analysis do not appear to be related to voice type or to sound pressure level. Furthermore, values for the same subject vary significantly.

Sources of error include the approximations and assumptions used for the vocal fold effective mass, the vocal fold length, and the contact area. The probe interference likely caused error in the human subject measurements but the extent of the probe's effect is unknown. These sources of error are discussed further in Section 5.2.

Table 4–1: Contact pressure ($P_{contact}$) measured with the probe microphone vs. location for two different subglottal pressures (P_{SG}), 730 Pa and 830 Pa. The greatest contact pressure was measured at the centre of the medial edge of the model vocal fold at ports 6, 9, and 12.

Port	$P_{contact}$ (Pa) for P_{SG}=730 Pa	$P_{contact}$ (Pa) for P_{SG}=830 Pa
1	72	115
2	338	356
3	160	544
4	84	135
5	94	204
6	433	894
7	310	83
8	260	311
9	587	1146
10	13	23
11	393	211
12	857	960
13	270	153
14	110	278
15	228	518

Table 4–2: Contact pressure ($P_{contact}$) values from high-speed video data and direct measurements in the physical model vocal fold. The measurements were made at port 9, in the centre of the medial edge of the vocal fold, for two different subglottal pressures (P_{SG}), 730 Pa and 830 Pa.

P_{SG} (Pa)	$P_{contact}$ (Pa) (images)	$P_{contact}$ (Pa) (direct measure)	% difference
730	593	588	1
830	1146	1122	2

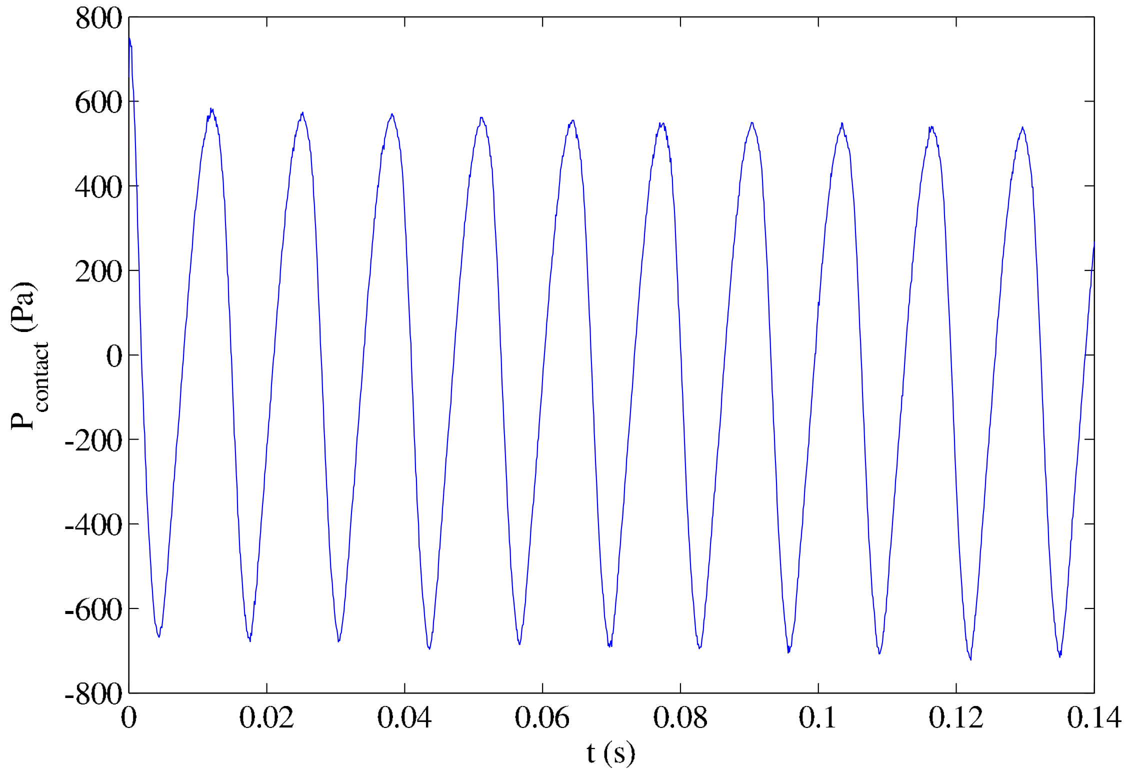


Figure 4–1: Contact pressure ($P_{contact}$) vs time (t) for the physical model vocal fold measured with a probe microphone. The measurement was made at port 9, in the centre of the medial edge of the vocal fold. The subglottal pressure was 731 Pa and the vocal fold was oscillating at approximately 80 Hz.

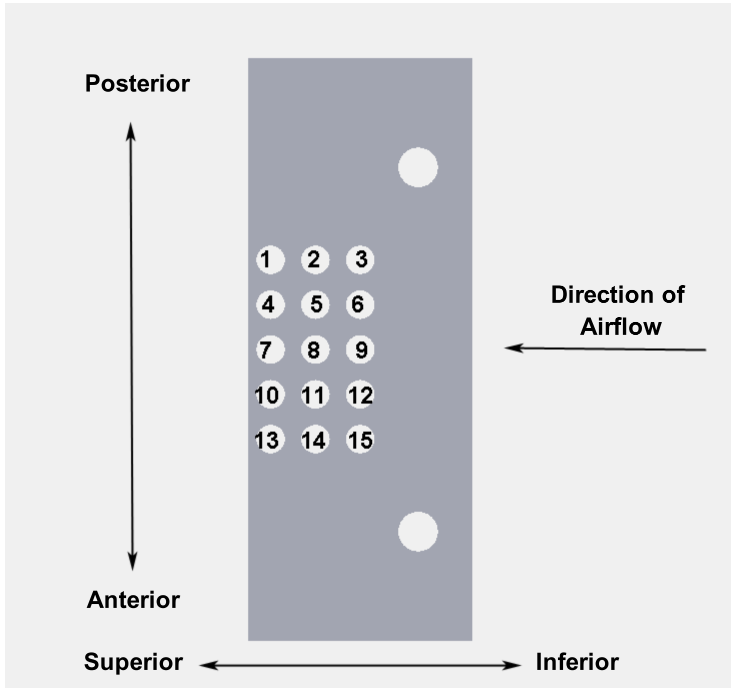


Figure 4–2: The contact pressure on the physical model vocal fold was measured at 15 different ports along the surface facing the medial edge of the model vocal fold. The ports are numbered 1 through 15.

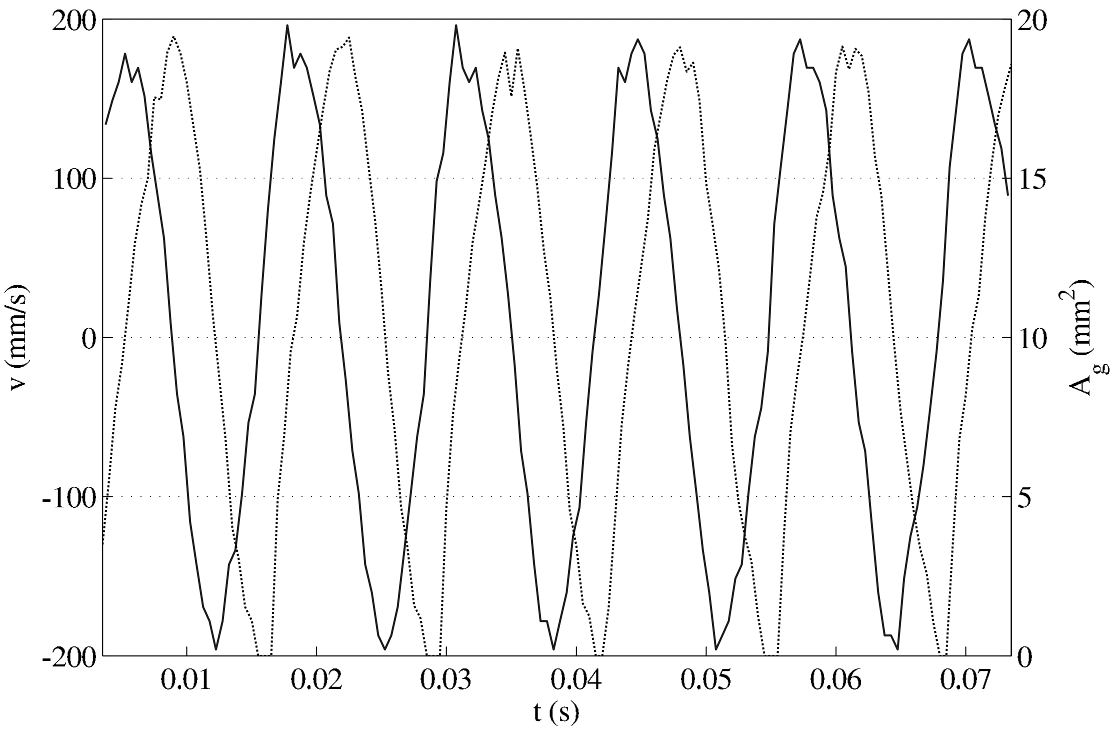


Figure 4–3: Edge velocity and glottal area as a function of times. −: edge velocity (v); ···: glottal Area (A_g); the glottal area and edge velocity at the centre of the physical model vocal fold. Positive edge velocity is in the lateral direction, or during the opening phase of the vocal folds.

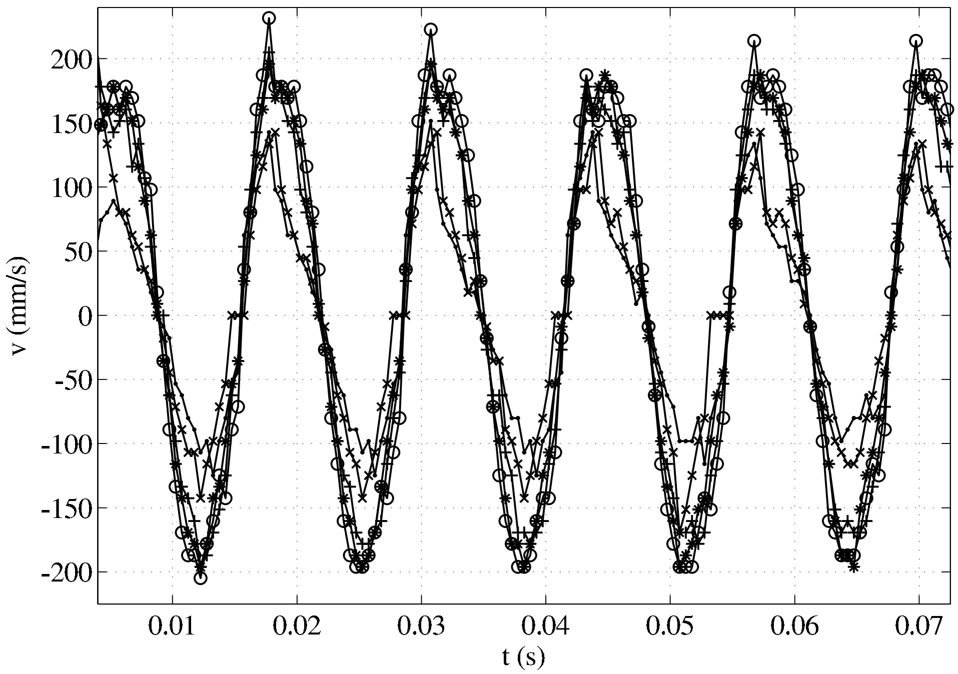


Figure 4–4: The edge velocity (v) vs. time of the physical model vocal fold at points aligned with each of the pressure ports. •: port 3; +: port 6; ∘: port 9; ∗: port 12; ×: port 15. Port 9 is in the centre of the medial edge of the physical model vocal fold, port 3 and port 5 are at the most posterior and anterior ports, respectively.

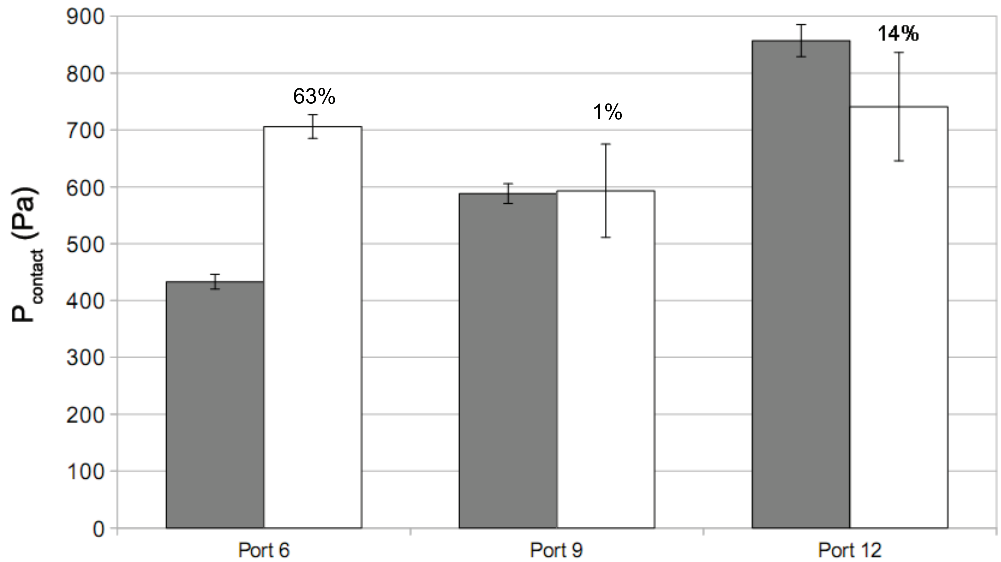


Figure 4–5: The contact pressure ($P_{contact}$) approximated from high-speed video data (white) with contact pressure measured with the probe microphone (gray) in the physical model vocal fold. The measurements were made at port 9, port 6, and port 12 at the same subglottal pressure (P_{SG} =730 Pa). The percent difference between the directly measured contact pressure values and the contact pressure values from image analysis is shown in the graph.

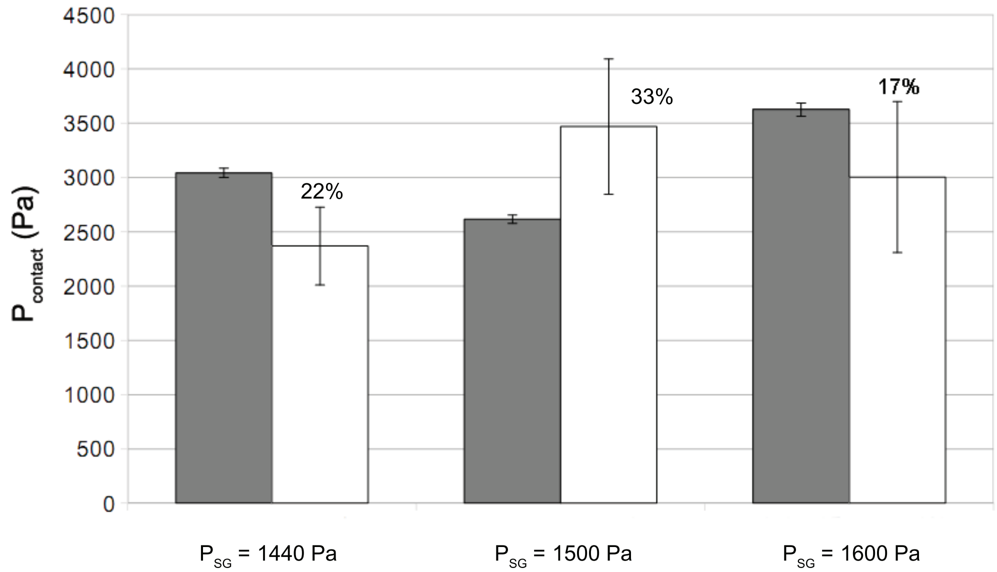


Figure 4–6: The contact pressure ($P_{contact}$) approximated from high-speed video data (white) with contact pressure measured with the probe microphone (gray) in the physical model vocal fold. The measurements were made at port 9 for three different subglottal pressures (P_{SG}), 1440 Pa, 1500 Pa, and 1600 Pa. The percent difference between the directly measured contact pressure values and the contact pressure values from image analysis is shown in the graph.

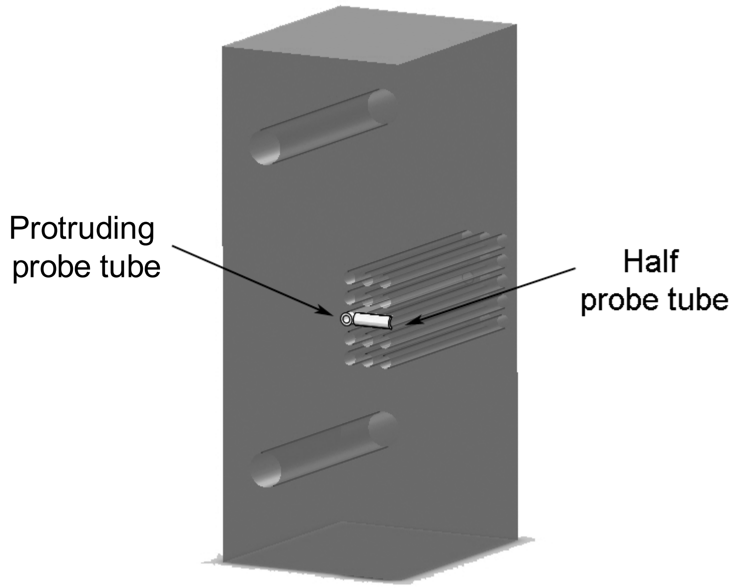


Figure 4–7: Modified contact area to mimic the influence of probe interference. The probe protruded from the surface by 0.8mm. A small tube cut in half was added to simulate the presence of the probe microphone.

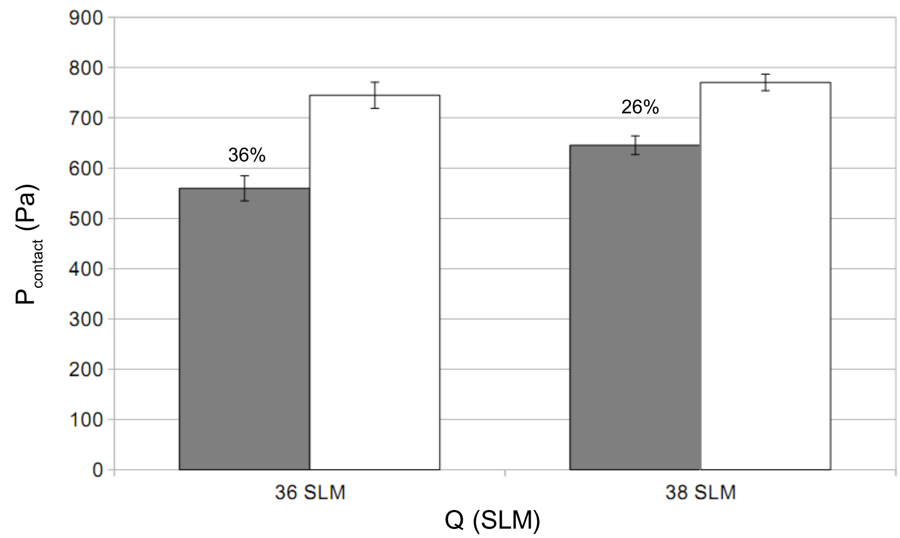


Figure 4–8: The measured contact pressure in the physical model for the case where the probe microphone is protruding (gray) and for the case where the probe microphone is recessed (white). The measurements were made for flow rates (Q) of 34 SLM and 37 SLM. The percent difference between the contact pressure measured with the probe protruding and the contact pressure measured with the probe recessed is indicated in the graph.

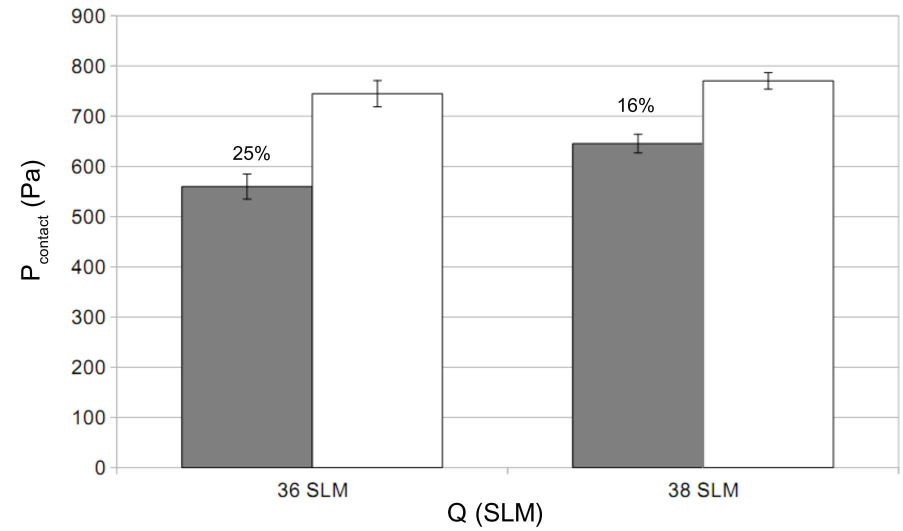


Figure 4–9: The measured contact pressure in the physical model for the case where the probe microphone is protruding (gray) and for the case where the probe microphone is recessed (white). The measurements were made for flow rates (Q) of 36 SLM and 38 SLM. The percent difference between the contact pressure measured with the probe protruding and the contact pressure measured with the probe recessed is indicated in the graph.

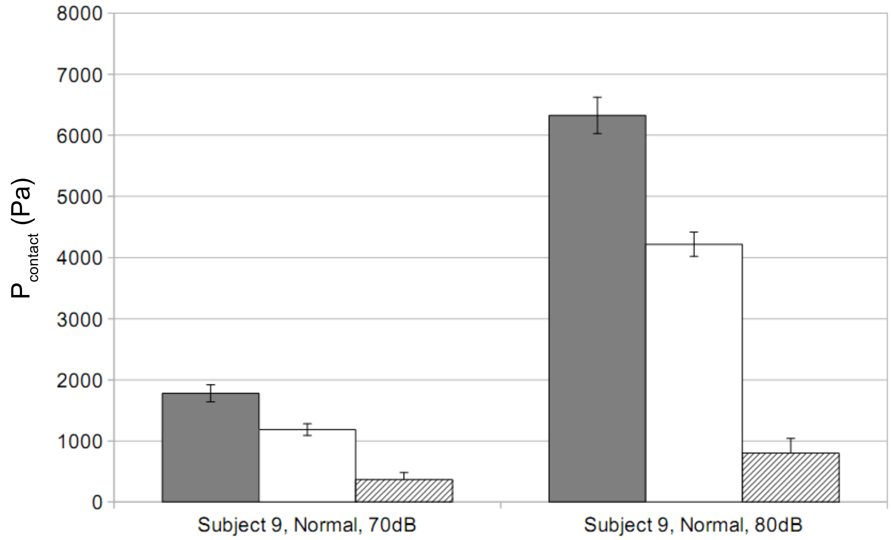


Figure 4–10: The contact pressure approximated from high-speed video data for vocal fold lengths of 15 mm (white) and 10 mm (gray) with contact pressure measured with the probe microphone (hatched) in Subject 9, for a normal voice, at two different sound pressure levels.

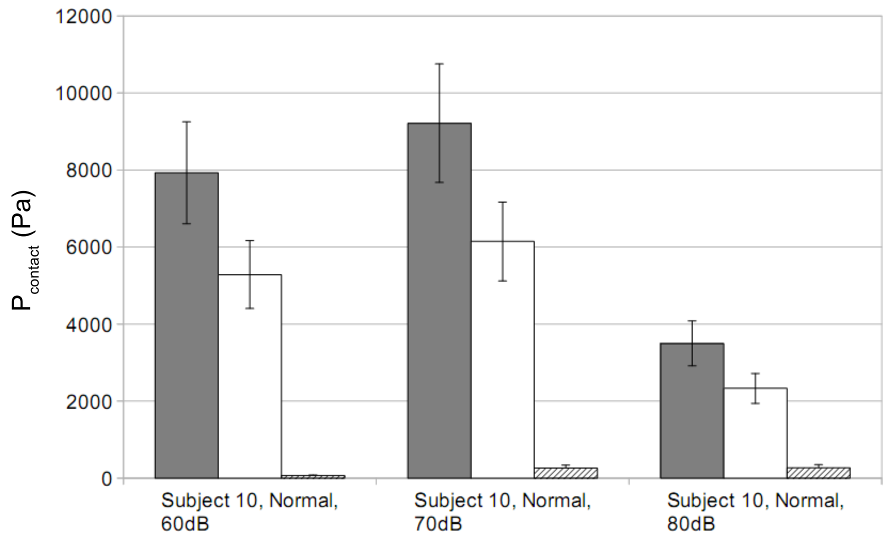


Figure 4–11: The contact pressure approximated from high-speed video data for vocal fold lengths of 15 mm (white) and 10 mm (gray) with contact pressure measured with the probe microphone (hatched) in Subject 10, for a normal voice, at three different sound pressure levels.

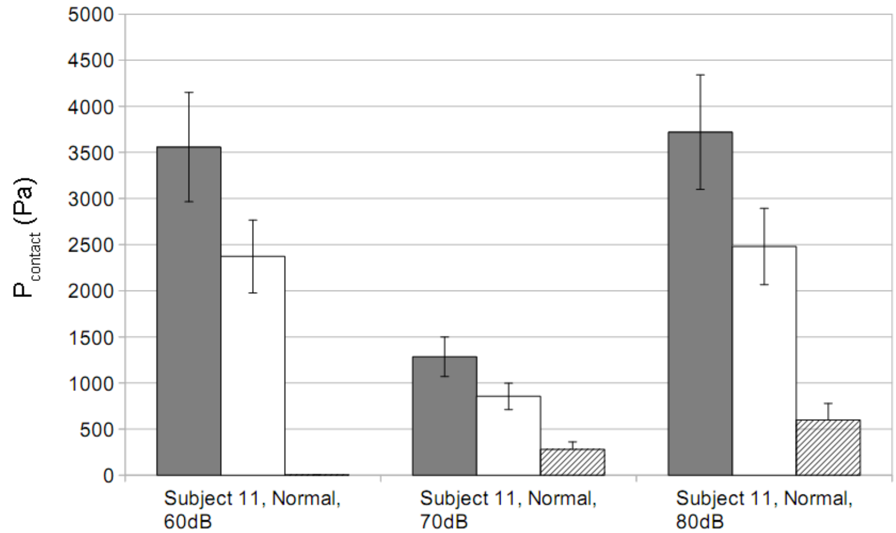


Figure 4–12: The contact pressure approximated from high-speed video data for vocal fold lengths of 15 mm (white) and 10 mm (gray) with contact pressure measured with the probe microphone (hatched) in Subject 11, for a normal voice, at three different sound pressure levels.

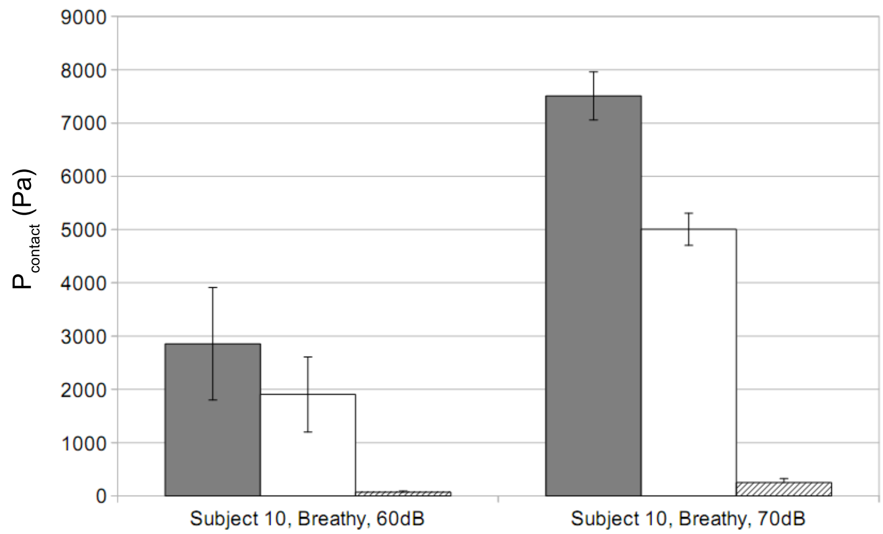


Figure 4–13: The contact pressure approximated from high-speed video data for vocal fold lengths of 15 mm (white) and 10 mm (gray) with contact pressure measured with the probe microphone (hatched) in Subject 10, for a breathy voice, at two different sound pressure levels.

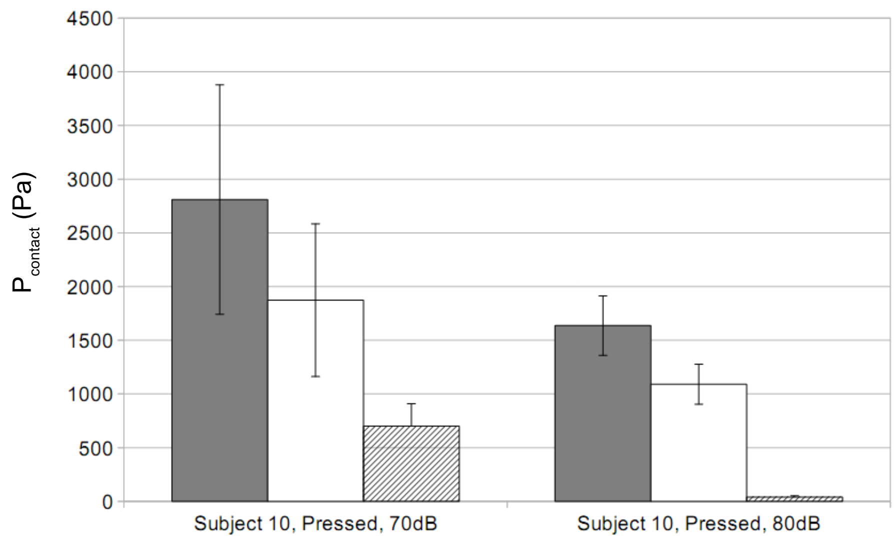


Figure 4–14: The contact pressure approximated from high-speed video data for vocal fold lengths of 15 mm (white) and 10 mm (gray) with contact pressure measured with the probe microphone (hatched) in Subject 10, for a pressed voice, at two different sound pressure levels.

CHAPTER 5
Discussion

5.1 Comparisons with Previous Studies

Contact pressure values measured with the probe microphone and those estimated from image analysis may be compared to values reported from previous studies. Previous human subject experiments were described in Section 2.3. The range of values reported by Hess et al. was from less than 1000 Pa to 4000 Pa [3]. Verdolini et al. reported values ranging from 400 Pa to 3200 Pa [4]. Jiang and Titze's canine hemilarynx experiment produced contact pressure values between 500 Pa and 5000 Pa [18]. Values reported in another study by Verdolini et al. were between 300 Pa and 5300 Pa [19]. Jiang et al. reported values between 303 Pa and 2620 Pa from another excised canine larynx experiment [20].

In the present study, the range of contact pressure values estimated from high-speed video data was between 600 Pa and 9200 Pa overall, and between 600 Pa and 6000 Pa for the lower bound estimate assuming the vocal fold length $L_l = 15$ mm. The range of contact pressure values measured with the probe microphone was between 5 Pa and 800 Pa. The values from image analysis were closer to what was expected based on results from previous studies.

5.2 Limitations and Sources of Error

Various error sources may have contributed to the discrepancy between the directly measured contact pressure values and those estimated from high-speed video data. The frame rate of the images for the human subject experiments (2000 frames per second) was less than ideal. While the same frame rate was used for the physical model experiments, the effect of the low sampling rate was lessened because the oscillation frequency of the physical model was about half that of the human subjects.

5.2.1 Approximations and Assumptions

The average contact area was assumed to be a rectangular shape, determined by the length and the thickness (or depth) of the vocal folds. It is very challenging to precisely measure the length of the vocal folds in human subjects. It is even more challenging to measure the thickness because endoscopy only shows the superior surface. The estimates of the vocal fold thickness ($\frac{1}{5}$ of the vocal fold length) in human subjects may not have been appropriate for all subjects, and has likely contributed to the error in the contact pressure estimation. The thickness of the physical model was measured directly, so this source of error would have contributed primarily to the human subject data.

The shape of the contact area may have been elliptical rather than rectangular [47]. The vocal folds open and close with a phase difference along the anterior-posterior direction and along the superior-inferior direction. This phase difference is described in more detail in Section 5.2.2. The simplification of the shape of the average contact area likely contributed to the error in the contact pressure approximation.

The value for the effective mass of the vocal folds was not determined experimentally for each subject. A single value for the effective mass of the vocal folds for all of the subjects was taken from the literature. The error in the contact pressure estimation would have likely been reduced if the effective mass of the vocal folds was determined experimentally.

The impulse-momentum approximation ignores the material properties and true geometry of the vocal folds. The collision of the vocal folds is assumed to be perfectly elastic. A coefficient of restitution of less than 1 could be used to account for the energy loss of the colliding vocal folds. It would be beneficial for future work to include experiments to determine the coefficient of restitution. Cooper and Titze performed experiments on excised bovine vocal folds to measure the change in temperature associated with viscous energy dissipation during vibration [44]. They measured an increase in temperature between 0.1°C and 0.8°C but cited a number of factors including blood flow and evaporation that weren't accounted for in their experiment. Error in the contact pressure approximation is likely due to the simplified analysis of the vocal fold vibration. This error may be reduced with a more comprehensive approach, requiring additional parameters to be determined experimentally, such as the coefficient of restitution.

Suggestions for determining the approximated parameters experimentally are outlined in Chapter 6.

5.2.2 Limitations of the Physical Model

The physical model used in the verification experiments does not perfectly replicate a human vocal fold. The geometry of the physical model is simplified. The model is uniformly made of isotropic silicone, in contrast with the layered structure of lamina propria in human vocal folds. The physical model does not include muscles that change its tension or length.

The vibration pattern of the physical model also differs from that of human vocal folds. Video frames reconstructing one cycle of a human subject's vocal fold vibration and one cycle of the physical model vibration are shown in Figure 5–1. The physical model does not exhibit the phase difference or "zipper-like" opening and closing motion of human vocal folds. The human vocal folds start opening at the posterior commissure and continue to open toward the anterior commissure. The physical model vocal fold opens symmetrically in the anterior-posterior direction. Furthermore, the physical model vibrates with a large displacement along the superior, or vertical, direction that is not observed in human vocal folds. The shortcomings of the physical model vocal fold, however, can not explain the the inconsistency in the human subject data, or more specifically, why the contact pressure measured with the probe microphone was lower than expected.

5.2.3 Limitations of the Probe Microphone

Section 4.2 showed the results from an experiment investigating the effect of the probe microphone on vocal fold oscillation. The influence of the probe microphone caused the measured contact pressure to be underestimated by 16% to 36%. The extent of the effect of the probe microphone in human subjects is not known.

The physical model verification tests showed that the area of maximum contact pressure is located near the centre of the medial edge of the vocal fold (see Section 4.1.1). Considering that the pressure ports were spaced 2.5 mm apart, the location of peak contact pressure has a relatively small area. The difference between measured contact pressure values was greater along the inferior-superior direction than along the anterior-posterior direction. In human subjects, it is most challenging to determine the position of the probe microphone along the inferior-superior direction. It is likely that the probe microphone was positioned a few millimeters superior or inferior to the area of maximum contact pressure, thus measuring the supraglottal or subglottal air pressure rather than the contact pressure.

It is more practical to verify the probe microphone position in the anterior-posterior direction in images taken from videoendoscopy. Figure 5–2(a) shows the probe microphone positioned near the posterior commissure, far from the region where the contact pressure is the greatest. Figure 5–2(b) shows the probe microphone located closer to the centre of the vocal folds, near the region where the contact pressure is hypothesized to be greatest.

In some cases, the probe microphone position was obscured by the epiglottis, or by the cannula used to guide the probe. In other cases it was not possible to accurately position the probe at the centre of the medial edge of the vocal fold. The position of the probe microphone is likely a significant source of error.

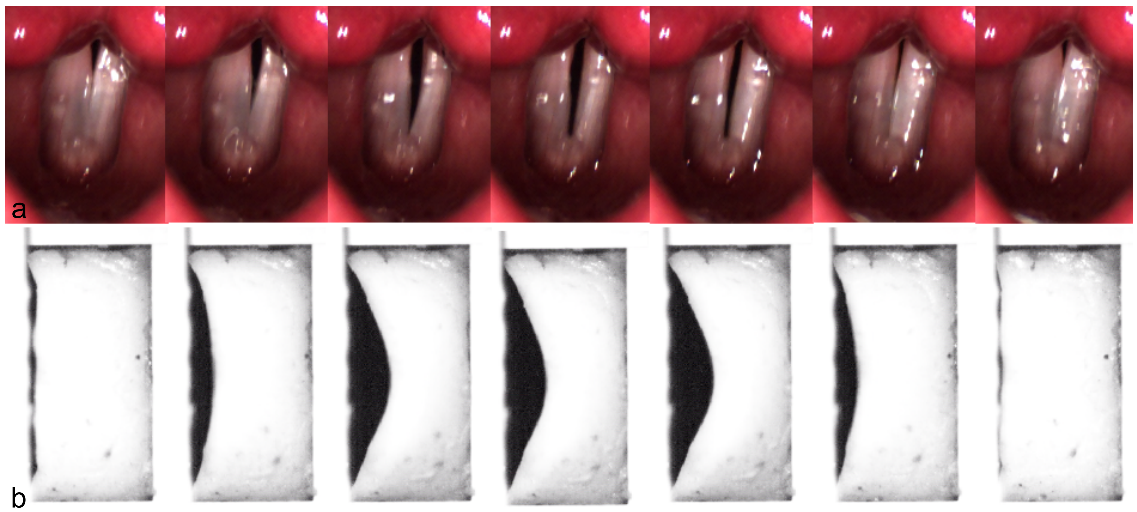
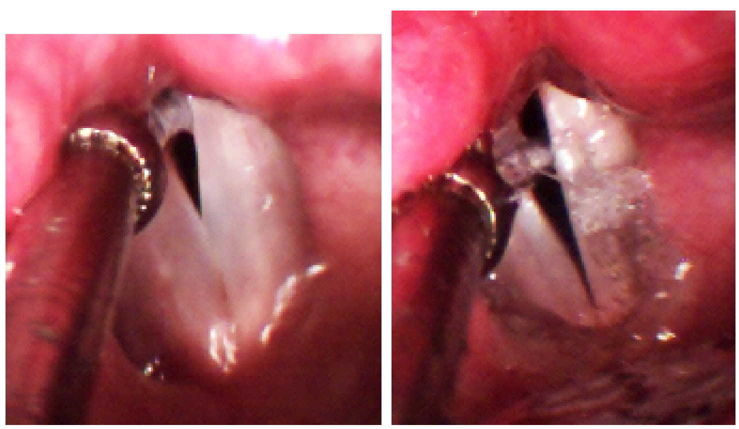


Figure 5–1: a) One cycle of vocal fold vibration from subject 11 and b) one cycle of physical model vocal fold vibration. The human subject vocal folds open from the posterior toward the anterior direction and close from the anterior toward the posterior direction. The physical model vocal folds open and close symmetrically in the anterior - posterior direction.



(a) (b)

Figure 5–2: In subject 11 a) the probe microphone is positioned at the posterior commissure, b) the probe microphone is positioned closer to the centre of the vocal folds.

CHAPTER 6
Conclusions and Future Work

The objective of this research was to estimate the contact pressure between the vocal folds in human subjects using high-speed videoendoscopy. A physical model was used to evaluate the contact pressure estimation method. The physical model experiments yielded comparable values from image analysis and from probe microphone measurement.

The image analysis method was then applied to human subjects experiments. The contact pressure values from high-speed video data were significantly greater than the values measured using the probe microphone, in contrast with the physical model verification experiments.

Limitations and sources of error for the contact pressure approximation were investigated. The assumptions for the effective mass and contact area likely contributed to the conflicting results in the human subject experiments. Additionally, the assumption of rigidity and averaging the contact force over time are simplifications that caused error in the analysis. Significant errors may have been due to the interference of the probe microphone in the human subjects. Inaccurate positioning of the probe microphone likely caused errors as well. Contact pressure values from the image analysis were more comparable to those reported in previous studies than the values from the probe microphone measurements.

Further investigation is necessary to determine the contact area and effective mass of the vocal folds. The contact area could be measured from a physical model or excised larynx in a hemilarynx configuration using high-speed video. The electroglottograph (EGG) is another tool that might be useful for measuring the vocal fold contact area. The EGG measures the electrical resistance across the larynx using two electrodes placed on the neck. When the vocal folds collide the electrical resistance is low, and when the vocal folds separate the resistance is high. The EGG is presently used for determining the portion of the glottal cycle for which the vocal folds are closed (the closed quotient) [19, 48]. It is also useful for determining the instants of glottal opening and closing [49]. While quantitative information can not be obtained directly from the amplitude of the EGG waveform [50], experiments could be conducted to determine if it might be possible to calibrate the EGG waveform to measure the vocal fold contact area.

In human subjects, a precise laser projection system may be used to calibrate endoscopic images and obtain accurate dimensions of the vocal folds. Three-dimensional information on the vocal fold dimensions may be obtained from MRI images. A finite element model based on the MRI geometry could be used along with material properties from the literature to determine the vocal fold contact area.

One possible approach to experimentally determining the effective mass of a vocal fold could be from its kinetic energy. The stress and strain on the superior surface of the vocal fold could be measured and used to determine the strain energy of the deformed vocal fold. The effective mass may be obtained from the kinetic energy, assuming that energy is conserved. Another possible approach is to use MRI images.

Images of a subject holding his breath may be compared to images of the subject's phonation to determine the volume of vibrating tissue. Values for the density of the tissue and the mode of vibration could be used to obtain the effective mass.

Future experiments should use high-speed video with a greater frame rate to more precisely capture the vocal fold vibration. Time interpolation methods could be used to improve the contact pressure approximations.

Finally, the development of an improved physical model vocal fold would help in advancing methods for contact pressure measurement. Physical model vocal folds that exhibit similar vibration pattern, layered structure, and material properties to human vocal folds would provide a better test platform for investigating vocal fold vibration.

Appendix A Edge velocity and glottal area from endoscopic video

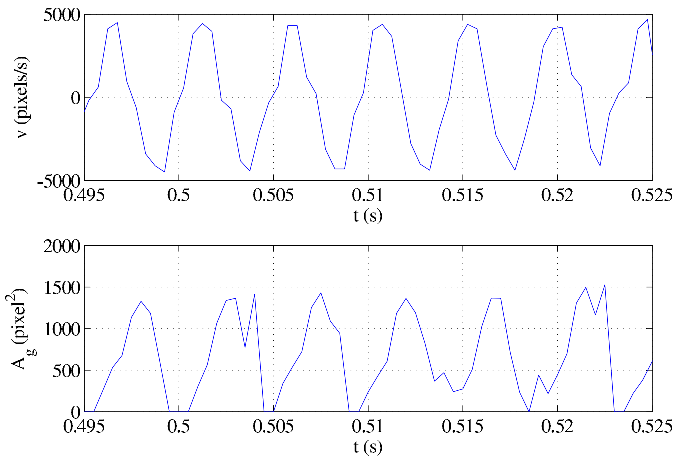


Figure A-1: Edge velocity (v) and glottal area (A_g) as a function of time from subject 9, normal voice, 70 dB.

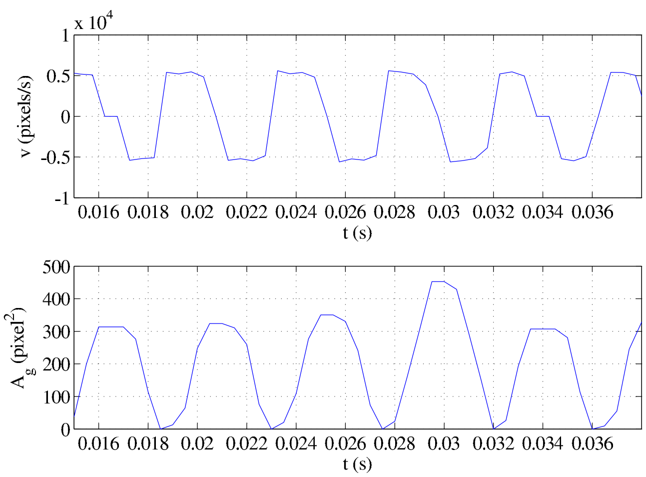


Figure A-2: Edge velocity (v) and glottal area (A_g) as a function of time from subject 9, normal voice, 80 dB.

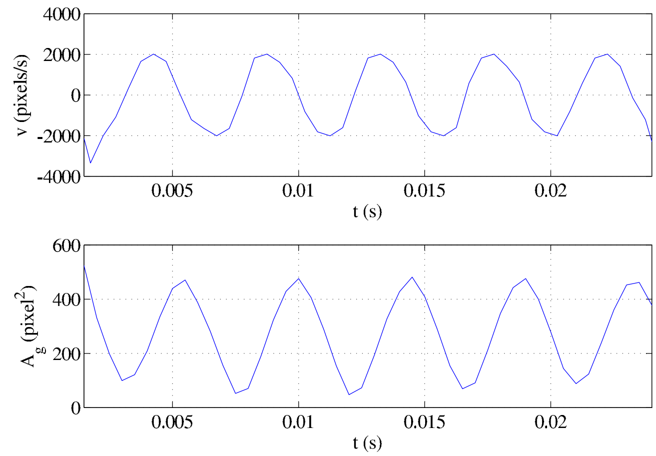


Figure A-3: Edge velocity (v) and glottal area (A_g) as a function of time from subject 10, breathy voice type, 60 dB.

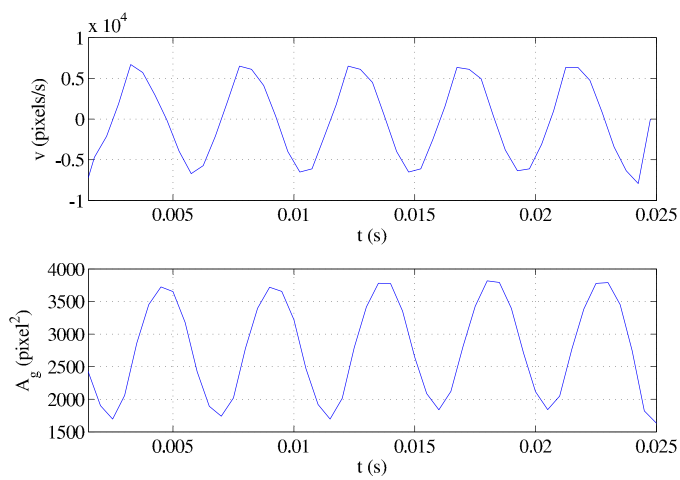


Figure A-4: Edge velocity (v) and glottal area (A_g) as a function of time from subject 10, breathy voice, 70 dB.

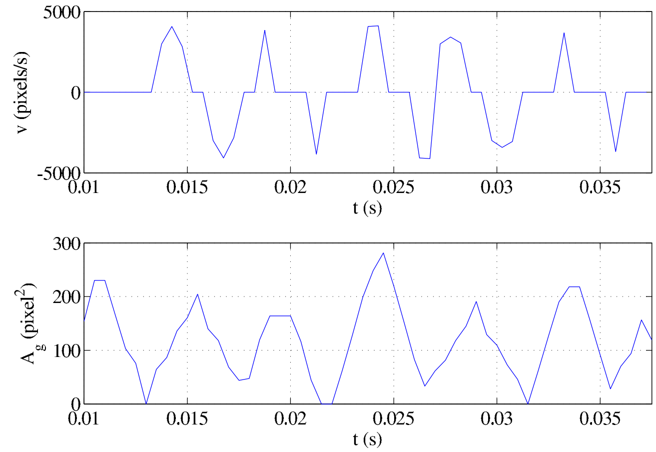


Figure A-5: Edge velocity (v) and glottal area (A_g) as a function of time from subject 10, normal voice, 60 dB.

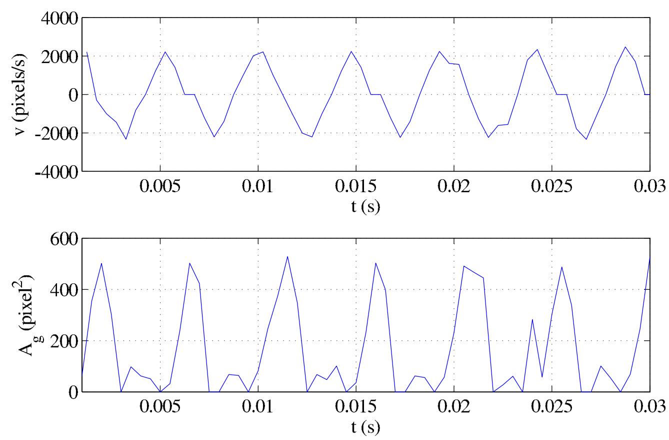


Figure A-6: Edge velocity (v) and glottal area (A_g) as a function of time from subject 10, normal voice type, 70 dB.

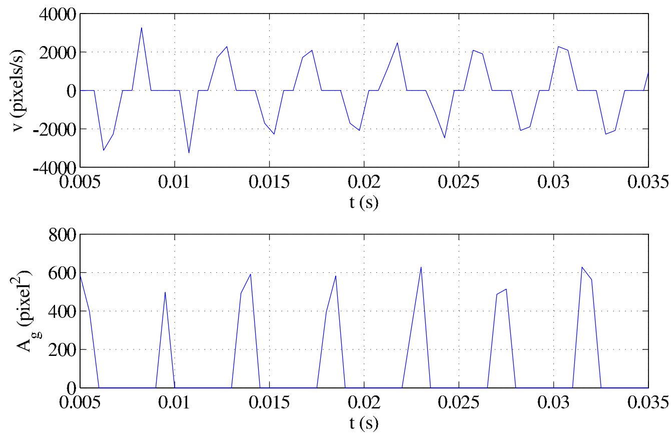


Figure A-7: Edge velocity (v) and glottal area (A_g) as a function of time from subject 10, normal voice, 80 dB.

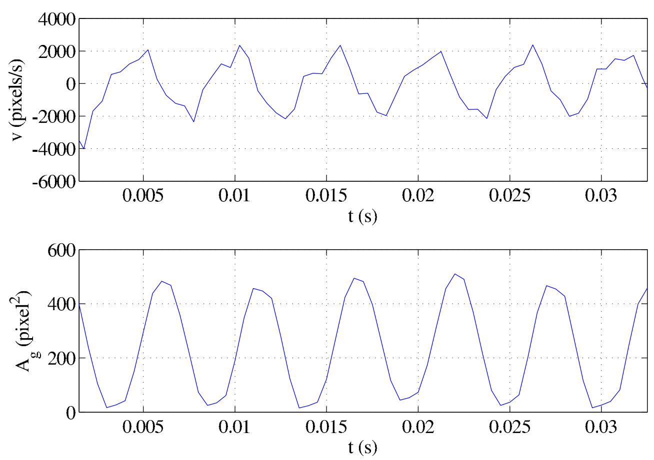


Figure A-8: Edge velocity (v) and glottal area (A_g) as a function of time from subject 10, pressed voice, 70 dB.

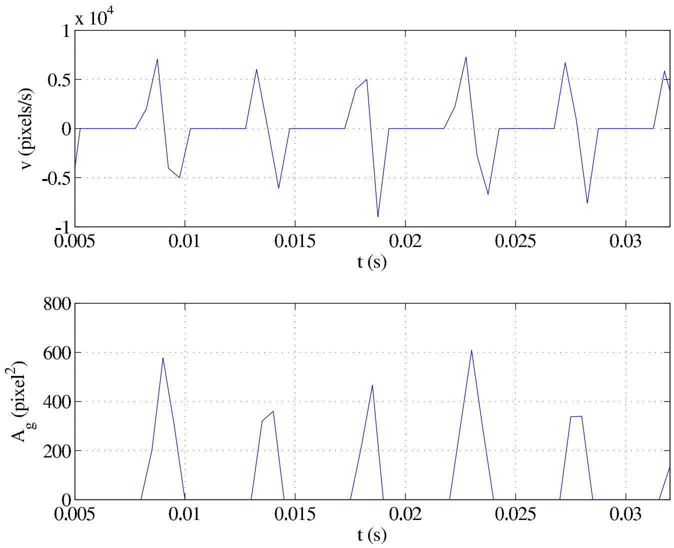


Figure A-9: Edge velocity (v) and glottal area (A_g) as a function of time from subject 10, pressed voice, 80 dB.

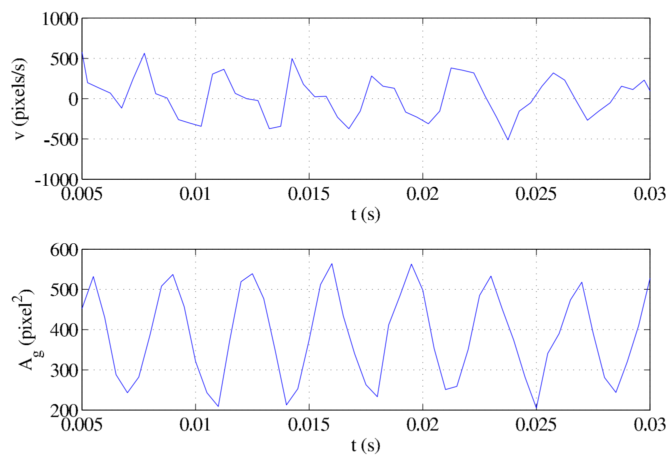


Figure A-10: Edge velocity (v) and glottal area (A_g) as a function of time from subject 11, normal voice, 60 dB.

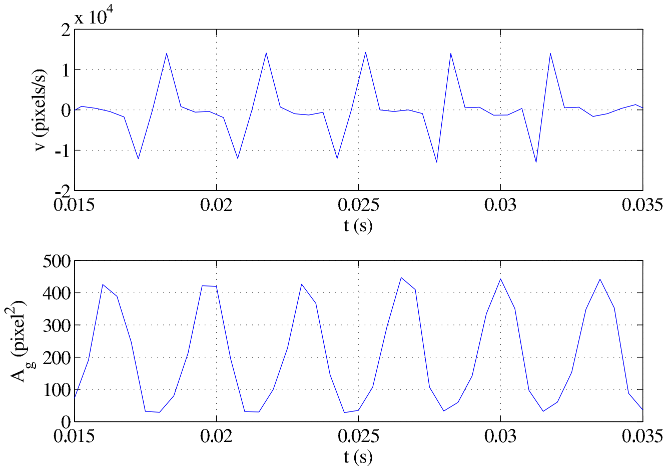


Figure A-11: Edge velocity (v) and glottal area (A_g) as a function of time from subject 11, normal voice, 70 dB.

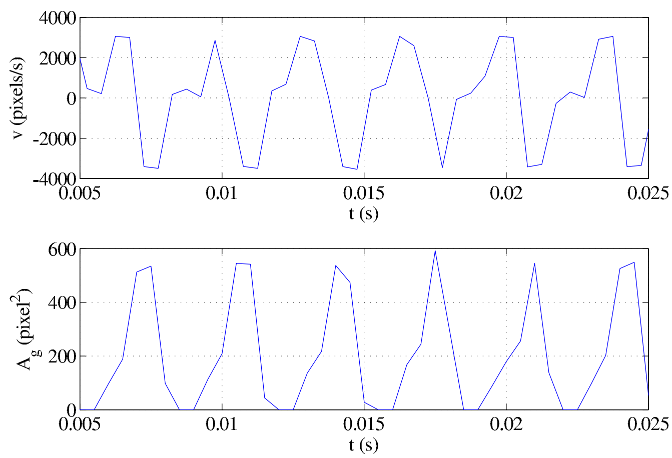


Figure A-12: Edge velocity (v) and glottal area (A_g) as a function of time from subject 11, normal voice type, 80 dB.

Appendix B MATLAB script for image segmentation

```
 %   JRobb Updated May 2010
clear all;clc; close all;

%% User-defined info enetered in this section
%   Location of images
index = strcat('C:\Users\Jenni\ThinkPad_MyDocuments\Pittsburgh\Data Pittsburgh July09\H
%   The first image to segment should be chosen to be the first image of
%   the opening phase in the glottal cycle
fframe = 1;             % The first image to segment
lframe = 50;             % The last image to segment

T = 15/255;              % Threshold for homogeneity criterion
LB = 0.001;     % The fraction of the total area that represents the lower
                 % bound of the glottal area
UB = 0.4;       % The fraction of the total area that represents the lower
                 % bound of the glottal area

%   Boundaries for cropping original image
%   Note y is numbered such that 0 is the top row

xLeftSeries = 56;
xRightSeries = 184;
yTopSeries = 19;
```

```
yBottomSeries = 250;

fps = 2000;

%  Unit conversion from pixels to mm for length, velocity, and area
%convertL = 0.054207071;  % mm/pixel length for physical model
convertL = 1;  % convertL = 1 for human subjects
convertV = convertL*fps;  % v(pix/frame)*2000(frame/s)*0.0xxxx(mm/pixel)
convertA = convertL.^2;   % area(#of pixels)*0.0xxxxx^2mm^2/pixel

%%

% Initializing arrays that grow in the main loop
%   Area vector contains the number of pixels defining each glottal area
area = zeros(lframe-fframe+1,2);
d = zeros(lframe-fframe+1,1);
P = zeros(lframe-fframe+1,5);
VFlength = zeros(lframe-fframe+1,1);
APpoints = 1;

for k = 1:(lframe-fframe+1)
    %disp('Iteration number')
    %disp(k)% Count each frame to analyze
    nframe = k+fframe-1;   % The frame number being processed in the loop
    if (nframe<=9)
      word=num2str(nframe);
      nameIP = strcat(index,'010-3-b-7000000',word,'.tif');
    elseif (10<=nframe && nframe<100)
```

```matlab
  word=num2str(nframe);
  nameIP = strcat(index,'010-3-b-700000',word,'.tif');
elseif (100<=nframe && nframe<1000)
  word=num2str(nframe);
  nameIP = strcat(index,'010-3-b-70000',word,'.tif');
elseif (1000<=nframe)
  word=num2str(nframe);
  nameIP = strcat(index,'010-3-b-7000',word,'.tif');
end

I = imread(nameIP);
f = rgb2gray(I);   % Convert to grayscale image

%   Cropping
f = (f(yTopSeries+1:yBottomSeries+1,xLeftSeries+1:xRightSeries+1));

%   Determine the first seed point interactively
if nframe == fframe
    % USER DEFINED SEED
    %S = roipoly(f);      % First seed point from user defined polygon
    %imwrite(S,strcat(index,'S', word,'.tif'), 'tif');
    % *OR* RE-USE SEED
    S = imread(strcat(index,'S',word,'.tif')); % reuse the poly image
    % for multiple runs

    Sopen = S;           % Keep the first seed point to use later
    TotalArea = size(f,1)*size(f,2) % The number of pixels in the image
```

```
    %  These are the locations where the edge velocity will be measured
    %  (for physical model only)
    port(3) = size(f,1)/2;
    port(2) = port(3) - 2.5/convertL;
    port(4) = port(3) + 2.5/convertL;
    port(1) = port(2) -2.5/convertL;
    port(5) = port(4) + 2.5/convertL;

end



%if nframe == 35
%    S = imread(strcat(index,'S35.tif')); % reuse the poly image
%end

%   Write the seed area to file (mostly for debugging, or first run)
% imwrite(S,strcat(index,'S', word,'.tif'), 'tif');

%   Determine glottal area using regiongrow
[g, NR, SI, TI] = regiongrow(f, S, T);

%   Glottal area (pixels)
area(k,1) = nframe;
area(k,2) = sum(sum(g))
```

```
%   Check that the glottal area was segemented properly by making sure
%   the number of segmented pixels (glottal area) is a reasonable value
if area(k,2) > LB*TotalArea && area(k,2) < UB*TotalArea
    disp(['The area is between: ', num2str(LB),' and ', num2str(UB),...
        ' times the total area']);

    g = imfill(g,'holes');  % Get ride of "speckles" in the segmented area
    area(k,2) = sum(sum(g));% Record the glottal area for this image
    S = g;                  % Determine seed point for next image
    %S = Sopen;
end

%   If he glottal area is improperly segmented
if area(k,2) > UB*TotalArea || area(k,2) < LB*TotalArea
     disp(['The area is less than: ', num2str(LB),' or greater than ', ...
         num2str(UB),' times the total area, try seed point from first image']);
    % Try regiongrow again using the seed point from the first image -
    % this step is especially usefull for catching the first image of
    % the opening phase after complete glottal closure
    S = imread(strcat(index,'S1.tif'));
    [g, NR, SI, TI] = regiongrow(f, S, T);
    g = imfill(g,'holes');
        if sum(sum(g)) > UB*TotalArea || sum(sum(g)) < LB*TotalArea
            disp(['The area is still less than: ', num2str(LB),' or greater than ',
            num2str(UB),' times the total area, set to zero']);
            % If that still doesn't work then it's probably a complete
            % glottal closure and the glottal area is zero
            g = zeros(size(f));
```

```
            area(k,2) = sum(sum(g));
            %disp('Complete glottal closure at frame number')
            %disp(word)
            S = Sopen ;
        else
            disp('using seed point from first image worked because g =')
            disp(sum(sum(g)))
            g = imfill(g,'holes');
            area(k,2) = sum(sum(g));
            S = g;



        end

 end

 %   Determine the boundary of the glottal area
 p = bwperim(g,4);            % Use 4 connectivity
 %   Determine the number or regions of the glottal area and label all
 %   the regions as 1
 L = bwlabel(g);
 ind = find(L);
 L(ind) = 1;



% if nframe == 23
%    g = zeros(size(f));
%    area(k,2) = sum(sum(g));
```

```
% end

% if nframe == 34
%    g = zeros(size(f));
%    area(k,2) = sum(sum(g));
% end

% if nframe == 35
%    g = zeros(size(f));
%    area(k,2) = sum(sum(g));
% end

 %   Determine the perimeter, major and minor axes, and points along
 %   medial edge for velocity measurement
 if area(k,2) ~= 0
     % still some issues with determine the upper edge and lower edge,
     % for now just use the minor axes for velocity points

     [s,midedge, portV, VFlength, APpoints] = diameter_modified(L, APpoints, ...
         k, VFlength, index, port);

     %disp(s.MajorAxis)
     %disp(s.MinorAxis)
     %disp(portV)

     %   Mark the velocity points with a "+" sign
     perimeter = markaxes(p == false, s.MajorAxis, midedge, portV);
     imwrite(perimeter,strcat(index,'perimeter', word,'.tif'), 'tif');
```

```
        %   Determine the distance from the velocity point to the midline
        d(k) = edgevelocity(midedge);

        for i = 1:5
            P(k,i) = edgevelocity(portV(:,:,i));
        end
    end

    % Convert segmented images to black on white (better for printing)
    gtemp = g == false;
    %SItemp = SI == false;
    %TItemp = TI == false;

    imwrite(gtemp,strcat(index,'g', word,'.tif'), 'tif');
    % imwrite(SI,strcat(index,'SI', word,'.tif'), 'tif');
    % imwrite(TItemp,strcat(index,'TI', word,'.tif'), 'tif');

    %disp(nframe)
    %disp(area(k,2))
end

%% Display results
area(:,2) = area(:,2)*convertA; % area in mm^2
disp('The glottal area for each image in mm^2 is:')
disp(area)

% Find the change in distance b/t minor axis coordinates after each frame
% Positive velocity is the left VF in the medial direction, v is in
```

```
% pixels/frame
v = zeros(length(d)-1,1);
vP = zeros(length(P)-1,5);
for i = 1:length(d)-1
v(i) = d(i+1)-d(i);
vP(i,:) = P(i+1,:)-P(i,:);
end

v = v*convertV; % v in mm/s
vP = vP*convertV; % v in mm/s
v = smooth(v);
for i = 1:5
    vP(:,i) = smooth(vP(:,i));
end
area(:,2)=smooth(area(:,2));

%disp('The edge velocity of one vocal fold in mm/s is:')
disp('The edge velocity of one vocal fold in pixels/s is:')
disp(v)

VFlength = VFlength*convertL;     % L in mm
%disp('The length of the vocal folds in mm is:')
disp('The length of the vocal folds in pixels is:')
disp(max(VFlength))

%   Time domain for edge velocity and area plot
domV = fframe/fps+1/(2*fps):1/fps:lframe/fps-1/(2*fps);
domT = fframe/fps:1/fps:lframe/fps;
```

```
%% Create figures
figure(1)
subplot(2,1,1)
plot(domV,v)
xlabel(' t (s) ')
%ylabel(' v (mm/s) ')   % Positive velocity in the lateral direction
ylabel(' v (pixels/s) ')   % Positive velocity in the lateral direction
grid on
subplot(2,1,2)
plot(domT,area(:,2))
xlabel(' t (s) ')
%ylabel(' Glottal area in mm^2 ')   % Positive velocity in the lateral direction
ylabel(' A_g (pixel^2) ')
grid on

figure(2)
%[AX,H1,H2]= plotyy(domV, vP(:,3), domT, area(:,2));
[AX,H1,H2]= plotyy(domV, v, domT, area(:,2));
set(get(AX(1),'Ylabel'),'FontSize', 12,'String','Edge Velocity (mm/s)')
set(get(AX(2),'Ylabel'),'FontSize', 12,'String','Glottal Area (mm^2)')
xlabel('Time (S)')
set(H1,'LineStyle','-')
set(H2,'LineStyle',':')
grid on



function [g, NR, SI, TI] = regiongrow(f, S, T)
```

```
%   Gonzalez, Woods and Eddins; Digital Image Processing using MATLAB,
%   2nd Edition, Gatesmark Pubslishing (2009);p.580

%REGIONGROW Perform segmentation by regiongrowing.
%   [G, NR, SI, TI] = REGIONGROW(F, S, T).  S can be an arragy (the same
%   size as F) with a 1 at the coordinates of every seed point and 0s
%   elsewhere.  S can also be a single seed value.  Similarly, T can be an
%   array (the same size as F) containing a threshold value for each pixel
%   in F.  T can also be a scalar, in which case it becomes a global
%   threshold.  All values in S and T must be in the range [0 ,1]
%
%   G is the result of region growing, with each region labeled by a
%   different integer, NR is the number of regions, SI is the final seed
%   image used by the algorithm, and TI is the image consisting of the
%   pixels in F that satisfied the threshold test, but before they were
%   processed for connectivity.

f = tofloat(f);
% If S is a scalar, obtain the seed image.
if numel(S) == 1
    SI = f == S;
    S1 = S;
else
    % S is an array.  Eliminate duplicate, connected seed locations to
    % reduce the number of loop executions in the following sections of
    % code.
    SI = bwmorph(S, 'shrink', Inf);
```

```
    S1 = f(SI);  % Array of seed values.
end

TI = false(size(f));
for K = 1:length(S1)
    seedvalue = S1(K);
    S = abs(f - seedvalue) <= T; % Re-use variable S.
    TI = TI | S;
end

% Use function imreconstruct with SI as the marker image to obtain the
% regions corresponding to each seed in S.  Function bwlabel assigns a
% different integer to each connected region.
[g, NR] = bwlabel(imreconstruct(SI, TI));



function [s, midedge, portV, VFlength, APpoints] = diameter_modified(L, APpoints, ...
    k, VFlength, index, port)
%

% Gonzalez, Woods and Eddins; Digital Image Processing using MATLAB,
% 2nd Edition, Gatesmark Pubslishing (2009);p.767-770

% %S = DIAMETER(L) computes the diameter, the major axis endpoints, the
% minor asic endpoints, and the basic rectangle of each labeled region in
% the label matrix L.  Positive integer elements of L correspond to
% different regions.  For example, the set of elements of L equal to 1
% corresponds to region 1; the set of elements of L equal to 2 corresponds
```

```
% to region 2; and so on.  S is a structure array of length max(L(:)).
The
% fields of the structure array include:
%
%   Diameter
%   MajorAxis
%   MinorAxis
%   BasicRectangle
%
% The Diameter field, a scalar, is the maximum distance between any two
% pixels in the corresponding region.
%
% The MajorAxis field is a 2-by-2 matrix.  The rows contain the row and
% column coordinates for the endpoints of the major axis of the
% correspontding region.
%
% The MinorAxis field is a 2-by-2 matrix.  The row contain the row and
% column coordinates for the endpoints of the minor axis of the
% corresponding region.
%
% The BasicRectangle field is a 4-by-2 matrix.  Each row contains the row
% and column coordinates of a corner of the region-enclosing rectangle
% defined by the major and minor axes.
%
% For more information about these measurements see Section 11.2.1 of
% Digital Image Processing, by Gonxalex and Woods, 2nd edition, Perntice
% Hall.
```

```
s = regionprops(L, {'Image', 'BoundingBox'});

for i = 1:length(s)
       [s(i).Diameter, s(i).MajorAxis, perim_r, perim_c] = ...
        compute_diameter(s(i));

       % disp('Original')
       % disp(s.MajorAxis)
        VFlength(k) = s.Diameter;
        if VFlength(k) == max(VFlength)
            APpoints = s.MajorAxis;
        end
        s.MajorAxis = APpoints;
       [s(i).BasicRectangle, s(i).MinorAxis] = ...
        compute_basic_rectangle(s(i), perim_r, perim_c);
    %disp(s.MinorAxis)
    midedge = velpoints(s.MinorAxis, s.BoundingBox, perim_r, perim_c);

   portV = zeros(2,2,length(port));
    for j = 1:length(port)
        portP = [port(j), s.MinorAxis(1,2); port(j) s.MinorAxis(2,2)];
        portV(:,:,j) = velpoints(portP, s.BoundingBox, perim_r, perim_c);
    end
end


end
```

```
function [ d, majoraxis, r, c] = compute_diameter(s)

%   Gonzalez, Woods and Eddins; Digital Image Processing using MATLAB,
%   2nd Edition, Gatesmark Pubslishing (2009);p.767-770

%   [D, MAJORAXIS, R, C] = COMPUTE_DIAMETER(S) computes the diameter and
%   major axis for the region represented by the structure S.  S must
%   contain the fields Image and BoundingBox.  COMPUTE_DIAMETER also
%   returns the row and column coordinates (R and C) of the perimeter
%   pixels of s.Image.

% Compute row and column coordinates of perimeter pixels.
[r, c] = find(bwperim(s.Image));
r = r(:);
c = c(:);

[rp, cp] = prune_pixel_list(r, c);

num_pixels = length(rp);
switch num_pixels
    case 0
        d = -Inf;
        majoraxis = ones(2,2);

    case 1
        d = 0;
        majoraxis = [rp cp; rp cp];
```

```
    case 2
        d = (rp(2)-rp(1))^2 + (cp(2)-cp(1))^2;%square root needed?
        majoraxis = [rp cp];

    otherwise
        % Generate all combinations of 1:num_pixels taken two at a time.
        % Method suggested by Peter Acklam.
        [idx(:, 2) idx(:, 1)] = find(tril(ones(num_pixels), -1));
        rr = rp(idx);
        cc = cp(idx);

        dist_squared = (rr(:, 1) - rr(:, 2)).^2 + ...
            (cc(:, 1) - cc(:, 2)).^2;
        [max_dist_squared, idx] = max(dist_squared);
        majoraxis = [rr(idx,:)' cc(idx,:)'];

        d = sqrt(max_dist_squared);

        upper_image_row = s.BoundingBox(2) + 0.5;
        left_image_col = s.BoundingBox(1) + 0.5;

        majoraxis(:, 1) = majoraxis(:, 1) + upper_image_row - 1;
        majoraxis(:, 2) = majoraxis(:, 2) + left_image_col - 1;
end


function [ basicrect, minoraxis] = ...
    compute_basic_rectangle(s,perim_r, perim_c)
```

```
%   Gonzalez, Woods and Eddins; Digital Image Processing using MATLAB,
%   2nd Edition, Gatesmark Pubslishing (2009);p.767-770

%   [BASICRECT,MINORAXIS] = COMPUTE_BASIC_RECTANGLE(S, PERIM_R,
%   PERIM_C) computes the basic rectangle and the minor axis
%   endpoints for the region represented by the structure S.  S must
%   contain the fields Image, BoundingBox, MajorAxis, and Diameter.
%   PERIM_R and PERIM_C are the row and column coordinates of perimeter
%   of s.Image.  BASICRECT is a 4-by-2 matrix, each row of which
%   contains the row and column coordinates of one corner of the basic
%   rectangle.

%Compute the orientation o the major axis
theta = atan2(s.MajorAxis(2, 1) - s.MajorAxis(1, 1),...
    s.MajorAxis(2, 2) - s.MajorAxis(1, 2));

%Form rotation matrix
T = [cos(theta) sin(theta); -sin(theta) cos(theta)];

%Rotate perimeter pixels
p = [perim_c perim_r];
ind = length(p)+1;

 majoraxis(:, 1) = s.MajorAxis(:, 1) - (s.BoundingBox(2) + 0.5) + 1;
 majoraxis(:, 2) = s.MajorAxis(:, 2) - (s.BoundingBox(1) + 0.5) + 1;

p(ind,1) = majoraxis(1,1);
```

```
p(ind,2) = majoraxis(2,1);
p = p * T';

%Calculate minimum and maxmum x- and y-coordinates for the rotated
%perimeter pixels
x = p(:, 1);
y = p(:, 2);
min_x = min(x);
max_x = max(x);
min_y = min(y);
max_y = max(y);

%imwrite(x, strcat(index,'xtest.tif.'), 'tif');
%imwrite(y, strcat(index,'ytest.tif.'), 'tif');

corners_x = [min_x max_x max_x min_x]';
corners_y = [min_y min_y max_y max_y]';

%Rotate corners of the basic rectangle
corners = [corners_x corners_y] * T;

%Translate according to the region's bounding box
upper_image_row = s.BoundingBox(2) + 0.5;
left_image_col = s.BoundingBox(1) + 0.5;

basicrect = [corners(:, 2) + upper_image_row - 1, ...
             corners(:, 1) + left_image_col - 1];
%Compute minor axis end-points, rotated
```

```
x = (min_x + max_x)/2;
y1 = min_y;
y2 = max_y;
endpoints = [x y1; x y2];

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% modified by jr 22OCT09
%xlow = (min_x + max_x)/3;
%xup = (min_x + max_x)*2/3;

%endpointslow = [xlow y1; xlow y2];
%endpointsup = [xup y1; xup y2];
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

%Rotate minor axis end-points back
endpoints = endpoints * T;

%endpointslow = endpointslow * T;
%endpointsup = endpointsup * T;

%Translate according to the region's bounding box
minoraxis = [endpoints(:, 2) + upper_image_row - 1, ...
    endpoints(:, 1) + left_image_col-1];

%pointslow = [endpointslow(:, 2) + upper_image_row - 1, ...
    %endpointslow(:, 1) + left_image_col-1];
%pointsup = [endpointsup(:, 2) + upper_image_row - 1, ...
    %endpointsup(:, 1) + left_image_col-1];
```

```
function [r, c] = prune_pixel_list(r, c)

%   Gonzalez, Woods and Eddins; Digital Image Processing using MATLAB,
%   2nd Edition, Gatesmark Pubslishing (2009);p.767-770

%   [R, C] = PRUNE_PIXEL_LIST(R,C) removes pixels from the vectors R
%   and C that cannot be endpoints of the major axis.  This elimination
%   is based on geometrical constraints described in Russ, Image
%   Processing Handbook, Chapter 8.

top = min(r);
bottom = max(r);
left = min(c);
right = max(c);

%Which points are inside the upper circle?
x = (left + right)/2;
y = top;
radius = bottom - top;
inside_upper = ((c - x).^2 + (r - y).^2) < radius^2;

%Which points are inside the lower circle?
y = bottom;
inside_lower = ((c - x).^2 + (r - y).^2) < radius^2;

%Which points are inside the left circle?
x = left;
y = (top + bottom)/2;
```

```
radius = right - left;
inside_left = ((c - x).^2 + (r - y).^2) < radius^2;

%Which points are inside the right circle?
x = right;
inside_right = ((c - x).^2 + (r - y).^2) < radius^2;

%Eliminate points that are inside all four circles
delete_idx = find(inside_left & inside_right & ...
                  inside_upper & inside_lower);
r(delete_idx) = [];
c(delete_idx) = [];



function [out, revertclass] = tofloat(in)

%   Gonzalez, Woods and Eddins; Digital Image Processing using MATLAB,
%   2nd Edition, Gatesmark Pubslishing (2009);p.806

%TOFLOAT Convert image to floating point
%   [OUT, REVERTCLASS] = TOFLOAT(IN) converts the input image IN to
%   floating-point.  If N is a double or single image, then OUT equals IN.
%   Otherwise, OUT equals IM2SINGLE(IN).  REVERTCLASS is a function handle
%   that can be used to convert back to the class of IN.

identity = @(x) x;
tosingle = @im2single;
```

```
table = {'uint8', tosingle, @im2uint8
    'uint16', tosingle, @im2uint16
    'int16', tosingle, @im2int16
    'logical', tosingle, @logical
    'double', identity, identity
    'single', identity, identity};

classIndex = find(strcmp(class(in), table(:, 1)));

if isempty(classIndex)
    error('Unsupported input image class. ');
end

out = table{classIndex, 2}(in);

revertclass = table{classIndex, 3};



function midedge = velpoints(MinorA, Box, perim_r, perim_c)



upper_image_row = Box(2) + 0.5;
    left_image_col = Box(1) + 0.5;
    perimcoords = [perim_r + upper_image_row - 1, ...
        perim_c + left_image_col-1] ;

    x1 = uint16(MinorA(1,1));
    x2 = uint16(MinorA(2,1));
```

```
y1 = uint16(MinorA(1,2));
y2 = uint16(MinorA(2,2));

midedge = zeros(2);
for i = 1:length(perimcoords)
    if perimcoords(i,1) == x1 && perimcoords(i,2) == y1
        midedge(1,:) = perimcoords(i,:);
    end
    if perimcoords(i,1) == x2 && perimcoords(i,2) == y2
        midedge(2,:) = perimcoords(i,:);
    end
end

NoMinor = find(midedge == 0);
if length(NoMinor) == 2
    if NoMinor == [1;3]
        indices = find(perimcoords(:,1)==x1);
        for i = 1:length(indices)
            ctemp(i) = perimcoords(indices(i),2);
        end
        closest = abs(double(ctemp)-double(y1));
        minc = min(closest);
        idx = closest == minc;
        midedge(1,:) = [x1 perimcoords(indices(idx),2)];
    end
    if NoMinor == [2;4]
        indices = find(perimcoords(:,1)==x2);
        for i = 1:length(indices)
```

```
            ctemp(i) = perimcoords(indices(i),2);
        end
        closest = abs(double(ctemp)-double(y2));
        minc = min(closest);
        idx = closest == minc;
        midedge(2,:) = [x2 perimcoords(indices(idx),2)];
    end
elseif length(NoMinor) == 4
        indices1 = find(perimcoords(:,1)==x1);
        indices2 = find(perimcoords(:,1)==x2);
        TF1 = isempty(indices1);
        TF2 = isempty(indices2);

        if TF1 == 0 && TF2 ==0
            for i = 1:length(indices1)
                ctemp1(i) = perimcoords(indices1(i),2);
            end
            closest1 = abs(double(ctemp1)-double(y1));
            minc1 = min(closest1);
            idx1 = closest1 == minc1;
            midedge(1,:) = [x1 perimcoords(indices1(idx1),2)];

            for i = 1:length(indices2)
                ctemp2(i) = perimcoords(indices2(i),2);
            end
            closest2 = abs(double(ctemp2)-double(y2));
            minc2 = min(closest2);
            idx2 = closest2 == minc2;
```

```
            midedge(2,:) = [x2 perimcoords(indices2(idx2),2)];
        end
    end


function [Boundary_Axes] = markaxes(Boundary,Major_Axis_Coordinates, ...
    Minor_Axis_Coordinates, portV)

 %Mark the major axis endpoints with a "+"
    Boundary_Axes = Boundary;
    p = Major_Axis_Coordinates(1,1);
    q = Major_Axis_Coordinates(1,2);
            Boundary_Axes(p,q) = 0;
         if p+1 > 0
            Boundary_Axes(p+1,q) = 0;
         end
         if (p-1)>0
            Boundary_Axes(p-1,q) = 0;
         end
            Boundary_Axes(p,q+1) = 0;
         if q-1 > 0
            Boundary_Axes(p,q-1) = 0;
         end
            Boundary_Axes(p+2,q) = 0;
         if p-2 > 0
            Boundary_Axes(p-2,q) = 0;
         end
            Boundary_Axes(p,q+2) = 0;
```

```
        if q-2 > 0
            Boundary_Axes(p,q-2) = 0;
        end
r = Major_Axis_Coordinates(2,1);
u = Major_Axis_Coordinates(2,2);
            Boundary_Axes(r,u) = 0;
        if (r-1)>0
            Boundary_Axes(r-1,u) = 0;
        else
             Boundary_Axes(1,u) = 0;
        end
            Boundary_Axes(r,u+1) = 0;

        if u-1 > 0
            Boundary_Axes(r,u-1) = 0;
        else
             Boundary_Axes(r,1) = 0;
        end
            Boundary_Axes(r+2,u) = 0;
        if r-2 > 0
            Boundary_Axes(r-2,u) = 0;
        else
             Boundary_Axes(1,u) = 0;
        end
            Boundary_Axes(r,u+2) = 0;
        if u-2 > 0
            Boundary_Axes(r,u-2) = 0;
        else
```

```
        Boundary_Axes(r,1) = 0;
    end

%Mark the minor axis endpoints with a "+"
if Minor_Axis_Coordinates ~= 0
p = int16(Minor_Axis_Coordinates(1,1));
q = int16(Minor_Axis_Coordinates(1,2));
        Boundary_Axes(p,q) = 0;
        Boundary_Axes(p+1,q) = 0;
        if p-1 > 0
        Boundary_Axes(p-1,q) = 0;
        end
        Boundary_Axes(p,q+1) = 0;
        if q-1 > 0
        Boundary_Axes(p,q-1) = 0;
        end

        Boundary_Axes(p+2,q) = 0;
        if p-2 > 0
        Boundary_Axes(p-2,q) = 0;
        end
        Boundary_Axes(p,q+2) = 0;
        if q-2 > 0
        Boundary_Axes(p,q-2) = 0;
        end
r = int16(Minor_Axis_Coordinates(2,1));
u = int16(Minor_Axis_Coordinates(2,2));
        Boundary_Axes(r,u) = 0;
```

```
            Boundary_Axes(r+1,u) = 0;
            Boundary_Axes(r-1,u) = 0;
            Boundary_Axes(r,u+1) = 0;
            if u-1 > 0
            Boundary_Axes(r,u-1) = 0;
            end
            Boundary_Axes(r+2,u) = 0;
            Boundary_Axes(r-2,u) = 0;
            Boundary_Axes(r,u+2) = 0;
            if u-2 > 0
            Boundary_Axes(r,u-2) = 0;
            end
    end

 %Mark the port points with a "+"
% for i = 1:5
%   if portV(:,:,i) ~= 0
%   p = int16(portV(1,1,i));
%   q = int16(portV(1,2,i));
%           Boundary_Axes(p,q) = 0;
%           Boundary_Axes(p+1,q) = 0;
%           Boundary_Axes(p-1,q) = 0;
%           Boundary_Axes(p,q+1) = 0;
%           if q-1 > 0
%           Boundary_Axes(p,q-1) = 0;
%           end
%
%           Boundary_Axes(p+2,q) = 0;
```

```
%           Boundary_Axes(p-2,q) = 0;
%           Boundary_Axes(p,q+2) = 0;
%           if q-2 > 0
%           Boundary_Axes(p,q-2) = 0;
%           end
%   r = int16(portV(2,1,i));
%   u = int16(portV(2,2,i));
%           Boundary_Axes(r,u) = 0;
%           Boundary_Axes(r+1,u) = 0;
%           Boundary_Axes(r-1,u) = 0;
%           Boundary_Axes(r,u+1) = 0;
%           %if u-1 > 0
%           Boundary_Axes(r,u-1) = 0;
%           %end
%           Boundary_Axes(r+2,u) = 0;
%           Boundary_Axes(r-2,u) = 0;
%           Boundary_Axes(r,u+2) = 0;
%           %if u-2 > 0
%           Boundary_Axes(r,u-2) = 0;
%           %end
%   end
 end
```

References

[1] Canadian Association of Speech-Language Pathologists and Audiologists, "Speech and hearing fact sheet." [Online]. Available: http://www.caslpa.ca. [Accessed: Oct. 1, 2009], 2005.

[2] I. R. Titze, "Mechanical-stress in phonation," *Journal of Voice*, vol. 8, no. 2, pp. 99–105, 1994.

[3] M. M. Hess, K. Verdolini, W. Bierhals, U. Mansmann, and M. Gross, "Endolaryngeal contact pressures," *Journal of Voice*, vol. 12, no. 1, pp. 50–67, 1998.

[4] K. Verdolini, M. M. Hess, I. R. Titze, W. Bierhals, and M. Gross, "Investigation of vocal fold impact stress in human subjects," *Journal of Voice*, vol. 13, no. 2, pp. 184–202, 1999.

[5] H. E. Gunter, R. D. Howe, S. M. Zeitels, J. B. Kobler, and R. E. Hillman, "Measurement of vocal fold collision forces during phonation: Methods and preliminary data," *Journal of Speech Language and Hearing Research*, vol. 48, no. 3, pp. 567–576, 2005.

[6] L.-J. Chen, *Investigation of mechanical stresses within vocal folds during phonation.* PhD thesis, Purdue Univeristy, 2009.

[7] I. R. Titze, *Principles of Voice Production.* Prentice Hall, 1994.

[8] K. Peterson, J. Barkmeier, K. Verdolinimarston, and H. Hoffman, "Comparison of aerodynamic and electroglottographic parameters in evaluating clinically relevant voicing patterns," *The Annals of Otology, Rhinology & Laryngology*, vol. 103, no. 5, pp. 335–346, 1994.

[9] N. Y.-K. Li, *Biosimulation of Vocal Fold Inflamation and Wound Healing.* PhD thesis, University of Pittsburgh, 2009.

[10] M. M. Johns, "Update on the etiology, diagnosis, and treatment of vocal fold nodules, polyps, and cysts.," *Curr Opin Otolaryngol Head Neck Surg*, vol. 11, pp. 456–61, Dec. 2003.

[11] P. D. Karkos and M. McCormick, "The etiology of vocal fold nodules in adults," *Current Opinion in Otolaryngology & Head and Neck Surgery*, vol. 17, no. 6, pp. 420–423, 2009.

[12] K. W. Altman, "Vocal fold masses," *Otolaryngologic Clinics of North America*, vol. 40, no. 5, pp. 1091 – 1108, 2007. The Professional Voice.

[13] J. Horacek, A. M. Laukkanen, and P. Sidlof, "Estimation of impact stress using an aeroelastic model of voice production," *Logopedics Phoniatrics Vocology*, vol. 32, no. 4, pp. 185–192, 2007.

[14] J. Horacek, P. Sidlof, and J. G. Svec, "Numerical simulation of self-oscillations of human vocal folds with hertz model of impact forces," *Journal of Fluids and Structures*, vol. 20, no. 6, pp. 853–869, 2005. 44 8th International Conference on Flow-Induced Vibrations (FIV 2004) Jul, 2004 Palaiseau, FRANCE.

[15] J. Horacek, A. M. Laukkanen, P. Sidlof, P. Murphy, and J. G. Svec, "Comparison of acceleration and impact stress as possible loading factors in phonation: A computer modeling study," *Folia Phoniatrica Et Logopaedica*, vol. 61, no. 3, pp. 137–145, 2009.

[16] M. Spencer, T. Siegmund, and L. Mongeau, "Determination of superior surface strains and stresses, and vocal fold contact pressure in a synthetic larynx model using digital image correlation," *Journal of the Acoustical Society of America*, vol. 123, no. 2, pp. 1089–1103, 2008.

[17] J. Young, "Measurement of impact stress on the superior surface of the vocal folds using digital imagery correlation," Master's thesis, McGill University, 2011.

[18] J. J. Q. Jiang and I. R. Titze, "Measurement of vocal fold intraglottal pressure and impact stress," *Journal of Voice*, vol. 8, no. 2, pp. 132–144, 1994.

[19] K. Verdolini, R. Chan, I. R. Titze, M. Hess, and W. Bierhals, "Correspondence of electroglottographic closed quotient to vocal fold impact stress in excised canine larynges," *Journal of Voice*, vol. 12, no. 4, pp. 415–423, 1998.

[20] J. J. Jiang, A. G. Shah, M. M. Hess, K. Verdolini, F. M. Banzali, and D. G. Hanson, "Vocal fold impact stress analysis," *Journal of Voice*, vol. 15, no. 1, pp. 4–14, 2001.

[21] M. Doellinger, "The next step in voice assessment: High-speed digital endoscopy and objective evaluation," *Current Bioinformatics*, vol. 4, no. 2, pp. 101–111, 2009.

[22] D. D. Deliyski and R. E. Hillman, "State of the art laryngeal imaging: research and clinical implications," *Current Opinion in Otolaryngology & Head and Neck Surgery*, vol. 18, no. 3, pp. 147–152, 2010.

[23] T. Wittenberg, R. Moser, M. Tigges, and U. Eysholdt, "Recording, processing, and analysis of digital high-speed sequences in glottography," *Machine Vision and Applications*, vol. 8, no. 6, pp. 399–404, 1995.

[24] D. D. Deliyski, P. P. Petrushev, H. S. Bonilha, T. T. Gerlach, B. Martin-Harris, and R. E. Hillman, "Clinical implementation of laryngeal high-speed videoendoscopy: Challenges and evolution," *Folia Phoniatrica Et Logopaedica*, vol. 60, no. 1, pp. 33–44, 2008.

[25] C. Tao, Y. Zhang, and J. J. Jiang, "Extracting physiologically relevant parameters of vocal folds from high-speed video image series," *IEEE Transactions on Biomedical Engineering*, vol. 54, no. 5, pp. 794–801, 2007.

[26] R. C. Gonzalez, R. E. Woods, and S. L. Eddins, *Digital Image processing using MATLAB*. [S.I.]: Gatesmark Publishing, 2009.

[27] Y. L. Yan, X. Chen, and D. Bless, "Automatic tracing of vocal-fold motion from high-speed digital images," *Ieee Transactions on Biomedical Engineering*, vol. 53, no. 7, pp. 1394–1400, 2006.

[28] J. Lohscheller, H. Toy, F. Rosanowski, U. Eysholdt, and M. Dollinger, "Clinically evaluated procedure for the reconstruction of vocal fold vibrations from endoscopic digital high-speed videos," *Medical Image Analysis*, vol. 11, no. 4, pp. 400–413, 2007.

[29] X. L. Qin, S. P. Wang, and M. X. Wan, "Improving reliability and accuracy of vibration parameters of vocal folds based on high-speed video and electroglottography," *IEEE Transactions on Biomedical Engineering*, vol. 56, no. 6, pp. 1744–1754, 2009.

[30] A. Skalski, T. Zielinski, and D. Deliyski, “Analysis of vocal folds movement in high speed videoendoscopy based on level set segmentation and image registration,” *Icses 2008 International Conference on Signals and Electronic Systems, Conference Proceedings*, pp. 223–226 565, 2008.

[31] S. Allin, J. Galeotti, G. Stetten, and S. H. Dailey, “Enhanced snake based segmentation of vocal folds,” *2004 2nd Ieee International Symposium on Biomedical Imaging: Macro to Nano, Vols 1 and 2*, vol. 1 and 2, pp. 812–815 1042, 2004.

[32] S.-Z. Karakozoglou, “Glottal source analysis: a combinatory study using high-speed videoendoscopy and electroglottography,” Master’s thesis, University of Paris-Sud XI, Computer Science Department University of Crete, Computer Science Department LIMSI-CNRS, Audio & Acoustics Group, June 2010.

[33] H. Larsson, S. Hertegard, P. A. Lindestad, and B. Hammarberg, “Vocal fold vibrations: High-speed imaging, kymography, and acoustic analysis: A preliminary report,” *Laryngoscope*, vol. 110, no. 12, pp. 2117–2122, 2000.

[34] J. G. vec and H. K. Schutte, “Videokymography: High-speed line scanning of vocal fold vibration,” *Journal of Voice*, vol. 10, pp. 201–205, June 1996.

[35] J. Lohscheller, U. Eysholdt, H. Toy, and M. Dollinger, “Phonovibrography: Mapping high-speed movies of vocal fold vibrations into 2-d diagrams for visualizing and analyzing the underlying laryngeal dynamics,” *Ieee Transactions on Medical Imaging*, vol. 27, no. 3, pp. 300–309, 2008.

[36] S. Schuberth, U. Hoppe, M. Dollinger, J. Lohscheller, and U. Eysholdt, “High-precision measurement of the vocal fold length and vibratory amplitudes,” *Laryngoscope*, vol. 112, no. 6, pp. 1043–1049, 2002.

[37] H. Larsson and S. Hertegard, “Calibration of high-speed imaging by laser triangulation,” *Logoped Phoniatr Vocol*, vol. 29, no. 4, pp. 154–61, 2004.

[38] M. Schuster, J. Lohscheller, P. Kummer, U. Eysholdt, and U. Hoppe, “Laser projection in high-speed glottography for high-precision measurements of laryngeal dimensions and dynamics,” *European Archives of Oto-Rhino-Laryngology*, vol. 262, no. 6, pp. 477–481, 2005.

[39] J. B. Kobler, D. I. Rosen, J. A. Burns, L. M. Akst, M. S. Broadhurst, S. M. Zeitels, and R. E. Hillman, “Comparison of a flexible laryngoscope with calibrated

sizing function to intraoperative measurements," *Annals of Otology Rhinology and Laryngology*, vol. 115, no. 10, pp. 733–740, 2006.

[40] A. Chan, "Vocal fold vibration measurements using laser doppler vibrometry," Master's thesis, McGill University, 2011.

[41] I. Titze, "Regulating glottal airflow in phonation: Application of the maximum power transfer theorem to a low dimensional phonation model," *Journal of the Acoustical Society of America*, vol. 111, pp. 367–376, JAN 2002.

[42] H. Larsson and S. Hertegard, "Vocal fold dimensions in professional opera singers as measured by means of laser triangulation," *Journal of Voice*, vol. 22, no. 6, pp. 734–739, 2008.

[43] M. Zanartu, L. Mongeau, and G. R. Wodicka, "Influence of acoustic loading on an effective single mass model of the vocal folds," *Journal of the Acoustical Society of America*, vol. 121, no. 2, pp. 1119–1129, 2007.

[44] D. S. Cooper and I. R. Titze, "Generation and dissipation of heat in vocal fold tissue," *Journal of Speech and Hearing Research*, vol. 28, pp. 207–215, 1985.

[45] S. L. Thomson, L. Mongeau, and S. H. Frankel, "Aerodynamic transfer of energy to the vocal folds," *Journal of the Acoustical Society of America*, vol. 118, no. 3, pp. 1689–1700, 2005.

[46] 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," *Journal of the Acoustical Society of America*, vol. 109, no. 4, pp. 1616–1630, 2001.

[47] D. G. Childers, D. M. Hicks, G. P. Moore, and Y. A. Alsaka, "A model for vocal fold vibratory motion, contact area, and the electroglottogram," *Journal of the Acoustical Society of America*, vol. 80, no. 5, pp. 1309–1320, 1986.

[48] K. Verdolini, D. G. Druker, P. M. Palmer, and H. Samawi, "Laryngeal adduction in resonant voice," *Journal of Voice*, vol. 12, no. 3, pp. 315–327, 1998.

[49] N. Henrich, C. d'Alessandro, B. Doval, and M. Castellengo, "On the use of the derivative of electroglottographic signals for characterization of nonpathological phonation," *Journal of the Acoustical Society of America*, vol. 115, no. 3, pp. 1321–1332, 2004.

[50] R. Baken and R. Orlikoff, *Clinical Measurement of Speech and Voice*. Singular Publishing Group, 2000.